Publicly Accessible Penn Dissertations

2010

Parts of a Whole: Distributivity as a Bridge Between Aspect and Measurement

Lucas Champollion
University of Pennsylvania

Parts of a Whole: Distributivity as a Bridge Between Aspect and Measurement

Abstract

Why can I tell you that I *ran for five minutes* but not that I **ran to the store for five minutes*? Why can you say that there are five pounds of books in this package if it contains several books, but not **five pounds of book* if it contains only one? What keeps you from using **sixty degrees of water* to tell me the temperature of the water in your pool when you can use *sixty inches of water* to tell me its height? And what goes wrong when I complain that **all the ants in my kitchen are numerous*?

The constraints on these constructions involve concepts that are generally studied separately: aspect, plural and mass reference, measurement, and distributivity. This work provides a unified perspective on these domains, connects them formally within the framework of mereological semantics, and uses this connection to transfer insights across unrelated bodies of literature. A generalized notion of distributivity is proposed and formalized as a parametrized higher-order property called stratified reference: a predicate that holds of a certain entity or event is required to also hold of its parts along a certain dimension and down to a certain granularity. The dimension parameter is a thematic role in the case of *each* and *all*, a measure function in the case of pseudopartitives, and time or space in the case of *for*-adverbials. The granularity parameter involves pure atoms in the case of *each*, pure and impure atoms in the case of all, and very small amounts of space, time, or matter in the cases of pseudopartitives and *for*-adverbials. Stratified reference is used to formulate a single constraint that explains each of the judgments above. The constraint is exploited to improve on existing characterizations of distributivity, atelicity, and monotonicity of measurement.

The framework results in a new take on the minimal-parts problem that occurs in the study of atelic predicates and mass terms. It scales up successfully from temporal to spatial aspect, and it explains why pseudopartitives and other distributive constructions are sensitive to the difference between intensive and extensive measure functions. It provides a fresh view on atomic and cover-based theories of quantificational distributivity. The framework is also used to account for the scopal behavior of *all* and of *for*-adverbials with respect to cumulative quantification and dependent plurals. Together with a novel theory of collective predication, the framework also provides an account of the differences between such predicates as *be numerous* and *gather* as they interact with *all*.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Linguistics

First Advisor
Cleo Condoravdi

Keywords
philosophy, religion and theology, language, literature and linguistics, distributivity, collectivity, telicity, pseudopartitives, mereology, dependent plurals

Subject Categories

Linguistics | Philosophy of Language

PARTS OF A WHOLE:

DISTRIBUTIVITY AS A BRIDGE BETWEEN ASPECT AND MEASUREMENT

Lucas Champollion

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2010

Cleo Condoravdi, Supervisor of Dissertation
Consulting Associate Professor of Linguistics, Stanford University

Eugene Buckley, Graduate Group Chair
Associate Professor of Linguistics

Committee Chair:

Aravind K. Joshi, Henry Salvatori Professor of Computer and Cognitive Science

Committee Members:

Maribel Romero, Professor of Linguistics, University of Konstanz
Florian Schwarz, Assistant Professor of Linguistics

PARTS OF A WHOLE:
DISTRIBUTIVITY AS A BRIDGE BETWEEN ASPECT AND MEASUREMENT

COPYRIGHT

2010

Lucas Champollion

The text in this document has been purged of a number of errata that were discovered after a hard copy of this dissertation was submitted to the University of Pennsylvania collections. Both the hard copy and this electronic version are set in single-spaced format, so it is possible to use either version for citation purposes.

The author can be reached at `champoll@gmail.com`. Any feedback on this work will be greatly appreciated.

Acknowledgements

shapes my own work more than I know. Like everyone else, I've looked on in amazement as she changed jobs, moved across the ocean, founded a family, became department chair, and kept an eye on my dissertation all at once. I thank her for her generosity and for her help as a committee member. There's no way to sneak a flawed linguistic argument past her, though I'm sure this work contains many that she has given up on trying to correct.

Florian Schwarz, who joined Penn and my committee only a short while ago, worked his way into the dissertation with remarkable speed and provided numerous detailed comments. I thank him for joining the committee without hesitation and for his enthusiastic approach to the enterprise. It was a great pleasure to have him on the committee. I have to be careful not to sing his praises too much, though, because already more of my fellow students seem to be eager to work with him than is really good for him.

Tony Kroch served on the dissertation committee for a good year or so. I'm grateful that he did, and I'm glad he stayed around even after he officially left. I thank him for giving me the benefit of the doubt early on, for encouraging me to be bold, and for not thinking that his efforts were lost on me. Tony and myself are easily the two most stubborn people I know in our department. He never made a secret of his opinion when he thought I was headed the wrong way, and given my personality, I think he deserves credit for that.

Once upon a time I was warned that graduate school and especially writing a dissertation can be a lonely business. Luckily, I found out that it doesn't have to be that way. I have a lot of people to thank for that, and for teaching me linguistics, engaging with my ideas, and making me feel at home at every station of my way. It would be impossible to thank everyone, and if you feel unjustly omitted from the following list, you most likely are.

The late Ellen Prince took me seriously when I was still green behind the ears, and I hear she put in a good word for me when I wanted to join Penn. I regret not having spent more time with her.

My Master's degree advisor, Mitch Marcus, has been supportive throughout my time at Penn, and it is largely thanks to his and Aravind's efforts that I could navigate the gap between linguistics and computer science.

When I arrived at Penn, my fellow first-years Keelan Evanini, Eva Florencio Nieto, Jonathan Gress-Wright, Lucy Lee, Laia Mayol, Jean-François Mondon, and Josh Tauberer plowed through the endless linguistics coursework with me, endured my island violations, and helped me laugh at it all from time to time. When I entered the Ph.D. program a year later, I got a bunch of new good friends and fellow first-years. Aviad Eilam and Catherine Lai have never lost patience with my lack of appetite for gossip and with my unending hunger for grammaticality judgments. In that category, I also have to thank Josh Tauberer again, whose intuitions about

v

Note. In the electronic version of this document, the entries in the table of contents and in the lists of tables and figures are hyperlinks that can be used to jump to the appropriate sections. In the running text, references to chapters, sections, examples, and bibliography entries also link to their respective sources.

ABSTRACT

PARTS OF A WHOLE:

DISTRIBUTIVITY AS A BRIDGE BETWEEN ASPECT AND MEASUREMENT

Lucas Champollion

Supervisor: Cleo Condoravdi

Why can I tell you that I *ran for five minutes* but not that I **ran to the store for five minutes*? Why can you say that there are *five pounds of books* in this package if it contains several books, but not **five pounds of book* if it contains only one? What keeps you from using **sixty degrees of water* to tell me the temperature of the water in your pool when you can use *sixty inches of water* to tell me its height? And what goes wrong when I complain that **all the ants in my kitchen are numerous*?

The constraints on these constructions involve concepts that are generally studied separately: aspect, plural and mass reference, measurement, and distributivity. This work provides a unified perspective on these domains, connects them formally within the framework of mereological semantics, and uses this connection to transfer insights across unrelated bodies of literature. A generalized notion of distributivity is proposed and formalized as a parametrized higher-order property called stratified reference: a predicate that holds of a certain entity or event is required to also hold of its parts along a certain dimension and down to a certain granularity. The dimension parameter is a thematic role in the case of *each* and *all*, a measure function in the case of pseudopartitives, and time or space in the case of *for*-adverbials. The granularity parameter involves pure atoms in the case of *each*, pure and impure atoms in the case of *all*, and very small amounts of space, time, or matter in the cases of pseudopartitives and *for*-adverbials. Stratified reference is used to formulate a single constraint that explains each of the judgments above. The constraint is exploited to improve on existing characterizations of distributivity, atelicity, and monotonicity of measurement.

The framework results in a new take on the minimal-parts problem that occurs in the study of atelic predicates and mass terms. It scales up successfully from temporal to spatial aspect, and it explains why pseudopartitives and other distributive constructions are sensitive to the difference between intensive and extensive measure functions. It provides a fresh view on atomic and cover-based theories of quantificational distributivity. The framework is also used to account for the scopal behavior of *all* and of *for*-adverbials with respect to cumulative quantification and dependent plurals. Together with a novel theory of collective predication, the framework also provides an account of the differences between such predicates as *be numerous* and *gather* as they interact with *all*.

ix

Contents

List of Tables

List of Figures

Chapter 1

Overview

The central claim of this work is that a unified theory of distributivity, aspect, and measurement for natural language is feasible and useful.

1.1   Introduction

I claim that a number of natural language phenomena from the domains of aspect,[1] plurality, cumulativity, distributivity and measurement, which are currently treated by separate theories, are in fact intimately related. Previous accounts of these phenomena either fail to generalize appropriately, or live on as limiting cases of a system presented here under the name of *strata theory*. This system is not a radical reorientation of the grammar. By subsuming and building on previous characterizations, strata theory retains much of what has been formerly gained, and provides a unified framework in which new correspondences are drawn between existing concepts.

The road to this claim starts with four semantic oppositions which are closely associated with the domains under consideration. These are the telic-atelic opposition, which is central to the study of aspect; the singular-plural opposition and the count-mass opposition, which are central to the study of plurality and measurement; and the collective-distributive opposition, which is central to the study of distributivity. These oppositions can be formally related to one another. This, in itself, is not a new insight. It has long been known that there are close

---

[1]*Aspect* is used in the literature to refer to many different things. Throughout this work, I use the term to refer to what has been variously called inner aspect, lexical aspect, temporal constitution, actionality, or aktionsart, as opposed to the phenomenon referred to as outer aspect, grammatical aspect, or viewpoint aspect. Broadly speaking, I understand inner aspect as referring to the telic-atelic opposition, and outer aspect as referring to the imperfective-perfective opposition. Outer aspect is not discussed in this work.

parallels between the singular-plural and the count-mass opposition (e.g. Link 1983) and, likewise, between the count-mass and the telic-atelic opposition (e.g. Bach 1986). That these formal parallels can be extended to encompass the collective-distributive opposition has not been explicitly mentioned as far as I know, but it is not difficult to do so.

The nature of the parallelism between all these oppositions can be described intuitively in terms of boundedness. Singular, telic, and collective predicates are delimited or bounded in ways that plural, mass, atelic, and distributive predicates are not. Making formal sense of the parallelism therefore amounts to characterizing the difference between boundedness and unboundedness. How to do this is one of the central questions which strata theory proposes to answer. I call it the *boundedness question.*

Answering the boundedness question amounts to specifying what it means for a predicate to be atelic, distributive, plural, or to have mass reference. It is not obvious that there should be a single property that is shared by all these predicates. As this work shows, however, it is indeed possible to isolate such a property. The identity of this property can be determined by analyzing a number of nominal and verbal constructions which all have one thing in common: each of them is sensitive to one of the semantic oppositions listed above. These constructions are *for*-adverbials, which distinguish atelic from telic predicates (1); pseudopartitives, which distinguish plurals and mass nouns from singular count nouns (2); and adverbial *each*, which distinguishes distributive from collective predicates (3). I refer to them collectively as *distributive constructions.*

(1)   a.   John **ran** for five minutes.   *atelic*
      b.   *John **ran to the store** for five minutes.   **telic*

(2)   a.   thirty pounds of **books**   *plural*
      b.   thirty liters of **water**   *mass*
      c.   *thirty pounds of **book**   **singular*

(3)   a.   The boys each **walked**.   *distributive*
      b.   *The boys each **met**.   **collective*

These three constructions form the empirical basis of this work. However, they probably represent only a small sample of distributive constructions. For example, true partitives and comparative determiners accept the same classes of nouns and of measure functions as pseudopartitives do (Schwarzschild 2006). For present purposes, it is enough to focus on the three distributive constructions above, firstly, because they cut across the domains of distributivity, aspect and measurement, and secondly because each of them is regarded as central to its domain in the sense that any theory of it must account for its behavior. More concretely, *for*-adverbials

are regarded as the prime diagnostic of atelicity (Verkuyl 1989); *each* is the standard example of a distributive item (Link 1987b); and pseudopartitives are arguably the most prominent place in which natural language shows its sensitivity to formal properties of measurement (Krifka 1998; Schwarzschild 2006).

The novel angle of this work consists in considering the constructions in (1) through (3) as parts of a whole. Previous work has produced separate theories to account for the behavior of each of these constructions and for the phenomena that they exemplify. The resulting theories are often more limited in scope than they could be. For example, work on distributivity has focused on how best to formalize distributive readings, rather than on extending the notion of distributivity. Likewise, the study of aspect has concentrated entirely on temporal phenomena, and the study of measurement in natural language has focused largely on mass terms, partitives, and comparatives. This development has obscured the view on the common properties of these constructions. However, this problem is not inherent in the approaches encoded in these theories. Once the connection between distributivity, aspect, and measurement is made formally explicit, it is easy to connect many existing theories to each other, and to extend them to domains beyond the ones in which they have traditionally been applied. One can then combine the strengths of each account, and synthesize them to extend their empirical coverage. This is the motivation behind the present work.

The presence of distributive constructions in every one of the domains of interest makes it possible to place strata theory on a solid empirical foundation, because these constructions allow us to operationalize the boundedness question. Instead of asking abstractly what it is that atelic and distributive and mass and plural predicates have in common with each other, we can search for the property that the bold constituents in the grammatical examples in (1a), (2a), (2b) and (3a) have in common, to the exclusion of the ungrammatical examples in (1b), (2c), and (3b).

The rest of this chapter outlines the intuition behind strata theory and gives an overview of things to come.

1.2 The central metaphor

The guiding idea behind this work is that the constructions illustrated in (1) through (3) exclude bounded predicates through a parametrized constraint which is introduced into distributive constructions through certain words such as *for*, *of*, and *each*. This constraint is formulated in terms of a higher-order property, *stratified reference*. This property requires a predicate that holds of a certain entity or event to also hold of its parts along a certain dimension and down to a certain granular-

ity. Dimension and granularity are understood as parameters which distributive constructions can set to different values.

The dimension parameter specifies the way in which the predicate in question is distributed. Different settings of this parameter allow one and the same predicate to be atelic but not distributive, or vice versa. When the dimension parameter is set to time, stratified reference applies to atelic predicates, as in (1). When it is set to a measure function like weight or volume, stratified reference applies to mass and plural predicates, as in (2). When it is set to a thematic role like agent, stratified reference applies to distributive predicates, as in (3).

The granularity parameter specifies that the parts in question must be either atomic or very small in size, as measured along the dimension. This parameter accounts for the differences between distributive constructions over discrete (count) domains, such as adverbial-*each* constructions, and those over domains involving continuous dimensions, such as *for*-adverbials and pseudopartitives.

The names *dimension*, *granularity*, and *stratified reference* are derived from a visual metaphor, which I develop here. I stress that I use this metaphor only for the purpose of conveying the intuitions behind strata theory. It does not have any formal status, it does not occur in the formulation of the theory, and it is not claimed to have any psychological or cognitive reality – unlike, for example, the diagrams in the cognitive grammar literature (Langacker 1986).

The metaphor is based on the idea that individuals, substances, and events occupy regions in an abstract space. The dimensions of this space include the familiar spatial and temporal dimensions as well as any measure functions and thematic roles that happen to be defined for the entity. (To understand a thematic role as a dimension, we assume that the individuals that correspond to these roles are ordered in an arbitrary but fixed canonical order, such as the alphabetical order given by their first and last names.) An object whose weight is large corresponds to a region with a large extent along the weight dimension. An event whose agent is a plural entity corresponds to a region with a large extent along the agent dimension, while an event whose agent is singular corresponds to a region which is not extended along the agent dimension at all. A temporally and spatially punctual event whose thematic roles are all singular entities corresponds to a point. A temporally and spatially punctual event that has plural entities as its agent and theme corresponds to an infinitely thin rectangle that is extended along the agent and theme dimensions.

Consider the old intuition that any atelic predicate has the subinterval property (Bennett and Partee 1972). This property says that whenever a predicate holds at an interval t, it also holds at every subinterval of t, all the way down to instants. Put in event semantic terms, a predicate like *run* is atelic because we can "zoom in" to any temporal part of a running event to find another running event. We cannot

do that with a telic predicate like *run to the store*, because the initial half of an event of running to the store does not itself qualify as running to the store. In the metaphor, the subinterval property translates to the following picture: any event in the denotation of a predicate that has the subinterval property can be divided into infinitely thin layers that run perpendicular to the time dimension and that are also in the denotation of this atelic predicate. This gives rise to the well-known "minimal-parts problem": strictly speaking, there are no instantaneous running events, for example. If the subinterval property is to have any viable chance, it must therefore be amended so that the event layers are constrained to be very thin, but do not have to be infinitely thin. Formally, this effect is achieved by adding a granularity parameter to the subinterval property and constraining this parameter to a low but nonatomic value. I call these layers *strata*. This name is chosen to remind the reader of geological strata, the layers of rock which can be observed in geological formations in places such as the Grand Canyon. A geological stratum can be just a few inches thick (though not infinitely thin) and extend over hundreds of thousands of square miles. This aspect is mirrored in the theory, where strata are constrained to be very thin along one dimension, but may be arbitrarily large as measured in any other dimension.

The metaphor I have used to describe the subinterval property involves layers or strata rather than points or pebbles, because the subinterval property does not constrain any dimensions other than time. This feature is not accidental. While the relevant parts of running events must be short, or thin, in the temporal dimension, they may have plural entities as agents or themes, they may be extended in space, and so on. This view leads to a natural generalization. Normally, geological strata are horizontal, but due to geological movement, they can also be oriented along another dimension. For example, they can run vertically. Similarly, I have introduced the concept of temporal strata as resulting from dividing an event along the temporal dimension, but we can also imagine spatial or "agental" strata – subevents that are constrained based on their spatial extent or based on their number of agents. Once this step is taken, the atelic-telic opposition can be related to the collective-distributive opposition in a Neo-Davidsonian setting. Distributive predicates require any event in their denotation to be divisible into strata that are constrained to have atomic thickness on the dimension of the appropriate thematic role. For example, any plural event in the denotation of a predicate like *smile* or *read a book* must be divisible into strata that have atomic agents and that belong to the denotation of the same predicate. Lexical predicates like *smile* have this property due to world knowledge, and phrasal predicates like *read a book* can acquire it through a modified version of the distributivity operators known from Link (1987b) and Schwarzschild (1996). Collective predicates like *be numerous* do not satisfy stratified reference on the thematic role of their subjects, because their

subjects can be plural entities whose parts are not themselves numerous.

1.3   Overview of things to come

Chapter 2, *The stage*, presents the framework on which strata theory builds. This framework builds on the work by Link (1998), Krifka (1998), Landman (2000), Lønning (1987), Zweig (2008), and others. Its mathematical foundation is classical extensional mereology, which is presented and discussed at length.

Chapter 3, *The cast of characters*, presents the three constructions that form the empirical basis for most of this work: adverbial-*each* constructions, *for*-adverbials, and pseudopartitives. Based on the background assumptions of Chapter 2, I provide skeletal LFs for these constructions.

Chapter 4, *The theory*, presents my answer to the boundedness question. The parallelism between the telic-atelic, collective-distributive, singular-plural, and count-mass oppositions is captured in a unified framework. The metaphor just presented is formalized through the higher-order property of stratified reference. This property is parametrized, which reflects the fact that unboundedness can be understood in more than one way. As the example *run to the store* shows, one and the same verb phrase can be distributive (unbounded with respect to agents) and telic (bounded with respect to runtime). As the example *kill* shows, one and the same verb can be collective on its agent position (bounded with respect to agents) and distributive on its theme position (unbounded with respect to themes). As the example *cable* shows, one and the same predicate can be divisive along its length (bounded with respect to length) and nondivisive along its diameter (unbounded with respect to diameter). Chapter 4 discusses these examples in detail.

Chapter 5, *Minimal parts*, considers the effect of the granularity parameter on the thickness of the strata. If the subinterval property is used as a model of atelicity, even infinitely thin strata (that is, strata which take place at instants) must be checked whether they belong to the denotation of the atelic predicate. I defend stratified reference as an appropriate modification of the subinterval property and I compare it with previous approaches to the minimal-parts problem.

Chapter 6, *Aspect and space*, extends the telic-atelic opposition to the spatial domain. This makes it possible to extend theories of aspect from temporal to spatial phenomena. I show that strata theory, in which the size of the subevents is constrained along the temporal or spatial dimension, is superior to an alternative view based on divisive reference, in which the size of the subevents is constrained along every dimension (Krifka 1998). Strata theory regards time and space as different settings of a parameter. This fact is used to account for differences in the distribution of temporal *for*-adverbials and their spatial counterparts.

Chapter 7, *Measure functions*, takes up the formal properties of measure functions and considers why one can say *thirty liters of water* but not **thirty degrees Celsius of water*. I show that strata theory subsumes and extends previous accounts of this phenomenon. A plural or mass entity to which a pseudopartitive applies (for example some water, or some books) is divided into strata which are very small as measured in the dimension determined by the pseudopartitive, but may extend arbitrarily in other dimensions. These strata are then required to be in the denotation of the noun. Singular count nouns always fail this test because the individuals in their denotation are atomic, and cannot be further subdivided into strata. Using this view, I explain why distributive constructions only allow extensive measure functions like volume, weight, or runtime, but not intensive ones like temperature or speed. For intensive measure functions like temperature, the parts of the entity generally have the same value as the whole entity, so it is not possible to find appropriate strata. The resulting theory is compared with previous accounts by Krifka (1998) and Schwarzschild (2006).

Chapter 8, *Distributivity and scope*, considers how verb phrases such as *build a sand castle* acquire their distributive interpretations, and integrates the D operator of Link (1987b) into strata theory. I compare atomic and nonatomic views of phrasal distributivity. I extend ideas from Schwarzschild (2006) into the nonatomic domain of time, and I study the interactions of *for*-adverbials with a temporal equivalent of his cover-based D operator.

Chapter 9, *For, each, and all*, brings together a number of threads. Distributivity is considered in connection with collectivity and cumulativity. Both *for*-adverbials and *all* restrict the availability of cumulative readings, and both license dependent plurals. To explain these facts, the word *all* is treated as a distributive item analogous to *for*. This is problematic at first sight, because *all* behaves differently from other distributive items like *each* and does not always give rise to distributive entailments (*All the boys gathered*). Given a novel theory of collectivity, strata theory allows us to trace the difference between *all* and *each* to different settings of the granularity parameter. The theory also explains why it is not possible to say **All the boys are numerous* on either a distributive or a collective interpretation. The resulting theory is compared with the account by Winter (2001).

Chapter 10 concludes with some suggestions for further research.

Chapter 2

The stage

2.1 Introduction

Since this work touches on a number of different semantic domains, it relies
on many background assumptions. This chapter presents and motivates these
assumptions. Readers who are familiar with the literature and who want to follow
the main narrative of this work are encouraged to skip or skim this chapter. It
is probably best not to read it from beginning to end, but to use it as a reference.
For this purpose, the following chapters include many backreferences to specific
sections in this one.

The textbooks by Partee, ter Meulen, and Wall (1990) and by Landman (1991)
provide introductions to many of the mathematical concepts described here. The
theory of the syntax-semantics interface follows standard assumptions of genera-
tive grammar and builds on the textbook by Heim and Kratzer (1998). To the extent
that it is not already contained in that textbook, the formal framework presented
in this chapter mainly builds on Krifka (1998), Link (1998), Landman (2000), and the
papers leading up to them. In most cases, I adopt the standard position when there
is one, and I state my choice explicitly when no standard position has evolved yet.

2.2 Notational conventions

I use the following typing conventions: t for propositions, e for ordinary objects
(meaning individuals and substances, Section 2.4.1), v for events (which include
states, Section 2.4.3), i for intervals (temporal and spatial entities, Section 2.4.4),
d for degrees (Section 2.4.5), and n for numbers (Section 2.4.6). I use the symbols
x, y, z, x', y', z' and so on for variables that range over ordinary objects, e, e', e'' for
events, t, t', t'' for intervals, n, n', n'' for numbers, and b, b', b'' for entities of any

basic type. I use P for predicates of type $\langle et \rangle$, V for predicates of type $\langle vt \rangle$, and f for functions of various types.

I assume that at least individuals, substances, events, and intervals are closed under mereological sum formation (Section 2.3.1). Intuitively, this means that these categories include plural entities. The lowercase variables just mentioned should therefore be taken to range over both singular and plural entities. In the literature on plurals, the distinction between singular and plural entities is often indicated by lowercase and uppercase variables. Since almost all the variables in my representations range over potentially plural entities, I do not follow this convention.

For acceptability judgments, I use the * (star) sign as indicating either syntactically or semantically unacceptable sentences. Some authors reserve * for syntactically unacceptable sentences and use another sign, such as # (hash), for semantically unacceptable ones. I do not follow this convention because most of the unacceptable sentences I mention are semantically rather than syntactically unacceptable. I use ? and ?? to indicate that a sentence has borderline status. Occasionally I give judgments for constituents which are not entire sentences. These judgments should be taken as indicating the status of appropriate sentences in which the constituent is used.

2.3 Mereology

This section provides a general introduction into the conceptual and mathematical underpinnings of mereology, defines the system known as classical extensional mereology, and relates it to set theory.

2.3.1 Foundations of algebraic semantics

In standard Montague semantics, the domain of discourse (the collection of things from which the denotations of words and larger constituents are built) is simply a collection of disjoint nonempty sets. Many model-theoretic accounts follow this tradition and assume no more than a set of individuals and a set containing the truth values; others also add sets representing such entities as events or possible worlds. Although this type of setup is very simple, it is a solid basis for formalizing the semantics of large areas of natural language. The textbook by Heim and Kratzer (1998) illustrates this approach.

Expressions like *John and Mary* or *the water in my cup* intuitively involve reference to collections of individuals or substances. The relation between these collections and their components is not modeled in standard Montague semantics,

Table 2.1: Some examples of wholes and parts

Whole	Part
some horses	a subset of them
a quantity of water	a portion of it
John, Mary and Bill	John
some jumping events	a subset of them
a running event from A to B	its part from A halfway towards B
a temporal interval	its initial half
a spatial interval	its northern half

which does not easily lend itself to capturing entailment relations such as the following:

(1)   a.   John and Mary sleep. ⇒
          John sleeps and Mary sleeps.
      b.   The water in my cup evaporated. ⇒
          The water at the bottom of my cup evaporated.

To formalize entailment relations like these ones, Link (1983, 1998) introduces the framework of *algebraic semantics*, which is also adopted here. In algebraic semantics, the domain of individuals also includes plural and mass entities. These entities are assumed to stand in a relation $\leq$, called *parthood*. This relation is added directly to the logic and model theory, and its properties are described by axioms.

Several extensions of algebraic semantics have been proposed to include other ontological domains such as intervals and events (Hinrichs 1985; Krifka 1986, 1998; Link 1987a; Landman 2000). I discuss my ontological assumptions concerning these entities in Section 2.4. This section focuses on the axiomatic underpinnings of algebraic semantics. Unlike for example the relation $\in$ in set theory, which is conceptually and mathematically well understood and for which a standard set of axioms is generally accepted, there is no consensus on the exact properties of the concept of parthood that $\leq$ expresses.

The philosophical theory of parthood relations is called *mereology*, and algebraic semantics is also called *mereological semantics*. Both in the philosophical and in the semantic literature, there is no consensus on which general principles should be taken to constrain the meaning of expressions like $a \leq b$, much less what they are supposed to capture conceptually. In this work, I take the position that the relation

$\leq$ is no more than a formal device in a semanticist's theory-building toolbox, a part of the metalanguage which should not be seen as subject to any constraints other than the ones the theory explicitly places on it. To help build intuitions, Table 2.1 gives a few examples of what I intend the relation $\leq$ to express.

The examples in this table have in common that there is a sense in which the "wholes" are collections without internal structure; they consist exactly of a bunch of things in the category of "parts", and their properties can be induced from these parts. In mereology, this concept is called *sum*. I call this parthood relation unstructured parthood, as opposed to the relation that holds between structured "wholes" and their "parts", which I call structured parthood. Examples of structured parthood are the relation between a broom and its handle, a table and its leg, an army and one of its soldiers, a ham sandwich and its ham slice, or a cat and its tail. These relations between structured wholes and their parts are qualitatively different from the relations between unstructured collections and their parts shown in Table 2.1. As an example of a structured whole, a ham sandwich is more than a piece of ham and two slices of bread put together, because its properties cannot be induced from those of its ingredients (Fine 1999). I am not claiming that it is easy to draw the precise boundary between structured and unstructured parthood. At this point I can only offer an intuition. Unstructured parthood corresponds to the relation between a plural or mass entity and the singular and plural entities that are its parts, whereas structured parthood corresponds to the relation between a singular entity and its parts. For discussion on related philosophical issues, see Simons (1987), Varzi (2010) and references therein.[2]

I use $\leq$ to model only unstructured parthood. I believe that this decision is in line with the mereological semantic literature. Whether or not natural language uses the word *part* for some relation, I do not consider this to be evidence for the nature of $\leq$. The relation $\leq$ should not be constrained by whatever is regarded as the meaning of the natural-language expression "part of", although it is obviously inspired by it. Not all semanticists agree on this point; see for example Moltmann (1997, 1998) for an opposing view, and Pianesi (2002) and Varzi (2006) for responses. However, there is an analogous practice in standard Montague semantics, where the relation $\in$ is not assumed to be constrained by anything other than the standard axioms of set theory. In particular, the relation $\in$ is not assumed to represent or even to be constrained in any way by the meaning of natural language expressions like "element of" or "member of". In algebraic semantics, the $\leq$ relation is fundamental in the same way as $\in$ is to set theory based approaches to semantics. Constraining it to conform to a natural-language expression with such a wide range of meanings as "part of" would make it more difficult to build

---

[2]I thank Alexander Williams for discussing this issue with me.

and evaluate theories such as this one. The natural language expression "part of" ranges over a wide range of notions. Varzi (2010) distinguishes the following concepts, for example:

(2)   a.   Spatial inclusion: The handle is part of the broom.
       b.   Temporal inclusion: The first act was the best part of the play.
       c.   Conceptual inclusion: Writing detailed comments is part of being a good referee.
       d.   Mixture composition: Gin is part of martini.

To make sure that the formal relation $\leq$ reflects the conceptual properties of the relevant notion of parthood, a number of axioms are imposed on it. Mereology does not have a standard axiom system. It is therefore important for any work that relies on mereology to specify which mereology is meant.

I adopt *classical extensional mereology* (CEM), a system which is probably the most widely used mereological system in philosophy and linguistics, although the linguistic literature is not always explicit about this. I now present CEM and lay out some of the intuitions behind the use of mereology for semantics. The discussion in this section is partly based on the excellent surveys in Simons (1987), Casati and Varzi (1999) and Varzi (2010).

One way of defining CEM is by taking parthood as a primitive relation and formulating axioms that impose constraints on it. The following axioms constrain parthood to be a partial order:

(3)   **Parthood (primitive relation)**
      $x \leq y$
      (x is part of y.)

(4)   **Axiom of reflexivity**
      $\forall x[x \leq x]$
      (Everything is part of itself.)

(5)   **Axiom of transitivity**
      $\forall x \forall y \forall z[x \leq y \ \wedge \ y \leq z \rightarrow x \leq z]$
      (Any part of any part of a thing is itself part of that thing.)

(6)   **Axiom of antisymmetry**
      $\forall x \forall y[x \leq y \ \wedge \ y \leq x \rightarrow x = y]$
      (Two distinct things cannot both be part of each other.)

These requirements are so basic that they are shared not only by CEM but by almost all mereological theories. Despite this, each of them has been criticized at times. Moltmann (1997, 1998) emphasizes that transitivity does not match the

natural language use of the *part of* construction. A similar point can of course be made for reflexivity. For example, (7) and (8) are odd, but they should be true if *part of* obeyed reflexivity and transitivity, respectively.

(7)   a.   This knob is part of this door, this door is part of the doors in my house, so this knob is part of the doors in my house.
      b.   John is part of himself.

These facts do not call the use of CEM into question since, as argued above, we do not want to model the everyday usage of the English *part of* construction. My assumptions are not consistent with the relation expressed in (7) because I assume that the elements in the denotations of singular count nouns like *knob*, *door*, *house* and the referents of proper names like *John* are *atomic*: they do not have any parts besides themselves (see Section 2.6). Reflexivity is imposed on the parthood relation mainly for technical convenience. We can define an irreflexive *proper-part* relation by restricting parthood to nonequal pairs:

(8)   **Definition: Proper part**
      $x < y \overset{\text{def}}{=} x \leq y \ \wedge \ x \neq y$
      (A proper part of a thing is a part of it that is distinct from it.)

Mereologies diverge in the way they formalize the concept of sum. It is important to choose the definition carefully to avoid Russell's paradox (Schein 1993, 2006; Link 1998, ch. 13) and to make sure that the axiom that uses it defines CEM as intended (Hovda 2009). I adopt the classical definition of sum in (9), which is due to Tarski (1929).

(9)   **Definition: Sum**
      $\text{sum}(x, P) \overset{\text{def}}{=} \forall y[P(y) \rightarrow y \leq x] \ \wedge \ \forall z[z \leq x \rightarrow \exists z'[P(z') \ \wedge \ z \otimes z']]$
      (A sum of a set P is a thing that consists of everything in P and whose parts each overlap with something in P.)

This definition uses the auxiliary concept of overlap, defined as in (10).

(10) **Definition: Overlap**
      $x \otimes y \overset{\text{def}}{=} \exists z[z \leq x \ \wedge \ z \leq y]$
      (Two things overlap if and only if they have a part in common.)

The definition in (9) can be understood as quantifying over all sets P, as suggested by my paraphrase, or as an axiom schema in each of whose axioms P is instantiated by a different formula. The difference between the two options bears on the completeness of the resulting Boolean algebras; the set interpretation places

the theory beyond first-order logic (Pontow and Schubert 2006; Varzi 2010).

The formulation of the definition reflects the intuitive fact that a sum may have other parts than just its immediate components. For example, the sum of (i) the referent of the conjoined term *John and Mary* and (ii) the referent of the proper name *Bill* has more parts than these two individuals: for example, it also has the referent of *John* among its parts, as well as the referent of *Bill and Mary*.

The following facts can be easily shown to follow from these definitions:

(11) **Fact**
$\forall x \forall y [x \le y \rightarrow x \otimes y]$
(Parthood is a special case of overlap.)

(12) **Fact**
$\forall x [\text{sum}(x, \{x\})]$
(A singleton set sums up to its only member.)

Different mereologies disagree on what kinds of collections have a sum, and whether it is possible for one and the same collection to have more than one sum. In classical extensional mereology, sums are unique, therefore two things composed of the same parts are identical. This is expressed by the following axiom:

(13) **Axiom of uniqueness of sums**
$\forall P [P \ne \varnothing \rightarrow \exists! z \ \text{sum}(z, P)]$
(Every nonempty set has a unique sum.)

The binary and generalized sum operators in (14) and (15) give us a way to refer explicitly to the sum of two things, and to the sum of an arbitrary set.

(14) **Definition: Binary sum**
$x \oplus y$ is defined as $\iota z \ \text{sum}(z, \{x, y\})$.
(The sum of two things is the thing which contains both of them and whose parts each overlap with one of them.)

(15) **Definition: Generalized sum**
For any nonempty set P, its sum $\bigoplus P$ is defined as $\iota z \ \text{sum}(z, P)$.
(The sum of a set P is the thing which contains every element of P and whose parts each overlap with an element of P.)

For example, we can write the plural individual denoted by the conjoined term *John and Mary* as "$j \oplus m$", and to the sum of all water as "$\bigoplus$ water". Since we can refer to this sum as *the water*, the $\bigoplus$ operator can be thought of as a definite description (see Section 2.8). In the semantic literature following Sharvy (1980), we often find the equivalent notation $\sigma x \ \text{P}(x)$ rather than $\bigoplus \text{P}$. I use $\bigoplus$ to make the

connection between binary and generalized sum clearer, and to avoid a name clash with the spatial trace function, which I also write σ (Section 2.5.2).

The pointwise sum operator (16) allows us to construct the sum of a relation. Its main use in this work is as an auxiliary concept in subsequent definitions.

(16) **Definition: Generalized pointwise sum**
For any nonempty n-place relation R_n, its sum $\bigoplus R_n$ is defined as the tuple $\langle z_1, \ldots, z_n \rangle$ such that each z_i is equal to
$\bigoplus \{x_i | \exists x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n[R(x_1, \ldots, x_n)]\}$.
(The sum of a relation R is the pointwise sum of its positions.)

Algebraic closure (17) extends a predicate P so that whenever it applies to a set of things individually, it also applies to their sum. Algebraic closure was initially introduced in the semantic literature in Link (1983) to describe the meaning of plural formation (see Section 2.6.2) and is also known as the "star operator".

(17) **Definition: Algebraic closure (Link 1983)**
The algebraic closure $^*P$ of a set P is defined as $\{x \mid \exists P' \subseteq P[x = \bigoplus P']\}$.
(The algebraic closure of a set P is the set that contains any sum of things taken from P.)

Note that the definition in (17) implies that $P' \neq \varnothing$ because $\bigoplus$ is only defined on nonempty sets. Similar provisos apply to the following definitions.

The following theorem clarifies the effect of algebraic closure:

(18) **Theorem**
$\forall P[P \subseteq {}^*P]$
(The algebraic closure of a set always contains that set.)

Proof: We need to show that $\forall P[P \subseteq \{x \mid \exists P' \subseteq P[x = \bigoplus P']\}]$, or equivalently, $\forall P \forall x[x \in P \rightarrow \exists P' \subseteq P[\text{sum}(x, P')]]$. This follows for $P' = \{x\}$, given Fact (12). End of proof.

Like the sum operation, algebraic closure can be easily extended to predicates of arbitrary arity (Vaillette 2001). The following definition is adapted from the concept of summativity in Krifka (1986, 1998). I use the symbol $\vec{x}$ to range over sequences. The "double-star operator" from the literature on cumulative readings corresponds to the special case for two-place relations.

(19) **Definition: Algebraic closure for relations**
The algebraic closure $^*R$ of a non-functional relation R is defined as
$\{\vec{x} \mid \exists R' \subseteq R[x = \bigoplus R']\}$
(The algebraic closure of a relation R is the relation that contains any sum

of tuples each contained in R.)

Finally, the following definition extends algebraic closure to partial functions. While functions are technically relations, it is still useful to treat them separately, because applying the previous definition to a function does not always yield a function. For example, let f_0 be the function $\{\langle x, a \rangle, \langle x \oplus y, b \rangle\}$. Applying definition (19) to it yields the relation $\{\langle x, a \rangle, \langle x \oplus y, b \rangle, \langle x \oplus y, a \oplus b \rangle\}$, which is not a function because $x \oplus y$ is related to two distinct values.

The following definition makes sure that the algebraic closure of a partial function is always a partial function. For the notation $\lambda x : \varphi \, . \, \psi$, see Section 4.7.

(20)   **Definition: Algebraic closure for partial functions**
The algebraic closure ${}^*f$ of a partial function f is defined as
$\lambda x : x \in {}^*\mathrm{dom}(f) . \bigoplus \{y | \exists z [z \leq x \ \wedge \ y = f(z)]\}$
(The algebraic closure of f is the partial function that maps any sum of things each contained in the domain of f to the sum of their values.)

To use the example from above, according to this definition, ${}^*f_0 = \{\langle x, a \rangle, \langle x \oplus y, a \oplus b \rangle\}$, which is a partial function.

Other definitions of these concepts exist based on the notion of closure under binary sum formation. For example, an alternative to (17) defines ${}^*P$ as the smallest set P' such that $P \subseteq P'$ and for any two members of P' their sum is also in P'. These definitions are equivalent when P has finite cardinality, but this assumption is not always warranted.[3] Take the set $\{x | x = x\}$, the set of all things. The axioms of mereology prohibit cycles in the parthood relation. For this reason, if there are only finitely many things, then the domain is *atomic*: every entity is the sum of a set of things which have no proper parts, as defined in (21). However, we do not want to restrict ourselves to atomic domains when we describe spatial and temporal intervals (see Section 2.4.4).

(21)   **Definition: Atom**
$\mathrm{Atom}(x) \stackrel{\mathrm{def}}{=} \neg \exists y [y < x]$
(An atom is something which has no proper parts.)

By themselves, the axioms of CEM do not specify whether atoms exist or not: among the models they describe, there are some in which everything is made up of atoms (in particular, this includes all models that contain finitely many elements), some in which there are no atoms at all, and intermediate cases. But there are domains for which we do not want to make this assumption. Each of the axioms (22) and (23), though not both at the same time, can be added to CEM to constrain

---

[3]I thank Adrian Brasoveanu for reminding me of this point.

it to one of the two limiting cases. Atomicity entails that everything is ultimately made up of atoms; atomlessness entails that everything is infinitely divisible.

(22) **Optional axiom: Atomicity**
$\forall y \exists x [x \leq_{Atom} y]$
(All things have atomic parts.)

(23) **Optional axiom: Atomlessness**
$\forall x \exists y [y < x]$
(All things have proper parts.)

Axiom (22) relies on the following definition, a useful shorthand.

(24) **Definition: Atomic part**
$x \leq_{Atom} y \stackrel{\text{def}}{=} x \leq y \ \wedge \ \text{Atom}(x)$
(Being an atomic part means being atomic and being a part.)

Following common practice in algebraic semantics, I assume that the mereology for the count domain is constrained to be atomic. The referents of proper names and the entities in the denotations of singular count nouns are then taken to be absolute mereological atoms. This is in line with my assumptions about the relationship between John and John's arm, or between a ham sandwich and its ham slice. The proper name *John* denotes an atom, and John's arm is not a part of John in the mereological sense. I return to this point in Section 2.6.

In general, however, mereologies do not need to be constrained to be either atomic or atomless. When neither of these axioms is added, the system remains underdetermined with respect to whether or not atoms exist. As Krifka (1998) notes, this underdetermination is one of the advantages of mereology: when we describe the domains of space, time, and mass substances, we do not need to take a stance on such questions as whether atomic events and atomic instants of time exist or whether mass substances can be infinitely subdivided.

One source of terminological confusion is whether atoms should be called sums. On the one hand, it is intuitive to speak of sums only when we consider entities that have more than one part. On the other hand, there is a sense in which atoms are sums: the sum of two atoms x and y can be an atom, namely just in case $x = y$. When I need to avoid ambiguities, I use the word *proper sum* to denote sums that are not atoms. I use the word *entity* to range over atoms and sums.

In Section 2.8, I adopt the notion of *impure atoms* from Landman (1989, 2000) in order to model certain instances of collective predication. Impure atoms are atomic entities which are derived from sums via a special group formation or "upsum" operator, written $\uparrow$. The operator introduces a distinction between the sum $a \oplus b$

whose proper parts are the individuals a and b, and the impure atom $\uparrow (a \oplus b)$, which has no proper parts. Group formation is a primitive function, that is, it is not defined in terms of other mereological relations. Although it is not strictly speaking part of CEM, I include it here for convenience.

(25) **Group formation (primitive function)**
$x = \uparrow (y \oplus z)$
(x is the upsum of $y \oplus z$.)

I call atoms which are not generated through the group formation operator *pure atoms*. The $\uparrow$ operator is constrained such that $\uparrow (a) = a$ if and only if a is an atomic individual. In other words, atoms are their own groups. The following definitions provide a handle on the pure/impure atom distinction. I assume that impure atoms do not occur in the denotations of singular count nouns, that is, these nouns apply neither to impure atoms nor to entities which have impure atoms as their parts. This assumption has the consequence that even group nouns like *committee* have pure rather than impure atoms in their denotations, as discussed in Section 2.6.4.

(26) **Definition: Impure atom**
$\text{ImpureAtom}(x) \overset{\text{def}}{=} \exists y[y \neq x \ \wedge \ x = \uparrow (y)]$
(An impure atom is an atom that is derived from a distinct entity through the group formation operation $\uparrow$.)

(27) **Definition: Pure atom**
$\text{PureAtom}(x) \overset{\text{def}}{=} \text{Atom}(x) \ \wedge \ \neg\text{ImpureAtom}(x)$
(A pure atom is an atom which is not impure.)

2.3.2 Higher-order properties

Algebraic semantics makes essential use of higher-order properties to describe predicate denotations and to model constraints on possible meanings. Various authors have identified the properties of mass nouns with the notions of *cumulative reference* (Quine 1960) and *divisive reference* (Cheng 1973) and the properties of singular count nouns with *quantized reference* (Krifka 1986), defined below. As described in Section 2.6, I do not rely on the assumption that mass nouns have divisive reference, but I do assume that singular count nouns have quantized reference. Cumulative reference has also been proposed as a property of plural count nouns. It is easy to show that $\text{CUM}(^*P)$ holds for any predicate P. I assume that plural count nouns as well as mass nouns have cumulative reference. Moreover, I assume that verbs have cumulative reference (Section 2.7.2).

(28)   **Definition: Cumulative reference**
       $\text{CUM}(P) \overset{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[P(y) \rightarrow P(x \oplus y)]]$
       (A predicate P is cumulative if and only if whenever it holds of two things,
       it also holds of their sum.)

(29)   **Definition: Divisive reference**
       $\text{DIV}(P) \overset{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[y < x \rightarrow P(y)]]$
       (A predicate P is divisive if and only if whenever it holds of something, it
       also holds of each of its proper parts.)

(30)   **Definition: Quantized reference**
       $\text{QUA}(P) \overset{\text{def}}{=} \forall x[P(x) \rightarrow \forall y[y < x \rightarrow \neg P(y)]]$
       (A predicate P is quantized if and only if whenever it holds of something, it
       does not hold of any its proper parts.)

2.3.3   Axiomatizations of classical mereology

CEM is completely described by the axioms of reflexivity (4), transitivity (5) and
antisymmetry (6) taken together with uniqueness of sums (13). In fact, this setup
makes axioms (4) and (6) redundant, because any transitive relation that satisfies
axiom (13) is provably reflexive and antisymmetric (see Hovda (2009) for discus-
sion). Models of CEM are essentially isomorphic to complete Boolean algebras
with the bottom element removed, or equivalently complete semilattices with their
bottom element removed. This result goes back to Tarski (1935); see also Pontow
and Schubert (2006) for qualifications. A bottom element is a "null thing" that
is a part of every other thing; the axioms of CEM do not permit the existence of
such a thing, apart from trivial models. In the semantic literature, models of CEM
are often called *complete join semilattices*. However, this use of the term *complete*
deviates from standard mathematical practice because of the absence of a bottom
element (Landman 1989). I refer to models of CEM as *mereologies*.

(31)   **Definition: Mereology**
       Let S be a set and $\leq$ be a relation from S to S. A pair $\langle S, \leq \rangle$ is called
       a mereology if and only if $\leq$ satisfies the axioms of transitivity (5) and
       uniqueness (13).

An example of a mereology is the powerset of a given set, with the empty set
removed, and with the partial order given by the subset relation. The empty set
must be removed because it is a subset of every other set and would therefore act
as a bottom element.

The parthood relation described by CEM has essentially the same properties as
the subset relation in standard set theory. For practical purposes one can therefore

often regard sums as sets, parthood as subsethood, and sum formation as union. Some correspondences between CEM and set theory are listed in Table 2.2. Readers who are unfamiliar with mereology might find this table useful to strengthen their intuitions about the properties of the parthood relation and of the other operations in CEM.

Table 2.2: Correspondences between CEM and set theory

	Property	CEM	Set theory
1	Reflexivity	$x \leq x$	$x \subseteq x$
2	Transitivity	$x \leq y \ \land \ y \leq z \to x \leq z$	$x \subseteq y \land y \subseteq z \to x \subseteq z$
3	Antisymmetry	$x \leq y \ \land \ y \leq x \to x = y$	$x \subseteq y \land y \subseteq x \to x = y$
4	Interdefinability	$x \leq y \leftrightarrow x \oplus y = y$	$x \subseteq y \leftrightarrow x \cup y = y$
5	Unique sum/union	$P \neq \varnothing \to \exists! z \ \text{sum}(z, P)$	$\exists! z \ z = \bigcup P$
6	Associativity	$x \oplus (y \oplus z) = (x \oplus y) \oplus z$	$x \cup (y \cup z) = (x \cup y) \cup z$
7	Commutativity	$x \oplus y = y \oplus x$	$x \cup y = y \cup x$
8	Idempotence	$x \oplus x = x$	$x \cup x = x$
9	Unique separation	$x < y \to$ $\exists! z[x \oplus z = y \land \neg x \otimes z]$	$x \subseteq y \to \exists! z[z = x - y]$

In the Tarski-style axiomatization of CEM adopted here, the sum operation $\oplus$ is defined in terms of the parthood relation $\leq$, which is taken as a primitive. Properties 2 (transitivity) and 5 (unique sum) in Table 2.2 are considered axioms, and the other properties follow from them as theorems. This is not the only possible way to go. For example, Krifka (1998) uses $\oplus$ instead of $\leq$ as a primitive and imposes the properties 4-9 in Table 2.2 as axioms (he calls 8 the remainder principle). The properties 1-3, which identify $\leq$ as a partial order, then follow as theorems. Other comparable systems are the logic of plurality defined by Link (1983, 1998), the part-of structures of Landman (1989, 1991, 2000), and the lattice sorts and part structures of Krifka (1990, 1998). These systems are clearly intended to describe CEM. For example, Link (1983) and Landman (1989) explicitly argue that modeling reference to plurals requires systems with the power of a complete Boolean algebra with the bottom element removed. However, careful review (Hovda 2009) shows that many of these axiomatizations contain errors or ambiguities, mostly in connection with the definition of sum, and therefore fail to characterize CEM as intended. This even extends to the systems given in the standard reference works (Simons 1987; Casati and Varzi 1999). I have tried to avoid these pitfalls by following Hovda's recommendations. Specifically, I follow

Hovda's advice in using Tarski's original definition of the sum concept over more modern alternatives.

2.3.4   Mereology versus set theory

Given the correspondences in Table 2.2 and the pitfalls uncovered in Hovda (2009), working directly with set theory instead of mereology might look like a better choice. After all, set theory is better known and more generally accepted than mereology. Indeed, early approaches to plural semantics adopted set theory (Hausser 1974; Bennett 1974), but sticking to mereology has a number of advantages.

First, mereology is used in much of the relevant literature. In the wake of Link (1983), most theories of plurals and mass nouns have been formulated in a mereological framework. This makes it easier to compare and integrate my proposal with existing accounts. (Link's own motivation for adopting mereology over set theory was for philosophical reasons. But Landman (1989), building on Cresswell (1985), argues convincingly that these reasons are independent of linguistic considerations.)

Second, representing plural individuals as proper sums rather than sets is a convenient way to keep them typographically and type-theoretically distinct from the sets that correspond to one-place predicates representing common noun denotations (Vaillette 2001). In this work, both singular and plural individuals are of type e.

Finally, moving from mereology to set theory requires some technical adjustments. In order to be able to apply the union operation (which is the counterpart of the sum operation, see Table 2.2) to singular individuals, these individuals need to be replaced by or identified with the singleton sets that contain them (Schwarzschild 1996). In addition, the standard axioms of set theory do not permit infinitely descending chains of set membership of the kind $\ldots e''' \in e'' \in e' \in e$. At the bottom of every set membership chain, there must be something that does not have any members itself. In mereological terms, this means that everything is ultimately composed of atoms. A restriction to atomistic mereologies causes problems for the modeling of events, mass entities, and spatiotemporal intervals, where one does not want to be forced to assume the existence of atoms. To avoid this problem, one needs to reject the relevant axioms of set theory. The result is what Bunt (1985) calls ensemble theory.

There is certainly nothing wrong with making these changes to set theory, but taken together, they result in a system that is neither as well known nor as universally accepted as standard set theory. Depending on the precise axiomatization chosen, its models are likely to turn out to be equivalent to those of CEM. In this sense, we would not gain much by moving from mereology to set theory.

2.4   Ontology

This section is devoted to the relatively extensive list of primitive objects that populate the "ontological zoo", as we might call it in analogy to the physicists' particle zoo. I assume that this zoo contains at least individuals and substances, spatial and temporal intervals, events, various sorts of degrees, and numbers. I do not introduce situations or possible worlds, since I do not consider any intensional constructions.

The treatment of the domains of individuals, spatial and temporal intervals, and events is largely symmetric. In particular, these domains are subject to partial orders which satisfy the axioms of classical extensional mereology. These axioms give rise to algebraic structures like those in Link (1998), Krifka (1998), and Landman (2000). As the last section has discussed, they are fairly standard. Degree scales represent a special case since they are assumed to be totally ordered, while mereologies are typically only partially ordered.

This is a nonreductionist account. I have not tried to reconstruct times from events (Kamp 1979; van Benthem 1983; Landman 1991), events from times (Pianesi and Varzi 1996), events from situations (Kratzer 2010), degrees from individuals (Cresswell 1976) or degrees from numbers (Hellan 1981). With respect to these enterprises, I agree with the sentiment expressed in Link (1983): reductionist considerations are "quite alien to the purpose of logically analyzing the inference structure of natural language ... [o]ur guide in ontological matters has to be language itself". A large number of primitives is like a large inventory of tools for theory building: it is easier to work with them, and it is easier to compare the result with other theories that use the same tools. While the framework I adopt here should be in principle amenable to reductionist reconstructions of the sort described above, I see them as a subsequent step to linguistic analysis. However, the proliferation of different kinds of entities should not be mistaken for lack of predictive power. Each of these kinds of entities is well motivated through the study of various phenomena. As in the case of the system in Krifka (1998), the predictive power of my system comes mainly from constraining the structural relationships between domains, and not from limiting the number of these domains.

Figure 2.1 illustrate the basic ontology I assume. I use the words *thing* and *entity* to refer to everything in this ontology, abstracting over individuals, events, and so on.

All these domains are assumed to be disjoint from one another. Nothing is both an individual and an event, or both a degree and an interval. This does not preclude a theory that reconstructs some of these entities from each other, as long as some condition similar to disjointness is maintained. Each of the domains consists of

Figure 2.1: The world (some details omitted)
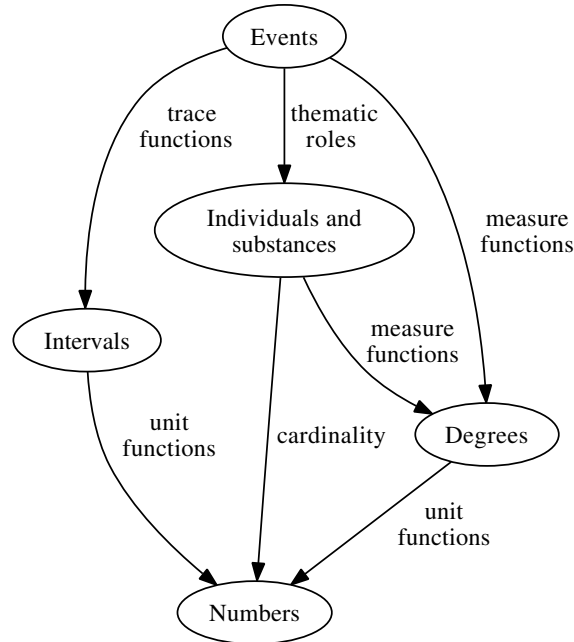
one or more sets or sorts, again disjoint. For example, the domain of degrees is subdivided into disjoint sets representing length, weight, temperature and so on, so that no degree does double duty; and following Link (1983), I assume that count and mass nouns take their denotations in disjoint mereologies. Where exactly to draw the boundaries between one sort and the other is a difficult question. Since I do not rely on the precise nature of these boundaries, I do not try to provide an answer to this question here. Following Krifka (1998) and against Bach (1986), I do not rely on sortal distinctions to model the telic-atelic opposition.

Strictly speaking, each of the mereologies in question comes with its own parthood relation, its own sum operation and so on. However, since we assume that these mereologies are disjoint, we can easily generalize over these different relations and operations by forming their unions. For example, we form the union of the parthood relation on individuals, the parthood relation on events and so on, and we let the symbol $\leq$ stand for this union. This is useful to make statements that generalize over mereologies.

I now describe the domains represented in Figure 2.1 in some detail.

2.4.1 Ordinary objects (type e)

The term *ordinary objects* includes anything which can be referred to by a proper name, or denoted by using a common noun, with the exceptions of nominalizations (which arguably involve reference to events) and measure nouns (which are treated here as involving reference to intervals or degrees, see Section 2.6.3). Intuitively, ordinary objects are either individuals or substances. Examples of individuals are firemen, apples, chairs, opinions, and committees (on the topic of group individuals, see Section 2.6.4). Examples of substances are portions of matter such as the water in my cup or the air which we breathe. A basic distinction between individuals and substances is apparently already made by infants, who know that individuals generally have boundaries, move along continuous paths, and survive collisions with each other, while substances do not (Spelke 1990).

Loosely following Quine (1960) and Link (1983), I draw a formal distinction between individuals and substances. The main purpose of this distinction in the present work is to express that individuals are built up from mereological atoms, while substances need not be. In connection with the assumption that count nouns always involve reference to individuals, this choice allows us to represent counting and distributing operations as involving the mereological atoms in a plural individual.

This is without doubt a simplifying assumption. For example, so-called fake mass nouns[4] like *furniture*, *clothing*, *jewelry*, *silverware*, *mail* and *offspring* provide evidence that at least some mass nouns also involve reference to inherently individuatable and countable entities. Barner and Snedeker (2005) show experimentally that fake mass nouns pattern with count nouns rather than with substance mass nouns in quantity judgment tests. Participants who were asked to compare amounts of matter referred to by either a count noun or an object mass noun relied on number rather than on weight or volume for their judgment. For these reasons, it is difficult to maintain that the referents of fake mass nouns are substances. If fake mass nouns involve reference to atoms like count nouns do, then another factor must be provided that blocks the application of numerals and of distributivity markers to them. I do not provide such a factor in this work. This leaves the potential for overgeneration: I do not have an explanation why one cannot say **three offspring* or **Her offspring each went to a boarding school* (Kratzer 2007). However, let me briefly sketch what such an explanation could look like. An account of fake mass nouns is proposed in Chierchia (1998a, 2010). The atoms involved in the denotation of fake mass nouns are there taken to be "unstable",

---

[4]This term comes from Chierchia (1998a, 2010). Other terms include *object mass nouns* (Barner and Snedeker 2005), *collective mass nouns* (Bunt 1985), *count mass nouns* (Doetjes 1997), and *individual mass nouns* (according to Chierchia (2010)).

that is, they have vague criteria of identity. Numerals are assumed to contain not only an atomicity test but also an anti-vagueness condition that rejects such unstable atoms. By adding the anti-vagueness condition to numerals and *each*, my account could conceivably be extended to prevent such overgeneration. That is, any lexical entries that use the predicate *Atom*, would be amended to use a predicate *StableAtom* instead. This predicate would be defined as in Chierchia (2010) and apply only to the individuals in the denotation of count nouns, but not to those in the denotation of fake mass nouns and other mass nouns. However, I do not pursue this route here.

2.4.2   Kinds

Kinds do not play a major role in this work, and their ontological status is debated. For this reason, I do not give them a type. Prototypical examples of kinds are plant and animal species, but the inventory of kinds might be much larger; arguably, (almost) all bare plural and mass noun phrases can be understood as involving reference to kinds. While I have categorized kinds as entities in their own right, following Carlson (1977), there have been attempts to formally relate them to other entities. In a mereological setting, it is tempting to associate kinds with sums. For example, the kind *potato* would be the sum of all potatoes. This idea causes problems in the case of kinds such as *Dodo* or *Phlogiston* that are not instantiated in the real world, a problem which can be countered to some extent by appealing to intensionality (Chierchia 1998b), though there are problems with assimilating kinds to sums too closely (Pearson 2009). I discuss kinds briefly again in Section 2.6.

2.4.3   Events (type v)

An event is a spatially and temporally bounded, ephemeral constituent of the world that has a single occurrence (Carlson 1998). The classical examples in the literature of events are Jones' buttering of the toast and Brutus' stabbing of Caesar. As described in the following, I adopt a standard view according to which sentences existentially quantify over events (Davidson 1967; Parsons 1990), and I assume that events are ordered in a mereological structure (Bach 1986; Krifka 1998). My purpose in making assumptions about events is to bring out the parallel between individuals and events. The null assumption is that there is no semantic distinction between nouns and verbs, only a syntactic distinction.

Just as individuals are "multidimensional" in the sense that they may have several thematic roles and can be measured along several dimensions, events are "multidimensional" in the sense that they exist in space and in time. For example,

Caesar's assassination occurred on March 15, 44BC, in the Roman Senate. In this sense, events are distinct from intervals, which are one-dimensional entities and live completely in time or in space (Section 2.4.4). The distinction between events and individuals is more difficult to draw. A popular assumption is that the distinction between events and individuals follows the lines of the 3D/4D controversy (Markosian 2009). Non-instantaneous events have temporal parts (they are 4D objects or perdurants), but all the proper parts of an individual are present at each point in time (they are 3D objects or endurants), or at least throughout the individual's lifetime. For example, the event in which John sleeps from sunset to midnight is a proper part of the event in which John sleeps from sunset to sunrise, but in each of these events, it is John as a whole who sleeps: any and all of his parts are present at every period of time during his sleep, and beyond. This assumption is implicitly present in much work on natural language semantics (e.g. Krifka 1998), and I adopt it too.

I use the term *event* in a wide sense, synonymous with *eventuality* (Bach 1986). I assume that all verbs without exceptions involve reference to events (see Section 2.7.1). Other authors use *event* in contradistinction to such things as *states* and sometimes *processes* (Mourelatos 1978; Bach 1986; Smith 1997). While Piñón (1995) argues that states, processes, and events are pairwise disjoint classes, I do not make use of this distinction. I assume, following the arguments in Parsons (1987, 1990, ch. 10), that any declarative sentence, including stative sentences, involves reference to an underlying entity that I call an event. Therefore, the category of events does not only include things like Brutus's stabbing of Caesar are events, but also things traditionally called "states", like John's being asleep, or the Golden Gate bridge's spanning from San Francisco to Marin County.

I assume that events form a mereology, meaning that they are closed under sum formation (Bach 1986; Krifka 1998). For example, if John runs to the store, this is an event; if he runs to the store several times, these are several events, but the sum of these events is again an event. Events may have parts which occupy less time, less space, or both. For example, the event e_0 in which John runs from the village to the store has as one of its parts the event e_1 in which John runs from the village halfway to the store. In general, it is not always the case that any part of an event which takes less time also occupies less space. For example, suppose that e_0 is a part of an event e_2 in which John runs from the village to the store and then back. Then e_0 takes less time than e_2 but occupies the same space.

Some authors assume that the domain of events is atomic (Landman 1996, 2000). Others assume that at least some kinds of events have mass denotations, which is especially plausible for events associated with atelicity (Mourelatos 1978). In line with Krifka (1998), I remain neutral on this question and I do not rely on either assumption.

2.4.4 Intervals (type i)

For the purpose of this work, intervals are one-dimensional entities that represent stretches of time and paths through space. Examples of intervals are the time from 1pm to 2pm today or the path along which Route 66 extends. I make the standard assumption (as discussed for example in Krifka (1998)) that intervals are at the basis of the denotations of temporal and spatial measure phrases like *three hours* and *three miles*, and of prepositional phrases like *from 3pm to 6pm* and *to the store* (see Section 2.9). Intervals are also used to specify the spatial and temporal extents of events and other entities. The functions that relate events to intervals are discussed in Section 2.5.2. Intervals also play a separate role in the literature on degree semantics (see Section 2.4.5), but I set aside this fact here. Also, while one of the main uses of interval semantics is the modeling of tense, I ignore the contribution of tense in this work.

An important distinction between intervals and events is that two distinct events may happen at the same place and/or time, while this is not possible for intervals. For example, a sphere can rotate quickly and heat up slowly at the same time (Quine 1985). Since the predicates *slowly* and *quickly* mutually exclude each other, I assume that they apply to different events, but then these are distinct events that happen at the same time and at the same place.

The standard conception of time in natural language semantics is based on intervals, which are generally conceived as being ordered by a subinterval relation and a temporal precedence relation. The relevant choices for axiomatizing temporal structures are thoroughly examined and discussed by van Benthem (1983). The algebra of paths and the contribution of directional prepositional phrases to lexical aspect has been studied by Zwarts (2005), among others. A recent summary of the literature can be found in Ursini (2006). Here I follow Krifka (1998), who proposes axioms for both temporal and spatial interval structures and integrates them in a mereological system, based on earlier ideas by Hinrichs (1985). Krifka's axioms ensure, among other things, that temporal precedence is irreflexive, asymmetric and transitive, and that any two intervals either overlap or one precedes the other.

The subinterval relation is equated with mereological parthood and written $\leq$; I write the temporal precedence relation as $\ll$. For example, if a is the interval from 2pm to 3pm today, b is the interval from 4pm to 5pm, and c is the interval from 1pm to 5pm, then we have $a \leq c$, $b \leq c$, and $a \ll b$. I assume that there are also directed spatial intervals. For example, the PP *to the store* involves reference to a directed interval whose end is the store (see Section 2.9). I use the self-explaining functions *start* and *end* to refer to the endpoints of directed intervals.

Since time is assumed to be a mereology, any two intervals have a unique sum. Since intervals can be arbitrarily summed, some of them are discontinuous. In the

example above, the interval $a \oplus b$ is discontinuous: it contains a gap from 3pm to 4pm. For this reason, the term "interval" is perhaps misleading, but I use it anyway following common practice.

I assume that the mereologies underlying intervals are nonatomic. According to von Stechow (2009), this is a standard assumption: most semanticists do not assume the existence of temporal atoms or instants.

2.4.5 Degrees (type d)

I assume that the ontological system contains a category of entities called degrees, which represent quantities assigned by measure functions such as height, weight, volume or beauty. Examples of degrees are John's weight, the thickness of the ice at the South Pole, and the speed at which John is driving his car now. Degrees have been used in semantic accounts of gradable adjectives (*tall, beautiful*) and measure nouns (*liter, hour*) as well as in accounts of the constructions that contain them, such as measure phrases (*three liters*), comparatives (*taller, more beautiful, more water, more than three liters*), and pseudopartitives (*three liters of water*). Degree semantic approaches to comparatives interpret *Mary is taller than John* as some variation of *Mary's height exceeds John's height*, where Mary's height is the degree to which she is tall. Similarly, *Mary is more beautiful than John* is interpreted as *Mary's beauty exceeds John's beauty*, where Mary's beauty is the degree to which she is beautiful. Degree semantic approaches to measure phrases interpret *three liters* as involving reference to a degree of volume. As we will see in Section 3.2, a pseudopartitive like *three liters of water* is then interpreted as involving reference to a water entity whose degree of volume is specified by *three liters*.

In the modern literature, semantic analyses that treat degrees as part of the ontology originate with Cresswell (1976). A different line of analysis treats degrees as contextual coordinates rather than as ontological entities in their own right (Lewis 1972). For comparisons of the two approaches, see Klein (1991), Kennedy (2007), and van Rooij (2008). I do not adopt the contextual-coordinate analysis because it is mainly concerned with gradable adjectives and does not, as far as I can see, provide an obvious way to represent the meaning of measure phrases.

Within analyses that treat degrees as part of the ontology, there is disagreement on what these degrees are. I model degrees as primitive entities. This is also the option chosen by Parsons (1970) and Cartwright (1975). Other authors model degrees as numbers (Hellan 1981; Krifka 1998) or as equivalence classes of individuals indistinguishable with respect to a relevant gradable property (Cresswell 1976; Ojeda 2003). There seems to be a consensus that degrees of a given sort are totally ordered, although it is a topic of debate whether degrees should be understood as points or initial intervals ("extents") on a scale. (Such intervals

must be carefully distinguished from the temporal and spatial intervals in Section 2.4.4. Temporal and spatial intervals are ordered by two relations: precedence and parthood. Degree-based intervals are assumed to be initial intervals on their scale and are therefore only ordered by a parthood relation.)

I adopt the view that degrees are totally ordered on scales. Following Fox and Hackl (2006), I assume that this order is dense. I represent it as $\leq$, just like the mereological parthood relation. I leave open the question whether degrees should be modeled as points or extents, as it does not seem to affect my proposal. See Krasikova (2009) for discussion of this point. While I generalize over degree order and mereological parthood, I remain noncommittal on the question of whether degree scales should indeed be understood to be special cases of mereologies. See Szabolcsi and Zwarts (1993) and Lassiter (2010a,b) for a proposal in this direction. While there is nothing wrong conceptually with assuming that the ordering relation on degrees of a scale is a special case of the mereological parthood relation, note that the former relation is usually assumed to be a total order while mereological parthood, in the general case, is only constrained to be a partial order. This has the consequence that the mereological sum operation does not have a natural counterpart on degree scales. Many ontological systems assume that there is a sum operation on degrees, which corresponds to arithmetic sum. This degree sum operation has different properties from mereological sum. For example, mereological sum operation is idempotent (Table 2.2), but this is not the case for the degree sum operation. Imagine two coins c_1 and c_2 which weigh one gram each. The degree sum of their weights is two grams. However, if the two coins are assumed to have identical weight degrees, then the mereological sum of their weights is one gram, because sum is idempotent. For more on this point, see Section 2.5.3.

Let me briefly explain why I do not derive degrees from other entities. Reconstructing degrees as equivalence classes does not help integrating degrees into a mereological system because it does not lead to any straightforward set-theoretic or mereological relations between different degrees. For example, Cresswell (1976) reconstructs the degrees *three pounds* and *six pounds* as the sets of all things whose weight is three pounds and six pounds, respectively. This means that *three pounds* is neither an element nor a subset of *six pounds*. The operation of adding *three pounds* to *three pounds* cannot be easily interpreted mereologically either, supposing that the result can be described as *six pounds*. It is different from mereological sum formation, since that operation sum is idempotent. It is also different from pointwise sum formation: although some of the elements of the set *six pounds* are mereological sums of two elements of the set *three pounds*, others might be atoms.

Reconstructing degrees as numbers would raise problems for modeling unit conversion. For example, six feet approximately equals 183cm, but 6 is not equal

to 183. It also makes it difficult to model the difference between different scales such as height or weight, since one set of numbers is used for all purposes. On the other hand, the total ordering and other properties of numbers is useful to model mathematical properties of degrees. A compromise proposed by Lønning (1987), which I adopt, is to introduce degrees as an intermediate layer that mediates between individuals and numbers. In Lønning's framework, one set of functions maps entities to their degrees of weight, length, temperature, and so on. Another set of functions maps these degrees to numbers or number-like entities. With the help of the intermediate layer, we can model degrees of weight, height, temperature and so on as different sorts. See Section 2.5.4 for discussion.

Having both degrees and intervals may seem redundant. Both time and space could be thought of as degree scales. However, I keep temporal and spatial intervals formally apart from degrees because time and space appear to be qualitatively different from degree scales. We can individuate temporal and spatial intervals, but we cannot do anything corresponding with degrees. For example, we feel that Thursday is a different temporal interval than Friday, even though both have the same length. But there is no sense in which we feel that two degrees that represent the same weight or temperature are different from each other. Intervals are subject to two partial orders: precedence and inclusion. For example, Thursday precedes Friday, and Thursday includes Thursday afternoon. Degrees only appear to be subject to one (total) order: four kilograms is less than five kilograms.

2.4.6 Numbers (type n)

I use numbers to represent the meanings of number words. For example, the word *three* denotes the number 3. Numbers also constitute the range of what I call unit functions, that is, functions like *hours* and *meters* (see Section 2.5.4). I use the type n to represent them. I assume that numbers include at least the rational numbers, so that a unit function like *days* can assign fractions of days to events with very short but not instantaneous runtimes. That is, the runtimes of these events match $\lambda t[\text{days}(t) \leq 1]$ even though their runtime is not instantaneous. See Fox and Hackl (2006) for related discussion.

Lønning (1987) replaces the domain of numbers by more abstract entities and considers which of the properties of the number scale must be imposed on these entities if one wants to avoid including all of mathematics in natural language. For example, if one wants to model the fact that people cannot always decide which of two very large numbers is bigger than the other, some numbers might be assumed not to be ordered with respect to one another. For ease of exposition, I do not following Lønning's suggestion and I continue instead to use actual numbers in the mathematical sense, along with the total order $\leq$ in which they stand.

2.5 Functions

As already shown in Figure 2.1, I assume that the domains of objects, events, intervals, degrees, and numbers presented in the last section are interconnected by a variety of relationships, which can be modeled as partial functions. Since all the functions in this work are partial, I use the word *functions* instead of the term *partial functions*. As described in this section, I assume that thematic roles map events to individuals; trace functions map events to their locations in space and time; measure functions map individuals and events to degrees; unit functions map degrees to numbers. In this section, I describe these functions in more detail. I have chosen those background assumptions that make these functions as similar as possible. In particular, I build on the concept of homomorphism, a special kind of function that preserves certain algebraic relationships such as sums and parts across domains. Homomorphisms constrain the structural relationships between domains and therefore, by extension, the domains themselves.

2.5.1 Thematic roles (type ⟨ve⟩)

Thematic roles are semantic relations that represent different ways in which entities participate in events (Parsons 1990; Dowty 1991). There are two common views of thematic roles. On the traditional view, due to Gruber (1965) and Jackendoff (1972), thematic roles encapsulate generalizations over shared entailments of argument positions in different predicates. Thematic roles in this sense include *ag* or agent (initiates the event, or is responsible for the event), *th* or theme (undergoes the event), *instrument* (used to perform an event), and sometimes also *location* and *time*. For example, if Brutus stabbed Caesar with a dagger in the Senate on March 15, then there is a stabbing event whose agent is Brutus, whose theme is Caesar, whose instrument is a dagger, whose location is the Senate and whose time is March 15. I discuss location and time in the next section. An alternative view sees thematic roles as verb-specific relations: Brutus is not the agent of the stabbing event but the stabber, Caesar is not its theme but its stabbee, and so on (Marantz 1984). This comes at the obvious price of missed generalizations, and it raises some technical problems. For example, it is not clear which role to assign to the subjects of coordinated sentences like *A girl sang and danced*. I adopt the more traditional, and relatively widespread, view that there is a small inventory of thematic roles that generalize over verbs. Since the semantic content of thematic roles is not an issue in this work, and since there does not seem to be a consensus on what to call the thematic role of the arguments of verbs like *own* or *be numerous*, I call the thematic role of the subject of any verb the agent, even if this is not always intuitive. Kratzer (1996) uses the world "holder" for the subject role of *own*, but she

notes that this is only for convenience.

Unfortunately, there is no consensus on the inventory of thematic roles. It is difficult to identify these roles in borderline cases, though not impossible: a wide-coverage database of English verbs with their thematic roles, extending prior work by Levin (1993), was made available in Kipper-Schuler (2005).

I assume that thematic roles as semantic relations have syntactic counterparts which relate verbs to their arguments (see Section 2.10 for an illustration). This is a standard assumption in generative grammar, at least as far as the agent role is concerned: the "little v" head is assumed to relate verbs to their external arguments, which are usually their agents (Chomsky 1995). Other thematic roles can be taken to have syntactic reflexes such as case marking and prepositions. Following Carnie (2006), I reserve the term *thematic role* for the semantic relation and I use the term *theta role* for its syntactic counterparts. However, some authors use these terms interchangeably.

The correspondence between theta roles and thematic roles is controversial. For example, Chomsky (1981) assumes that each argument is limited to one role. In many cases, though, it seems appropriate to assign more than one thematic role to the same argument: for example, the subject of a verb like *fall* can be regarded both as the agent and the theme of the event (Parsons 1990).

I adopt the assumption that each event has at most one agent, at most one theme, and so on. This assumption is generally called the Unique Role Requirement (URR) or thematic uniqueness. Arguments for the existence of thematic roles and for the URR are reviewed in Carlson (1984, 1998), Parsons (1990) and Landman (2000). Formally, I assume that thematic roles are functions of type $\langle v, e \rangle$ rather than just relations of type $\langle v, et \rangle$. This choice is not essential because either option leads to type mismatches. I assume they are resolved by type-shifting (see Section 2.10).

The URR can be integrated in a mereological setting where events are closed under sum formation. I assume for this purpose that thematic roles are their own algebraic closures (Krifka 1986, 1998; Landman 2000). This property is also called cumulativity or summativity of thematic roles. The following definition uses algebraic closure on functions, as defined in (20).

(32)   **Cumulativity assumption for thematic roles**
        For any thematic role θ it holds that $\theta = {}^*\theta$. This entails that
        $\forall e, e', x, y[\theta(e) = x \ \wedge \ \theta(e') = y \ \rightarrow \ \theta(e \oplus e') = x \oplus y]$

For example, if e is a talking event whose agent is John and e' is a talking event whose agent is Mary, then there is an event $e \oplus e'$ whose agent is the sum of John and Mary. Thus, $e \oplus e'$ has a unique entity as its agent, even though this entity is a

proper sum.

As a reminder of the cumulativity assumption, I include the algebraic closure operator in the typographical representation of thematic roles. For example, instead of writing th for the meaning of the theme role, I write *th.

The URR is a controversial assumption. Krifka (1992) rejects it because, as he claims, one "can see a zebra and, with the same event of seeing, see the mane of the zebra as well" and one "can touch a shoulder and a person with the same event of touching". This view on events is problematic: it is not contradictory to say that somebody saw a zebra consciously and its mane unconsciously, or that he touched a man intentionally but touched his shoulder unintentionally. But on the plausible view that such manner adverbs are predicates over events, this means that the two events in each case must be considered distinct (though one may be considered part of the other). Other arguments against the URR are discussed and refuted in Landman (2000).

It is sometimes useful to talk about the entity to which a certain event e is mapped under a thematic role θ, while abstracting over the identity of θ. I refer to this entity as the θ of e. For example, if x is the agent of e then x is the θ of e for $\theta = $ agent.

As a consequence of (32), thematic roles are homomorphisms, or structure-preserving maps, with respect to the $\oplus$ operation (33).

(33)   **Fact: Thematic roles are sum homomorphisms**
For any thematic role θ, it holds that $\theta(e \oplus e') = \theta(e) \oplus \theta(e')$.
(The θ of the sum of two events is the sum of their θs.)

Again, not everyone accepts this assumption. Suppose with Kratzer (2003) that there are three events e_1, e_2, e_3 in which Al dug a hole, Bill inserted a rosebush in it, and Carl covered the rosebush with soil. Then, in virtue of these three events, one can say that there is also an event in which Al, Bill, and Charles planted a rosebush. Let e_4 be this event. Do we consider e_4 equal to the proper sum event $e_1 \oplus e_2 \oplus e_3$? If we do, this scenario is a counterexample to the cumulativity assumption, as Kratzer notes. The themes of e_1, e_2, e_3 are the hole, the rosebush, and the soil, and the theme of e_4 is just the rosebush. The theme of e_4 is not the sum of the themes of e_1, e_2, and e_3. This violates cumulativity.

I respond to this challenge following my general strategy of not assuming that the mereological parthood relation should model all parthood relations that can be intuitively posited (see Section 2.3.1). In this case, I assume that e_4 is not actually the sum of e_1, e_2, and e_3. Even though the existence of e_4 can be traced back to the occurrence of e_1, e_2, and e_3, nothing forces us to assume that these three events are actually parts of e_4, just like we do not consider a plume of smoke to be part

of the fire from which it comes, even though its existence can be traced back to the fire. Without the assumption that e_4 contains e_1 through e_3 as parts, Kratzer's objection against cumulativity vanishes. For more discussion of this point, see Williams (2009).

2.5.2   Trace functions: runtime and location (type $\langle vi \rangle$)

Trace functions map events to those intervals that represent their temporal and spatial locations. Since some events may not be situated in space and/or time, trace functions are partial. I assume the existence of two such functions, temporal trace or *runtime* and spatial trace or *location*. Following common usage in the literature, I also write these functions as τ and σ respectively. Since intervals correspond to spatiotemporal locations, trace functions do not simply indicate the amount of space and time that an event takes, but also the precise location in space and time. For example, suppose John sings from 1pm to 2pm and Mary sings from 2pm to 3pm. Although each event takes the same time, their runtimes are different since they take place at different times. The temporal trace function mirrors this by mapping John's running and Mary's running to distinct intervals. These intervals are not directly related to each other, though they are mapped to the same numbers by the unit functions *hour*, *minute* and so on (see Section 2.5.4).

Following Link (1998) and Krifka (1998), I assume that trace functions are sum homomorphisms: the runtime of the sum of two events is the sum of the runtimes. For example, in the scenario just mentioned, the runtime of the sum event that combines John's and Mary's running is the interval from 1pm to 3pm, and this is the sum of the intervals from 1pm to 2pm and from 2pm to 3pm. This is formally expressed in the following equations:

(34)   **Trace functions are sum homomorphisms**
       σ is a sum homomorphism: $\sigma(e \oplus e') = \sigma(e) \oplus \sigma(e')$
       τ is a sum homomorphism: $\tau(e \oplus e') = \tau(e) \oplus \tau(e')$
       (The location of the sum of two events is the sum of their locations, and similarly for their runtimes.)

This assumption is parallel to the cumulativity assumption on thematic roles described in Section 2.5.1. In general, trace functions are not very different from thematic roles. The main differences are that trace functions map events to intervals rather than to individuals and that they are typically expressed by adjuncts rather than by arguments.

Trace functions cross the bridge between interval semantics and event semantics, since they allow us to realize the intuition that a predicate holds at a certain

interval. Interval semantics assumes the existence of a relation AT that evaluates a proposition at an interval. It is useful to be able to refer to this relation also in event semantics, where instead of propositions we can relate event predicates to intervals. The following definition fixes the relation between AT and τ.

(35) **Definition: Holding at an interval**
      $\text{AT}(V, i) \stackrel{\text{def}}{=} \exists e[V(e) \ \wedge \ \tau(e) = i]$
      (An event predicate V holds at an interval i if and only if it holds of some event whose temporal trace is i.)

Since events are closed under sum, the homomorphism assumption has the consequence that two events whose runtimes are not adjacent have as their sum an event whose runtime is not continuous. See Section 2.5.4 for more discussion on noncontinuous intervals.

For more on temporal and spatial trace functions, see Hinrichs (1985), Lasersohn (1988), Krifka (1998), Link (1998), as well as Zwarts (2005, 2006), where the spatial and temporal trace functions are combined into one.

2.5.3   Measure functions (types $\langle ed \rangle$ and $\langle vd \rangle$)

I use the term *measure function* to denote a mapping between a class of physical objects and a degree scale that preserves a certain empirically given ordering relation, such as "be lighter than" or "be cooler than". Typical measure functions are height, weight, speed, and temperature. It seems plausible to ascribe at least some of them to events rather than individuals. I have already mentioned the example due to Quine (1985) of a sphere which can at the same time rotate quickly and heat up slowly. Provided that the rotating and the heating up are two separate events, and that *quickly* and *slowly* are event modifiers, we can avoid the undesirable conclusion that the sphere is both quick and slow by treating speed as a function that applies to events rather than individuals. See also Kratzer (2001) for an analysis that ascribes weights to states.

Some authors do not distinguish between trace functions and measure functions (e.g. Kratzer 2001). I prefer to treat them differently for two reasons. First, trace functions map into intervals and measure functions into degrees. Section 2.4.5 explains my reasons not to conflate these two categories. Second, the sum homomorphism assumption (34) must be limited to trace functions (and thematic roles). Extending it to measure functions would have counterintuitive consequences and would not be compatible with the standard conception of degrees as totally ordered entities. For example, suppose that we have two coins, c_1 and c_2, which each weigh one gram. Following Lønning (1987), we represent this fact as follows: the measure

function *weight* to maps c_1 to its weight w_1 and c_2 to its weight w_2, and the unit function *grams* maps both w_1 and w_2 to 1. If we wanted to extend assumption (34) to weight, it would follow that the weight function maps $c_1 \oplus c_2$, the sum of the two coins, to $w_1 \oplus w_2$. To model the fact that the two coins together weigh two grams, we assume that the function *grams* maps this entity to the number 2. From these assumptions, it follows that $w_1 \neq w_1 \oplus w_2$, that is, the weight of c_1 is not equal to the weight of $c_1 \oplus c_2$, which is plausible. From this fact it follows that $w_1 \neq w_2$, due to idempotence (Table 2.2). In other words, the two coins are mapped to different weight degrees, even though intuitively they "weigh the same". More generally, whenever a measure function maps the sum of two objects to a different degree than it maps one of them, it also maps both objects to different degrees. As far as I can tell, no empirical reasons speak against this possibility.[5] It is even consistent with the rest of the assumptions in this work, as long as one is willing to relax the standard assumption that degrees of the same sort are totally ordered (otherwise one needs to insist that w_1 precedes w_2 or vice versa). However, it is certainly conceptually unattractive to assume this proliferation of degree entities. It is easier to limit the sum homomorphism assumption to trace functions and to exclude measure functions from it. Concretely, I assume that trace functions are sum homomorphisms but measure functions need not be. This motivates the conceptual distinction between trace functions and measure functions.

2.5.4 Unit functions (types $\langle dn \rangle$ and $\langle in \rangle$)

As discussed in Section 2.4.5, I adopt the proposal of Lønning (1987), according to which degrees occupy an intermediate layer betwen individuals and numbers. I refer to the functions that relate entities to degrees as *measure functions* and to the functions that relate degrees to numbers as *unit functions*. This use of the term *measure function* differs from Krifka (1998), whose framework lacks the intermediate layer of degrees. Krifka's measure functions map entities directly onto numbers (see Section 7.5). I use the term *degree* exclusively to refer to the things in the range of the measure functions, and I use the term *numbers* exclusively to refer to the things in the range of the unit functions.

As an illustration, suppose John weighs 150 pounds (68 kilograms) and measures six feet (183 centimeters). These facts are represented as follows. The measure function *weight* maps John to a degree which the unit functions *pounds* and *kilogram* map to the numbers 150 and 68, respectively. The measure function *height* maps John to a degree which is mapped by the unit function *feet* to the number 6 and by the unit function *centimeters* to the number 183:

---

[5] I thank Roger Schwarzschild for discussing this issue with me.

(36) a. pounds(weight(john))=150
      b. kilograms(weight(john))=68
      c. feet(height(john))=6
      d. centimeters(height(john))=183

I assume that the total order of degrees mirrors the order of the range of the corresponding unit functions. For example:

(37) height(john) ≤ height(mary)
      ⇔ meters(height(john)) ≤ meters(height(mary))

Splitting up the semantics of measurement into measure functions and unit functions introduces an additional degree of freedom in the semantic representation. This degree of freedom is needed to model the fact that pseudopartitives do not always specify the way in which the measure noun is related to the substance noun. For example, as Schwarzschild (2002) observes, a pseudopartitive like (38) can be understood as involving reference to an amount of oil that is either three inches deep, or (given the right context, for example when talking about a growing puddle), three inches in diameter. The two interpretations of the pseudopartitive can lead to different truth conditions. This can be modeled by the following two interpretations:

(38) three inches of oil

      a. $\lambda x[\text{oil}(x) \ \wedge \ \text{inches}(\text{height}(x)) = 3]$
      b. $\lambda x[\text{oil}(x) \ \wedge \ \text{inches}(\text{diameter}(x)) = 3]$

As we see, this difference can be expressed straightforwardly by Lønning's split: the interpretations differ in their measure function, but not in their unit function. Of course, not every pseudopartitive is ambiguous in this way. For example, (39) unambiguously involves reference to a quantity of oil whose volume is three liters (39a). It does not give rise to readings in which, say, the height of the oil is asserted to be three liters (39b). Since this reading is nonsensical, I assume that the unit function *liters* is a partial function that is only defined on degrees in the range of the measure function *volume* but not on those in the range of the measure function *height*. Many pseudopartitives are unambiguous as a consequence, even though in a pseudopartitive, only the unit function and not the measure function is overt.

(39) three liters of oil

      a. $\lambda x[\text{oil}(x) \ \wedge \ \text{liters}(\text{volume}(x)) = 3]$ *OK*
      b. $\lambda x[\text{oil}(x) \ \wedge \ \text{liters}(\text{height}(x)) = 3]$ *nonsensical*

For more motivation of the split between unit functions and measure functions,

see Lønning (1987) and Schwarzschild (2002, 2006).

Following Krifka (1998), I assume that spatiotemporal intervals are related to numbers by functions like *hours* and *meters*. I call these functions unit functions as well, even though I make a formal difference between degrees and intervals. Intervals can be discontinuous (see Section 2.4.4). This poses a problem for the definition of unit functions. For example, if a is the interval from 2pm to 3pm, and b is the interval from 4pm to 5pm, then we have hours(a) = hours(b) = 1, but what about hours$(a \oplus b)$? Is it 3 or just 2? Let us call the first option the *generous construal* and the second option the *stingy construal*. Evidence for either option is not hard to come by if we assume (see Section 3.3) that temporal *for*-adverbials use unit functions to map the runtimes of events to numbers. The relevant observations go back to Dowty (1979). For example, sentence (40) is coherent and favors the generous construal:

(40)   John worked in New York for four years but he usually spent his weekends at the beach.   (Dowty 1979, p. 334)

However, (41) can be true even if John served four non-consecutive one-year terms, and therefore favors the stingy construal.

(41)   John served on that committee for four years.   (Dowty 1979, p. 334)

The generous-stingy opposition should probably really be seen as a continuum, since the interpretation of following sentence appears to be intermediate between both construals:

(42)   From 2002 to 2008, John lived in New York for a total of four years, but he usually spent his weekends in Maine.

In the examples I consider in this work, stingy construals of *for*-adverbials do not play a role. For this reason, I assume that unit functions are always interpreted generously. That is, I assume that unit functions map any discontinuous interval i to the same number as they map the smallest continuous interval containing i. In the example above, this means that hours$(a \oplus b)$ = 3. Since hours(a) + hours(b) = 2, this assumption also means that unit functions are not sum homomorphisms.

Note that the present setup does not impose a choice between being generous and being stingy. We could equally well generalize unit functions to "unit relations", and constrain them in such a way that both hours$(a \oplus b, 3)$ and hours$(a \oplus b, 2)$ hold in the above example. As discussed in Section 2.6.3, I assume that measure nouns have relational types like $\langle n, it \rangle$ and that their translations are based on unit functions. The meaning of measure nouns like *hour* could therefore easily be changed to accommodate unit relations:

(43)   a. $[\![$hour$]\!]$ (using unit functions) $= \lambda n \lambda t [\text{hours}(t) = n]$
        b. $[\![$hour$]\!]$ (using unit relations) $= \lambda n \lambda t [\text{hours}(t, n)]$

I do not pursue this route, however.

2.5.5   The cardinality function (type $\langle en \rangle$)

I assume that there is a partial function that maps sums which consist of singular individuals onto the number of singular individuals of which they consist. For example, the function maps a sum of three boys to the number three. I assume that it is only defined on entities in the extension of count nouns. Following standard practice, I call this function cardinality, and I define it in terms of the set-theoretic notion of cardinality.

(44)   **Definition: Cardinality**
        For any sum x such that $^*\text{Atom}(x)$, the *cardinality* of x, written $|x|$, is defined as $|\{\, y \mid y \leq_{Atom} x \,\}|$.

I assume that singular individuals are atomic (see Section 2.6.1). This means that singular individuals always have cardinality one. An alternative is found in Krifka (1989), who assumes that there is a partial *natural unit* function, NU, which assigns plural individuals a number that intuitively represents their cardinality. NU is a primitive concept, so it does not require the atomicity assumption.

In order to be able to give a single lexical entry to numerals whether they occur in measure phrases or in run-of-the-mill noun phrases, I give number words the type n and I let them denote actual numbers, as in Krifka (1989), Hackl (2001) and Landman (2004). For example, *three* denotes the number 3 rather than a generalized quantifier or a predicate over pluralities with three atoms. To represent numeral noun phrases, I assume that a silent head [many] introduces the cardinality function, as in Figure 2.2 (see Hackl (2009) for a similar proposal). This represents the sum reading of the numeral *three boys* (see Section 2.8). The lexical entry for *boys* is justified in Section 2.6.2. In measure phrases, the number word combines with the measure noun, which is assumed to be of type $\langle n, dt \rangle$ (see Section 2.6.3), as shown in Figure 2.3. In this case, the [many] head is not necessary.

In my entries for [many] and for unit functions, I use an *exactly* interpretation rather than an *at least* interpretation. That is, I use an equality sign rather than a $\geq$ sign. This decision keeps things compatible with Krifka (1998), who assumes that predicates like *eat three apples* are quantized and therefore do not apply to events in which more than three apples are eaten. There is no consensus in the literature on whether the literal meaning of *three boys* is *at least three boys* (e.g. Horn 1972, 1989; Barwise and Cooper 1981; Levinson 2000; van Rooij and Schulz 2006) or *exactly*

three boys (e.g. Harnish 1976; Sadock 1984; Partee 1987; Carston 1998; Geurts 2006; Breheny 2008). For a review of this literature, see Nouwen (2006) and Kennedy (2009). In an event semantics, it does not matter much whether numerals are given an *exactly* or an *at least* reading, since the maximality conditions involved in *exactly* readings have to be realized through other mechanisms anyway (Krifka 1999; Landman 2000; Robaldo 2010; Brasoveanu 2010).

Figure 2.2: LF of the noun phrase *three boys*

$$\lambda x[|x| = 3 \ \wedge \ {}^*\text{boy}(x)]$$

$$\lambda x[|x| = 3] \qquad\qquad \text{boys}$$
$$\lambda x[{}^*\text{boy}(x)]$$

$$\text{three} \qquad [\text{many}]$$
$$\langle n \rangle \qquad \langle n, et \rangle$$
$$3 \qquad \lambda n \lambda x[|x| = n]$$
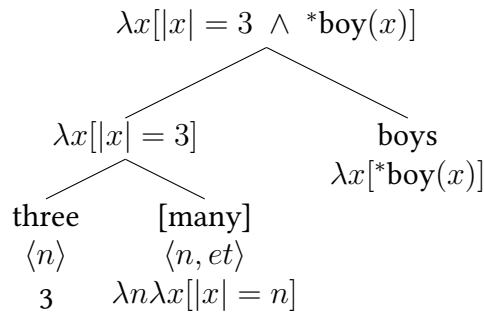
Figure 2.3: LF of the noun phrase *three liters*

$$\lambda d[\text{liters}(d) = 3]$$

$$\text{three} \qquad \text{liters}$$
$$\langle n \rangle \qquad \lambda n \lambda d[\text{liters}(d) = n]$$
$$3$$
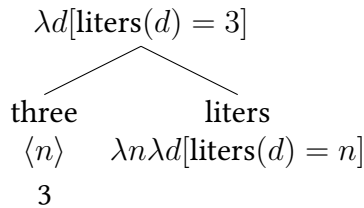
2.6 Nouns

In this section, I describe my take on the count-mass opposition, the singular-plural opposition, and the status of measure nouns and group nouns.

I analyze nouns as one-place predicates over entities, or equivalently, as denoting sets of entities, except for measure nouns, which I assume to denote unit functions. The entities in the denotations of most nouns are ordinary objects (Section 2.4.1), but I assume that derived nouns like *running* and, by extension, *three hours of running*, involve reference to events.

Following standard usage in the literature, I assume that the terms *mass noun* and *count noun* are defined with reference to morphological and syntactic properties rather than semantic properties. Mass nouns are compatible with the quantifiers *much* and *little* and reject quantifiers such as *each, every, several, a/an, some*

and numerals (Bunt 2006; Chierchia 2010). Most mass nouns are incompatible with plural morphemes, and in the absence of these morphemes only mass nouns are able to form noun phrases by themselves.

It is well known that many nouns can be used as count nouns or as mass nouns, as shown by example (45).

(45)   a.   Kim put an apple into the salad.
          b.   Kim put apple into the salad.

This raises the issue of whether the two occurrences of *apple* represent two different words or two different senses of one word. The differences between these positions are discussed in Pelletier and Schubert (2002). I assume that the two occurrences correspond to two different lexical entries, one singular count noun and one mass noun. The denotations of these entries are taken to have different formal properties, as described in the following sections. For example, the former entry only applies to mereological atoms, but the latter need not.

Altough my discussion is phrased in terms of nouns, similar considerations also apply to nominals, that is, noun phrases without determiners. The assumptions I make about the meanings of nouns carry over to nominals. This means that instead of talking about count and mass nouns, more exact terms would be count and mass expressions (Pelletier and Schubert 2002). However, for simplicity I use the term *nouns* rather than nominals or expressions.

2.6.1   Singular count nouns

I adopt the standard assumption that singular count nouns denote sets of individuals. For example, *table* denotes the set of all tables. I call the entities in the denotation of a count noun *singular individuals*. As described in Section 2.3.1, I assume that all singular individuals are pure mereological atoms. The notion of mereological part must be distinguished from the intuitive notion of part. For example, the leg of a table is not a mereological part of the table. Since all entities in the denotation of a count noun are atoms, all count nouns have quantized reference (Section 2.3.2).

This assumption is standard in accounts that are not primarily concerned with the count-mass distinction, but it is not self-evident (Kratzer 2010; Casati and Varzi 1999, pp. 112-115). It has been challenged by Zucchi and White (2001) on the basis of words like *twig*, *rock*, and *sequence*. Intuitively, a twig may have a part that is again a twig, a rock may have a part that is again a rock, and so on. Following my general stance described in Section 2.3.1, I assume that the relation between a twig and its part is not mereological parthood, just like we are not committed to the assumption

that the relation between John and his arm is mereological parthood. Regarding *twig*-type nouns, I assume that the meaning of these nouns is partially specified by context, and that when the context is fixed, each of these nouns denotes a quantized set. This assumption seems justified given that these nouns can be used with numerals and quantifiers just like other count nouns (*a twig, two rocks, every sequence*). An example of a formal implementation of this assumption can be found in Chierchia (2010), Section 5.2, with respect to the meaning of the similar noun *quantity*. For more discussion, see Rothstein (2010).

2.6.2 Plural count nouns

I assume that plural count nouns denote sets of entities and that a plural count noun is the algebraic closure of its singular counterpart:

(46) a. $[\![\text{cat}]\!] = \{c_1, c_2, c_3\}$
      b. $[\![\text{cats}]\!] = {}^*[\![\text{cat}]\!] = \{c_1, c_2, c_3, c_1 \oplus c_2, c_2 \oplus c_3, c_1 \oplus c_3, c_1 \oplus c_2 \oplus c_3\}$

These assumptions are common but not uncontroversial. Different analyses of plural count nouns result in different answers to the question of whether their denotations include singular entities and whether they include proper parts of singular entities. As we will see, dependent plurals pose a special challenge for any analysis, because they show that plural count nouns are interpreted differently depending on the polarity of their context.

In most mereological approaches to the semantics of count nouns, a singular noun denotes a quantized set, and a plural count noun denotes either the algebraic closure of the set denoted by its singular form, or a certain subset of this closure. This is a controversial point. The issue is whether the denotation of a plural noun also contains entities denoted by its singular form. In the terminology of Farkas and de Swart (2010), the two views on this question are termed exclusive and inclusive. Their own view is a combination of the two and can be called mixed. These three views can be summarized as follows.

- *Exclusive view:* Singular and plural forms of a count noun N denote disjoint sets. The plural form denotes the algebraic closure of its singular form with the singular individuals removed, as described by the equation in (47). The plural form N_{pl} essentially means the same as *two or more N*. For example, the plural count noun *apples* applies to sums consisting of two or more apples (Link 1983; Chierchia 1998a).

  (47) $[\![N_{pl}]\!] = {}^*[\![N_{sg}]\!] - [\![N_{sg}]\!]$

- *Inclusive view:* The plural form of a count noun N denotes a proper superset of the singular form, namely its algebraic closure, as described by the equation in (48). The plural form essentially means the same as *one or more N* (Krifka 1986; Sauerland 2003; Sauerland, Anderssen, and Yatsushiro 2005; Chierchia 2010). For example, the plural count noun *apples* applies to sums consisting of one or more apples. When singular reference is intended, singular and plural forms are in competition, and inclusive theories usually appeal to pragmatic notions in order to explain why the singular form blocks the plural form.

(48) $[\![N_{pl}]\!] = {}^*[\![N_{sg}]\!]$

- *Mixed view:* The plural form of a count noun N is ambiguous between the meanings it has on the exclusive and on the inclusive view (Farkas and de Swart 2010).

If singular count nouns are taken to apply to mereological atoms, the exclusive view holds that plural count nouns apply to proper sums, and the inclusive view holds that they apply both to atoms and to proper sums.

Environments such as downward-entailing contexts and questions present a challenge for the exclusive view because in these environments, plurals can be verified by singular individuals (for this argument and its sources, see Schwarzschild (1996), page 5). For example, suppose that there is only one doctor. On the exclusive view, the plural form *doctors* denotes the empty set, from which it is not possible to compositionally retrieve any information about the intended referent (the doctor). Therefore, the exclusive view has major problems deriving the meanings of sentences like (49) compositionally.

(49)   a.   No doctors are in the room.
         b.   Are there doctors in the room?

Bare plurals can occur in a particular configuration in which their semantic contribution cannot be modeled completely by either the inclusive view or by the exclusive view. In this configuration, which has been studied extensively by Zweig (2008, 2009), the bare plural occurs as a verbal argument and is c-commanded by another argument or adjunct of the same verb whose head noun is typically plural as well. Consider for example the bare plural *kites* in examples (50a-d), modeled after examples from Zweig.

(50)   a.   Boys flew kites.
         b.   Some boys flew kites.

    c.   Several boys flew kites.
    d.   Five boys flew kites.

These sentences have a reading in which the semantic contribution of the bare plural *kites* consists of two components. First, each of the boys in question flew one or more kites. I call this the *distributivity component.* Second, the total number of kites flown was two or more (though as Cohen (2005) argues, the total number implicated may be even higher than two). Loosely following Zweig, I call this the *multiplicity component.* Bare plurals which contribute these two components are called *dependent plurals* (de Mey 1981).

An inclusive analysis by itself would wrongly predict that *kites* should be synonymous with *one or more kites.* In (50), this prediction does not explain the multiplicity component. An exclusive analysis would predict that *kites* is synonymous with *two or more kites*, and it would therefore make the right predictions in sentences with dependent plurals like (50).[6] However, we have already seen above that an exclusive analysis is not viable for bare plurals in downward entailing contexts. As Zweig shows, these bare plurals never contribute a multiplicity component, even when they occur in the same syntactic position as dependent plurals. For example, the bare plural *kites* in sentences (51) is synonymous with *one or more kites*, because their subjects quantify over boys who flew one or more kites. This is contrary to the exclusive analysis, according to which they should quantify over boys who flew two or more kites.

(51)   a.   No boys flew kites.
        b.   Few boys flew kites.

In a nutshell, the puzzle is that sentences like (50) are incompatible with an inclusive analysis, while sentences like (49) and (51) are incompatible with an exclusive analysis. However, bare plurals are not simply ambiguous between an inclusive and an exclusive reading, because all these sentences are unambiguous.

I adopt Zweig's solution to this puzzle. In a nutshell, Zweig proposes that bare plurals are interpreted both inclusively and exclusively, and he splits their contribution into two parts. The truth-conditional meaning of a bare plural is inclusive and provides the distributivity contribution, and its exclusive part is a scalar implicature. Following Chierchia (2006), Zweig assumes that scalar implicatures can be added to truth-conditional meanings, but only if this strengthens the truth-conditions of the entire sentence. This is the case in upward-entailing

---

[6]Zweig assumes that the subjects of all sentences with dependent plurals are interpreted in situ and do not distribute over their verb phrases. That bare plurals are interpreted distributively in the sentences considered here is a result of the lexical semantics of the verb, not of a scopal mechanism. See Chapter 9.

contexts: for example, *five boys flew two or more kites* is stronger than *five boys flew one or more kites*, so the former meaning is retained. But it is not the case in downward-entailing contexts: *no boys flew two or more kites* is weaker than *no boys flew one or more kites*, so the latter meaning is retained. In this way, Zweig's account explains why bare plurals are interpreted exclusively in upward-entailing contexts and inclusively in downward-entailing contexts.

I return to dependent plurals in Chapter 9, where I extend Zweig's theory to *all* and *for*-adverbials.

2.6.3 Measure nouns

I assume that measure nouns like *liter*, *kilogram*, and *year* have a separate analysis from other count nouns, namely, they have the type $\langle n, dt \rangle$ (relations between degrees and numbers) or $\langle n, it \rangle$ (relations between intervals and numbers) and their denotation makes reference to unit functions like *liters* and *kilograms* in the sense of Section 2.5.4. I assume that singular and plural forms of measure nouns have identical denotations:

(52)   a. $[\![\text{liter}]\!] = [\![\text{liters}]\!] = \lambda n \lambda d[\text{liters}(d) = n]$
       b. $[\![\text{year}]\!] = [\![\text{years}]\!] = \lambda n \lambda t[\text{years}(t) = n]$

I conjecture that my analysis of measure nouns could also be extended to collection nouns like *bunch*, and that the ontological difference between individuals and unit functions might provide a handle on the problem of distinguishing collection nouns from group nouns. See Pearson (2009) and the next section.

In connection with my assumption that number words denote numbers (Section 2.4.6), this translation of measure nouns has the consequence that measure phrases like *three liters* have predicative type, similarly to other weak quantificational noun phrases such as *three boys* (Section 2.4.6). Different versions of the predicative analysis of measure phrases are defended in Zwarts (1997) and Schwarzschild (2006), among others.

Motivation for this analysis of measure nouns comes from the ambiguity between measure and individuating readings of container pseudopartitives (Rothstein 2009, and references therein):

(53)   three glasses of wine

       a.   *Measure reading:* a quantity of wine that corresponds to three glassfuls
       b.   *Individuating reading:* three actual glasses containing wine

Only the individuating reading, but not the measure reading, entails the existence of actual glasses. This fact can be explained if certain nouns like *glass* are

taken to be ambiguous between an ordinary and a measure noun interpretation. The measure noun interpretation can be taken to give rise to the measure reading and the latter to the individuating reading:

(54) three glasses of wine
    a. *Measure reading:* $\lambda x[\text{wine}(x) \ \wedge \ \text{glasses}(x) = 3]$
        (a quantity of wine that corresponds to three glassfuls)
    b. *Individuating reading:* $\lambda x[|x| = 3 \ \wedge \ ^*\text{glass}(x) \ \wedge \ \text{contains}(x, \text{wine})]$
        (three actual glasses containing wine)

In the individuating reading, *wine* appears as the argument of the relation *contains*. Rothstein treats *wine* as a kind in this case (see Section 2.4.2). Another possibility would be to treat it as the set of all wine entities or its characteristic function, corresponding to its occurrence in the measure reading. Since I ignore individuating readings, I leave this issue open.

In contrast to container pseudopartitives, measure pseudopartitives like *three liters of water* only have the measure reading, which indicates that the noun *liter* only has a measure noun interpretation. In this work, I concentrate on measure pseudopartitives, since I am not interested in individuating readings.

2.6.4   Group nouns

Group nouns are nouns like *committee, army,* and *league.* Barker (1992) defines group nouns as count nouns that can take an *of* phrase containing a plural complement but not a singular complement (*the group of armchairs/*armchair, a committee of woman/*women, an army of children/*child* etc.). By contrast, non-group nouns either take no *of* complements (**a book of page/pages*) or take *of*-complements which can be singular and/or plural (*a piece of cookie/*cookies, a picture of a horse/horses*).

Although at first sight, we might not be ready to view the individuals in the denotation of group nouns as atomic individuals, I assume that the entities in the denotation of singular group nouns are indeed mereological atoms, like other singular count nouns. This analysis is defended in Barker (1992), Schwarzschild (1996), and Winter (2001) against previous proposals such as Bennett (1974), in which group nouns involve reference to pluralities. I also follow Barker (1992) in assuming that there is no semantic operation that would allow us, for example, to recover the members of a committee from the atom in the denotation of the noun *committee*. In particular, the relation between a committee and its members is not the same as the relation between the sum and group interpretations of a plural definite like *the boys* (see Section 2.8). In terms of Section 2.3.1, I assume that the entities in the denotation of a group noun are pure atoms.

I use the name *group partitives* for the *of*-constructions that Barker uses in his test. Group partitives are superficially similar to pseudopartitives like *a liter of water* or *a box of cookies*. According to Barker's test, we might classify container words like *box* as group nouns (*a box of cookies/*cookie*). My account of pseudopartitives is not intended to apply to group partitives. I propose to distinguish them from each other as follows. Pseudopartitives based on container words like *three glasses of wine* are ambiguous between a measure reading, in which only the quantity of the wine is reported, and an individuating reading, which involves reference to actual glasses (53). Following Rothstein (2010), I assume that the two readings correspond to different syntactic configurations (see Section 2.6.3 for details). By contrast, group partitives only have an individuating reading (55).

(55)   a committee of women

    a.   *Measure reading (unavailable):* a set of women whose number corresponds to the membership of a committee

    b.   *Individuating reading:* an actual committee whose members are women

To avoid the problem of pseudopartitives, I propose to extend Barker's diagnostic for group nouns as follows: a group noun N is a count noun that can take an *of* phrase containing a plural complement but not a singular complement, *provided that the resulting interpretation entails the presence of actual Ns.* This prevents *box* from being incorrectly diagnosed as a group noun, because *a box of cookies* has a measure reading that does not entail the presence of an actual box.

This amended diagnostic faces a potential problem with nouns like *bunch* and *pile*, for whose referents it is difficult to establish independent existence criteria. As Pearson (2009, 2010) shows, these nouns behave differently from what we might call "true" group nouns like *committee* and *army*. Pearson calls the former *collection nouns* and the latter *committee nouns*. I conjecture that the *of*-constructions in which collection nouns appear are pseudopartitives and not group constructions. I leave open whether the predictions in this work extend to pseudopartitives with collection nouns.

2.6.5   Mass nouns

Following Bealer (1979) and Krifka (1991), we can distinguish three analyses of mass nouns. According to the *general term analysis*, the denotation of a mass noun is a set of entities. For example, on the general term analysis, the mass noun *gold* denotes the set of all gold entities, similarly to a count noun. According to the *singular term analysis*, the denotation of a mass noun is a sum. On the singular term analysis, *gold* denotes the sum of all gold entities, similarly to a proper

name. A variant of the singular term analysis is the *kind reference analysis*, on which the mass noun *gold* denotes the kind *gold* (see Section 2.4.2). Some authors also advocate a dual analysis according to which mass nouns are systematically ambiguous between a reading on which they denote sums or kinds, and a reading on which they denote sets.

I assume that all mass nouns can be translated as sets of entities. This is compatible with both the general term analysis and the dual analysis. I do not assume that mass terms have divisive reference, because of the problem of fake mass nouns like *furniture* (see Section 2.4.1). However, I assume that mass nouns that any entity in the denotation in a mass noun N that is larger than a certain threshold $\varepsilon(C)$ can be divided into parts which are again in the denotation of N. This idea is sketched in Krifka (1986) and can be formalized as follows:

(56)   **Definition: Approximate divisive reference**
       $\text{ADIV}_{\varepsilon(C)}(P) \overset{\text{def}}{=} \forall x[P(x) \; \wedge \; \varepsilon(C)(x) \rightarrow \forall y[y < x \; \wedge \; \varepsilon(C)(y) \rightarrow P(y)]]$
       (A predicate P is approximate divisive with respect to a threshold $\varepsilon(C)$ if and only if whenever it holds of something, it also holds of each of its proper parts, excluding entities below the threshold.)

I assume that the threshold is set depending on a context C, and that in the case of fake mass nouns it is large enough to exlude putatively atomic entities like pieces of furniture. Approximate divisive reference avoids the problem of minimal parts, which is treated in detail in Chapter 5.[7]

2.6.6   Kind-referring readings

So far, I have concentrated on the object-referring reading of nouns. Let me briefly comment on another reading, the so-called kind-referring reading (Krifka et al. 1995; Krifka 2004; Delfitto 2005). Both readings occur in bare NP uses of count and mass nouns and can be contrasted in the following example (Milsark 1974; Carlson 1989):

(57)   Typhoons arise in this part of the Pacific.
       a.   *Kind-referring reading:* It is a property of typhoons in general that they arise in this part of the Pacific.
       b.   *Object-referring reading:* This part of the Pacific has the property that some typhoons arise in it.

---

[7]Anticipating the discussion in that chapter, if a mass noun has approximate divisive reference with respect to a threshold $\varepsilon(C)$, then for any dimension d, it has stratified reference with respect to $\varepsilon(C)$ and d.

Existing analyses of the kind-object opposition can be ordered in three categories, similarly to the different analyses of mass nouns (Krifka 1991; Krifka et al. 1995; Lasersohn 2010): a) all bare NPs denote predicates over entities; b) all bare NPs denote special entities, such as maximal sums or kinds; c) kind-referring NPs denote entities, object-referring NPs denote sets of entities, and the two are related by type-shifting operators. The ambiguity in (57) suggests that two uses of NPs must be distinguished semantically, as expected on analysis c).

For the purpose of this work, in analogy to my assumptions about mass nouns, I assume that all plural count nouns can be translated as sets of entities. This is compatible both with analyses b) and c).

The question of how to model the kind-object opposition touches on my account of substance nouns in pseudopartitives and of dependent plurals. I assume that substance nouns in pseudopartitives are object-referring, since on this view they denote sets, which makes the formulation of strata theory easier. This is perhaps the prevailing view, though not the only one: a kind-referring analysis of pseudopartitives is found in Ionin, Matushansky, and Ruys (2006). In the case of dependent plurals, I follow Zweig (2008, 2009) in adopting an object-referring analysis, as described above.

2.7 Verbs

2.7.1 The Neo-Davidsonian position and its alternatives

There is no agreement in the formal semantic literature on how to represent the meanings of verbs. Early work, as well as some modern authors, simply represents the meaning of a verb with n syntactic arguments as an n-ary relation. A transitive verb, for example, is assumed to denote a two-place relation. Against this, Davidson (1967) argued that verbs denote relations between events (see Section 2.4.3) and their arguments, so that a transitive verb denotes a three-place relation. Once events have been introduced, we can express the relationship between events and their arguments by thematic roles (see Section 2.5.1). This is the so-called neo-Davidsonian position (e.g. Carlson 1984; Parsons 1990; Schein 1993). Finally, Kratzer (2000) argues for an asymmetric position, according to which only agents are represented as thematic roles.[8] The positions are illustrated in Table 2.3.

For the purpose of this work, I adopt the neo-Davidsonian position and I assume that verbs denote sets of events. The motivation for my choice is technical.

---

[8]I do not think that Kratzer's arguments for the asymmetric position are convincing, because I suspect that the asymmetries she tries to model correlate with syntactic positions and not with thematic roles. See Champollion (2010) for details.

Table 2.3: Approaches to verbal denotations

Position	Verbal denotation	Example: Brutus stabbed Caesar
Traditional	$\lambda y \lambda x [\text{stab}(x,y)]$	$\text{stab}(b,c)$
Classical Davidsonian	$\lambda y \lambda x \lambda e [\text{stab}(e,x,y)]$	$\exists e [\text{stab}(e,b,c)]$
Neo-Davidsonian	$\lambda e [\text{stab}(e)]$	$\exists e [\text{stab}(e) \wedge \text{ag}(e,b) \wedge \text{th}(e,c)]$
Asymmetric	$\lambda y \lambda e [\text{stab}(e,y)]$	$\exists e [\text{ag}(e,b) \wedge \text{stab}(e,c)]$

It is easier to state generalizations across the categories of nouns and verbs if these categories both denote sets, and constraints on thematic roles can be stated more easily if these roles are reified. This is not to say that this work could not be reformulated in one of the other frameworks. Such a reformulation would probably not be perspicuous, though. See Bayer (1997) for a study of what it would take to reformulate a similar framework into eventless semantics.

Since the identity of thematic roles is not important for my purposes, I make the simplifying assumption that the subject of a sentence in active voice is always its agent, as mentioned in Section 2.5.1. Although I assume that the predicates denoted by verbs range over events and not over individuals, I say that a verb "applies" to an individual as a shorthand for stating that the individual is the agent of an event to which the verb applies. This is simply a matter of convenience and should not be confused with adopting a traditional or classical Davidsonian position.

2.7.2 Lexical cumulativity

In Section 2.6.1, I called individuals in the denotation of singular nouns *singular individuals*. Similarly, I call an event whose thematic roles map it exclusively to singular individuals a *singular event*, and an event that is not a singular event a *plural event*. Choosing these terms allows me to sidestep the question of whether singular events should be represented as mereological atoms. Different views exist on this question. Landman (2000) assumes that all singular events are atoms because this allows him to treat distributivity and plurality as the same thing. For Krifka (1998), this is not the case. For example, the event in which John reads a certain book has proper parts in which John reads part of the book. In fact, Krifka leaves open whether atomic events exist at all. (The discrepancy between the two accounts is likely related to the fact that Landman does not try to model the aspectual phenomena that motivate Krifka's assumptions.) I take Krifka's view

and make no assumptions about whether singular events are atomic.

In mereological event semantics, the sum of any two events is itself an event (see Section 2.4.3). In general, the sum of two singular events is a plural event. For example, let e_1 be the event in which John (j) lifts a certain box b and e_2 the event in which Mary (m) lifts a certain table t. The sum $e_1 \oplus e_2$ is itself an event. The agent of e_1 is j and the agent of e_2 is m. Since I assume that thematic roles are sum homomorphisms (see Section 2.5.1), the agent of the sum event $e_1 \oplus e_2$ is $j \oplus m$, the sum of their agents. Since entity $j \oplus m$ is not a singular individual, the event $e \oplus e_2$ is a plural event.

While it might seem intuitive to assume that verbs only apply to singular events, I take the opposite view: verbs may apply both to singular and to plural events. More specifically, whenever two events are in the denotation of a verb, so is their sum. This is a common and well-motivated assumption (Scha 1981; Schein 1986, 1993; Lasersohn 1989; Krifka 1986, 1992; Landman 1996, 2000). Following Kratzer (2007), I call this assumption *lexical cumulativity*. Within a Neo-Davidsonian framework, lexical cumulativity is complemented by the analogous assumption that all thematic roles are closed under sum formation (see Section 2.5.1). In the context of other approaches to verbal denotations where verbs denote relations, lexical cumulativity can be implemented as assuming that these relations are closed under pointwise sum formation (see Section 2.3.1).

In the previous example, lexical cumulativity has the consequence that the verb *lift* applies not only to the event e_1 and to the event e_2, but also to their sum $e_1 \oplus e_2$. In general, verbs can be said to have plural denotations, in the sense that their denotation obeys the same equation (58) as plural count nouns on the inclusive view (48), repeated in (59):

(58) $[\![V]\!] = {}^*[\![V]\!]$

(59) $[\![N_{pl}]\!] = {}^*[\![N_{sg}]\!]$

Following Kratzer (2007), I include the algebraic closure operator in the typographical representation of verb meanings as a reminder of the lexical cumulativity assumption. For example, instead of writing $\lambda e[\text{lift}(e)]$ for the meaning of the verb *lift*, I write $\lambda e[^*\text{lift}(e)]$. I do the same for thematic roles.

The lexical cumulativity assumption is motivated by the entailments in (60) and (61) (Krifka 1989, 1992). Because of the parallelism between (58) and (59), the explanation of these entailments is completely analogous to the explanation of the entailment in (62), which motivated the treatment of plurality in Link (1983).

(60)   a.   John slept.
        b.   Mary slept.

  c. ⇒ John and Mary slept.

(61) a. John lifted box b.
   b. Mary lifted table t.
   c. ⇒ John and Mary lifted box b and table t.

(62) a. John is a boy.
   b. Bill is a boy.
   c. ⇒ John and Bill are boys.

As Kratzer (2007) notes, another motivation for lexical cumulativity comes from iterative interpretations of verbs. In English, semelfactive verbs such as *cough* and achievements such as *find* are systematically ambiguous or underspecified between a "punctual" and an iterative sense. (I use the term "punctual" with caution because I do not want to suggest that it takes no time at all to cough or to find something. Instants are not part of my temporal ontology anyway, see Section 2.4.4.) For example, the sentence *John coughed* can be understood as saying that John coughed once, or that he coughed several times (Carlson 2006), and the sentence *John found fleas* can be understood as saying that there was one punctual event in which John found some fleas, or that he found them over an ongoing period of time, possibly one by one. The lexical plurality assumption means that a predicate like $\lambda e[\text{*cough}(e) \wedge \text{*ag}(e) = j]$ applies both to singular and to plural coughing events. This fact captures the systematic ambiguity of semelfactive predicates.

Iterative interpretations of semelfactives typically involve an alternation of times at which the predicate is true and times at which it is false. The lexical cumulativity assumption does not semantically entail this alternation, because sum events with continuous runtimes fall under the denotation of the event predicate as well as sum events with discontinuous runtimes. This is justified if the alternation is a pragmatic implicature, as opposed to a semantic entailment, as argued by Egg (1995). For more on iterativity, see Lasersohn (1995) and van Geenhoven (2004). These authors model iterativity by dedicated verb-level iterativity operators, but these operators are more difficult to motivate independently than lexical cumulativity (Kratzer 2007). Since iterativity appears to be possible with all English predicates, an operatorless approach is ceteris paribus preferable to an operator-based approach, as Kratzer notes.

While lexical cumulativity does not entail that all verb phrases have cumulative reference (for example, the sum of two events in the denotation of the verb phrase *carry exactly two pianos* is not again in its denotation, because it involves four rather than two pianos), there are some verb phrases that do obtain cumulative reference as a consequence of lexical cumulativity. For example, lexical cumulativity has

the effect that the lexical predicate *carry* has cumulative reference, and applies not only to singular carrying events but also to sums of such events. The effect of this assumption on the denotation of the verb phrase *carry the piano* is that it denotes a predicate which applies not only to singular carrying events whose theme is the piano, but also to sums of such events (63a). As a consequence, the entire phrasal predicate *carry the piano* has cumulative reference, and it has the same denotation as it would have if the star operator were applied to the entire phrasal predicate (63b). In other words, (63a) and (63b) are equivalent.

(63) a. $\lambda e[^*\mathrm{carry}(e) \,\wedge\, {}^*\mathrm{th}(e) = \iota x.\mathrm{piano}(x)]$
      b. $^*\lambda e[\mathrm{carry}(e) \,\wedge\, {}^*\mathrm{th}(e) = \iota x.\mathrm{piano}(x)]$

The concept of lexical cumulativity is a theoretical assumption which is implemented by adopting the premise that every verb denotes a predicate that satisfies cumulative reference. This implementation of lexical cumulativity affects the denotations of all verbs without exception. A consequence of this assumption is that is not possible to use cumulative reference to model the telic-atelic opposition, as is sometimes proposed (e.g. Egg 1995; Zwarts 2005). Lexical cumulativity entails that even certain telic predicates have cumulative reference, such as achievement verbs with definite objects. For example, *reach the summit* is telic but cumulative.

2.8   Noun phrases

I assume a treatment of noun phrases along the lines of Landman (1996, 2000). In Landman's account, definite and indefinite noun phrases are ambiguous between sum and group interpretations. This ambiguity is motivated by cumulative and collective readings.

Cumulative readings involve two plural entities A and B and a relation R that holds between the members of these plural entities in a certain way. In canonical examples of cumulative readings, A and B are introduced by two plural definite or indefinite arguments of a verb that is distributive on both these arguments, and R is introduced by this verb. A cumulative reading with a distributive predicate licenses the inference that the relation R relates each singular individual in A to at least one singular individual in B, and vice versa. For example, (64a) involves reference to a plural individual consisting of 600 Dutch firms (A), another one consisting of 5000 American computers (B), and a relation of using (R). The most prominent reading of (64a), paraphrased in (64b), is a cumulative reading.

(64) a. 600 Dutch firms use 5000 American computers.   (Scha 1981)
      b. 600 Dutch firms each use at least one American computer and 5000

American computers are each used by at least one Dutch firm.

Cumulative readings are modeled as scopeless relations: Even when the plural entities A and B are introduced by scope-taking elements such as indefinites, none of them takes scope over the other. For example, the cumulative reading of Sentence (64a) is modeled as follows:

(65) $\exists e[^*\text{use}(e) \wedge {}^*\text{dutch.firm}(^*\text{ag}(e)) \wedge |^*\text{ag}(e)| = 600 \wedge$
$^*\text{american.computer}(^*\text{th}(e)) \wedge |^*\text{th}(e)| = 5000]$

This representation is equivalent to the representation in (66) given the following assumptions: verbs and thematic roles are closed under sum (see Sections 2.7.2 and 2.5.1), and *own* is distributive down to atoms on both its agent and theme arguments (see Section 4.5). While the representation in (66) is probably more readable, the one in (65) is easier to derive compositionally (see Section 2.10) because it keeps the two arguments neatly apart. This observation goes back to Schein (1986) and Krifka (1986), and is developed in detail in Krifka (1999) and Landman (2000).

(66) $\exists x \exists y[^*\text{dutch.firm}(x) \wedge |x| = 600 \wedge {}^*\text{american.computer}(y) \wedge |y| = 5000$
$\wedge \forall x' \leq_{Atom} x \,\exists y' \leq_{Atom} y \,\exists e[\text{use}(e) \wedge \text{ag}(e) = x' \wedge \text{th}(e) = y']$
$\wedge \forall y'' \leq_{Atom} y \,\exists x'' \leq_{Atom} x \,\exists e[\text{use}(e) \wedge \text{ag}(e) = x'' \wedge \text{th}(e) = y'']]$

I assume that cumulative readings are a separate kind of interpretation and must be distinguished from collective readings (cf. Landman 1996, 2000). The term *collectivity* is understood in this section in the sense of what I call *thematic collectivity* in Section 9.5.1. Instances of thematic collectivity are characterized by non-inductive generalizations. Landman's cases of collectivity all represent thematic collectivity.

Typical examples of collective readings involve a noun phrase in agent position and entail that the group to which the noun phrase refers bears collective responsibility for the event, above and beyond the individual responsibilities of its members. Typical examples of cumulative readings do not give rise to this entailment.

For example, (67) only has a collective reading, not a cumulative reading. The interpretation of this sentence is not cumulative because it does not entail that each cowboy sent an emissary to some Indian, nor that each Indian was sent an emissary, and it is collective because the cowboys bear collective responsibility for the action of sending an emissary.

(67) The cowboys sent an emissary to the Indians.

By contrast, example (68) only has a cumulative reading, not a collective

reading, firstly because it does entail that each man in the room is married to a girl across the hall and vice versa, and secondly because there is no sense in which the men have collective responsibility for being married to the girls above and beyond their individual responsibilities.

(68)   The men in the room are married to the girls across the hall. (Kroch 1974)

Collective readings also occur when there is only one noun phrase in the sentence:

(69)   Three boys carried a piano upstairs.

Following Landman (2000), I assume that this sentence has the following readings and representations:

(70)   Different-pianos distributive reading:
$\exists x[{}^*\text{boy}(x) \wedge |x| = 3 \wedge \forall y \leq_{Atom} x \exists e[{}^*\text{carry.upstairs}(e) \wedge {}^*\text{ag}(e) = y \wedge \text{piano}({}^*\text{th}(e))]]$
(Three boys each carried a piano upstairs.)

(71)   Same-piano distributive reading:
$\exists e[{}^*\text{carry.upstairs}(e) \wedge {}^*\text{boy}({}^*\text{ag}(e)) \wedge |{}^*\text{ag}(e)| = 3 \wedge \text{piano}({}^*\text{th}(e))]$
(There is a piano that three boys each carried upstairs.)

(72)   Collective reading:
$\exists e[{}^*\text{carry.upstairs}(e) \wedge \exists x[{}^*\text{boy}(x) \wedge |x| = 3 \wedge {}^*\text{ag}(e) =\uparrow (x) \wedge \text{piano}({}^*\text{th}(e))]]$
(There is a piano that three boys, as a group, carried upstairs.)

These representations rely on the assumption that *carry.upstairs* is distributive on at least its agent predicate. That is, whenever a sum of indidivuals is the agent of a carrying-upstairs event e, its mereological parts are the agents of carrying-upstairs events that are parts of e. The notion *distributive predicate* is taken up again in Section 4.5.

The collective reading blocks this entailment by relating the piano to an atom that represents the boys as a group. Following Landman (2000), I represent this reading by shifting the sum individual x that represents three boys to a group individual $\uparrow (x)$ that represents these three boys, taken as a group (see Section 2.3.1). Since this group is taken to be a mereological atom, the relation between a group and its members is distinct from the mereological relation between a proper sum and its parts. Following Landman, I call atoms "impure" when they are formed through the group formation operator. As described in Section 2.3.1, I assume that atomicity checks can distinguish between pure and impure atoms. I assume that

55

the atoms in the denotation of singular count nouns are all pure atoms (see Section 2.6.1).

To derive the readings above, I assume that definites and indefinites are ambiguous between a sum and a group interpretation. Some sample translations of verb phrases with definites and indefinites are given below (see Section 2.10 for the compositional process that underlies them). The $\bigoplus$ notation for plural definite descriptions is equivalent to the alternative notation that uses σ, as explained in Section 2.3.1.

(73)   a. $[\![\text{a boy}]\!] = [\![\text{boy}]\!] = \lambda x[\text{boy}(x)]$
          b. $[\![\text{boys}]\!] = \lambda x[{}^*\text{boy}(x)]$
          c. $[\![\text{three boys}_{sum}]\!] = \lambda x[{}^*\text{boy}(x)] \; \wedge \; |x| = 3$
          d. $[\![\text{three boys}_{group}]\!] = \lambda x \exists y[x =\uparrow (y) \; \wedge \; {}^*\text{boy}(y) \; \wedge \; |y| = 3]$
          e. $[\![\text{the boys}_{sum}]\!] = \bigoplus \text{boy}$
          f. $[\![\text{the boys}_{group}]\!] = \uparrow (\bigoplus \text{boy})$

As these translations illustrate, I assume that numerals have an *exactly* interpretation. As discussed in Section 2.5.5, this assumption is not crucial because in this work, I only consider upward-entailing quantifiers anyway. The generalization to non-upward entailing quantifiers is compatible with the approach presented here but requires additional technical devices such as maximization operators (Landman 2000; Krifka 1999; Brasoveanu 2010).

Example (74) illustrates the standard translation of quantificational noun phrases, which I adopt.

(74) $[\![\text{every boy}]\!] = \lambda P \forall x[\text{boy}(x) \rightarrow P(x)]$

There are two different quantifying-in operations that can apply to definites and indefinites. NQI, or non-scopal quantifying-in, corresponds to classical Montague quantifying-in or quantifier raising. SQI or scopal quantifying-in gives a sum noun phrase wide scope and ensures that it is interpreted distributively by replacing the plural entity introduced by a noun phrase with a quantifier over every one of its atoms. (Applying SQI to a group noun phrase has no effect. If the noun phrase introduces a group, this is vacuous, since the only atom of a group is the group itself.) SQI is used to produce the different-pianos distributive reading (70). In both cases, the raised quantifier is replaced by a trace of type e. For more details, see Landman (1996, 2000).

Section 2.4.6 gives a compositional derivation of a numeral noun phrase on its sum interpretation. I assume that the other types of noun phrases have similar derivations. For example, the group interpretation can be obtained through an optional type shifting mechanism. The details do not matter. See Landman (2000)

for a concrete proposal.

2.9   Prepositional phrases

There is only one prepositional phrase that plays a role in this work, namely the directional prepositional phrase *to the store*. I assume a semantic representation in the style of Zwarts (2006) that involves a directed spatial interval or path whose end is (the location of) the store.

(75)   [[to the store]] = $\lambda V_{\langle vt \rangle} \lambda e[\text{end}(\sigma(e)) = \text{the.store}]$

See also Hinrichs (1985), Krifka (1998), Zwarts (2005, 2006) for the treatment of paths in aspectual composition and Section 2.4.4 for the notion of spatial interval.

2.10   The compositional process

Following Parsons (1990) and Krifka (1992), I assume that verbs and verbal projections denote sets of events, type $\langle vt \rangle$ (Section 2.7.1). For noun phrases, as described in Section 2.8, I follow the standard assumption that they can have individual type (e), predicative type $\langle et \rangle$ or quantificational type $\langle et, t \rangle$. I assume that the application of a predicative or referential noun phrase to a verbal projection amounts to intersecting two sets of events. This idea goes back at least to Carlson (1984). For example, the interpretation of *John loves Mary* amounts to intersecting the set of events whose agent is John, the set of loving events, and the set of events whose theme is Mary. At the end of the derivation, a sentence mood operator applies. Since I only consider declarative sentences, I make the standard assumption that this operator only binds the event variable with an existential quantifier. I call this operator existential closure or simply [∃].

Following Landman (2000), I assume that referential and predicative noun phrases can be interpreted in situ or by quantifier raising (see Section 2.8). The latter option can also account for distributive readings of these noun phrases, given certain assumptions about the quantifier raising operation. Since the relevant readings do not play a role in this work, I refer the reader to Landman (2000) for details. Quantificational noun phrases like *every boy* could also be interpreted in situ if we make certain adjustments to their translations and/or to the compositional process (Champollion 2010). However, I prefer to follow Landman (2000) in assuming that they always undergo quantifier raising. This assumption is necessary for the account of dependent plurals as formulated by Zweig (2008, 2009), on which Chapter 9 relies. As for the noun phrase *all the boys*, I assume in that chapter that

it has referential rather than quantificational type, and is interpreted in situ.

Implementing Carlson's idea requires shifting the types of verbal projections and noun phrases. There does not seem to be a best practice of how to resolve this kind of type mismatch; instead, authors follow their own conventions. For example, Kratzer (1996), who only uses the thematic role corresponding to the external argument, assumes an operation of "event identification" that connects the thematic role with the verb phrase, while Landman (2000) assumes that the compositional process can lift the type of verbal projections so that they combine with noun phrases through function application.

I assume, similarly to Landman, that the type mismatch is resolved by type shifters. My setup is different from his because he assumes neither that verbs denote sets of events, nor that thematic roles are introduced by separate lexical entries. I assume that the following type shifters can apply freely to thematic roles, verbal projections, and noun phrases:

(76)   Predicative type shifters:

   a.   VP, then NP: $\lambda\theta_{\langle ve \rangle}\lambda V_{\langle vt \rangle}\lambda P_{\langle et \rangle}\lambda e[V(e) \ \wedge \ P(\theta(e))]$
   b.   NP, then VP: $\lambda\theta_{\langle ve \rangle}\lambda P_{\langle et \rangle}\lambda V_{\langle vt \rangle}\lambda e[V(e) \ \wedge \ P(\theta(e))]$

(77)   Referential type shifters:

   a.   VP, then NP: $\lambda\theta_{\langle ve \rangle}\lambda V_{\langle vt \rangle}\lambda x_{\langle e \rangle}\lambda e[V(e) \ \wedge \ \theta(e) = x]$
   b.   NP, then VP: $\lambda\theta_{\langle ve \rangle}\lambda x_{\langle e \rangle}\lambda V_{\langle vt \rangle}\lambda e[V(e) \ \wedge \ \theta(e) = x]$

For concreteness, I assume that these type shifters always apply first to the thematic role, and then in any order to the verbal projection and to the noun phrase. This flexible order allows me to remain noncommittal about whether noun phrases c-command thematic role heads or are c-commanded by them. The former assumption is generally made for the little-v head, which I represent as [ag], while the latter assumption makes thematic role heads analogous to silent prepositions. These choices are not essential for the present work.

The compositional process is illustrated in Figure 2.4. The sentence used here is *John and Mary saw thirty zebras*. Based on our background assumptions, the LF generates the cumulative reading of this sentence: there is a sum of seeing events whose agents sum up to John and Mary and whose themes sum up to thirty zebras. This reading does not entail that each zebra was seen by both John and Mary, nor that the zebras were seen simultaneously. I return to this point in Chapter 4.

The same LF could also model the sentence *John and Mary looked at thirty zebras* if we assume that *see* and *look* have the same lexical entries and that *at* is the overt counterpart of [th]. This is the main motivation for the syntactic position of [th].

58

Figure 2.4: The compositional process
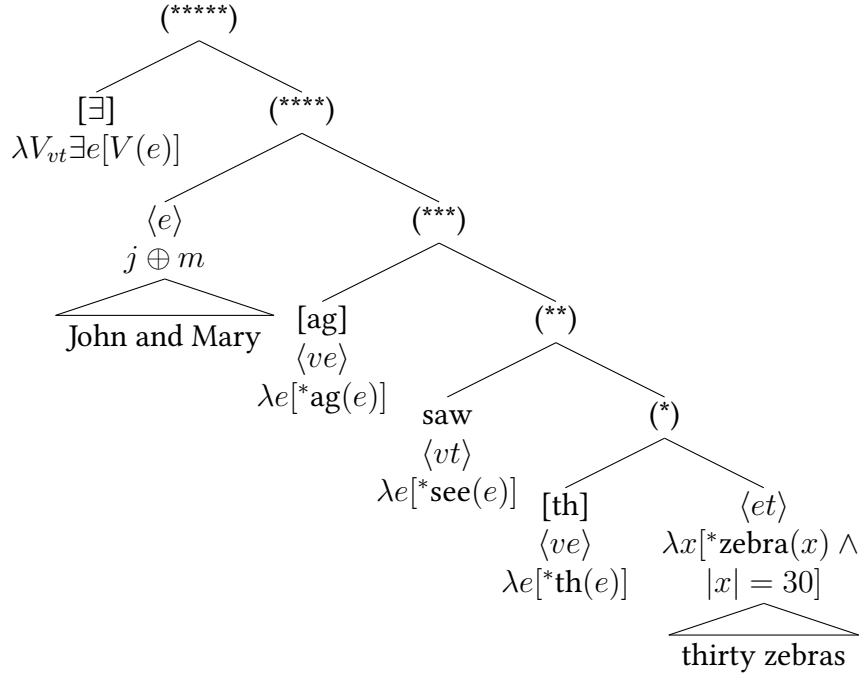
(*) Predicative type shifter, NP first: $\lambda\theta_{\langle ve\rangle}\lambda P_{\langle et\rangle}\lambda V_{\langle vt\rangle}\lambda e[V(e)\ \wedge\ P(\theta(e))]$
  Result: $\lambda V_{\langle vt\rangle}\lambda e[^*\text{zebra}(^*\text{th}(e))\ \wedge\ |^*\text{th}(e)|=30]$

(**) Function application
  Result: $\lambda e[^*\text{see}(e)\ \wedge\ {}^*\text{zebra}(^*\text{th}(e))\ \wedge\ |^*\text{th}(e)|=30]$

(***) Referential type shifter, VP first: $\lambda\theta_{\langle ve\rangle}\lambda V_{\langle vt\rangle}\lambda x_{\langle e\rangle}\lambda e[V(e)\ \wedge\ \theta(e)=x]$
  Result: $\lambda x\lambda e[^*\text{ag}(e)=x\ \wedge\ {}^*\text{zebra}(^*\text{th}(e))\ \wedge\ |^*\text{th}(e)|=30]$

(****) Function application
  Result: $\lambda e[^*\text{see}(e)\ \wedge\ {}^*\text{ag}(e)=j\oplus m\ \wedge\ {}^*\text{zebra}(^*\text{th}(e))\ \wedge\ |^*\text{th}(e)|=30]$

(****) Existential closure
  Result: $\exists e[^*\text{see}(e)\ \wedge\ {}^*\text{ag}(e)=j\oplus m\ \wedge\ {}^*\text{zebra}(^*\text{th}(e))\ \wedge\ |^*\text{th}(e)|=30]$

Chapter 3

The cast of characters

3.1 Introduction

Three constructions are at the center of this work: *for*-adverbials, pseudopartitives, and constructions with *each*.

(1)   *For*-adverbials:

    a.   John **ran** for five minutes.   *atelic*

    b.   *John **ran to the store** for five minutes.   **telic*

(2)   Pseudopartitives:

    a.   thirty pounds of **books**   *plural*

    b.   thirty liters of **water**   *mass*

    c.   *thirty pounds of **book**   **singular*

(3)   Constructions with *each*:

    a.   The boys each **walked**.   *distributive*

    b.   *The boys each **met**.   **collective*

As mentioned in Chapter 1, these constructions all have a constituent which is subject to a certain restriction. The predicate modified by a *for*-adverbial must be atelic, the substance noun of a pseudopartitive must be either mass or plural, and the verb phrase of an *each*-construction must be distributive. In the examples above, I have marked these constituents in bold.

In this chapter, I present a plausible baseline theory for the syntax and semantics of these constructions and their constituents, relying on the assumptions in Chapter 2. I keep things symmetric across domains as much as seems reasonable, so that the parallels which I draw in the subsequent chapters are not obscured more than necessary. More refined theories might require giving up the symmetries I assume

here. Doing so should not affect the insights underlying the theory presented in the subsequent chapters, but it might make it necessary to reformulate the technical implementation in a less conspicuous way.

3.2 Pseudopartitives

Pseudopartitives, also called measure constructions, are noun phrases that involve reference to an amount of some substance. Both the amount and the substance involved are specified with a noun; in English, the nouns are separated by the word *of*:

(4)   three liters of drinkable water

The term *pseudopartitive* was introduced by Selkirk (1977) to distinguish this construction from true partitives such as *three liters of the water*. I refer to the noun that comes to the right of *of* as the *substance noun*, and to the substance noun together with any of its modifiers the *substance nominal*. The noun *liters* in this example is a *measure noun* (see Section 2.6.3). In example (4), the substance noun is *water*, and the substance nominal is *drinkable water*. As explained below, I treat *three liters* as a constituent, namely a *measure phrase*.

Pseudopartitives with measure nouns are called *measure pseudopartitives*. Pseudopartitives can also be formed in other ways, for example with container nouns (*three glasses of wine*) or classifier nouns (*three heads of cattle*). For a discussion of the syntactic and semantic properties of the different kinds of pseudopartitives, see Keizer (2007), Chapter 6. Group nouns like *committee* can head a construction that is at least superficially similar to pseudopartitives (*a committee of women*). As discussed in Section 2.6.4, I do not consider this construction a pseudopartitive.

The claims in this work only concern the measure reading of pseudopartitives. Pseudopartitives with measure nouns only have a measure reading; pseudopartitives with container nouns are ambiguous between a measure reading and an individuating reading (see Section 2.6.3). As the term *measure phrase* suggests, I assume that the measure noun of a measure pseudopartitive forms a constituent with the determiner that precedes it (5a). An alternative to this analysis would involve the right-branching structure (5b). I adopt (5a) to reflect the semantic parallel between pseudopartitives and other distributive constructions (see Table 4.3). However, it would not be difficult to reformulate this work in a way that is consistent with structure (5b). My compositional implementation in Section 4.7 places the essential machinery into the lexical entry of *of*. The only relevant difference between the two structures is whether *of* combines with *two* and *pounds* at once or one at a time. It is an easy technical exercise to rewrite the lexical entry

of *of* to mirror this difference.

(5)   a.   two   pounds   of   tomatoes   b.   two   pounds   of   tomatoes
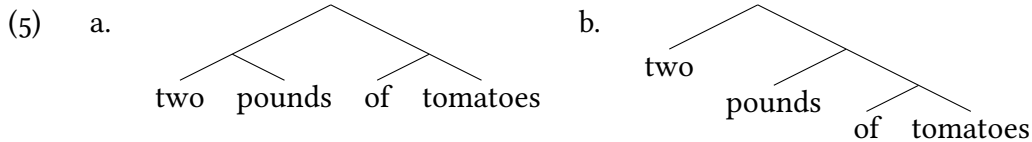
The following coordination test shows that my syntactic assumption is at least initially plausible.[9] While the string *two pounds* can be coordinated without problems (6), many people reject sentences in which the string *pounds of tomatoes* is part of a coordination structure (7).

(6)   John bought two pounds and two ounces of tomatoes.

(7)   *John bought two pounds of tomatoes and grams of saffron.

More arguments in favor of the symmetric structure are found in Gawron (2002). However, the choice between the two structures is by no means settled. The symmetric structure is also adopted by Akmajian and Lehrer (1976), Guéron (1979), Gawron (2002), Schwarzschild (2002, 2006), while the right-branching structure is adopted by Stickney (2008), Chierchia (2008), Bale (2009), among others. Landman (2004) and Rothstein (2009) make the plausible claim that the symmetric structure corresponds to the measure reading and the right-branching structure to the individuating reading.

I represent the meaning of a measure pseudopartitive as follows:

(8)   three liters of water
       $\lambda x[\text{water}(x) \wedge \text{liters}(\text{volume}(x)) = 3]$

I write $\text{liters}(\text{volume}(x)) = 3$ and not simply $\text{liters}(x) = 3$ because I assume that there is a layer of degrees that mediates between substances and natural numbers (Section 2.4.5). I use the term *measure function* for functions from entities to degrees, and the term *unit function* for functions from degrees to numbers (see Sections 2.5.3 and 2.5.4).

To derive the representation in (8) compositionally, I assume the lexical entries and the LF in Figure 3.1. I represent the source of the measure function *volume* as a silent lexical item, which I write [volume]. This is one possible way to express the fact that the overt part of a pseudopartitive does not always uniquely specify this measure function (see Section 2.5.4). It would also be possible to introduce the measure function as a free variable. As far as I can tell, this choice does not interact with any other assumptions. The position of [volume] in the tree does not

---

[9]I thank Alan Bale for discussing this issue with me.

matter either.

Figure 3.1: Skeletal LF of an ordinary pseudopartitive

$$\langle et \rangle$$
$$\lambda x[\text{water}(x) \ \wedge \ \text{liters}(\text{volume}(x)) = 3]$$

three $\langle n \rangle$ — 3

liters $\langle n, dt \rangle$ — $\lambda n \lambda d[\text{liters}(d) = n]$

[volume] $\langle ed \rangle$ — $\lambda x[\text{volume}(x)]$

of $\langle \alpha t, \langle \alpha \beta, \langle \beta t, \alpha t \rangle \rangle \rangle$ — $\lambda S \lambda M \lambda K \lambda b$ $[S(b) \ \wedge \ K(M(b))]$

water $\langle et \rangle$ — $\lambda x[\text{water}(x)]$
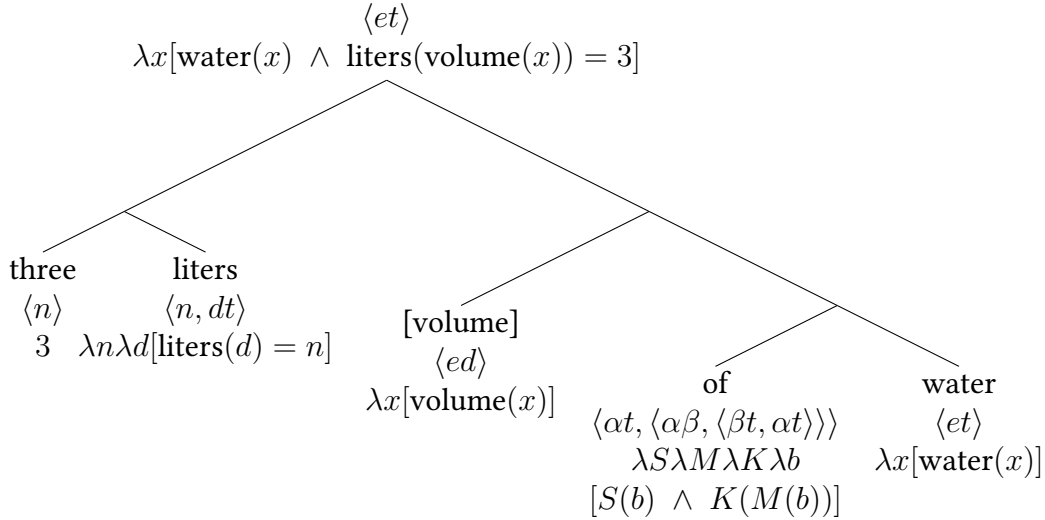
The arguments of the translation of *of* are represented with the letters S, M, K, and b. The uppercase letters can be taken as standing for predicates of arbitrary predicative type here. In Section 4.4, they resurface as mnemonics for the terms Share, Map, and Key. The letter b is a variable over entities of any basic type. The type of *of* is polymorphic because I assume that events rather than ordinary objects and intervals rather than degrees can also be used. The Greek letters α and β stand for arbitrary types. In this particular LF, α is resolved to e and β to d. But this may be different in other cases. For example, I assume that a pseudopartitive like *three hours of walking* denotes a set of events rather than individuals, and that the domain of the unit function introduced by the word *hours* contains intervals rather than degrees (see Section 2.5.4).

Figure 3.2 shows an LF for *three hours of walking*. The star in front of the translation of *walk* is a reminder that the denotations of verbs are assumed to be their own algebraic closures, see Section 2.7.2.

This type of representation is an incomplete skeleton. Alone, it does not provide any explanation for the constraints on pseudopartitives described in Chapters 1 and 7, because it does not contain any constraints itself. For example, there is no explanation why *three pounds of book* cannot be interpreted as follows:

(9)   three pounds of book
        $\lambda x[\text{book}(x) \ \wedge \ \text{pounds}(\text{weight}(x)) = 3]$

Figure 3.2: Skeletal LF of an event pseudopartitive

$\langle vt \rangle$
$\lambda e[^*\text{walk}(e) \;\wedge\; \text{hours}(\tau(e)) = 3]$

three
$\langle n \rangle$
3

hours
$\langle n, it \rangle$
$\lambda n \lambda t[\text{hours}(t) = n]$

[runtime]
$\langle vi \rangle$
$\lambda e[\tau(e)]$

of
$\langle \alpha t, \langle \alpha \beta, \langle \beta t, \alpha t \rangle \rangle \rangle$
$\lambda S \lambda M \lambda K \lambda b$
$[S(b) \;\wedge\; K(M(b))]$

walking
$\langle vt \rangle$
$\lambda e[^*\text{walk}(e)]$
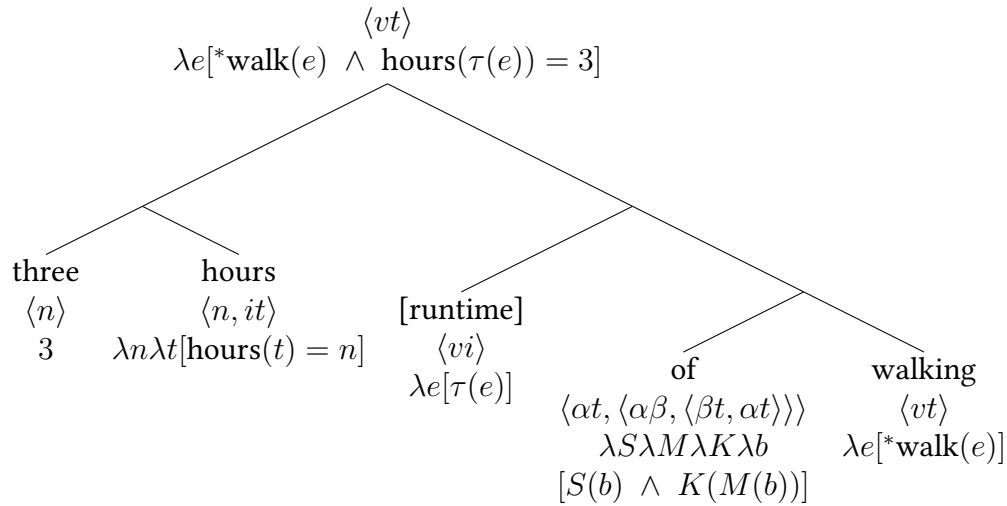
A constraint that rules out this pseudopartitive, along with the other unaccept-able examples of the constructions discussed in Chapter 1, is introduced in Chapter 4. It is further developed in Chapter 7.

3.3 *For*-adverbials

For-adverbials have also been called *measure adverbials*, *aspectual adverbials*, and (when they are temporal rather than spatial) *durative adverbials*. Based on examples like (10), they have been used as a diagnostic of atelicity since at least Verkuyl (1972). In fact, Verkuyl (1989) considers them the most reliable indicator of atelicity.

(10) a. John ran for five minutes / for three hours / for miles.
      b. *John ran to the store for five minutes / for three hours / for miles.

As in the case of the count-mass distinction (see Section 2.6), the telic-atelic distinction shows a certain amount of elasticity. To some extent, *run to the store* may be reinterpreted or "coerced" as an atelic predicate with the meaning "run towards the store". I have nothing to say about the phenomenon of aspectual coercion. I also exclude from consideration the so-called result state related interpretation of *for*-adverbials (Piñón 1999b), which is discussed by Dowty (1979) under the name of internal interpretation. Example (11) is ambiguous between interpretations (12a) and (12b). Dowty attributes this observation to Binnick (1969).

(11)   The Sheriff of Nottingham jailed Robin Hood for four years.

(12)   a.   The Sheriff of Nottingham spent four years bringing it about that
              Robin Hood was in jail.
        b.   The Sheriff of Nottingham brought it about that for four years Robin
              Hood was in jail.

The result state related interpretation must be controlled for when we use
for-adverbials as an atelicity diagnostic, because it allows *for*-adverbials to combine
with telic and punctual predicates (*John opened the window for five minutes*). On
this interpretation, which corresponds to (12b), *for*-adverbials are compatible with
accomplishments (e.g. *John opened the window for five minutes*), and therefore do
not diagnose telicity. Many languages, such as German, have different words for
the standard and the result state interpretation of *for*-adverbials (Piñón 1999b).

For-adverbials stand in near-complementary distribution with *in*-adverbials,
which reject atelic predicates and accept telic predicates. As Section 4.3.1 discusses,
the entailment pattern in (13) shows that sentences with *for*-adverbials, but not
those with *in*-adverbials, are distributive constructions:

(13)   a.   John ran for five minutes.
              ⇒ John ran for four minutes.
              ⇒ John ran for three minutes.
        b.   John ran to the store in five minutes.
              ⇏ John ran to the store in four minutes.

I do not discuss *in*-adverbials in this work.

Regarding the syntax of *for*-adverbials, there is no consensus on whether they
attach below or above the subject. This issue becomes relevant in connection with
the interaction of *for*-adverbials and the Perfect. An overview of the relevant issues
and literature is found in Rathert (2004). The issue hinges on whether sentences
like the two following ones are synonymous:

(14)   a.   John has been in Boston for four years.
        b.   For four years, John has been in Boston.

Dowty (1979) reports that (14a) does not specify whether John is still in Boston,
but he judges (14b) to entail that he is still in Boston. Researchers who do assume an
ambiguity often attribute it to different attachment sites: in (14a), the *for*-adverbial
can attach only above the subject, in (14b), it can also attach below the subject.
However, the accuracy of Dowty's judgment has been challenged (Abusch and
Rooth 1990; Rathert 2004).

In any case, the premise of such arguments is that *for*-adverbials always take semantic scope where they take syntactic scope. Given this premise, another argument that the higher attachment site is available comes from the contribution of the subject to the aspectual properties of the sentence. Verb phrases that are otherwise unacceptable with a *for*-adverbial can be rescued by an unbounded subject (Verkuyl 1972):

(15) a. For years, water / *a liter of water came in from the rock.
      b. For hours, people / *this person walked out of the house.
      c. Patients here / *These two patients died of jaundice for months.

Conversely, the premise can also be used to argue for a possible attachment site below the subject by using examples of verb phrase coordination:

(16) Yesterday one of my friends worked for eight hours, had dinner, and then slept for five hours.

In this sentence, the two *for*-adverbials semantically, and therefore arguably syntactically, modify the individual verb phrases rather than the entire sentence.

Figure 3.3 shows my LF for a verb phrase to which a *for*-adverbial has attached. As in the case of pseudopartitives, this is an incomplete account that serves as a canvas for what follows. Nothing in the entry of the *for*-adverbial presented here rules out telic predicates.

Figure 3.3: Skeletal LF of a *for*-adverbial
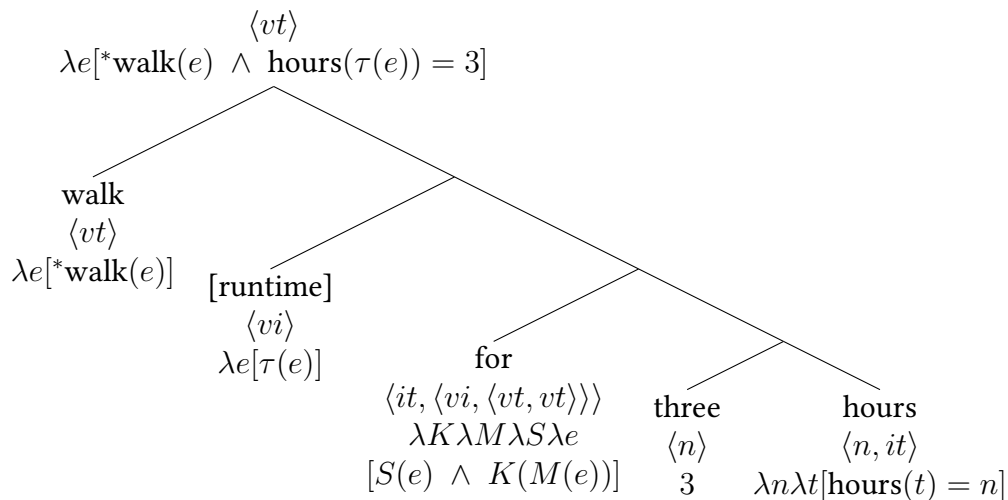
As Figures 3.2 and 3.3 show, I assume that there is no semantic difference between an event pseudopartitive like *three hours of running* and a verb phrase

modified by a *for*-adverbial like *run for three hours*. Of course, the two phrases do not have the same distribution. However, since I assume that all verbal projections involve reference to events, I have not used polymorphic types in my translation of the word *for*.

The parallel between *for*-adverbials and event-based pseudopartitives that I assume here is analogous to the well-known parallel between sentences and action nominalizations. On the assumption that both involve reference to events, they can be given identical semantic representations:

(17) a. The Romans destroyed a city.
      b. the Romans' destruction of a city
      c. $\lambda e[^*\text{destroy}(e) \wedge {}^*\text{ag}(e) = \text{the.romans} \wedge \text{city}(^*\text{th}(e))]$

3.4 Adverbial-*each* constructions

The distributive item *each* can occur in different positions: as a determiner-quantifier (18a), as an adverbial modifier (18b), and as a "distance-distributive" item (18c). The adverbial modifier is also called "floated" *each* in Choe (1987). Distance-distributive *each* has also been called "shifted" (Postal 1974; Choe 1987), "anti-quantifier" (Choe 1987), "binominal" (Safir and Stowell 1988), or "adnominal" (Zimmermann 2002).

(18) a. Each man carried two suitcases.
      b. The men each carried two suitcases.
      c. The men carried two suitcases each.

In this work, I focus on the adverbial modifier use of *each* and I call the type of sentences that contain it adverbial-*each* constructions. I use *each* mainly for the purpose of building a bridge between distributivity and the other phenomena considered here. To this end, I show how to model its distributivity effect in an event-semantic framework. I have nothing to add to the specific problems brought about by distance-distributive *each*. See Zimmermann (2002) for a study of this use of *each* and its equivalents across many languages.

I assume that adverbial *each* modifies a verb phrase. As discussed in Section 2.10, I assume that thematic role heads occur between noun phrases and the verbal projections with which they combine. For the agent role head, this assumption is widespread in the syntactic literature. The usual name for this head is v or "little v". Since I omit category labels in syntactic representations, I represent it instead as [ag]. The square brackets are a reminder of the parallelism I assume between thematic roles, trace functions like [runtime], and measure functions like [height].

I assume that thematic roles are functions as well (see Section 2.5.1).

Like *of* in pseudopartitives and *for* in *for*-adverbials, adverbial *each* can only be associated with predicative noun phrases (**every boy each walked*). As in the case of *of* and *for*, I write this fact directly into the lexical entry of *each*. To accommodate the fact that referential noun phrases such as *the boys* or *John and Mary* are also compatible with *each*, I assume that any noun phrase N whose translation is of type e can be shifted to a noun phrase of type $\langle et \rangle$ which denotes the singleton set $\lambda x[x = [\![N]\!]]$. This is the *Quine* operator of Partee and Rooth (1983). For a similar assumption regarding the definite article, see Winter (2001, p. 153).[10]

Figure 3.4 shows my LF for the sentence *three boys each walked* before existential closure applies (Section 2.10). For lack of space, I have not decomposed *boys* into *boy* plus plural. See Section 2.6.2 for my assumptions on the meaning of plural count nouns. As in the previous sections, this is an incomplete LF which does not yet rule out collective predicates like *met* and which does not make sure that predicates like *build a raft* are interpreted distributively. These possibilities will be ruled out in Chapter 4, when a constraint is added to this LF.

Figure 3.4: Skeletal LF of an adverbial-*each* construction

---
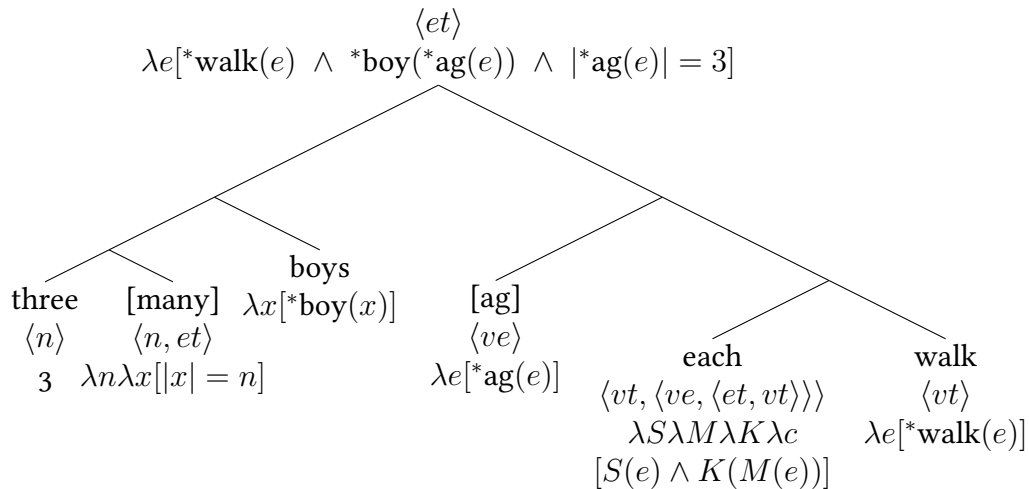
[10]The *Quine* operator could also be used to cut down the number of type shifters in Section 2.10 by half, at the expense of producing slightly less readable logical translations.

Chapter 4

The theory

4.1 Introduction

This chapter motivates and develops strata theory. I have already introduced its main idea in Chapter 1. Here, I develop it in full formal detail against the backdrop of the ontological and semantic assumptions presented in Chapter 2 and the LFs presented in Chapter 3.

As described in the introduction, strata theory capitalizes on cross-categorial parallels between atelic aspect, mass reference, plural reference, and distributivity. Many authors have made analogies between some of these properties and put them to use for drawing semantic generalizations across constituents that have these properties. For example, the parallel between mass and plural reference is exploited in Link (1983) and the one between mass and atelic reference in Bach (1986). Rather than modeling these parallels through syntactic features, these authors take the important step of relating them to semantic properties of the denotations of the constituents in question. In a mereological framework, these properties can be formalized through higher-order properties such as cumulative and divisive reference. This is also the strategy I use here.

As this chapter will show, the higher-order properties that Link and Bach used to characterize mass, plural and atelic reference do not adequately characterize the restrictions that distributive constructions impose on their constituents. For this reason, I introduce a new higher-order property called stratified reference. This property has parameters which provide additional degrees of freedom compared to more well-known properties like divisive and cumulative reference. These parameters make it possible to use stratified reference to express similarities and differences between distributive constructions in a uniform way.

4.2 What is distributivity?

The use of the word *distributivity* generally indicates the application of a predicate to the members or subsets of a set, or to the parts of an entity. This application is diagnosed by the presence of certain entailments I call distributive entailments. While the word *distributivity* and derived terms are widespread in the semantic literature, there are no standard definitions, and a number of related concepts can be distinguished for which the word is used. Distributivity can be a seen as property of quantifiers (Section 4.2.1), a relation between two constituents (Section 4.2.2), a property of predicates (Section 4.2.3), or as I propose in Section 4.3, a property of constructions. All these concepts are useful, but it is important to delimit them from each other. Partly building on an overview by Tsoulas and Zweig (2009), this section reviews existing notions of distributivity and provides an overview of the relevant results from the semantic literature.

4.2.1 Quantificational distributivity

We speak of quantificational distributivity in connection with quantificational noun phrases headed by determiners like *every* or *each* (e.g. Scha 1981). The truth conditions of these noun phrases involve application of the verbal predicate to each member of their witness set. This is probably the least controversial type of distributivity, and it does not play a major role in this work. An exception is the word *all* in its adnominal use. There is no consensus on whether adominal *all* should be classified as a distributive quantifier, because there is a class of predicates with which it gives rise to collective readings (I will discuss them below under the name *gather-type predicates*). Chapter 9 presents arguments for treating *all* as a distributive quantifier and draws a parallel between its behavior and other distributive items.

4.2.2 Relational distributivity

When two different constituents contribute to the content of a distributive entailment, they are said to stand in a distributive relation. I refer to the presence of a distributive relation generally as *relational distributivity*. In standard examples like (1a), the constituents involved in a distributive relation are a subject and a verb phrase, but this need not be so. A theory of distributivity that relies heavily on the concept of a distributive relation is developed in Choe (1987). By contrast, I do not assume that distributive relations are reified as syntactic or semantic links.

(1)   a.   Al and Bill each ate a pizza.

b.   Al and Bill ate a pizza.

Sentence (1a) has the distributive entailments that Al ate a pizza, and that Bill ate a pizza. A distributive relation can be obligatory or optional. For example, sentence (1b) leaves open whether one or two pizzas were eaten. In (1a), the relation between the subject and the verb phrase is obligatorily distributive. In (1b), it is optionally distributive.

A distributive relation can be indicated by distributive markers such as *each* in (1a) (see Gil (1982) for a cross-linguistic survey). In the absence of overt markers, certain sentences such as (1b) can exhibit a distributive relation optionally.

4.2.3   Predicative distributivity

Distributivity understood as a property of predicates is generally set in opposition to collectivity. These notions are based on the behavior of predicates when they occur with plural definites, noun phrases headed by *every*, and coordinated noun phrases. Predicates such as *smile* or *sing* lead to (near-)equivalent sentences when these different kinds of arguments are used (2). These predicates are classified as distributive. The class of collective predicates is formed by those predicates for which this pattern breaks down because the combination with *every* and with singular proper names leads to a category mistake (3).

(2)   **Distributive predicates**
   a.   The children smiled. ⟺ Every child smiled.
   b.   John and Mary sang. ⟺ John sang.

(3)   **Collective predicates**
   a.   The children gathered. ⇎ *Every child gathered.
   b.   John and Mary gathered. ⇎ *John gathered.

This distinction between distributive and collective predicates has been criticized by Winter (2001, 2002) as not very useful and hard to justify. Winter notes that the patterns in (2) and (3) are only valid if one abstracts away from a number of factors: conventionalized coordinations, nonmaximality effects on definite plurals, and effects related to group nouns. Winter's concern about conventionalized coordinations like *Simon and Garfunkel*, if I understand him correctly, is that they do not always give rise to entailments like (2b), as shown in (4). The biconditional in (2a) is only valid to the extent that the referent of the definite plural includes every member of its complement noun, but this is not the case if the definite plural has a nonmaximal interpretation (5). Furthermore, the test in (3) is only reliable as long as its nouns are not replaced by group nouns and noun phrases like *committee*

and *Committee A* (6).

(4)   Simon and Garfunkel are performing in Central Park.
      ⇏ Simon is performing in Central Park.   (Winter 2001)[11]

(5)   At the end of the press conference, the reporters asked the president questions.
      ⇏ Every reporter asked the president a question.   (Dowty 1987)

(6)   a.   Every committee gathered. *vs.* *Every child gathered.
      b.   Committee A gathered. *vs.* *John gathered.

Winter concludes from these problems that the standard distributive-collective classification is not tenable. He proposes to replace the traditional test in (3) with an alternative test that does not use definite plurals and conjunctions and works even when group nouns like *committee* are used in them. Winter's test leads to an alternative classification based on whether or not a predicate is sensitive to the distinction between singular quantificational determiners like *every* and plural ones like *all*. Distributive predicates like *smile* are compatible with both kinds of determiners and lead to equivalent interpretations. Winter calls this class atom predicates (7). Some collective predicates, like *be numerous* show the same behavior as distributive predicates like *smile*, while others like *gather*, which he calls set predicates (8), distinguish between both.

(7)   **Atom predicates**
      a.   All the children smiled. ⇔ Every child smiled.
      b.   All the committees are numerous. ⇔ Every committee is numerous.

(8)   **Set predicates**
      a.   All the children gathered. ⇏ *Every child gathered.
      b.   All the committees gathered. ⇏ Every committee gathered.

As shown in Table 4.1, Winter's test draws the boundary at a different place than the traditional distributive-collective criteria. For this reason, it is not useful as a characterization of distributive predicates, which it is not meant to be. On the other hand, by placing the boundary within the traditional class of collective predicates, Winter's test introduces a new and useful distinction within that class.

The categories I use are also shown in Table 4.1. They represent a synthesis of both the traditional categories and those of Winter. Distributive predicates are retained as a category, and collective predicates are split into numerous-type and gather-type predicates. Section 9.5 proposes to account for the differing behavior

---

[11]Winter attributes a similar example to Landman (p.c.).

of these two classes of predicates in terms of the sum/group distinction. For now, I leave collective predicates aside and concentrate on the notion of a distributive predicate.

Table 4.1: Comparison of the distributive-collective and atom-set typologies

Example	Traditional	Winter	This work
smile	distributive	atom predicate	distributive
be numerous	collective		numerous-type
gather		set predicate	gather-type

Winter does not make any distinction between distributive and collective predicates because he does not consider this distinction well-motivated. At least for the purpose of this work, however, it is useful to have an operational definition of a distributive predicate. To develop such a definition, it is necessary to address Winter's concerns about the reliability of the traditional tests. I propose to do that by slightly reformulating the tests in order to control for effects related to nonmaximal interpretations, conventionalized coordinations, and group noun effects. As Malamud (2006b) observes, indefinite numerals do not give rise to nonmaximal interpretations. Sentence (9a) is from Malamud and sentence (9b) is from Landman (1996, p. 435).

(9) a. The women in Bogoduhov gave birth to only seven children.
     b. *Fifteen women gave birth to only seven children.

In reformulating the traditional test, we must take care not to rely on constructions involving coordinations (because they might be conventionalized) and definite plurals (because they might be nonmaximal). Based on this reasoning, I propose the following operational definition of a distributive predicate:

(10) **Operational definition: Distributive predicate**
      A distributive predicate is a predicate for which (11a) and (11b) are acceptable and mutually entail each other when it is substituted for PRED.

(11) a. Three people PRED.
      b. Three people each PRED.

The noun *people* may be replaced by another noun if necessary to avoid selectional restrictions such as animacy requirements. To address Winter's concern about group nouns, we restrict the test by agreeing that this noun may not be replaced by a group noun. I propose an operational definition of group nouns in Section 2.6.4, which is based on a test proposed by Barker (1992).

To mention a few examples, the predicates *sleep*, *run*, *sneeze*, *get up*, *wear a dress*, and *take a breath* are all distributive predicates according to the definition in (10) because the entailment from (11a) to (11b) is obligatory with them. The predicates *eat a pizza*, *build a raft* and *ask a question* are not distributive because the entailment from (11a) to (11b) is not obligatory with them. The predicates *meet* and *be numerous* are not distributive because at least (11b) is not acceptable except when we replace the word *people* by a group noun such as *committee*, which by convention is not allowed.

Following standard usage in the literature, I have described predicative distributivity as a property of intransitive predicates. The notion can be generalized to describe transitive predicates, but in this case it needs to be relativized to an argument position or thematic role. For example, *kill* is distributive on its theme role but not on its agent role, since a group of people can only be killed if each of them is killed, but a group of people can kill a person without each of them killing that person (Lasersohn 1988; Landman 1996). This is illustrated in the following scenario. The two outlaws Bonnie and Clyde were killed by a posse of six police officers, which included Sheriff Jordan. Given this background knowledge, (12a) entails (12b) but not (12c).

(12)   a.   The police officers killed Bonnie and Clyde.
       b. $\Rightarrow$ The police officers killed Bonnie.
       c. $\nRightarrow$ Sheriff Jordan killed Bonnie and Clyde.

The idea that distributivity needs to be parametrized for a thematic role to account for such phenomena is taken up again in Section 4.5.

4.3   Distributive constructions

Section 4.2 has shown how distributivity can be seen as a property of predicates, of quantifiers, and of pairs of constituents. We can see distributivity in yet another way, namely as a property of entire constructions. By understanding distributivity in this way, we can abstract away from individual sentences and predicates. This allows us to generalize the concept of distributivity to encompass restrictions on aspectual properties and measure functions.

For this purpose, I introduce the following terminology. A *distributive con-*

struction is a lexicosyntactic configuration that imposes an obligatory distributive relation between two of its constituents. A sentence *instantiates* a distributive construction when the sentence exemplifies the syntactic configuration of that construction. We have already seen that the adverbial-*each* construction obligatorily contains a distributive relation. Sentences with *each* are therefore uncontroversial examples of distributive constructions.

A sentence like (13) does not instantiate a distributive construction, even though it involves a distributive predicate (its verb phrase) and this predicate stands in a distributive relation to the subject. The sentence does not instantiate a distributive construction because there are other sentences with the same syntactic configuration, such as (14), which do not contain a distributive relation.

(13)   Al and Bill took a deep breath.

(14)   Al and Bill met.

This section provides initial motivation for treating adverbial-*each* constructions, *for*-adverbials, and pseudopartitives as members of a natural class, namely the class of distributive constructions.

4.3.1   *For*-adverbials are distributive constructions

For-adverbials can be classified as distributive constructions because they obligatorily involve a distributive relation. On the assumptions in Section 3.3, a sentence like (15a) involves reference to an event e_0 whose runtime is an interval of five-minute length. This sentence entails (15b) and (15c).

(15)   a.   John ran for five minutes.
         b. $\Rightarrow$ John ran for four minutes.
         c. $\Rightarrow$ John ran for three minutes.

In the present event semantic framework, the entailments in (15) involve the existence of an event e_1 in which John ran for four minutes, an event e_2 in which he ran for three minutes, and so on. The sentence features a distributive relation because two of its constituents – the *for*-adverbial and the verbal or sentential predicate it modifies – jointly determine this entailment and because on the semantic assumptions in Chapter 2, the events $e_1, e_2, \ldots$ that represent these entailments are parts of the event e_0.

4.3.2 Pseudopartitives are distributive constructions

Pseudopartitives, such as the subject of (16), can also be classified as distributive constructions.

(16) Three liters of water are sufficient.

Because the word *water* occurs in (16), the sentence gives rise to the entailment that the parts of the quantity in question consist themselves of water. This entailment only involves one constituent, namely *water*. However, the sentence also has another distributive entailment: among these parts there exist one liter of water, two liters of water, and so on. This existence entailment is present despite the fact that the sentence has a nondistributive verb phrase, *be sufficient*, so the entailment must be due to the pseudopartitive rather than the verb phrase. It is a distributive entailment because it concern the parts of an entity, namely the water entity to which the pseudopartitive refers, and it involves a distributive relation because the meanings of both the constituent *water* and the constituent *three liters* contribute to it. It is easy to see that the presence of this distributive relation is not accidental, but that it is required by the construction, because every pseudopartitive gives rise to similar entailments. Since pseudopartitives obligatorily involve a distributive relation, I classify them as distributive constructions.

4.4 The components of a distributive relation

The previous section has established that adverbial-*each* constructions, *for*-adverbials and pseudopartitives are distributive constructions because they involve obligatory distributive relations between their constituents. Here, I introduce terminology with which we can express generalizations over the three constructions.

As Zimmermann (2002, p. 23) observes, there is considerable terminological confusion in the literature concerning the components of a distributive relation. Table 4.2, which is expanded from that source, gives an overview of the terminology.

For conciseness, I adopt the terms *Key* and *Share*, as in Gil (1989) and Choe (1991).[12] Intuitively, the term *Share* refers to the constituent whose denotation is distributed over the parts of the referent of the other constituent, which is called the

---

[12]The term "Sorting Key" in Choe (1987) is taken from Kuno (1982), who uses it in describing answers to multiple wh-questions. There are small differences not only in terminology but also in its application. Choe, and Safir and Stowell, use their terms *Distributed Share* and *DistNP* differently than other authors: they apply it to the noun phrase *a breath*, as opposed to the entire verb phrase *took a breath*. Link uses the term *Distributional Key* only for Keys that also exhibit quantificational distributivity, like *every*.

Table 4.2: Terms for the components of a distributive relation

Author	Name for *"The boys"*	Name for *"(took) a breath"*
Link (1986)	distributional domain	Distributive Share (DstrShr)
Choe (1987)	Sorting Key (SrtKy)	Distributed Share (DstrShr)
Safir and Stowell (1988)	Range NP	Distributing NP (DistNP)
Gil (1989), Choe (1991)	Key	Share
Zimmermann (2002)	DistKey	DistShare
Blaheta (2003)	Dist phrase	Range
This work	**Key**	**Share**

Key. For example, the property of reading a book is distributed over the individual boys, the property of being water is distributed over the liters, and the property of being an event of pushing a cart is distributed over the hours. My use of the terminology is illustrated in Table 4.3. The constituent structure presupposed by this use is justified in Chapter 3.

Table 4.3: A bridge from distributivity to aspect and measurement

Construction	Example	Key	Share	Map
Adverbial *each*	Three boys each laughed	three boys	laugh	agent
For-adverbial	John ran for three hours	three hours	John run	runtime
Pseudopartitive	three liters of water	three liters	water	volume

The intuitive criterion behind the assignment of the terms Key and Share just described is perhaps not completely clear. I also use a formal criterion, which is more dependent on my specific choice of background assumptions, but which coincides with the intuitive criterion and makes this assignment clearer. As discussed in Section 2.5, I adopt a number of standard assumptions to the effect that constituents are related by certain covert functions: thematic roles, trace functions, and measure functions. These functions coincide with distributive relations in a particular way: they always map entities associated with the Share to entities associated with the Key. For example, in *Three boys each laughed*, the function *ag* maps events associated with the Share to entities associated with the Key. In *John ran for three hours*, the function *runtime* maps events associated with the Share to intervals associated with the Key, and in *three liters of water*, the function *volume* maps substances associated with the Share to degrees associated

with the Key. It can be useful to refer to these functions independently of the construction involved. As shown in Table 4.3, I use the term *Map* for this purpose.

4.5   The constraints on distributive constructions

Recall that *each*-constructions only combine with distributive predicates, *for*-adverbials only combine with atelic predicates, and pseudopartitives accept only mass nouns and plural count nouns as substance nouns. Put in the terms just introduced, these facts all concern Shares: *each*-constructions accept only distributive predicates as Shares; *for*-adverbials accept only atelic predicates as Shares; and pseudopartitives accept only mass nouns and plural count nouns as Shares.

I now present the traditional explanations of these facts, which are all formulated separately, and then update each of them in successive refinements to reflect my own assumptions. In each case, the updates result in practically the same concept. On the basis of this convergence, I then formulate a single constraint that bridges the differences between the three constructions.

I now propose a higher-order property, stratified distributive reference, which formally captures the notion of a distributive predicate that was operationally defined in Section 4.2.3. All distributive predicates can then be modelled as having stratified distributive reference. Some predicates like *smile* or *sneeze* have this property by virtue of their lexical semantics, and others like *take a breath* may acquire this property through a D operator, to be introduced in Chapter 8. Stratified distributive reference provides a useful generalization that abstracts away from these different ways in which a predicate can be distributive. It encapsulates what it means to be a distributive predicate. Isolating this property is a crucial step on the way towards unifying it with its siblings in the domains of aspect and measurement.

I introduce stratified distributive reference in several steps. We start with a relatively simple idea: we require that the only kinds of events in a distributive predicate are those which consist only of subevents to which the predicate also applies and whose agent is a pure atom. This formulation is chosen in order to exclude impure atoms from consideration. Impure atoms are used to represent collective interpretations of plural definite and indefinite noun phrases (see Section 2.8). Pure atoms represent the entities in the denotations of singular nouns (see Section 2.6). The following preliminary definition encapsulates the idea just described.

(17)   **Stratified distributive reference (preliminary definition 1)**
$$\mathrm{SDR}_{ag}(P) \overset{\text{def}}{=} \forall e[P(e) \rightarrow \forall e' \leq e \begin{pmatrix} P(e') \wedge \\ \mathrm{PureAtom}(^*\mathrm{ag}(e')) \end{pmatrix}]$$

(An event predicate P has stratified distributive reference with respect to the thematic role *ag* if and only if every event e to which P applies only has subevents to which P also applies and whose agents are pure atoms.)

This preliminary definition does not work correctly, however. Even lexical predicates like *sneeze*, which are distributive according to the operational definition, do not satisfy it. For the definition to be true, each subevent of $e_1 \oplus e_2$ must have an atomic entity as its agent. The definition interferes with the assumption of lexical cumulativity (Section 2.7.2). This assumption has the consequence that the sum of two sneezing events is again in the denotation of *sneeze*. These sum events cause the definition to break down, because by cumulativity of thematic roles (Section 2.5.1), their agents are generally not atoms. These sum events have subevents whose agents are not atoms, namely themselves, because the subevent relation $\leq$ is modeled by the mereological parthood relation, and is therefore reflexive (Section 2.3.1). For example, if e_1 is a sneezing event whose agent is John and e_2 is a sneezing event whose agent is Mary, then event $e_1 \oplus e_2$ has itself as a subevent, but its agent, $j \oplus m$, is not atomic, contrary to the requirement of the definition.

A first attempt at fixing the definition would consist in replacing $\leq$ (parthood) with $<$ (proper parthood) in order to avoid the problem caused by reflexivity. However, this does not work either: if the toy model is extended by a third sneezing event e_3, lexical cumulativity causes the event $e_1 \oplus e_2 \oplus e_3$ to be a sneezing event, and the event $e_1 \oplus e_2$ is a proper part of this event. Since $e_1 \oplus e_2$ does not have an atomic agent, the definition again does not apply to *sneeze*.

Another attempt would consist in using the relation $<_{Atom}$ (proper atomic parthood). Since e_1 and e_2 each have atomic agents, this attempt causes *sneeze* to have stratified distributive reference as desired, assuming that e_1 and e_2 are absolute atoms. This assumption is plausible because sneezing events are punctual. But there is no guarantee that distributive predicates always apply exclusively to events that consist of atomic events. Section 2.4.3 has already pointed out that it is not necessary to assume that singular events are atomic. It is in fact implausible to make this assumption, because non-punctual events in the extension of atelic predicates are plausibly not atomic. Some of these atelic predicates, such as *run*, are also distributive, but they would not be predicted to be distributive under this attempt.

These two attempts at extending the concept of a distributive predicate to an event-based framework have failed because of problems in selecting the right subevents with atomic agents. These problems can be avoided in a definition that remains neutral on the issue of how to identify these subevents. Schematically, instead of requiring for any event e that *every* subevent (or proper subevent, or proper atomic subevent) of e has an atomic agent, we require that there is *some*

way of dividing every event e into subevents with atomic agents. In (18), this is expressed by the star operator. To understand how this definition works, note that $e \in {}^{*}\lambda e'.C(e')$ is true if and only if there is a way to divide e into one or more possibly overlapping parts that are in C. To use a term from Schwarzschild (1996), C must essentially act as a cover of e. See Section 5.4.2 for more discussion of this point.

(18) **Stratified distributive reference (preliminary definition 2)**

$$\text{SDR}_{ag}(P) \stackrel{\text{def}}{=} \forall e[P(e) \rightarrow e \in {}^{*}\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \text{PureAtom}({}^{*}\text{ag}(e')) \end{array} \right)]$$

(An event predicate P has stratified distributive reference with respect to a thematic role ag if and only if every event e to which P applies can be exhaustively divided into one or more subevents ("strata") to which P also applies and whose agent is a pure atom.)

On this definition, *sneeze* is correctly predicted to have stratified distributive reference with respect to the agent role. This is so because e_1, e_2 and their sum $e_1 \oplus e_2$ can each be divided into one or more sneezing events with an atomic agent. The definition in (18) does not involve the notion of an atomic event, which makes it possible for arguably nonatomic distributive predicates like *run* to have stratified distributive reference.

The definition is still preliminary because it is tied to the thematic role *ag*. In a next step, we generalize it to arbitrary thematic roles θ. Definition (19) does so by replacing *ag* with θ.

(19) **Final definition: Stratified distributive reference**

$$\text{SDR}_{\theta}(P) \stackrel{\text{def}}{=} \forall e[P(e) \rightarrow e \in {}^{*}\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \text{PureAtom}(\theta(e')) \end{array} \right)]$$

(An event predicate P has stratified distributive reference with respect to a thematic role θ if and only if every event e to which P applies can be exhaustively divided into one or more subevents ("strata") to which P also applies and whose θ is a pure atom.)

I call this definition of stratified distributive reference "final" because I no longer change it under this name. Instead, it will be subsumed under a more general concept along with atelicity and monotonicity of measurement. Before doing so, I apply the concept of stratified distributive reference to adverbial-*each* sentences. With the help of this concept, we can now formulate the constraint on Shares in adverbial-*each* sentences.

This generalization makes it easy to account for the inferential behavior of predicates like *kill* which are distributive on only one of their arguments, in this

case the theme. We can formalize this behavior as meaning postulates:

(20) **Meaning postulates: *kill* is distributive with respect to themes but not agents**
  a. $\text{SDR}_{th}(\text{kill})$
  b. $\neg\text{SDR}_{ag}(\text{kill})$

These postulates, together with the background assumptions on event semantics and lexical cumulativity, correctly account for the inferential behavior of *kill*: together with the translation of (21a), they entail the translation of (21b) but not that of (21c). For ease of exposition, I represent the agents of (21a) and (21b) as sums rather than groups. The difference does not matter here since there are no inferences on the agents in these sentences.

(21) a. Officers 1, 2, and 3 killed Bonnie and Clyde.
         $\exists e[\text{kill}(e) \ \wedge \ {}^*\text{ag}(e) = o_1 \oplus o_2 \oplus o_3 \ \wedge \ {}^*\text{th}(e) = b \oplus c]$
     b. $\Rightarrow$ Officers 1, 2, and 3 killed Bonnie.
         $\exists e'[\text{kill}(e') \ \wedge \ {}^*\text{ag}(e') = o_1 \oplus o_2 \oplus o_3 \ \wedge \ {}^*\text{th}(e') = b]$
     c. $\nRightarrow$ Officer 1 killed Bonnie and Clyde.
         $\exists e''[\text{kill}(e'') \ \wedge \ {}^*\text{ag}(e'') = o_1 \ \wedge \ {}^*\text{th}(e'') = b \oplus c]$

4.5.1   Adverbial *each*

The Share of a sentence with adverbial *each* must be a distributive predicate. This is so by definition if we accept the operational definition in (10). With the help of stratified distributive reference, we can now express this fact formally:

(22) **Constraint on Shares of adverbial-*each* sentences**
     A sentence with adverbial *each* whose Key is K, whose Share is S and whose Map is M is acceptable only if S has stratified distributive reference with respect to M (formally: $\text{SDR}_M(S)$).

The following examples illustrate how this constraint works:

(23) Three boys each laughed.
  a. Key: Three boys
  b. Map: agent
  c. Share: laugh

Applied to example (23), constraint (22) predicts that the sentence is acceptable only if the condition $\text{SDR}_{ag}(\text{laugh})$ is fulfilled. This condition expands as follows:

(24) $SDR_{ag}(\text{laugh})$
$\Leftrightarrow \forall e[\text{laugh}(e) \rightarrow e \in {}^*\lambda e' (\text{ laugh}(e') \wedge \text{PureAtom}({}^*\text{ag}(e')))]$
(Every laughing event can be divided into one or more laughing events ("strata") whose agent is a pure atom.)

This condition is fulfilled by the predicate *laugh* on the assumption that stratified distributive reference accurately models predicative distributivity, that is, on the assumption that distributive predicates (according to the operational definition) have stratified distributive reference (according to the formal definition). This assumption is independently needed to explain the inference from (25a) to (25b).

(25) a. John and Mary laughed.
      b. ⇒ John laughed.

The following example illustrates the case of a predicate that is not distributive:

(26) *Three boys each met.
      a. Key: Three boys
      b. Map: agent
      c. Share: meet

Applied to example (26), constraint (22) predicts that the sentence is acceptable only if the condition $SDR_{ag}(\text{meet})$ is fulfilled. This condition expands as follows:

(27) $SDR_{ag}(\text{meet})$
$\Leftrightarrow \forall e[\text{meet}(e) \rightarrow e \in {}^*\lambda e' (\text{ meet}(e') \wedge \text{PureAtom}({}^*\text{ag}(e')))]$
(Every meeting event can be divided into one or more meeting events whose agent is a pure atom.)

This condition is not fulfilled by the predicate *meet*. A plural individual corresponding to two boys can meet, and this meeting event does not have any parts, let alone any parts whose agent is a pure atom. This assumption is independently needed to explain the lack of inference from (28a) to (28b).

(28) a. John and Mary met.
      b. ⇏ *John met.

4.5.2 Temporal *for*-adverbials

The Share of a temporal *for*-adverbial must be an atelic predicate. As Dowty (1979) shows, the entailment properties of atelic predicates with respect to a large range of other phenomena, including tense, the progressive, and aspectualizers, can be

represented as a first approximation by modeling the denotations in terms of the *subinterval property*. This is an old idea. For example, Vendler (1957) suggests that activity predicates like *run* have the property that "any part of the process is of the same nature as a whole". He contrasts this property with accomplishment predicates like *run a mile*, which "proceed toward a terminus which is logically necessary to their being what they are". The model-theoretic implementation of the subinterval property is usually attributed to Bennett and Partee (1972). In their original formulation, matrix verb phrases with this property are true at every subinterval, including every moment, of the interval at which the sentence that contains them is true. In our event-based framework, the subinterval property can be defined with the help of the relation AT from Section 2.5.2:

(29)   **Definition: Subinterval property**
$\text{SUB}(P) \stackrel{\text{def}}{=} \forall i[\text{AT}(P, i) \rightarrow \forall j[j < i \rightarrow \text{AT}(P, j)]]$
$\Leftrightarrow \forall i \forall e[P(e) \wedge \tau(e) = i \rightarrow \forall j[j < i \rightarrow \exists e'[P(e') \wedge \tau(e') = j]]]$
(An event predicate P has the subinterval property if and only if whenever it holds at an interval, it also holds at every one of its subintervals.)

It is well known that the subinterval property is too strong. This is known as the minimal-parts problem (to be discussed in more detail in Chapter 5). For example, *waltz* is an atelic predicate, but a subinterval that is so short that it contains less than three steps makes it hard to determine whether these steps qualify as waltzing (Dowty 1979). For this reason, it is difficult to know whether the subinterval property in (29) applies to *waltz*, and this makes it unsuited for modeling atelicity.

Definition (30) implements a weaker version of the subinterval property that avoids the minimal-parts problem:

(30)   **Definition: Stratified subinterval reference**
$\text{SSR}_{\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall e[P(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \varepsilon(K)(\tau(e')) \end{array} \right)]$
(An event predicate P has stratified subinterval reference with respect to a Key K if and only if every event in its denotation can be divided exhaustively into very small parts ("strata") that are each in P. A very small part is one whose runtime satisfies the predicate $\varepsilon(K)$.)

The exact formulation of this definition is motivated in Chapter 5, but the basic intuition behind it is the following:

- Instead of testing whether the predicate holds at every subinterval of the interval i, the definition only tests whether this interval can be divided into

sufficiently small subintervals.[13]

- The definition of sufficiently small is specified through the threshold parameter $\varepsilon(K)$. I assume that ε is a predicate which takes a comparison class K and an entity x and returns true if and only if x is very small with respect to the comparison class K. For concreteness, I represent $\varepsilon(K)$ with a very small value, for example as $\lambda i.seconds(i) \leq 1$ for a one-second threshold. This method avoids explicit reference to atomic events or intervals, similarly to the definition of stratified distributive reference in (19).

We can now formulate the constraint on Shares in temporal *for*-adverbials with the help of stratified subinterval reference. This constraint links the threshold $\varepsilon(K)$ to the Key (the temporal noun phrase) of the *for*-adverbial to express the fact that what counts as very small is vague but depends on the Key.

(31)   **Constraint on Shares of temporal *for*-adverbials**
        A temporal *for*-adverbial whose Key is K and whose Share is S is acceptable only if S has stratified subinterval reference (formally: $\text{SSR}_{\varepsilon(K)}(S)$), where the threshold $\varepsilon(K)$ holds of intervals whose runtime is very short with respect to K.

The following examples illustrate how this constraint works:

(32)   John and Mary waltzed for an hour.
        a.   Key: an hour
        b.   Threshold: e.g. $\lambda i.\text{seconds}(i) \leq 3$
        c.   Share: waltz

Applied to example (32), constraint (31) predicts that the sentence is acceptable only if the condition $\text{SSR}_{\lambda i.\ seconds(i) \leq 3}(\text{waltz})$ is fulfilled (I have specified a somewhat arbitrary value of $\varepsilon(K)$ for concreteness). This condition expands as follows:

(33) $\text{SSR}_{\lambda i.\ seconds(i) \leq 3}(\text{waltz})$

$$\Leftrightarrow \forall e[\text{waltz}(e) \rightarrow e \in {}^*\lambda e' \left(\begin{array}{c} \text{waltz}(e') \wedge \\ \text{seconds}(\tau(e)) \leq 3 \end{array} \right)]$$

(Every waltzing event can be divided into one or more waltzing events whose runtime is at most three seconds.)

---

[13]More precisely, the definition tests whether the event can be divided into subevents with sufficiently small subintervals, but this difference is not crucial. Since the runtime function τ is a sum homomorphism (Section 2.5.2), the interval $\tau(e)$ is guaranteed to be the sum of these small intervals of the parts of e that are at most as big as $\varepsilon(K)$.

This condition is fulfilled by the predicate *waltz*, as shown by the inference patterns like the following (Dowty 1979).

(34)   a.   John and Mary waltzed from 3pm to 4pm.
       b. $\Rightarrow$ John and Mary waltzed from 3pm to 3:30pm.

The following example illustrates the case of a telic predicate:

(35)   *John built a sand castle for an hour.
       a.   Key: an hour
       b.   Threshold: e.g. $\lambda i.\text{seconds}(i) \leq 3$
       c.   Share: build a sand castle

Applied to example (35), constraint (31) predicts that the sentence is acceptable only if the condition $\text{SSR}_{\lambda i.\ seconds(i)\ \leq\ 3}(\text{build.a.sand.castle})$ is fulfilled (again, my choice of value for $\varepsilon(K)$ is only for concreteness). This condition expands as follows:

(36) $\text{SSR}_{\lambda i.\ seconds(i)\ \leq\ 3}(\text{build.a.sand.castle})$

$\Leftrightarrow \forall e[\text{build.a.sand.castle}(e) \rightarrow e \in {}^{*}\lambda e' \left(\begin{array}{c} \text{build.a.sand.castle}(e') \land \\ \text{seconds}(\tau(e')) \leq 3 \end{array} \right)]$

(Every sand-castle-building event can be divided into one or more sand-castle-building events whose runtime is at most three seconds.)

This condition is not fulfilled by the predicate *build.a.sand.castle* (Krifka 1998). More generally, no sand-castle-building event can be divided into sand-castle-building evens whose runtime is very short with respect to one hour. My specific choice for the value of the threshold does not affect the validity of the argument.

4.5.3   Pseudopartitives

The Share (the substance noun) of a pseudopartitive must be a mass term or a plural count term. As is often noted (e.g. Bach 1986), the subinterval property, which has been used to characterize atelicity, is similar to the concept of divisive reference, which has been used to characterize mass terms and plurals.[14]   The definition of divisive reference is repeated here from Section 2.3.2:

(37)   **Definition: Divisive reference**
       $\text{DIV}(P) \stackrel{\text{def}}{=} \forall x[P(y) \rightarrow \forall y[y < x \rightarrow P(y)]]$

---

[14]I follow Krifka (1989) in calling this concept divisive reference. Some authors refer to this property as distributive reference, but I have already used a similar term above in (19) with a different meaning.

(A predicate P over ordinary objects has divisive reference if every part of any object in its denotation is also in its denotation.)

As a characterization of Shares in pseudopartitives, there is a serious problem with divisive reference: all mass terms are acceptable Shares in pseudopartitives, but the near consensus view in the semantic literature is that they do not all have divisive reference. The evidence for this view is reviewed in Section 5.2.2. The question then arises what kind of property can characterize the Shares of pseudopartitives.

An example of a mass noun that does not have divisive reference is the mass noun *cable*.[15] It does not have divisive reference because not every part of a cable segment qualifies itself as a cable segment. Linear segments of a cable segment do, but slices which run the length of the spool or which are taken through the middle along the length of the cable segment do not.

Schwarzschild (2002, 2006) observes that these facts correlate with the interpretation of pseudopartitives whose substance noun is *cable*: (38) can only be interpreted as being about a cable whose length is three inches, not as a cable of whose diameter is three inches.

(38)   three inches of cable

As discussed in Chapter 3 and above, I adopt the standard assumption that the measure and substance noun of a pseudopartitive are related by a covert measure function, which I call the Map. The different potential interpretations of (38) are assigned different logical representations which differ only in the choice of this measure function:

(39) $\lambda x[\text{cable}(x) \; \wedge \; \exists d[\text{length}(x) = d \; \wedge \; \text{inches}(d) = 3]]$
        (available)

(40) $\lambda x[\text{cable}(x) \; \wedge \; \exists d[\text{diameter}(x) = d \; \wedge \; \text{inches}(d) = 3]]$
        (unavailable)

These logical representations give us an initial handle on the cable problem. The task consists of finding a constraint that rules out (40) but not (39). Note that the use of the measure function *diameter* is not itself the problem, as there

---

[15]I refer to the mass noun *cable*, in the sense in which it appears in *a spool of cable* or *a lot of cable*. I refer to entities in the denotation of the mass noun *cable* as cable segments. The count noun use of *cable* is irrelevant here because singular count nouns do not occur as Shares of pseudopartitives. As described in Section 2.6, I assume that there are two lexical entries for the count noun and the mass noun use of words like *cable*, and that singular count nouns have quantized reference. As described in Section 2.6.1, a contextual parameter specifies the set of cable segments which are in the denotation of the count noun *cable*.

are other pseudopartitives which allow for an interpretation in which the covert measure function is diameter. Schwarzschild (2002) considers the scenario of a growing pool of oil seeping out of the ground. He observes that we can report its progress by declaring there to be first ten inches of oil, then fifteen inches of oil, and so on. As long as the measure function diameter is clearly retrievable from the context of the utterance, these pseudopartitives are easily understood as involving reference to the diameter of the pool.

I now propose a constraint that rules out (40), and formulate it in such a way that it also excludes mass nouns and plural count nouns. Consider a cable segment whose length is ten feet and whose diameter is ten inches. Among its parts, there are segments whose length is three feet, others whose length is six feet, and so on. These segments each qualify as cable, and their length is always lower than the length of the entire segment. Their diameter, however, is always the same as the diameter of the entire segment.

These facts can be modeled either by considering only certain subsets of the parthood relation among the parts of the cable segment, or by loosening the requirement of divisive reference so as to take a dimension parameter into account. The former route is taken by Schwarzschild (2002, 2006), but it involves making the parthood relation dependent on context. I argue against this strategy in Chapter 7.

For this reason, I take the latter route, which relies only on the mereological parthood relation. The idea is to relativize the concept of divisive reference with respect to a certain dimension or measure function such as length or diameter. Intuitively, the mass noun *cable* is divisive along its length, but not along its diameter. The following concept realizes this intuition by adding a parameter μ for measure functions to the concept of divisive reference.

(41) **Stratified measurement reference (preliminary definition)**

$$\text{SMR}_\mu(P) \overset{\text{def}}{=} \forall x[P(x) \to \forall d[d < \mu(x) \to \exists y \begin{pmatrix} y < x \ \wedge \ P(y) \\ \wedge \ \mu(y) = d \end{pmatrix}]]$$

(A predicate P has stratified measurement reference relative to a measure function μ if and only if for any object x to which P applies, every value smaller than the μ-value of x is the μ-value of some part of x to which P again applies.)

The following equations illustrate the application of this definition to the cable example.

(42) $\text{SMR}_{length}(\text{cable}) \Leftrightarrow$

$$\forall x[\text{cable}(x) \to \forall d[d < \text{length}(x) \to \exists y \begin{pmatrix} y < x \ \wedge \ \text{cable}(y) \\ \wedge \ \text{length}(y) = d \end{pmatrix}]]$$

(Every cable segment of length 10in has as part a cable segment of length

87

9in, another one of length 8in, etc.)

The fact that cable segments do not contain cable segments with a smaller diameter can also be expressed:

(43) $\neg\text{SMR}_{diameter}(\text{cable}) \Leftrightarrow$

$\neg\forall x[\text{cable}(x) \rightarrow \forall d[d < \text{diameter}(x) \rightarrow \exists y \begin{pmatrix} y < x \ \wedge \ \text{cable}(y) \\ \wedge \ \text{diameter}(y) = d \end{pmatrix}]]$

(Not every (in fact, no) cable segment of diameter 10in has as part a cable segment of diameter 9in, etc.)

While the preliminary definition of SMR above successfully links the reference properties of *cable* with the available measure functions in pseudopartitives, it does not yet address the fundamental problem of divisive reference. As discussed in Chapter 5, the substances in the denotation of many mass nouns have parts outside of their denotations. This problem is analogous to the minimal-parts problem of atelic predicates that we encountered in the previous section. For example, cable is not strictly speaking divisive because very short slices taken from a cable segment may not qualify as cable segments. The solution to this problem is also analogous to the case of *for*-adverbials: instead of testing that *all* values below the μ-value of the substance are themselves μ-values of parts of the substance, we restrict our attention to the merely very small (but not arbitrarily small) values. The concept of stratified measurement reference is therefore redefined in (44). This definition tests whether there is a way of dividing every substance in the denotation of a predicate P into parts which have a very small μ-value. As in the case of the stratified subinterval property, a threshold parameter $\varepsilon(K)$ is used to provide a way to define what counts as very small.

(44)   **Final definition: Stratified measurement reference**

$\text{SMR}_{\mu,\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow x \in {}^*\lambda y \begin{pmatrix} P(y) \ \wedge \\ \varepsilon(K)(\mu(y)) \end{pmatrix}]$

(A predicate P over ordinary objects has stratified measurement reference with respect to a function μ and a threshold $\varepsilon(K)$ if and only if there is a way of dividing every entity in its denotation exhaustively into parts ("strata") that are each in P and which have a very small μ-value. Very small μ-values are those that are less than or equal to a value that satisfies $\varepsilon(K)$.)

For the sake of the example, assume that the minimum length at which it is clear that a part of a cable segment qualifies as cable is 0.5in. Then the statement that cable has stratified measurement reference with respect to length amounts

to the statement that any cable segment can be divided into parts of 0.5in length that each qualify as cable. The semantics of the star operator allows these parts to overlap, which makes sure that the account extends to cable segments whose length in inches is not a multiple of 0.5.

We can now formulate the constraint on Shares in pseudopartitives with the help of stratified measurement reference. This constraint links $\varepsilon(K)$ to the Key (the measure nominal) of the pseudopartitive to express the fact that what counts as very small can vary from pseudopartitive to pseudopartitive.

(45)   **Constraint on Shares of pseudopartitives**
A pseudopartitive whose Key is K, whose Share is S and whose Map is M is acceptable only if S has stratified measurement reference with respect to M (formally: $\text{SMR}_{M,\varepsilon(K)}(S)$), where the threshold $\varepsilon(K)$ is very small with respect to K.

The following examples illustrate how this constraint works. The first example assumes that the covert Map is length:

(46)   three inches of cable
  a. Key: three inches
  b. Share: cable
  c. Map: length
  d. Threshold: e.g. $\lambda d.\text{inches}(d) \leq 0.5$

Applied to example (46), constraint (45) predicts that a sentence containing (46) is acceptable only if the condition $\text{SMR}_{length,\ \lambda d.inches(d)\ \leq\ 0.5}(\text{cable})$ is fulfilled. This condition expands as follows:

(47) $\text{SMR}_{length,\ \lambda d.inches(d)\ \leq\ 0.5}(\text{cable})$
$$\Leftrightarrow \forall x[\text{cable}(x) \rightarrow x \in {}^*\lambda y \left(\begin{array}{l} \text{cable}(y)\ \wedge \\ \text{inches}(\text{length}(y)) \leq 0.5 \end{array} \right)]$$
(Every cable segment can be divided into one or more cable segments ("strata") whose length is at most 0.5 inches.)

This condition is intuitively fulfilled by the predicate *cable*. We have assumed above that *cable* has divisive reference relative to length. We then concluded that this assumption is too strong because arbitrarily short parts of a cable segment do not always qualify as cable segments. The new definition corrects this problem.

The following example illustrates why *cable* cannot be used in a pseudopartitive in connection with a diameter Map:

(48)   three inches of cable

a. Key: three inches
b. Share: cable
c. Map: diameter
d. Threshold: e.g. $\lambda d.\text{inches}(d) \leq 0.5$

Applied to example (48), constraint (45) predicts that the sentence is acceptable only if the condition $\text{SMR}_{diameter,\ \lambda d.inches(d)\ \leq\ 0.5}(\text{cable})$ is fulfilled. This condition expands as follows:

(49) $\text{SMR}_{diameter,\ \lambda d.inches(d)\ \leq\ 0.5}(\text{cable})$

$\Leftrightarrow \forall x[\text{cable}(x) \to x \in {}^*\lambda y \left(\begin{array}{l} \text{cable}(y)\ \wedge \\ \text{inches}(\text{diameter}(y)) \leq 0.5 \end{array} \right)]$

(Every cable segment can be divided into one or more cable segments ("strata") whose diameter is at most 0.5 inches.)

As discussed above, this condition is not fulfilled, because when a cable segment is cut lengthwise, the resulting parts do not generally qualify as cable segments.

4.6   Unifying the constraints

The fundamental parallel across distributive constructions emerges when we put the properties that were defined in the previous section side by side:

(50)   **Definition: Stratified distributive reference**

$\text{SDR}_\theta(P) \stackrel{\text{def}}{=} \forall e[P(e) \to e \in {}^*\lambda e' \left(\begin{array}{l} P(e')\ \wedge \\ \text{PureAtom}(\theta(e')) \end{array} \right)]$ = (19)

(51)   **Definition: Stratified subinterval reference**

$\text{SSR}_{\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall e[P(e) \to e \in {}^*\lambda e' \left(\begin{array}{l} P(e')\ \wedge \\ \varepsilon(K)(\tau(e')) \end{array} \right)]$ = (30)

(52)   **Definition: Stratified measurement reference**

$\text{SMR}_{\mu,\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall x[P(x) \to x \in {}^*\lambda y \left(\begin{array}{l} P(y)\ \wedge \\ \varepsilon(K)(\mu(y)) \end{array} \right)]$ = (44)

We arrived at these three properties from very different starting points: stratified distributive reference through the attempt to model predicative distributivity in an event-based framework; stratified subinterval reference as a refinement of the subinterval property to take the minimal-parts problem into account; and stratified measurement reference from an attempt to explain why *three inches of cable* cannot refer to a three-inch-thick cable segment.

Although the three properties were independently motivated, they are very similar to each other. They can now be subsumed under one common property, which I simply call stratified reference.

(53) **Definition: Stratified reference**

$$\text{SR}_{f,\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow x \in {}^*\lambda y \left(\begin{array}{c} P(y) \wedge \\ \varepsilon(K)(f(y)) \end{array} \right)]$$

(A predicate P has stratified reference with respect to a function f and a threshold $\varepsilon(K)$ if and only if there is a way of dividing every entity in its denotation exhaustively into parts ("strata") which are each in P and which have a very small f-value. Very small f-values are those that satisfy $\varepsilon(K)$.)

The symbols for constants and variables in this definition should be taken to be unsorted, since the definition is intended to subsume properties of event predicates and properties of predicates over ordinary objects.

With the help of stratified reference, we can now unify the constraints on distributive constructions which were formulated separately above, and which are repeated here for convenience.

(54) **Constraint on Shares of adverbial-*each* sentences** = (22)
A sentence with adverbial *each* whose Key is K, whose Share is S and whose Map is M is acceptable only if S has stratified distributive reference with respect to M (formally: $\text{SDR}_M(S)$).

(55) **Constraint on Shares of temporal *for*-adverbials** = (31)
A temporal *for*-adverbial whose Key is K and whose Share is S is acceptable only if S has stratified subinterval reference (formally: $\text{SSR}_{\varepsilon(K)}(S)$), where the threshold $\varepsilon(K)$ holds of intervals whose runtime is very short with respect to K.

(56) **Constraint on Shares of pseudopartitives** = (45)
A pseudopartitive whose Key is K, whose Share is S and whose Map is M is acceptable only if S has stratified measurement reference with respect to M (formally: $\text{SMR}_{M,\varepsilon(K)}(S)$), where the threshold $\varepsilon(K)$ is very small with respect to K.

With the understanding the adverbial-*each* constructions, pseudopartitives and *for*-adverbials are distributive constructions, these constraints can now be replaced by the following statement:

(57) **Distributivity Constraint**
A distributive construction whose Key is K, whose Share is S and whose Map is M is acceptable only if S has stratified reference with respect to M

(formally: $\text{SR}_{M,\varepsilon(K)}(S)$).

This constraint makes reference to the threshold function ε, which we can now define as follows.

(58)   **Definition:** ε
For any predicate K, if $K \subseteq {}^*\text{PureAtom}$, then $\varepsilon(K) = \text{PureAtom}$.
Otherwise, $\varepsilon(K) = \lambda x[M(x) \leq k]$, where k is some constant that is very small with respect to the entities in K.

This definition implements the insight that distributivity in *each* constructions is always over atoms, while distributivity in *for*-adverbials and pseudopartitives are over vague granularity levels. The effect of this definition is that for the adverbial-*each* construction, the threshold $\varepsilon(K)$ always resolves to *PureAtom*, no matter what value K takes; otherwise it resolves to a predicate that applies to degrees or intervals which are very small given K as a comparison class. The definition does not specify exactly how what counts as very small is determined. I assume that what counts as very small is a vague notion. There are many ways to model vagueness (see Williamson (1994), Keefe (2000), and Kennedy (2010) for overviews) and my interest here is not in them, so I simply assume that ε applies to certain intervals below a cutoff point, which is represented through the constant k.

4.7   Compositional implementation

The Distributivity Constraint does not specify anything about compositional implementations. This section shows one possible way to incorporate it into a compositional framework, namely by using the LFs presented in Chapter 3. I implement the constraint as a selectional restriction on the words that occur necessarily as part of a distributive construction. Following McCawley (1968) and Singh (2007), I assume that selectional restrictions are a kind of lexical presupposition. For example, in *for*-adverbials, the presupposition is attached to the word *for*; in *each*-constructions, to the word *each*; in pseudopartitives, to the word *of*. In a framework that uses partial truth values in the style of Karttunen and Peters (1979), Muskens (1996) or Heim and Kratzer (1998), a lexical presupposition is expressed as a restriction on the possible input values of a function. The expression $\lambda x : \varphi \cdot \psi$ represents the partial function that is defined for all x such that its lexical presupposition φ holds, and that returns ψ wherever the function is defined. Sentences are interpreted as pairs of propositions: an assertion and a presupposition. These presuppositions must be fulfilled in order for the sentence to have a truth value. I assume that the global presupposition of a sentence is the conjunction of all the

lexical presuppositions associated with its lexical items. That is, all presuppositions project straight to the top. Sentences whose global presupposition is true have the same truth value as their assertion; sentences where it is false lack a truth value. Denotations of lexical items that carry a presupposition are represented as partial functions that are undefined whenever this lexical presupposition is false.

This account is obviously too simple to model presupposition projection and accommodation in a serious way. I ignore the finer details of presupposition projection for two reasons. First, none of the examples I consider in this work involve more than one clause, so there would be little room for presuppositions to project other than going to the top. Second, the kinds of presuppositions I model are not based on contingent factual information. This would make it difficult to interpret them in any other way than by projecting them to the top even in multiclause contexts. For example, *John thinks that Bill ran a mile for an hour* has the presupposition that *run a mile* is atelic. If *run a mile* could be locally satisfied, we would expect this sentence to have the interpretation *John thinks that Bill ran a mile in an hour and that running a mile is atelic*. But it is not a contingent fact that the predicate *run a mile* is telic. Believing that it is atelic would require the hearer to believe that as soon as one starts running a mile, one has already run a mile. If John knows the meaning of *run* and of *a mile*, he should not be able to hold this kind of belief.

One might similarly ask whether it is possible to accommodate the kinds of presuppositions I assume. I do not have much to say about this question. However, in Chapter 8, we will consider various flavors of distributive operators. As we will see, insertion of such an operator in the scope of a presupposition-carrying distributive item can be seen as a repair strategy, somewhat similar to presupposition accommodation.

The following lexical entries all embody the Distributivity Constraint. They are identical to the skeletal entries in Chapter 3, except that I have added a lexical presupposition. The letters K, S, and M stand for Key, Share, and Map respectively. For clarity, the entry for *each* anticipates the effect of Definition (58) and replaces $\varepsilon(K)$ with *PureAtom*.

(59) $[\![of]\!] = \lambda S_{\langle at \rangle} \lambda M_{\langle \alpha\beta \rangle} \lambda K_{\langle \beta t \rangle} \lambda b_{\langle \alpha \rangle} : \mathrm{SR}_{M,\varepsilon(K)}(S) \, . \, S(b) \, \wedge \, K(M(b))$

(60) $[\![for]\!] = \lambda K_{\langle it \rangle} \lambda M_{\langle vi \rangle} \lambda S_{\langle vt \rangle} \lambda e : \mathrm{SR}_{M,\varepsilon(K)}(S) \, . \, S(e) \, \wedge \, K(M(e))$

(61) $[\![each]\!] = \lambda S_{\langle vt \rangle} \lambda M_{\langle ve \rangle} \lambda K_{\langle et \rangle} \lambda e : \mathrm{SR}_{M, \, PureAtom}(S) \, . \, S(e) \, \wedge \, K(M(e))$

The skeletal LFs from Chapter 3 are repeated in Figures 4.1 through 4.4 with these entries added. These full LFs are my official proposal. The types of these entries reflect the specific order in which they combine with their constituents according to these LFs. This order is not essential and could be easily changed

without consequences for the theory.

For ease of reference, here is the translation of a temporal *for*-adverbial:

(62) $[\![\text{for an hour}]\!]$
$= \lambda P_{\langle vt \rangle} \lambda e : \text{SR}_{\tau, \varepsilon(\lambda t[\text{hours}(t)=1])}(P).$
$P(e) \wedge \text{hours}(\tau(e)) = 1$

This translation is obtained as a result of combining the entry for *for* in (60) with *an hour* and with a [runtime] head as shown in Figure 4.3.

Also for ease of reference, I spell out and expand the results of the computations in Figures 4.1 through 4.4.

(63) $[\![\text{three liters of water}]\!]$ =
$\lambda x : \text{SR}_{volume,\ \varepsilon(\lambda d[liters(d) = 3])}(\lambda x[\text{water}(x)]).$
$[\text{water}(x) \wedge \text{liters}(\text{volume}(x)) = 3]$
(This function is defined if and only if *water* has stratified reference with respect to dimension parameter *volume* and granularity parameter set to a very small value compared with three liters. If defined, it is true of any water amount whose volume measures three liters.)

(64) $[\![\text{three hours of walking}]\!]$ =
$[\![\text{walk for three hours}]\!]$ =
$\lambda e : \text{SR}_{\tau,\ \varepsilon(\lambda t[hours(t) = 3])}(\lambda e[^*\text{walk}(e)]).$
$[^*\text{walk}(e) \wedge \text{hours}(\tau(e)) = 3]$
(This function is defined if and only if *walk* has stratified reference with respect to dimension parameter τ and granularity parameter set to a very small value compared with three hours. If defined, it is true of any walking event whose runtime measures three hours.)

(65) $[\![\text{three boys each walked}]\!]$ =
$\lambda e : \text{SR}_{ag,\ PureAtom}(\lambda e[^*\text{walk}(e)]).$
$[^*\text{walk}(e) \wedge {^*}\text{boy}(^*\text{ag}(e)) \wedge |^*\text{ag}(e)| = 3]$
(This function is defined if and only if *walk* has stratified reference with respect to dimension parameter *agent* and granularity parameter *PureAtom*. If defined, it is true of any walking event whose plural agent is three boys.)

The definedness conditions of these functions expand as in (66) through (68). I assume that these definedness conditions are all fulfilled as a matter of world knowledge or lexical semantics. The expansions in (66) through (68) can be understood as meaning postulates. Their paraphrases should make it clear that this is a plausible assumption.

(66) $\text{SR}_{volume,\ \varepsilon(\lambda d[\text{liters}(d) = 3])}(\lambda x[\text{water}(x)])$

$\Leftrightarrow \forall x[\text{water}(x) \to x \in {}^*\lambda y \left(\begin{array}{l} \text{water}(y)\ \wedge \\ \varepsilon(\lambda d[\text{liters}(d) = 3])(\text{volume}(y)) \end{array} \right)]$

(There is a way of dividing every water amount exhaustively into parts ("strata") which are water amounts and whose volumes are very small compared to three liters.)

(67) $\text{SR}_{\tau,\ \varepsilon(\lambda t[\text{hours}(t) = 3])}(\lambda e[{}^*\text{walk}(e)])$

$\Leftrightarrow \forall e[{}^*\text{walk}(e) \to e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\text{walk}(e')\ \wedge \\ \varepsilon(\lambda t[\text{hours}(t) = 3])(\tau(e')) \end{array} \right)]$

(There is a way of dividing every walking event exhaustively into parts ("strata") which are walking events and whose runtimes are very small compared to three hours.)

(68) $\text{SR}_{ag,\ PureAtom}(\lambda e[{}^*\text{walk}(e)])$

$\Leftrightarrow \forall e[{}^*\text{walk}(e) \to e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\text{walk}(e')\ \wedge \\ \text{PureAtom}(\text{ag}(e')) \end{array} \right)]$

(There is a way of dividing every walking event exhaustively into parts ("strata") which are walking events and whose agents are pure atoms.)

4.8 Summary

Chapter 1 has presented the parallelism between the telic-atelic, collective-distributive, singular-plural, and count-mass oppositions in terms of boundedness. Intuitively, singular, telic, and collective predicates are delimited or bounded in ways that plural, mass, atelic, and distributive predicates are not. I pointed out that making formal sense of this parallelism amounts to answering what I called the *boundedness question*: How can the the difference between boundedness and unboundedness be formally characterized?

This chapter has presented my answer to the boundedness question. The higher-order property of stratified reference characterizes what it means to be unbounded. This property is parametrized, which reflects the fact that unboundedness can be understood in more than one way. As the example *run to the store* shows, one and the same verb phrase can be distributive (unbounded with respect to agents) and telic (bounded with respect to runtime). As the example *kill* shows, one and the same verb can be collective on its agent position (bounded with respect to agents) and distributive on its theme position (unbounded with respect to themes). As the example *cable* shows, one and the same predicate can be divisive along its length (bounded with respect to length) and nondivisive along its diameter (unbounded with respect to diameter).

The picture presented in this chapter is idealized in several respects. For example, I have not discussed the fact that distributive entailments do not always literally hold as far down as they can. *for*-adverbials are compatible with predicates like *waltz* and *sleep in the attic* that do not satisfy the subinterval property (Dowty 1979). Similarly, pseudopartitives are compatible with heterogeneous mass nouns like *fruit cake* (Taylor 1977). The next chapter motivates the granularity parameter of stratified reference. As we will see, this parameter "makes room" for such predicates in distributive constructions.

Figure 4.1: Full LF of an ordinary pseudopartitive

$\langle et \rangle$

$\lambda x : \mathrm{SR}_{volume,\ \varepsilon(\lambda d[liters(d)\ =\ 3])}(\lambda x[\mathrm{water}(x)]).$
$[\mathrm{water}(x)\ \wedge\ \mathrm{liters}(\mathrm{volume}(x))\ =\ 3]$

three
$\langle n \rangle$
3

liters
$\langle n, dt \rangle$
$\lambda n \lambda d[\mathrm{liters}(d)\ =\ n]$

[volume]
$\langle ed \rangle$
$\lambda x[\mathrm{volume}(x)]$

of
$\langle \alpha t, \langle \alpha \beta, \langle \beta t, \alpha t \rangle \rangle \rangle$
$\lambda S \lambda M \lambda K \lambda b :$
$\mathrm{SR}_{M, \varepsilon(K)}(S)\ .$
$[S(b)\ \wedge\ K(M(b))]$

water
$\langle et \rangle$
$\lambda x[\mathrm{water}(x)]$
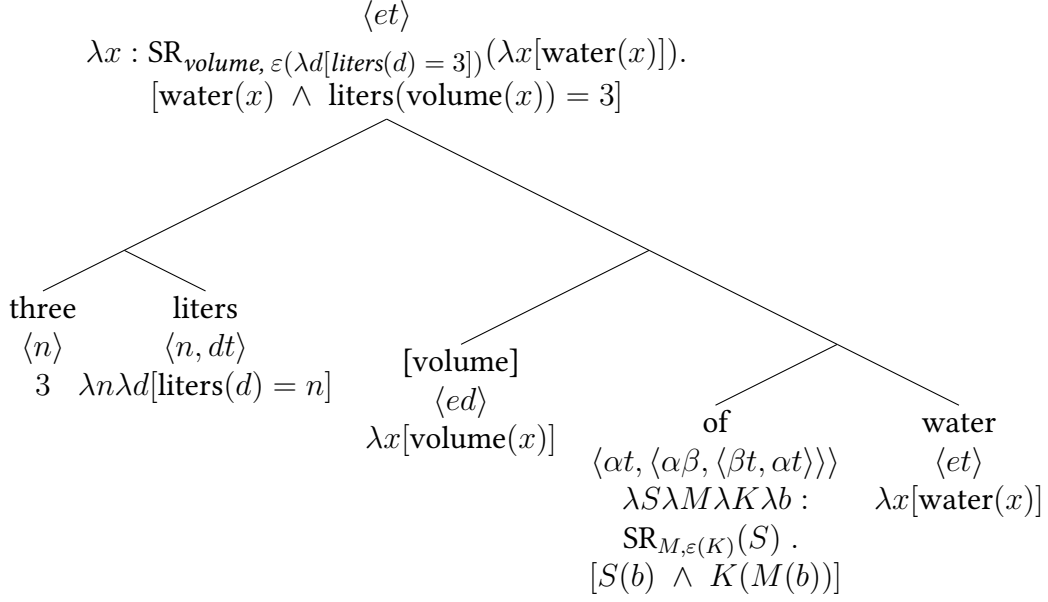
Figure 4.2: Full LF of an event pseudopartitive

$\langle vt \rangle$

$\lambda e : \mathrm{SR}_{\tau,\ \varepsilon(\lambda t[hours(t)\ =\ 3])}(\lambda e[^*\mathrm{walk}(e)]).$
$[^*\mathrm{walk}(e)\ \wedge\ \mathrm{hours}(\tau(e))\ =\ 3]$

three
$\langle n \rangle$
3

hours
$\langle n, it \rangle$
$\lambda n \lambda t[\mathrm{hours}(t)\ =\ n]$

[runtime]
$\langle vi \rangle$
$\lambda e[\tau(e)]$

of
$\langle \alpha t, \langle \alpha \beta, \langle \beta t, \alpha t \rangle \rangle \rangle$
$\lambda S \lambda M \lambda K \lambda b :$
$\mathrm{SR}_{M, \varepsilon(K)}(S)\ .$
$[S(b)\ \wedge\ K(M(b))]$

walking
$\langle vt \rangle$
$\lambda e[^*\mathrm{walk}(e)]$

97
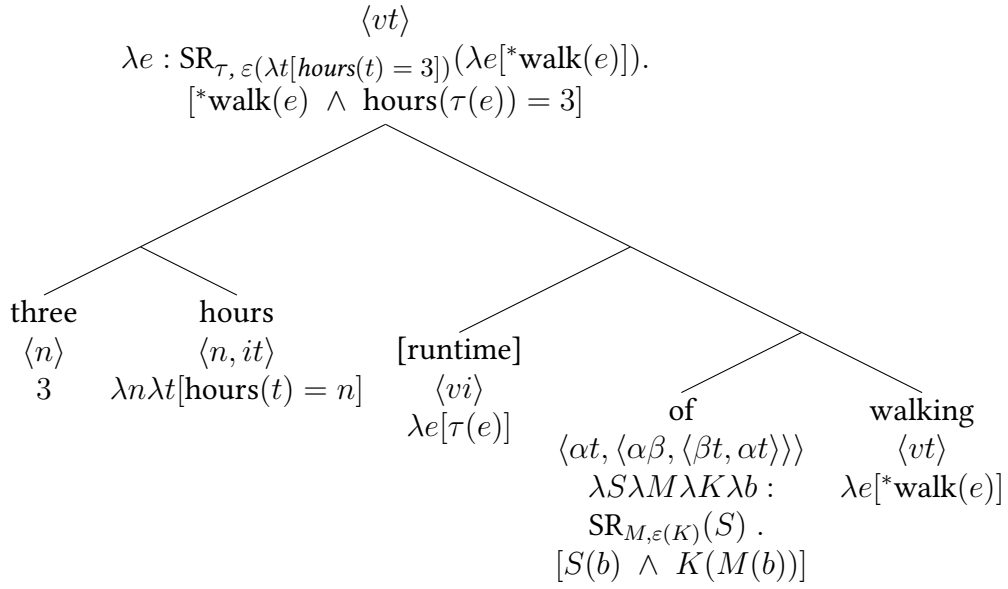
Figure 4.3: Full LF of a *for*-adverbial

$\langle vt \rangle$
$\lambda e : \text{SR}_{\tau,\, \varepsilon(\lambda t[\mathit{hours}(t) = 3])}(\lambda e[^{*}\text{walk}(e)]).$
$[^{*}\text{walk}(e) \;\wedge\; \text{hours}(\tau(e)) = 3]$

walk
$\langle vt \rangle$
$\lambda e[^{*}\text{walk}(e)]$

[runtime]
$\langle vi \rangle$
$\lambda e[\tau(e)]$

for
$\langle it, \langle vi, \langle vt, vt \rangle \rangle \rangle$
$\lambda K \lambda M \lambda S \lambda e :$
$\text{SR}_{M,\varepsilon(K)}(S)\; .$
$[S(e) \;\wedge\; K(M(e))]$

three
$\langle n \rangle$
3

hours
$\langle n, it \rangle$
$\lambda n \lambda t[\text{hours}(t) = n]$
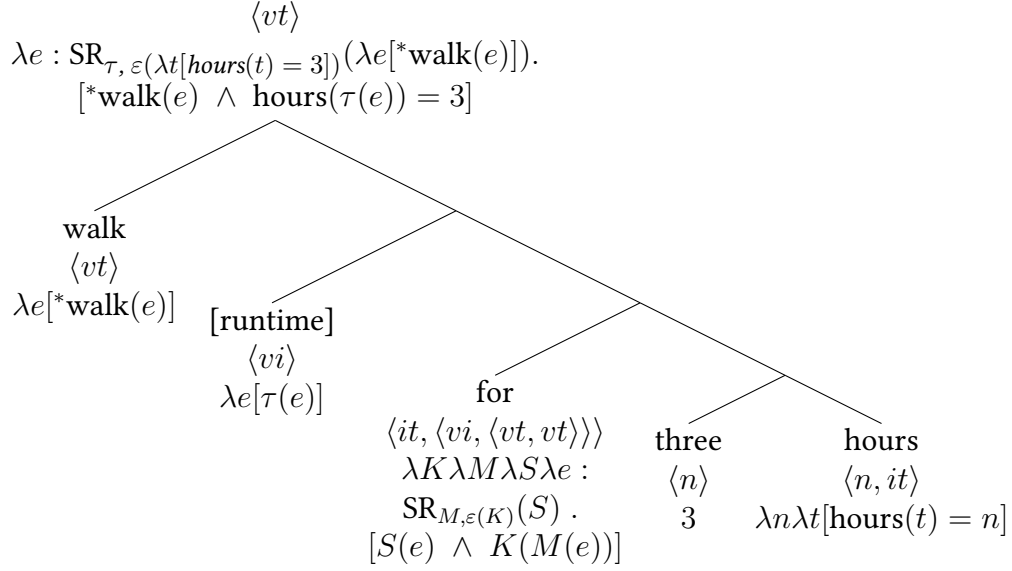
Figure 4.4: Full LF of an adverbial-*each* construction

$\langle et \rangle$
$\lambda e : \text{SR}_{ag,\, \mathit{PureAtom}}(\lambda e[^{*}\text{walk}(e)]).$
$[^{*}\text{walk}(e) \;\wedge\; {^{*}}\text{boy}(^{*}\text{ag}(e)) \;\wedge\; |^{*}\text{ag}(e)| = 3]$

three
$\langle n \rangle$
3

[many]
$\langle n, et \rangle$
$\lambda n \lambda x[|x| = n]$

boys
$\lambda x[^{*}\text{boy}(x)]$

[ag]
$\langle ve \rangle$
$\lambda e[^{*}\text{ag}(e)]$

each
$\langle vt, \langle ve, \langle et, vt \rangle \rangle \rangle$
$\lambda S \lambda M \lambda K \lambda e :$
$\text{SR}_{M,\, \mathit{PureAtom}}(S)\; .$
$[S(b) \wedge K(M(e))]$

walk
$\langle vt \rangle$
$\lambda e[^{*}\text{walk}(e)]$
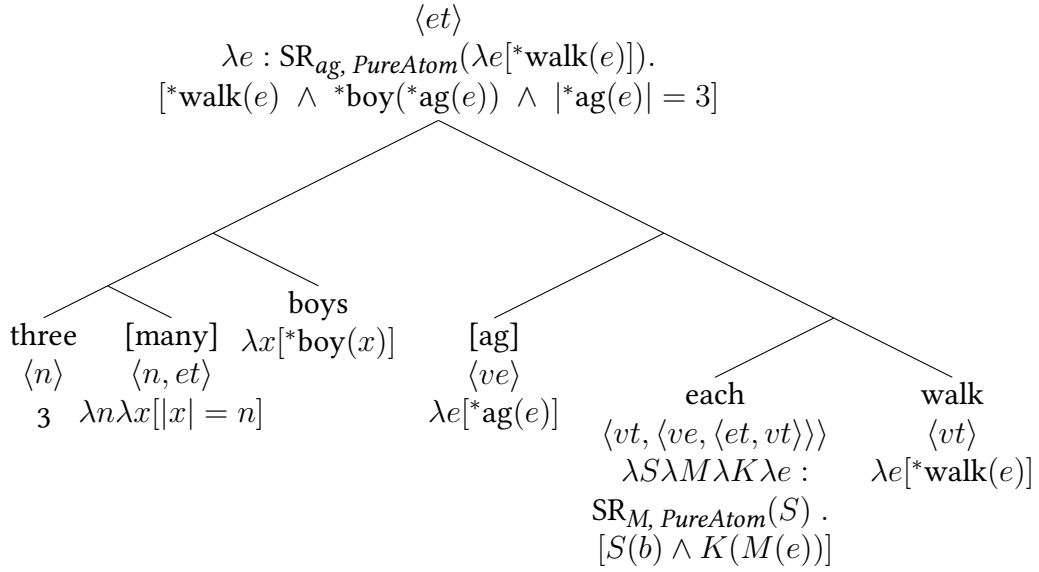
Chapter 5

Minimal parts

5.1 Introduction

In strata theory, as presented in Chapter 4, the Shares of adverbial-*each* construc-
tions, *for*-adverbials, and pseudopartitives are characterized with the help of a novel
higher-order property called stratified reference. The Shares of these construc-
tions are distributive predicates, atelic predicates, and mass or plural predicates,
respectively. With respect to adverbial-*each* constructions, I have argued that
stratified reference is an adequate way to transfer traditional notions of predicative
distributivity into an event-based framework. With respect to *for*-adverbials and
pseudopartitives, I have briefly argued that stratified reference surpasses the subin-
terval property and divisive reference, which are the properties commonly used
for the purpose of characterizing atelic, plural and mass reference. In this chapter,
I present the case against the subinterval property and against divisive reference in
more detail. The case is based on what has been known at least since Dowty (1979)
as the minimal-parts problem: entities and events that satisfy an atelic or mass
predicate may have parts that are too small to still satisfy the predicate. I compare
my implementation based on stratified reference with previous accounts of the
minimal-parts problem (Dowty 1979; Hinrichs 1985; Moltmann 1989, 1991; Link 1991).
I argue that these previous accounts suffer from various technical deficiencies and
that stratified reference captures the essence of the minimal-parts problem in a
more accurate way.

5.2 The minimal-parts problem

The minimal-parts problem occurs in the domains of verbal and nominal predicates
(Taylor 1977; Dowty 1979). I discuss each of them in turn.

5.2.1 The minimal-parts problem for verbal predicates

As discussed in Section 4.5.2, the entailment properties of atelic predicates with respect to a large range of other phenomena can be represented as a first approximation in terms of the subinterval property (Bennett and Partee 1972; Dowty 1979). A predicate P has the subinterval property if and only if whenever it holds at an interval, it also holds at every one of its subintervals. The minimal-parts problem consists in the fact that the subinterval property only applies to certain atelic predicates if we are willing to disregard subintervals below a certain threshold. For example, *waltz* is an atelic predicate, but waltzing takes at least three steps, and there is no consensus on whether or not a given event whose runtime is substantially shorter than this three-step threshold still qualifies as waltzing. For this reason, it is difficult to know whether the subinterval property applies to *waltz*, and it is not clear whether this property is really an adequate characterization of atelicity. We can perhaps already conclude from this fact that the subinterval property must be given up. This conclusion is sometimes objected to on the grounds that that the problem is not part of linguistics and should be ignored. Such an objection assumes that we do not really know what kinds of events occur at instants or even at very short subintervals, or in any case that these events should be seen as irrelevant for the purposes of semantics because they do not enter our daily experience. Once such events are excluded the subinterval property can still technically be assumed to hold of all atelic predicates.

I think this kind of objection misses the point. There is no benefit in choosing the underlying assumptions of semantic theory with the sole purpose of being able to continue using the subinterval property, because one might as well avoid the problem by using a different property than the subinterval property to begin with. The extent to which our intuitions about short events would have to be changed in order to maintain that atelic predicates have the subinterval property is considerable. For example, in order to give the predicate *pass on* in (1) the subinterval property, even generations would have to be considered as having minimal time:

(1)   The Chinese people have created abundant folk arts, such as paper-cuttings, acrobatics, etc., passed on from generation to generation for thousands of years.[16]

It does not seem practical to maintain that the predicate *pass on from generation to generation* really has the subinterval property, but it is clearly atelic since it can

---

[16]Attested example (http://www.twinbridge.com/detail.aspx?ID=315). Accessed November 15, 2010.

be modified by a *for*-adverbial. I conclude that atelic predicates do not necessarily have the subinterval property. This conclusion is not new. As we will see, Dowty (1979) already recognized that the subinterval property is an idealization, and many subsequent authors have avoided using this property in their formalizations (see Section 5.3 below).

5.2.2   The minimal-parts problem for mass nouns

An analogous concept to the subinterval property, divisive reference, is sometimes used following Cheng (1973) as a defining semantic propety of mass nouns: any part of something denoted by a mass noun is assumed to be denoted by the same mass noun. The definition of divisive reference is repeated here from Section 2.3.2:

(2)   **Definition: Divisive reference**
      $\mathrm{DIV}(P) \overset{\text{def}}{=} \forall x[P(x) \to \forall y[y < x \to P(y)]]$
      (A predicate P is divisive if and only if whenever it holds of something, it also holds of each of its proper parts.)

The assumption that divisive reference holds of mass denotations runs into a well-known problem with very small parts of mass substances. This problem is analogous to the one that arises in connection with the subinterval property, and concerns the atomic nature of matter: arguably hydrogen atoms do not qualify as water, so if we consider hydrogen atoms to be part of water, then not every part of water is water (Quine 1960). A different form in which this problem occurs comes from heterogeneous mass predicates like *fruit cake*, *pea soup*, or *succotash* (e.g. Taylor 1977): for example, a portion of fruit cake may contain sultanas, but these sultanas do not themselves qualify as fruit cake. This type of argument is arguably stronger than Quine's argument, because in this case, the minimal parts that do not qualify are directly accessible to the senses. While it might be possible to save the assumption that mass nouns have divisive reference by assuming that we conceptualize even predicates like *fruit cake* as having divisive reference, this assumption is even more difficult to justify in the case of fake mass nouns like *furniture*. Psycholinguistic experiments suggest that the cognitive structures underlying fake mass nouns are more similar to those of count nouns than to those of other mass nouns (see Section 2.4.1). It would be implausible both in light of the factual reality and in light of our cognitive modeling of it to assume that fake mass nouns have having divisive reference (Barner and Snedeker 2005; Chierchia 2010).

The conclusion that mass nouns do not always have divisive reference appears to be generally accepted. Gillon (1992) notes that "[w]hile some semanticists retain the divisivity of reference as a criterion to distinguish mass nouns from count

nouns ... only Bunt (1979, 1985) has attempted to justify the retention." (p. 598). Gillon gives counterarguments to Bunt's reasons and argues that the grammar is mute on whether or not mass terms have divisive reference. This is also the position I adopt here.

5.3 Previous accounts of the minimal-parts problem

The significance of the minimal-parts problem lies in the challenges it poses for formalizing atelicity checks and checks for mass reference. For example, a proposal that explains why *for*-adverbials reject telic predicates by assuming that *for*-adverbials check for the subinterval property faces the minimal-parts problem. Solutions to the problem consist of reformulations of the subinterval property in a way that retains their capability to distinguish between atelic and telic predicates while correctly categorizing predicates like *waltz*.

This section reviews a representative sample of previous approaches to the minimal-parts problem that gives an idea of the space of possibilities. I discuss the promising aspects and the technical problems of these proposals. I have not attempted to include every approach known to me. For example, I do not discuss the interesting accounts in Vlach (1993), Piñón (1999a), van Geenhoven (2004), or Landman and Rothstein (2009). Other reviews of approaches to the minimal-parts problem are found in Krifka (1986) and Mollá-Aliod (1997). I concentrate on the domain of aspect rather than on mass reference, because this is where most solutions seem to have been proposed.

5.3.1 Dowty (1979)

The first formal proposal for a model-theoretic translation of *for*-adverbials is found in Dowty (1979, p. 333). His analysis is prefigured in Dowty (1972). Dowty models *for*-adverbials as universal quantifiers ranging over the subintervals of some time interval. In other words, he implements the subinterval property directly, and as a part of the asserted content rather than as a presupposition. Here is Dowty's proposal in its original form.

(3) a. $\llbracket \text{for} \rrbracket = \lambda P_t \lambda P \lambda x [P_t\{n\} \ \wedge \ \forall t[t \subseteq n \rightarrow \text{AT}(t, P\{x\})]]$
    b. $\llbracket \text{an hour} \rrbracket = \lambda t[\text{an-hour}(t)]$

This entry is used in an intensional system where formulas are interpreted relative to a world-time index $\langle w, i \rangle$. The notation $P\{x\}$ is an abbreviation for $\check{}P(x)$, the result of applying the extension of P to x. Under any world-time index $\langle w, i \rangle$, the indexical constant n denotes the time index i. Phrases like *an hour* and

six weeks are interpreted as predicates of times. For example, *an hour* denotes the set of all contiguous temporal intervals whose duration is an hour. $AT(t, p)$ evaluates a proposition p at a time index t.

In order to translate the sentence *John slept for an hour*, Dowty combines *for* with *an hour*, *sleep*, *john*, and with an operation that introduces an existential quantifier over past times. I abstract away from various details of his analysis not relevant here. In particular, his analysis is embedded in an intensional system, but this fact plays no role for his analysis of *for*-adverbials. Without its irrelevant parts, Dowty's translation of this sentence can be represented as follows (for the original analysis, see Dowty (1979), p. 334):

(4)   ⟦John slept for an hour⟧
      = $\exists t_1[\text{an-hour}(t_1) \;\wedge\; \forall t_2[t_2 \subseteq t_1 \rightarrow \text{sleep}(j, t_2)]]$
      (There is a time interval t_1 which lasts an hour and John sleeps at each of its subintervals.)

The representation (4) essentially implements the subinterval property one-to-one, because it quantifies over all subintervals of the interval t_1. Dowty is aware of the minimal-parts problem and suggests that to account for it, quantification should not be over literally all subintervals, as in this representation, but only over "all subintervals large enough to be minimal intervals for the activity in question"; he notes that how to do this is unclear. For simplicity, he decides to leave his analysis as it stands.

For many purposes, Dowty's analysis of aspect is adequate even though (or because) it ignores the minimal parts problem. It is adopted by many authors who are concerned with the interaction of *for*-adverbials with other semantic components, in particular the Perfect, rather than with the meaning of *for*-adverbials themselves. This line of work includes Richards (1982), Heny (1982), Mittwoch (1988), Parsons (1990), Abusch and Rooth (1990), Kamp and Reyle (1993), Hitzeman (1997), Iatridou, Anagnostopoulou, and Izvorski (2001), and Rathert (2004). The latter contains an excellent review of this part of the literature. Here, I ignore these approaches and I focus on the minimal-parts problem.

5.3.2   Hinrichs (1985)

Hinrichs (1985), a dissertation supervised by Dowty, assumes a Davidsonian event-based semantics. In order to avoid the minimal-parts problem, Hinrichs relaxes the requirement that *for*-adverbials place on their predicates. Hinrichs' proposal is complicated by the fact that he adopts Carlson (1977)'s three-tiered ontology of kinds, objects, and stages, which Carlson proposed in order to account for the

properties of bare plurals and generic sentences. Abstracting away from this and other complexities, here is how *John slept for an hour* comes out on his proposal:

(5) $[\![$John slept for an hour$]\!]$
        $= \exists e \exists l[\text{hour}(l) \ \wedge \ l \leq \tau(e) \ \wedge \ \text{sleep}(j)(e)] \ \wedge$
        $\forall l'[l' < l \rightarrow \exists e'[e' < e \ \wedge \ l' \leq \tau(e') \ \wedge \ \text{sleep}(j)(e')]]$

Hinrichs comments on his translation as follows (p. 235):

> The translation in [(5)] requires that …there has to be a spatio-temporal location l with the property of being one hour long such that the entire process of John's sleeping spatio-temporally contains l and for each proper sublocation l' of l there has to be a proper subprocess of John's sleeping containing l'. … [W]e don't require that each sublocation denoted by the temporal measurement phrase has to *be* a subprocess itself. The requirement that each proper sublocation be contained in a proper subprocess of the maximal process making up the event, rules out that for each sublocation we could simply pick the maximal process itself.

The way Hinrichs avoids the minimal-parts problem does not follow Dowty's suggestion of quantifying over less than literally all subintervals. Instead of requiring that a predicate like *John sleep* be true at each or most subintervals, Hinrichs requires that the runtime of every subinterval of John's sleeping must be *contained in* (and not necessarily equal to) that of a sleeping event. This sleeping event must be a proper part of the event that the sentence describes. This is possible because *sleep* is not quantized; if we assume that all telic predicates are quantized, they are correctly ruled out. The assumption that all telic predicates are quantized is not available to us because it is incompatible with the assumption of lexical cumulativity (see Section 2.7.2). But Hinrichs (1985) does not assume lexical cumulativity.

Krifka (1986) (p. 150) criticizes Hinrichs' approach as arbitrary and notes that the proper-part requirement does not explain anything, since it does not serve any purpose other than that of excluding telic predicates. There is also another problem for Hinrichs' account, which causes it to run into something very similar to the minimal-parts problem. The variable l' in (5) ranges over subintervals rather than (just) over instants. On the assumption that time is dense, it ranges over subintervals that are only minimally shorter than l. For example, it ranges over intervals of length 58, 59, 59½…minutes. For each of these intervals, the definition requires there to be a proper part of the event of John's sleeping which lasts at least as long as the interval and which qualifies as John's sleeping. This gives the events

in the denotation of atelic predicates a dense structure. Moreover, depending on the structure of the underlying mereology, there will generally be a unique complement that results from removing the smaller sleeping event from the larger sleeping event. In classical extensional mereology (CEM), this is a consequence of Unique Separation (see Section 2.3.3). Depending on how large a sleeping event one separates off, the runtime of its complement may become arbitrary small. It follows on Hinrichs' account that only events with arbitrarily short parts can be described by a sentence with a *for*-adverbial, certainly not an intuitive assumption.

At first sight it might look like Hinrichs' approach could be salvaged by restricting l' to range only over instants and not intervals. However, this requires the assumption that time is atomic, otherwise there are no instants. According to von Stechow (2009), this assumption is rejected by most semanticists, and I do not rely on it myself (see Section 2.4.4). Even if we adopt the assumption, the resulting modification of Hinrichs' approach becomes too weak, as it ends up merely requiring that the main-clause event can be divided into at least two possibly overlapping parts that fulfill the main-clause predicate. That is, *P for an hour* is predicted to be true of an event e if and only if P applies to two or more events $e_1, \ldots, e_n$ which are distinct from e and whose sum is e. These events may overlap and so their individual runtimes may be arbitrarily close to the runtime of e.

The analysis of Hinrichs (1985) is used in slightly adapted form by Abusch and Rooth (1990), but their adaptations do not prevent the problem from arising. A more significant reformulation of Hinrichs' proposal is found in Rathert (2004), who renders it using the following representation:

(6)   ⟦John ran for two hours⟧(t) = 1
       if and only if $|t| = 2$ hours $\land \forall m \subseteq t \exists n [m \subseteq n \subseteq t \land$ ⟦John ran⟧$(n) = 1]$,
       with n being a minimal run-event (two steps, done faster than a walk)

This reformulation runs into problems of its own: for example, since m is not required to be an instant, it ranges over some subintervals which are too big to contain the runtimes of any *minimal* run-events. In fairness, I note that Rathert is less interested in the meaning of *for*-adverbials than in how they interact with tense and with the Perfect. Still, we see that the proposal of Hinrichs (1985) is technically flawed and has resisted several attempts to repair it.

5.3.3 Moltmann (1989, 1991)

Moltmann (1989, 1991) recasts Dowty's analysis in an event semantic framework, and follows it quite closely. Her representation of an atelic predicate is as follows:

(7) 〚John slept for an hour〛
= $\exists t_1[\text{an-hour}(t_1) \ \wedge \ \forall t_2[t_2 P t_1 \rightarrow \exists e \ \text{sleep}(e, \text{john}) \ \wedge \ \tau(e) = t']]$
"There is a time interval t_1 which lasts an hour and John sleeps at each of its relevant parts."

Where Dowty has the relation $\subseteq$ and where the present framework would use the mereological relation $\leq$, Moltmann uses a relation she calls P. The most recent presentation of the formal model that describes the properties of this relation P is found in Moltmann (1997, 1998). This framework departs in significant ways from the basic axioms of CEM (see Section 2.3): The relation P is not assumed to satisfy transitivity, and the domains of individuals and events are not assumed to be closed under sum (see also the brief discussion of her views in Section 2.3.1). With respect to *for*-adverbials, Moltmann (1991) summarizes the relevant aspects of her model as follows:

> The part structure of an interval cannot be taken as being strictly divisive in a mathematical or physical sense. Rather, it appears that semantics involves a coarser part structure and a notion of relevant or contextually determined part, namely the relation P. Depending on the type of event, the part structure of the interval must have smallest subintervals of a certain minimal length. This is required, for instance, when the event is a process such as writing (not any physical part of a writing event is considered as writing) or a repetitive event (not any part of a repetitive revolving is a revolving). Therefore, the intended meaning of 'P' is the relation 'is a relevant part of', a relation which does not involve any subinterval of the measuring interval. ...P has to be understood not as a part relation in a strict mereological sense, but rather as a contextually determined relation that may be coarser than the mereological part relation, as the relation 'is relevant part of' ...One and the same entity may have different part structures depending on the respective context. For instance, a time interval may be conceived of as consisting of smallest subintervals of different length in different contexts – depending, for instance, on the type of events that are under consideration."

As Moltmann observes, there is a connection between the acceptability of sentences with *for*-adverbials and the size of the runtimes of the individual events of which the main-clause predicate holds relative to the duration indicated by the *for*-adverbial. As she puts it, the predicate modified by a *for*-adverbial must involve reference to an event that consists of "sufficiently many" subevents. I will refer to this observation as the **sufficiently-many-events observation**. An example is the

following minimal pair, taken from Moltmann (1991) with the question mark added according to the judgment described in her text.

(8)   **Context**: John can draw two pictures per hour.

    a.   ?For one hour John drew pictures.

    b.   For ten hours John drew pictures.

Moltmann notes that "[g]iven that John draws two pictures per hour, [(8a)] seems less acceptable than [(8b)]. Apparently, the partition of the interval into smallest parts in [(8a)] and [(8b)] depends on the relative size of the interval." In other words, the reason for the contrast between (8a) and (8b) appears to be that the ratio of half an hour (the time in which John draws a picture) to one hour is not sufficiently big, while the ratio of half an hour to ten hours is.

Moltmann intends the sufficiently-many-events observation to follow from her idea that the *for*-adverbial quantifies over "relevant or contextually determined" parts, and writes (pp. 635 and following):

> Concerning divisivity, [my analysis] requires only that the parts of the measuring entity, e.g. an interval or a region, be 'matched' with appropriate events. Thus, divisivity of the event predicate is only induced by the divisivity of the part structure of the measuring entity, e.g. the divisivity of the part structure of an interval or a region. Generally, there seems to be sufficient evidence that natural language semantics deals with contextually determined part structures, which are coarser than the part structures a physicist would ascribe to an interval or a region. Therefore arguably, divisivity is required and - in the acceptable cases - satisfied - relative to the contextual individuation of parts.

Moltmann's departure from CEM seems to me too radical a step to take, especially because the formal system by which she replaces it is not nearly as constrained and well understood as CEM. This makes it difficult to evaluate the predictions of her theory other than those intended by her. As Zucchi and White (2001) put it: "Since Moltmann does not tell us much about what *relevant* parts are, it is unclear to what extent her formulation actually solves the minimal parts problem." For example, nothing in her account actually entails her sufficiently-many-events observation. A more thorough criticism of Moltmann's position from the point of mereology is found in Pianesi (2002). See also Varzi (2006) for related discussion.

However, Moltmann's sufficiently-many-events observation, and her appeal to context, point towards the solution to the problem. While other accounts of

the minimal-parts problem seem to me like patches that have one single purpose, namely to rescue accounts that otherwise ignore the problem, Moltmann's observation provides motivation for modeling *for*-adverbials as sensitive to a more coarse-grained and context-sensitive property than divisive reference. Whatever is done to account for the contrast in (8) is likely to also account for the minimal-parts problem. My own account, to be presented in Section 5.4, rejects Moltmann's notion that context determines the part-whole relation and preserve the power of mereology, but I keep her notion that context plays a role in determining what counts as sufficiently many events.

Another possible way to make sense of Moltmann's suggestions could be the proposal in Link (1987b), who sketches a way to resolve the minimal-parts problem by replacing the lattice of events with a partial order of lattices E_i, with each of them standing for a certain granularity of events: intuitively, if $j \leq k$ then the events in E_j are more fine-grained than the events in E_k. A family of homomorphisms h_{ij} between the lattices relates events to events in such a way that the part relation between these events is partially collapsed in coarser models but otherwise preserved. Formally, h_{ii} is the identity map on E_i, and for any three lattices E_i, E_j, E_k such that $i \leq j \leq k$, the map is constrained by $h_{ik} = h_{ij} \circ h_{jk}$. This allows for the possibility that "a certain event can be atomic in a coarser domain and still be the image under a h_{ij} of a complex sum of events in a more fine-grained E_j." Link proposes to treat the granularity of events as a discourse parameter.

Unfortunately, Link's proposal is difficult to evaluate because he does not go into more detail than that. It seems to me that a great deal of complexity is added to the model by appealing to a family of lattices. A later version of his proposal is described in the following section.

5.3.4 Link (1991)

Link (1991) does not deal with *for*-adverbials per se, but he proposes a modification of the subinterval property that excludes from consideration temporal intervals below a certain threshold. He calls this threshold the *granularity parameter*. The modified subinterval property is designed to be able to hold of activity predicates despite the minimal-parts problem.

Link's formulation of the subinterval property is shown in (9):

(9) $\text{SUB}_{Link}(P) =_{def}$
        $\forall e, t[P(e) \wedge t \leq \tau(e) \wedge |t| \geq \gamma(e) \rightarrow \exists e'(e' \leq e \wedge P(e') \wedge \tau(e') = t)]$
        (A predicate P has the **Link-style subinterval property** iff for any e of
        which it holds, every part of the runtime of e whose length is at least $\gamma(e)$ is

the runtime of a part of e of which P holds as well.)

Link calls $\gamma(e)$ the *granularity of e* and comments that this parameter "expresses the observation that some time stretches might be too small to be a trace of the event e; thus $\gamma(e)$ fixes the minimal length that a time stretch must have to serve as a trace for e" (p. 217). Link's implementation therefore quantifies over all temporal intervals whose length is $\gamma(\theta)$ or larger. This implementation essentially requires every partition of the runtime of an event into minimal-size subintervals to be a set of runtimes of P-events.

While the granularity intuition is on the right track, this condition seems too strict. Take a sentence such as (10), repeated from (8b) above:

(10)   For ten hours John drew pictures.

Suppose that John draws one picture per hour and that he draws only one picture at a time. Then we can imagine (10) to be true of an event e_0 in a scenario in which John drew ten pictures in a row, and we would want to express that the predicate P = *John draw pictures* has the Link-style subinterval property with granularity one hour. But this is not possible in Link's implementation. We would expect P to have the subinterval property with respect to e_0 and we set $\gamma(e_0) = 1hr$. Suppose this event has a runtime i that lasts from noon to 10pm. Link's definition requires every subinterval of i that lasts one hour or more to be the runtime of an event in which John draws pictures. The problem is that not every subinterval of i satisfies this condition. For example, the subinterval from 12:30pm to 1:30pm lasts one hour, but it is not the runtime of an event that qualifies as *John draw pictures*, because during this time John finishes one picture and begins another.

Maribel Romero (p.c.) notes that the problem Link's account faces is analogous to a problem that Kratzer (1989) discusses in connection with what she calls lumping semantics. When we quantify over individuals or situations, we do not do so arbitrarily, but in any given context we pick out a salient partition of the world. Metaphorically speaking, we lump together individual bits and pieces of our surroundings and package them as entities. For example, when we stand in front of a cupboard filled with china and we say "most of the things in this cupboard are mine", we typically only quantify over individual pieces of china, and not over the left half of this teapot or the shards of china that would result if that plate fell apart. For count domains, I have made the simplifying assumption that they consist of mereological atoms. Already in the case of fake mass nouns like *china*, we see that this treatment is simplistic. A more general notion of individuation that takes this kind of lumping or packaging phenomena into account might conceivably be extensible to fix Link's account. For example, one might extend the group

formation operator ↑ used for collectivity effects (see Section 2.8) so that it can also group together temporal intervals in a way that packages them for universal quantification by something like the subinterval property. However, it is not easy to justify such an account, because a sentence like (10) does not provide much support for the idea that the *for*-adverbial quantifies over individuated lumps or groups of time. For this reason, I prefer not to add temporal groups or lumps to the ontology, and I do not pursue this approach further.

5.3.5 Summing up

To summarize, the account by Dowty (1979) ignores the minimal-parts problem, though Dowty himself is aware of it; the account by Hinrichs (1985) seeks to avoid it by a mathematical trick but unintentionally reintroduces it; Moltmann (1991) makes the "sufficiently many events" observation and recognizes that context plays a role but, in my eyes at least, throws the baby out with the bathwater by giving up mereology in favor of a very minimal theory of part structures; and the account by Link (1991) is too strict because it requires the main-clause predicate to be true at every subinterval of a certain length.

In addition to these technical problems, most of these accounts except for Moltmann's are limited in scope because they restrict themselves to atelic predicates, which is just one place where the minimal-parts problem arises. This exposes them at least in principle to the problem of being descriptions rather than explanations. Here I follow a simple concept of explanation suggested in von Stechow (1984) for the purposes of comparing semantic theories: "If a number of highly complex and apparently unrelated facts are reducible to a few simple principles, then these principles explain these facts."[17] The next section motivates my formulation of stratified reference, which was already presented in Chapter 4, and shows that this formulation avoids the minimal-parts problem.

5.4 My take on the minimal-parts problem

In this section, I show that *for*-adverbials are parallel to variant frequency adverbs like *occasionally* (Stump 1981) and to frequentative aspect (van Geenhoven 2004) in that they introduce several sources of variation relating to temporal distribution. I use this fact to further motivate the formulation of stratified reference introduced in Chapter 4 and I discuss its connection with the concept of cover.

---

[17]Von Stechow takes it that this concept of explanation is also present in syntactic work of generative grammar in the Chomskian tradition, but he cites Chomsky (1982) as rejecting it for semantics.

5.4.1 Parallels between *for*-adverbials and frequency adverbs

Stump (1981) observes that certain frequency adverbs introduce several sources of variation relating to temporal distribution. Stump distinguishes between fixed frequency adverbs like *daily, weekly* and variant frequency adverbs like *frequently, occasionally*.[18] First, he observes that the average period length associated with a variant frequency adverb like *occasionally* is partly dependent on the length of the interval of the entire sentence, as shown by his examples (11) and (12). I refer to this as the **relative length observation**.

(11)   During the commercial, the fly occasionally landed on my foot.

(12)   During their westward migration, the Celts occasionally encountered hostages.

Second, Stump notes that "even under an interpretation at a given interval, the average period length associated with a variant frequency adverb is not fixed. For example, [(11)] is vague as to the exact number of times the fly landed on my foot." I refer to this as the **vagueness observation**.

Stump illustrates his third observation with the following sentence:

(13)   An occasional sailor strolled by.

He observes that "even under an interpretation at a given interval i, the individual periods into which a variant frequency adverb 'segments' i may vary in length; for example, if, during the interval from 1:00 to 2:00, a sailor strolled by at 1:07, 1:20, 1:29, and 1:51, [(13)] might be true, even though there is no fixed frequency with which *A sailor strolls by* is true from 1:00 to 2:00." I refer to this as the **varying frequency observation**.

As noted in van Geenhoven (2004), the varying frequency observation needs to be qualified for certain adverbs like *regularly*. For example, *Last year, John jogged in the park regularly* cannot describe a situation in which John jogged many times from January to April and for the rest of the year jogged only once or twice every other month. So frequency may vary but only to a limited degree. I refer to this as the **regularity observation**.

With respect to his observations, Stump concludes that these "sorts of variability must be built into the truthconditions for variant frequency operators, however

---

[18]Stump actually uses the word *adjectives* rather than *adverbs* because his main point is to establish the connection between sentences like *An occasional sailor strolled by* and *Occasionally, a sailor strolled by*. This point is not of concern here.

these are to be stated."[19] Even though Stump does not explicitly mention it, van Geenhoven's regularity observation can be included in this conclusion.

Although *for*-adverbials were not in the scope of Stump's work, all of these observations also apply to them. First, take the relative length observation. The following examples show that *for*-adverbials are compatible with frequencies ranging from a few fractions of nanoseconds to one generation.

(14) The Chinese people have created abundant folk arts, such as paper-cuttings, acrobatics, etc., passed on from generation to generation for thousands of years.   = (1)

(15) Ded'leg says: How i stop a macro for 1sec? Cog says: By creating a script that will loop for 3600 milliseconds ... depending on how long you think it will take for both your computer, the server and internet lag to affect the macro.[20]

The length interval that counts as very small for the purpose of the *for*-adverbial varies relative to the length of the interval in the denotation of the Key. In (14), repeated from (1), that interval is thousands of years long, and a generation already counts as a very small interval. In (15), the interval in the denotation of the Key is only 3600 milliseconds long, and the very small intervals are even shorter.

By themselves, these examples do not show that *for*-adverbials presuppose varying frequencies, just that they are compatible with varying frequencies. That the varying frequencies are indeed part of the presupposition of *for*-adverbials is suggested by the combined evidence from these examples and the contrast in (8), repeated here:

(16) **Context**: John can draw two pictures per hour.
      a. ?For one hour John drew pictures.
      b.   For ten hours John drew pictures.

The contrast in (16) shows a correlation between acceptability and the ratio of the durations of the individual events (half an hour) to the length of the global interval (one hour or ten hours). Specifically, this ratio is high (1:2) in (16a) and low (1:20) in (16b). Since the absolute duration of the individual events is the same (half an hour) in both cases, the pair suggests that the *for*-adverbial are sensitive to a low relative length rather than a low absolute length. Examples (14) and (15) support

---

[19]In his footnote 5, Stump (1981) proposes a solution in terms of delineation coordinates in the sense of Lewis (1972). This solution, like mine, involves adding contextual parameters. I leave open whether it could be extended to other distributive constructions.

[20]Attested example (`http://www.wowmacroswarcraft.com/faq/`). Accessed Nov 15, 2010.

this reasoning because they show that *for*-adverbials are in principle compatible with both high and low absolute lengths of the individual events (a generation vs. a few milliseconds), while the relative length is low in both cases (given that scripts typically take at most a few milliseconds to perform one loop). This is analogous to Stump's relative length observation visible in (11) and (12).

Stump's vagueness observation also has a correlate in *for*-adverbials. Just as (11) is vague as to the exact number of times the fly landed on my foot, (14) and (15) are vague as to the exact number of generations and script loops involved.

As for a correlate of Stump's varying length observation, sentence (14) does not require there to be a fixed frequency with which one generation passes on the knowledge to the next generation. At the same time, the regularity observation also holds in *for*-adverbials, though perhaps to a lesser degree than with the word *regularly*. A habitual sentence like (17) is an odd way to describe a situation in which John jogged many times from January to April and then waited until December to jog again.

(17)   John jogged in the park for a year.

The next subsection presents stratified reference again and demonstrates that it formally captures the pattern empirically described by Stump's and van Geenhoven's observations.

5.4.2   Motivating stratified reference

Since frequency adverbs give rise to similar phenomena as *for*-adverbials, I use their meaning as a source of inspiration in order to formalize a notion that can capture the aspectual properties of *for*-adverbials. The idea is the following. Many of the solutions to the minimal-parts problem in Section 5.3 share the intuition that one must carve out an exception for very small entities in the higher-order property that is used to characterize atelicity and mass reference. For example, instead of quantifying over literally all subintervals as the subinterval property does, the solution by Link (1991) avoids quantifying over certain subintervals which are below a certain granularity threshold. The discussion of Link's implementation in Section 5.3.4 has shown that it does not yield the right results to simply fix a threshold, exclude all subintervals below it, and quantify over all the ones above it. A sentence like *For ten hours John drew pictures* does not require every sufficiently long subinterval of a ten-hour interval to be the runtime of a picture-drawing event. Rather, this sentence is true in virtue of the fact that this ten-hour interval can be subdivided into sufficiently long subintervals at each of which John draws a picture. In the present framework, sentences do not apply directly to intervals

but to events, so we may instead say that the *for*-adverbial requires the sentence *John drew pictures* to apply to an event that can be divided into sufficiently long subevents each of which is a picture-drawing event by John. In mereological semantics, dividing an event into a given set of subevents amounts to testing whether the event is the sum of these subevents. We may therefore equivalently say that the *for*-adverbial requires the sentence to apply to an event that is the sum of sufficiently long subevents which are each picture-drawing events by John.

To understand the formalization to be proposed, suppose for a moment that it is already known how to determine whether a subevent counts as sufficiently long and that there is a "threshold predicate" C that encapsulates that knowledge. C is analogous to the granularity parameter γ in Link (1991). We can now write $\lambda e[C(\tau(e))]$ for the set of all sufficiently long subevents, and $\lambda e[P(e) \wedge C(\tau(e))]$ for the intersection of this set with a predicate P. Intuitively, P stands for the predicate whose telicity we want to test. Then the algebraic closure ${}^*\lambda e[P(e) \wedge C(\tau(e))]$ is the set of all sums of sufficiently long subevents which are each in P. We can then describe the requirement that a *for*-adverbial imposes on a predicate P by stating that any event to which this predicate applies must also be such that the predicate ${}^*\lambda e'[P(e') \wedge C(\tau(e'))]$ applies to it. This is expressed in the following preliminary definition:

(18)   **Preliminary definition: Stratified subinterval reference**

$$\text{SSR}_C(P) \stackrel{\text{def}}{=} \forall e[P(e) \rightarrow e \in {}^*\lambda e' \begin{pmatrix} P(e') \wedge \\ C(\tau(e')) \end{pmatrix}]$$

(An event predicate P has stratified subinterval reference if and only if every event in its denotation can be divided exhaustively into very small parts ("strata") that are each in P. A very small part is one whose runtime satisfies the predicate C.)

I will update this definition in a moment by refining the unanalyzed predicate C. The result was already anticipated in Chapter 4. Before doing so, let us strengthen the intuitions of what it means for an event to have a starred predicate like ${}^*\lambda e'[P(e') \wedge C(\tau(e'))]$ apply to it. For this purpose, it is useful to consider the relation between this kind of starred predicate and the concept of cover that figures prominently in the work of Gillon (1987) and Schwarzschild (1996). In a set-based representation of plural individuals, covers are partitions of a set whose cells are allowed to overlap. For example, the set $\{\{a, b\}, \{b, c\}\}$ is a cover of the set $\{a, b, c\}$. This notion can be defined as follows:

(19)   **Definition: Cover (set-theoretic)**

$$\text{Cov}(C, P) \stackrel{\text{def}}{=} \varnothing \notin C \wedge \bigcup C = P$$

(C is a cover of a set P if and only if C is a set of nonempty subsets of P

whose union is P.)

As is often observed (e.g. Heim 1994), this notion can be translated into mere-ological terms. One possible way to do so is the following definition (see also Section 2.3.3 for correspondences between set theory and mereology):

(20)   **Definition: Cover (mereological)**
       $\text{Cov}(C, x) \stackrel{\text{def}}{=} x = \bigoplus C$
       (C is a cover of a mereological object x if and only if C is a set whose sum is x.)

As the following theorem illustrates, there is a close connection between the concept of cover and starred predicates:

(21)   **Theorem:**
       $\forall x[x \in {}^*\lambda y[C(y)] \Leftrightarrow \exists C' \subseteq C[\text{Cov}(C', x)]]$
       (The algebraic closure of C applies to x if and only if there is a subset of C that covers x.)

The proof of this theorem is simple. Using the definition of algebraic closure in Section 2.3.1, we rewrite $x \in {}^*\lambda y[C(y)]$ as $\exists C' \subseteq C[x = \bigoplus C']$. By definition (20), this is equivalent to $\exists C' \subseteq C[\text{Cov}(C', x)]$. For more discussion of this connection between algebraic closure and cover-based approaches, see Vaillette (2001).

Given this connection, my proposal amounts to describing the atelicity require-ment of a *for*-adverbial in terms of a set C that must cover any given event in the denotation of the predicate. Since C covers these events, the cells of C must contain their parts. Given C, an atelic predicate can be defined as one such that each cell of C contains an event that is again in the denotation of the predicate.

So far, I have not said anything about the nature of this predicate C. To do so, I rely on the analogy between the requirement imposed by frequency adverbs on the "average period length" at which the predicates they modify must hold, and the requirement of a *for*-adverbial on the predicate it modifies. I propose that C represents the average period length imposed by a frequency adverb or *for*-adverbial on its predicate. Just like in the case of frequency adverbs, the value of C for any given *for*-adverbial is not fixed, but vague. At the same time, its value is not chosen randomly, but relative to the length of the interval expressed by the Key (the measure phrase of a *for*-adverbial). As a consequence, the size of the cells in C is determined differently from case to case. I assume that the ratio of the size of these cells to the length of the Key interval must always be low. For example, in a sentence like (16b), the Key is ten hours, and C applies to events whose runtime is at most half an hour. By contrast, sentence (16a), whose Key is

one hour, is ruled out, and I assume that this is because C in this case only applies to events whose runtime is significantly less than half an hour. For a sentence like (14), where the *for*-adverbial is *for thousands of years*, C applies to events whose runtime is a generation. For a sentence like (15), where the *for*-adverbial is *for 3600 milliseconds*, C only applies to events whose runtime is a few milliseconds.

To suggest that the threshold in a *for*-adverbial depends on the value of the Key, instead of writing C for this threshold, I write $\varepsilon(K)$. To recapitulate the definition of ε from Section 4.6, I assume that ε is a function that maps an argument K to a predicate over intervals. K is a mnemonic for the Key, and I assume that it is always supplied by the Key of the distributive construction in question (as is stated in the Distributivity Constraint, see Section 4.6). For example, in sentence (10), the function ε maps a value of K = *ten hours* to a predicate like $\lambda t[\text{hours}(t) \leq 0.5]$; in sentence (14), it maps K = *thousands of years* to a predicate like $\lambda t[\text{generations} \leq 1]$. These precise predicates should be thought of as stand-ins for vague predicates with similar meanings. They play the role of covers in my system, but note that I do not think of covers as primitives; a predicate may come to act as a cover of an event by virtue of being included in the scope of a star operator whose output is applied to this event.

The updated definition of stratified subinterval reference, in which C is replaced by $\varepsilon(K)$, is shown here:

(22)   **Definition: Stratified subinterval reference**

$$\text{SSR}_{\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall e[P(e) \rightarrow e \in {}^{*}\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \varepsilon(K)(\tau(e')) \end{array} \right)]$$

(An event predicate P has stratified subinterval reference with respect to a Key K if and only if every event in its denotation can be divided exhaustively into very small parts ("strata") that are each in P. A very small part is one whose runtime satisfies the predicate $\varepsilon(K)$.)

Section 4.5.2 has shown how this definition can be used to identify *waltz* as an atelic predicate, despite the minimal-parts problem. Section 4.6 has generalized the definition to stratified reference, repeated here as (23).

(23)   **Definition: Stratified reference**

$$\text{SR}_{f,\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow x \in {}^{*}\lambda y \left(\begin{array}{c} P(y) \wedge \\ \varepsilon(K)(f(y)) \end{array} \right)]$$

(A predicate P has stratified reference with respect to a function f and a threshold $\varepsilon(K)$ if and only if there is a way of dividing every entity in its denotation exhaustively into parts ("strata") which are each in P and which have a very small f-value. Very small f-values are those that satisfy $\varepsilon(K)$.)

Instead of testing whether the predicate to which it is applied holds at every subinterval, stratified reference only tests whether the entity of which it holds can be divided into entities which are sufficiently small. I call these entities "strata" as a reminder of the fact that they must have very small values as measured in one dimension, but may be arbitrarily large as measured in any other dimension. As discussed in Chapter 1, the term *strata* is chosen as an allusion to the geological formations which are visible in places such as the Grand Canyon. A geological stratum can be just a few inches thick and yet extend over hundreds of thousands of square miles. The dimension along which definition (51) constrains strata is specified by the parameter f; their "maximum thickness" is specified via the threshold parameter $\varepsilon(K)$. This parameter can be realized as a very small value, for example as $\lambda i.seconds(i) \leq 1$ for a one-second threshold in the case of temporal *for*-adverbials. This method avoids explicit reference to atomic events or intervals.

I now sketch an account of examples like Moltmann's (16). Stratified reference allows us to express the fact that the predicate denoted by *John draw pictures* "almost" has the subinterval property, that is, whenever it is true of a sum event, it can be divided into strata whose are very short and which also qualify as *John draw pictures*.

The sentences in (8) are predicted to presuppose (24):

(24) $\text{SR}_{\tau,\varepsilon(K)}(\llbracket\text{John draw pictures}\rrbracket)$

where $\varepsilon(K)$ is a function that applies to intervals whose runtime is very small in relation to the Key. In the case of (16b), whose Key is *ten hours*, assume that such intervals can last 30 minutes at most. In this case, (24) expands to this:

(25) $\forall e[\llbracket\text{John draw pictures}\rrbracket(e) \rightarrow$
       $e \in {}^*(\lambda e'.[\text{minutes}(\tau(e')) \leq 30 \ \wedge \ \llbracket\text{John draw pictures}\rrbracket(e')])]$

This means that any event e in the denotation of *John draw pictures* can be divided up into events each of which has the following properties: (i) its runtime is at most thirty minutes; (ii) it is in the denotation of *John draw pictures*. This presupposition is compatible with the context, which predicts correctly that (16b) is acceptable.

The presupposition of (16a) is predicted to be like (25) except that the threshold $\varepsilon(K)$ is lower than 30 minutes, since its Key is not *ten hours* but *one hour*. Suppose for example that five minutes is the threshold:

(26) $\forall e[\llbracket\text{John draw pictures}\rrbracket(e) \rightarrow$
       $e \in {}^*(\lambda e'.[\text{minutes}(\tau(e')) \leq 5 \ \wedge \ \llbracket\text{John draw pictures}\rrbracket(e')])]$

Since it takes John 30 minutes to draw a picture, there exists no event that

verifies (27).

(27) $\exists e[[\![\text{John draw pictures}]\!](e) \;\wedge\; \text{minutes}(\tau(e)) \leq 5]$

While there may exist events whose runtime is five minutes or less in which John draws a proper part of a picture, I assume that these events normally do not themselves qualify as *John draw pictures*. Therefore, the presupposition (26) is violated, so sentence (16a) is unacceptable. Florian Schwarz (p.c.) notes that this assumption may be too strong in certain cases, since he judges *John drew pictures for ten hours* to be true even in case John draws complete pictures for nine hours and forty-five minutes, and then starts a new picture but does not finish it by the hour. I leave this problem aside here.

More generally, stratified reference captures several aspects of the presupposition of sentences such as (16b) that correspond to the observations by Dowty, Stump and Moltmann:

- the minimal-parts problem is accounted for by not requiring the threshold $\varepsilon(K)$ to be infinitely low;

- the "sufficiently many events" and the relative-length observation are captured in the requirement that the threshold $\varepsilon(K)$ be very low relative to the Key;

- the vagueness observation and the varying-length observation are expressed in the fact that the threshold $\varepsilon(K)$ is assumed to be vague and pragmatically determined and that it is allowed to hold of intervals of varying length (e.g. intervals of thirty minutes or less);

- the regularity observation is captured by the semantics of the star operator, which requires that there is a way to divide the runtime of the main-clause predicate into intervals no longer than the threshold $\varepsilon(K)$ at which the main-clause predicate holds.

The proposals reviewed in Section 5.3 do not capture these aspects adequately. Dowty (1979) is to be credited for connecting the minimal-parts problem with *for*-adverbials, but by his own admission he does not account for it because he lets *for*-adverbials quantify over all subintervals. This also fails to capture the other observations listed above. Hinrichs (1985)'s semantics suffers from its own problems and fixing it does not capture the "sufficiently many events" observation. Moltmann (1989, 1991) deserves credit for making that observation, but her own analysis relies on an unanalyzed parthood notion that is weaker than mereological

parthood. Finally, Link (1991) imposes too strict conditions and, as a result, is unable to capture these aspects.

Let me point out what I believe to be the main limitation and the main strength of the present account. On the negative side, in the absence of a theory of vagueness, any statements about the function ε used to determine the threshold are themselves vague. I think that this is the best we can do in a paradigm that is designed to make categorical predictions about acceptability, because the relative acceptability of sentences with *for*-adverbials may as well be continuous as categorical. Consider again Moltmann's example with additions in (28): At which point and how does acceptability drop off as we increase the ratio between length of individual event and length of main-clause event? To what extent does acceptability depend on whether the context specifies an exact value (e.g. John can draw exactly two pictures per hour) or an inexact value (it is not known how long John takes to draw a picture, or it varies)?

(28)   **Context**: John can draw two pictures per hour.
         For *half an /?one/.../ten hour(s) John drew pictures.

On the positive side, the main strength of the present account lies in the novel connection it draws between *for*-adverbials and other distributive constructions. This makes it possible to formulate the solution to the minimal-parts problem in a completely general way that does not refer to *for*-adverbials specifically. We can look for equivalents of the observations by Dowty, Stump, and Moltmann in other distributive constructions. The next subsection shows that the reformulation of the Distributivity Constraint in terms of stratified reference is independently justified by the behavior of pseudopartitives.

5.4.3   Parallels between *for*-adverbials and pseudopartitives

Chapter 4 has proposed that stratified reference, via the Distributivity Constraint, is operative not only in *for*-adverbials but also in pseudopartitives. The immediate prediction is that a similar set of effects to the ones already discussed in this chapter is observable in pseudopartitives. For example, we expect there to be a minimal-parts problem and an analogue to the sufficiently-many-events requirement. To the extent that this prediction is true, it provides independent justification for stratified reference. In the sense of von Stechow (1984), the present account can then be considered a solution to the minimal-parts problem: a number of complex and apparently unrelated facts are reducible to a few simple principles.

Some of these effects are indeed attested. My starting point is the observation christened the "paradox of grams" by Bale (2009). Bale's central example is the

following:

(29)   a.   Give me 500 grams of apples.
       b.   ?Give me 50 grams of apples.
       c.   ??Give me one gram of apples.

This example shows that a pseudopartitive as a whole cannot be predicated of an entity that is too small to be in the denotation of the substance nominal. For example, (29b) and (29c) are ruled out on the assumption that it is impossible for an apple to weigh 50 grams or even one gram.

The generalization just mentioned is actually a special case of a phenomenon that is analogous to the sufficiently-many events requirement: the sum entity to which the pseudopartitive applies must have sufficiently many parts that each fall into the denotation of the substance nominal. This is illustrated by the paradigm in (30).[21]

(30)   a.   *twelve pounds of twelve-pound weights
       b.   ?twelve pounds of six-pound weights
       c.   twelve pounds of four-pound weights
       d.   twelve pounds of three-pound weights
       e.   twelve pounds of two-pound weights
       f.   twelve pounds of one-pound weights

I generalize from this paradigm that native speakers tend to reject pseudopartitives where the Map value of the substance nominal atoms is low relative to that of the measure nominal. This is analogous to the relative-length observation and it is predicted by the Distributivity Constraint. For example, according to the classification in Chapter 4 as applied to the pseudopartitives in (30), the Share is *n-pound weights*, the Map is *weight*, and the Key is *twelve pounds*. As for the threshold value $\varepsilon(K)$, it is predicted to be very low relative to the Key. Since the Key is the same in (30a) through (30f) and since the judgments for this paradigm are made independently of context, we can assume that the value of $\varepsilon(K)$ is the same in all these cases. To derive the cutoff between (30b) and (30c), I assume that $\varepsilon(K)$ picks out those degrees which correspond to four pounds or less on the weight scale. Although the discussion here is preliminary and does not meet the standards of a quantitative study, the following observation is suggestive: The 1:2 ratio between six pounds and twelve pounds in the "strained" example (30b) corresponds to the 1:2 ratio in Moltmann's *for*-adverbial example (8a), which she judged "less acceptable". To the extent that this correspondence holds up to

_____

[21]This paradigm came up in a discussion with Eytan Zweig.

scrutiny, it may provide independent evidence for choosing four pounds as the threshold value in the case of (30). But we should not expect there to be an exact threshold value: the vagueness observation and the varying-length observation also hold in pseudopartitives (since *ten pounds of books* does not presuppose that all the books in question have the same weight.

Strata theory predicts that sentences (30a) through (30f) are only acceptable if (31), which expands to (32), is true. Here I write n-pound weights to abbreviate *twelve-pound weights*, *six-pound weights*, etc. corresponding to the examples in (30).

(31) $\text{SR}_{weight,\lambda d.\text{pounds}(d)\leq 4}([\![\text{n-pound weights}]\!])$

(32) $\forall x[[\![\text{n-pound weights}]\!](x) \rightarrow$
$x \in {}^*(\lambda y.\text{pounds}(\text{weight}(y)) \leq 4 \ \wedge \ [\![\text{n-pound weights}]\!](y))]$
(Every entity in the denotation of *n-pound weights*, that is, every sum consisting of one or more such weights, consists of n-pound weights that each weigh four pounds or less.)

In sum, the Distributivity Constraint can be regarded as an independently justified explanation of the minimal-parts problem because it predicts that Stump's observations about frequency adverbs have parallels not only in *for*-adverbials but also in pseudopartitives.

5.5 Summary

This chapter has proposed a solution to the minimal-parts problem in terms of stratified reference. An atelic predicate P is defined as one such that every event in its denotation can be divided exhaustively into very small parts ("strata") that are each in P. A very small part is one whose runtime satisfies a threshold which is determined in relation to the Key. Since the Key – the measure phrase of a *for*-adverbial – differs across *for*-adverbials, the result is a partly context-dependent notion of atelicity. The minimal-parts problem is avoided by not considering the properties of those parts whose runtimes are shorter than the threshold, which enters stratified reference through its granularity parameter. An implementation in terms of algebraic closure was shown to be related to the concept of cover known from work by Schwarzschild (1996) and others on distributivity, and to be superior to previous formulations which use universal quantification.

The existence of a threshold, and the need to determine it in relation to the Key, is independently argued for through an observation from Moltmann (1991). This observation was illustrated with the following minimal pair:

(33)   **Context**: John can draw two pictures per hour.
   a.   ?For one hour John drew pictures.
   b.   For ten hours John drew pictures.

Since stratified reference is assumed to govern the behavior of *for*-adverbials and pseudopartitives alike, an analogous observation is predicted to be found in pseudopartitives. I argued that this is indeed the case due to an observation by Bale (2009), as illustrated by the following minimal triple:

(34)   a.   Give me 500 grams of apples.
   b.   ?Give me 50 grams of apples.
   c.   ??Give me one gram of apples.

This parallel points towards a general and construction-independent concept that subsumes both atelicity and mass reference. The parametrized higher-order property of stratified reference, whose formulation was motivated on independent grounds in Chapter 4 and again in this chapter based on analogies to observations by Stump (1981) and van Geenhoven (2004) about frequency adverbials, provides a way to capture this concept. Its granularity parameter was introduced in Chapter 4 to account for the differences between distributivity in count and noncount domains, and it has been used in the present chapter to provide a way for different threshold values to influence acceptability judgments.

Chapter 6

Aspect and space

6.1 Introduction

Most previous work on aspect, as well as this work so far, has focused on properties of temporal *for*-adverbials such as *for an hour* and *in an hour*, and on temporal properties of verbal predicates. This chapter starts from the observation that the telic-atelic opposition is not confined to the temporal domain, but is neatly replicated in space. Following Gawron (2004), I speak of temporal aspect and spatial aspect. Not all theories of aspect survive the extension from the temporal to the spatial domain equally well. This chapter shows that accounts of *for*-adverbials which are based on divisive reference, such as Krifka (1998), cannot be easily extended to the spatial domain. Alternative accounts based on the subinterval property, in the style of Dowty (1979), can be extended to the spatial domain without problems. Such accounts also include the present work, since stratified reference can be seen as a generalization of the subinterval property.

Going from time to space takes aspect and telicity beyond their traditional usage. For example, Garey (1957), the source of the terms *telic* and *atelic*, defines a verbal predicate V to be atelic if and only if when one is interrupted while Ving, one can be said to have Ved.[22] The classical definition of aspect in Comrie (1976), "aspects are different ways of viewing the internal temporal constituency of a situation", makes explicit reference to time. Likewise, the famous categorization by Vendler (1957) of verbal predicates into the four basic groups *state, activity, achievement, accomplishment* is based on time-related criteria such as interaction with temporal *for*- and *in*-adverbials and with the progressive.

As a consequence of extending the study of aspect to space, it becomes nec-

---

[22]This test is modeled on the distinction between *kinesis* and *energeia* in Aristotle's *Metaphysics*, though the correspondence is not straightforward (Graham 1980).

essary to refine Vendler's classical categorization. In modern implementations of this system, the difference between the four classes is partly determined by telicity: states and actitivies are atelic, achievements and accomplishments are telic (Verkuyl 1989).[23] Splitting up telicity into spatial and temporal telicity means that we have to duplicate at least part of Vendler's categorization, and talk about spatial activities and accomplishments as Gawron (2005) does. The same is true for any of the other classification schemes that compete with Vendler's.[24] I do not dwell on this issue, because I regard Vendler classes as descriptive labels rather than analyses in their own right, in agreement with Verkuyl (1989, 2005) on this point. Vendler classes do not appear as primitives in this work, any more than they do in contemporary analyses of aspect such as Krifka (1998) and Verkuyl (1989, 1993).

While the traditional tests must be understood as diagnostics of temporal aspect, spatial aspect can be diagnosed by analogous means. Evidence for temporal and spatial aspect comes from a class of adverbial adjuncts which Moltmann (1991) calls measure adverbials. Moltmann applies this term to items like those in (1).

(1) a. John drank wine / *a glass of wine *for two hours.*
     b. Mary drew pictures / *the picture *until noon.*
     c. Children / *1000 children suffer from hunger *worldwide.*
     d. *Throughout the country* women / *a hundred women protested against the abortion law.

Measure adverbials in this sense include temporal as well as spatial *for*-adverbials. Example (2) is adapted from Gawron (2005).

(2) a. The road widens for 100 meters.
     b. *The road widens nearly two meters for 100 meters.

There is an irrelevant sense on which (2b) is acceptable: it can mean there is a 100-meter stretch of the road in which it is two meters wider than outside of this stretch. This appears to be the spatial counterpart of the result state related interpretation of *for*-adverbials (see Section 3.3). The point here is that (2b) cannot mean that the road widens for 100 meters and at the end of the 100 meters it is nearly two meters wider than at the beginning.

The examples in (2) involve the degree achievement verb *widen*. In general, *widen* can be understood in two ways: either on a temporal interpretation, as involving reference to events in which an object such as an air balloon becomes wider and wider over time, or on an atemporal interpretation, as involving reference to

---

[23]The other distinctions are as follows: States differ from activities in that they are taken to be "static", and achievements differ from accomplishments in that they are taken to be punctual.

[24]Tatevosov (2002) has compiled a useful list of 16 such classifications.

events in which the width of an object increases gradually along a spatial axis. In this case, the object need not change its shape over time. To avoid confusion in connection with the word *event*, note that I also use this word in reference to states (see Section 2.4.3). Only the atemporal interpretation of *widen* and of other degree achievement verbs is relevant here. I have chosen the subject of the sentences in (2), *road*, to exclude any temporal interpretation that these sentences may have. Such an interpretation is implausible because roads do not normally change their shape over time.

Talmy (1996) analyzes examples such as *The scenery was rushing past me* in terms of "fictive motion". While these examples are superficially similar to the sentences in (2), I agree with Gawron (2005) that fictive motion is not involved in this case. For example, the sentences in (2) cannot be placed into the progressive form, unlike Talmy's example. In any case, implementing the concept of fictive motion in formal semantics seems more cumbersome to me than the straightforward extension of mereological approaches that I advocate here, and it would not capture the symmetrical behavior of temporal and spatial aspect.

Spatial aspect is not confined to degree achievement verbs, but is also exhibited by predicates like *meander* and *end* (Gawron 2005). As diagnosed by spatial adverbials, the former can be regarded as spatially atelic and the latter as spatially telic:

(3) a. *The road meanders in a mile.
     b.   The road meanders for a mile.

(4) a.   The road ends in a mile.
     b. *The road ends for a mile.

As noted by Gawron (2005), sentences like *The road meanders* and *The road ends* are both stative, and applying the traditional, temporal-based notions of telicity to them makes them both temporally atelic, so the contrast between them cannot be described in terms of *temporal* telicity. However, what goes wrong in the spatial cases (1c), (1d), (2b), (3a) and (4b) does seem analogous to what goes wrong in the temporal cases (1a), (1b). For example, (2a) does not just convey that the road is wider at the end of the 100-meter stretch than at the beginning, but also that this widening is gradual and distributed over the length of the 100-meter stretch, in such a way that each part of the road within that stretch widens. Similarly, it is plausible to link the unacceptability of (2b) to the fact that if the 100-meter stretch of the road widens by two meters in total, then the parts of the road within that stretch do not each widen by two meters. This is analogous to the explanation for (1a) in terms of aspectual composition, as found for example in Krifka (1998): the parts of an event of drinking wine are themselves events of drinking wine,

but the parts of an event of drinking a glass of wine are not themselves events of drinking a glass of wine. This is so even though there are events which are both in the denotation of *drink wine* and in the denotation of *drink a glass of wine.* Again, there is an analogy with (2): among the states to which *The road widens* applies, there are states to which *The road widens nearly two meters* applies.

Gawron views spatial aspect as involving change over a spatial rather than temporal axis. While I agree with the attempt to generalize boundedness from time to space, I will not adopt Gawron's treatment itself here. First, it would be difficult to integrate his analysis of aspect to the mereological framework which I have adopted. Second, Gawron attempts to characterize telicity in terms of cumulativity, but this is incompatible with the lexical cumulativity assumption which I need for independent purposes. On this assumption, all verbs have cumulative reference regardless of telicity (see Section 2.7.2).

To sum up this section, spatial measure adverbials provide us with a reason to dissociate the notion of aspect from time. Research into aspect must take seriously the analogous behavior of temporal and spatial *for*-adverbials, and any theory of aspect should account for both spatial and temporal *for*-adverbials in order to capture this parallel.

The rest of this chapter is organized as follows. Section 6.2 lays out the space of possibilities for a unified account of temporal and spatial (a)telicity: either we characterize the two notions by using a property that tests for application of the predicate in arbitrarily confined subregions of spacetime, such as cumulative or divisive reference, or we use a notion in which dimension (space or time) is a parameter and in which we test for application of the predicate in regions of spacetime that are constrained in only one dimension and can extend arbitrarily in the other dimensions. Based on the aspectual properties of predicates with one bounded and one unbounded argument, the section makes the case for the second option, which I call the strata-based approach.

Section 6.3 presents Krifka (1998) as an example of the first option, which I call the subregion-based approach, and contrasts it with Dowty (1979) and my own account as two examples of the strata-based approach.

Section 6.4 shows first intuitively and then in formal detail how only the second option can account for the difference between spatial and temporal aspect. This work implements the second option because the higher-order property it uses, stratified reference, can be parametrized for time or space through its dimension parameter.

Section 6.5 concludes.

126

6.2 Subregions versus strata

This section distinguishes two kinds of algebraic models of telicity and classifies them under the terms *subregion-based approach* and *strata-based approach*. These terms are chosen with reference to the metaphor introduced in Chapter 1. In this metaphor, events and other entities are thought of as occupying regions in an abstract space whose dimensions specify, among others, their spatial and temporal extent. Events can be both spatially and temporally extended at the same time (see Section 2.4.3). This opposes them to intervals, which correspond to one-dimensional lines (see Section 2.4.4). While the rest of this work considers thematic roles and measure functions to be among the dimensions of this abstract space, this part of the metaphor can be ignored here. Only the spatial dimensions and time matter for the present purpose, and the relevant subspace of the abstract space can therefore be thought of as spacetime. In fact, we will never have to consider more than one spatial dimension at a time, so the relevant subspace has only two dimensions and can be drawn on paper. I will make use of this fact when I provide illustrations.

By "strata" I mean parts of an event that are constrained to be very short along one of these dimensions, and which extend up to the boundaries of the event along all the other dimensions. In a three-dimensional space, a stratum is formed like a layer, and the thinner it is, the closer it approximates a plane. In a two-dimensional space, the thinner a stratum is, the closer it comes to a line. By "subregions" I mean parts of an event that occupy arbitrary regions of spacetime, including regions that are constrained in all dimensions. For example, consider an event whose location extends fifty meters in one direction, and whose runtime is an interval of fifty minutes. A stratum of this event is either a part whose runtime is very short but whose spatial extent is fifty meters, or a part whose spatial extent is very small but whose runtime is fifty minutes. A subregion is an arbitrary part of this event. Both the runtime and the spatial extent of a subregion could be very small, but they do not have to be. It follows that any stratum of an event is also one of its subregions, but not vice versa.

I do not define strata and subregions formally because they are only used to build intuitions. They do not play a formal role in any of the theories I discuss, including my own, but the ideas they represent informs the following classification.

- I use the term **subregion-based approach** for implementations that characterize the difference between telicity and atelicity based on divisive reference or related notions. Divisive reference holds of a predicate P if every part of any event in P is also in P (see Section 2.3.2). In our visual metaphor, divisive reference requires the event predicate to hold of every subregion of

any event to which it applies. A faithful implementation of the subregion-based approach that is based on divisive reference would predict that it is impossible for one and the same predicate to be both spatially bounded and temporally unbounded, or vice versa, because an event predicate cannot be at the same time divisive and not divisive.

- The **strata-based approach** characterizes the difference between telicity and atelicity in a way that is parametrized on a dimension such as time or space. My own account of aspect implements the strata-based approach. The notion of stratified reference motivated in Chapters 4 and 5 abstracts over time and space via its dimension parameter. In our visual metaphor, stratified reference requires any event to which the event predicate applies to be subdivisible into strata that are constrained along the dimension specified by the dimension parameter, and run perpendicularly to it. The event predicate is then required to hold of every stratum of any event to which it applies. I will illustrate this in more detail in the following discussion. Note that on the strata-based approach, the dimension parameter provides an additional degree of freedom compared with the subregion-based approach. This approach therefore makes it possible for predicates to be both spatially bounded and temporally unbounded or vice versa. For example, it is logically possible for one and the same predicate to have stratified reference with respect to runtime but not to location, or vice versa.

Each of the two characterizations of atelicity outlined above has been implemented in various ways by different authors, in the shape of different lexical entries for *for*-adverbials. For example, the subregion-based approach is implemented in the work of Krifka (1998), while accounts that implement the strata-based approach include Dowty (1979), Moltmann (1991) as well as the present work. Here, I contrast prototypical examples of the two characterizations, namely Vendler (1957) and Bennett and Partee (1972). The next section contrasts the specific proposals of Krifka, Dowty, and the present work. While neither Vendler nor Bennett and Partee proposed entries for *for*-adverbials, the simplicity of their proposals makes it useful for expository purposes to discuss them before moving on to the more advanced implementations.

The essence of the subregion-based approach is present in Vendler (1957)'s suggestion that (atelic) activities have the property that "any part of the process is of the same nature as a whole". Vendler contrasts this property with (telic) accomplishment predicates, which "proceed toward a terminus which is logically necessary to their being what they are". The property of atelic predicates that Vendler identifies corresponds to the model-theoretic notion of divisive reference.

Although Vendler speaks of processes rather than events and does not consider whether processes have spatial equivalents, the point is that he considers all the parts of a process, regardless of time.

This feature is absent from the strata-based approach, a characterization that explicitly makes reference to the temporal and spatial properties of predicates. The essence of this view dates back at least to Bennett and Partee (1972). In connection with atelicity, Bennett and Partee single out the class of predicates which "have the property that if they are the main verb phrase of a sentence which is true at some interval of time I, then the sentence is true at every subinterval of I including every moment of time in I". This property is commonly referred to as the *subinterval property*. Because the subinterval property makes reference to a temporal interval, it might seem from my description that it can characterize only temporal but not spatial atelicity. This is not the case. As shown in Chapter 4, stratified reference can be parametrized for time or space. The resulting notion can be put to use to formulate identical entries for temporal and spatial uses of *for* that check for the subinterval property with respect to the interval denoted by their complement. That interval can either be temporal, as in the case of *for three hours*, or spatial, as in *for three miles*.

Of the authors I cite here, only two actually consider spatial aspect. One of them is Gawron (2005), whom I have already discussed. The other one is Moltmann (1991), who recognizes that Dowty (1979)'s analysis can be generalized to spatial measure adverbials of the kind she considers in (1c) and (1d).[25] Moving from temporal to spatial aspect not only allows us to increase the empirical testing ground of theories of *for*-adverbials, but as we will see in the next section, it allows us to rule out the subregion-based approach.

Both the strata-based and the subregion-based approach must be able to model the different distribution of spatial and temporal *for*-adverbials. This distribution can be observed in connection with predicates which contain a bounded and an unbounded argument. See Beavers (2008) for discussion of this type of predicate. Here is an example:[26]

(5) a.   John pushed carts to the store for fifty minutes.
     b.   *John pushed carts to the store for fifty meters.

---

[25]Several authors, e.g. Hinrichs (1985) and Krifka (1998), use a model of space to represent the aspectual properties of motion predicates such as *walk to the store*. By spatial aspect I mean the contrast exhibited in (2), i.e. the distribution of spatial measure adverbials, not the influence of spatial predicates on temporal measure adverbials.

[26]To prevent a partitive coercion of *to the store* into an unbounded predicate with a meaning similar to *towards the store*, the examples in this chapter can be made stronger, if less conspicuous, by adding the modifier *all the way* before *to the store*. *All the way* is a "true delimiter" in the sense of Smollett (2001).

The predicate *John pushed carts to the store* is compatible with a temporal *for*-adverbial on an interpretation where John went back and forth and pushed the carts in question to the store one by one, or at least little by little (5a).[27] This type of reading has been called *frequentative* or *iterative* (van Geenhoven 2004). As discussed in Section 2.7.2, I assume that iterativity is a consequence of lexical cumulativity. The same predicate, however, is incompatible with a spatial *for*-adverbial, irrespectively of whether it is interpreted iteratively or not (5b).

The fact that temporal and spatial *for*-adverbials do not have the same distribution shows that there is no single property that can account for the distribution of all *for*-adverbials. If we want to characterize what happens in (5a) as well as (5b) in terms of (a)telicity, as was argued in the previous section, any approach to aspect needs to contain the equivalent of a parameter that can be set to space or time. The property of divisive reference does not contain such a parameter, so the subregion-based approach is only viable if it relies on a suitably amended verison of divisive reference, or on another property altogether. The property of stratified reference contains such a parameter. I have called it the dimension parameter. Chapter 4 has motivated this parameter independently of aspect-related considerations.

The contrast in (5) cannot be explained by appealing to the lexical semantics of the verb alone, since *push* is compatible with spatial *for*-adverbials in other sentences:

(6)   John pushed carts towards the store for fifty meters.

This means that the distribution of spatial *for*-adverbials is determined by properties of phrasal constituents and not just by verbs. Spatial aspect is therefore compositionally determined, as is temporal aspect (Verkuyl 1972).

Both the subregion-based and the strata-based approach presuppose that events and/or intervals have parts, so they can both be implemented within the subregion-based mereological framework presented in Chapter 2. Krifka (1986, 1989, 1992, 1998) and Kratzer (2007) both implement the subregion-based approach in such a setting. Moreover, both theories can be formulated in an event semantic setting. Dowty (1979), an early implementation of the strata-based approach, is formalized using neither event semantics nor mereological semantics, but subsequent

---

[27]A sentence like *John pushed carts to the store* does not entail that John pushed more than one cart at a time; he could have pushed one cart at a time. This is consistent with the inclusive view on plurals, for which I have presented arguments in Section 2.6.2. In addition, the sentence entails that John pushed more than one cart in total. The bare plural *carts* therefore acts as a dependent plural. That *for*-adverbials license dependent plurals can be seen from sentences like *John wore yellow neckties for a week*, which do not give rise to the conclusion that John wore several neckties at once. In Chapter 9, I present an account of dependent plurals in *for*-adverbials.

implementations are, including Hinrichs (1985) and Moltmann (1991).

6.3 Implementations of both approaches

This section presents concrete implementations of the subregion-based approach and the strata-based approach. The next section compares them with respect to how well they succeed in accounting for the different distribution of spatial and temporal measure adverbials. I present the proposal in Krifka (1998) as an example of the subregion-based approach. The strata-based approach is illustrated by Dowty (1979) and by my own proposal. For the purpose of this chapter, and unlike Chapter 5, the differences between Dowty's proposal (once it has been generalized to space) and my own are irrelevant, since both accounts are strata-based. Metaphorically speaking, these differences only concern the thickness of the strata: Dowty considers arbitrarily thin strata and I consider "very thin" but not arbitrarily thin ones.

6.3.1 Dowty (1979) and my proposal

Dowty (1979) essentially models *for*-adverbials as universal quantifiers ranging over the subintervals of some time interval. For example, a sentence like *John ran for an hour* is predicted to be true if John ran at each subinterval of a certain hour-long interval. I have reviewed Dowty's analysis of *for*-adverbials in Section 5.3.1. Since many of the details of his analysis are not relevant for the present purpose, I will not use Dowty's original entry, but a simplified version.

In Dowty's analysis, events and part structures are unnecessary, and Dowty himself does not make use of them. However, for purposes of comparison with Krifka's theory, it is useful to couch his analysis in a mereological subregion-based framework anyway (cf. Moltmann 1991). To give all implementations a fair chance to model spatial aspect, I extend the background assumptions of the accounts I compare in this chapter in the following way: I assume the existence of a trace function τ that relates events to their runtime, and I assume that events are also mapped to spatial intervals by a location function σ that mirrors τ in all important respects. In particular, both σ and τ are assumed to be homomorphisms with respect to the sum operation. For example, the runtime of the sum of two events is the sum of their runtimes. For more details on these assumptions, see Section 2.5.2.

Casting both theories into the same framework is not only beneficial for comparison. Since I argue against the subregion-based approach, adopting the framework in which this theory was formulated makes sure that my criticism is targeted against the original theory and not against a potentially inaccurate

translation. At the same time, my criticism against the subregion-based approach is not tied to the specific details of Krifka's implementation. It is targeted against the implicit premise of the subregion-based approach and so it is likely to equally apply to other implementations of the subregion-based approach.

The following is a simplified version of Dowty's entry that retains his basic insight, namely, *for*-adverbials universally quantify over subintervals.

(7) ⟦for an hour⟧ (Dowty, simplified)
$= \lambda P \exists t [\text{hours}(t) = 1 \ \wedge \ \forall t' \subseteq t[\text{AT}(P, t')]]$

In Dowty's framework, the relation AT evaluates a proposition at an interval. In event semantics, AT can be reformulated in terms of τ. This definition is repeated here from Section 2.5.2.

(8) **Definition: Holding at an interval**
$\text{AT}(V, t) \overset{\text{def}}{=} \exists e[V(e) \ \wedge \ \tau(e) = t]$
(An event predicate V holds at an interval t if and only if it holds of some event whose temporal trace is t.)

Given this definition, we can rewrite the translation in (7) as follows:

(9) ⟦for an hour⟧ (Dowty, simplified)
$= \lambda P \exists t [\text{hours}(t) = 1 \ \wedge \ \forall t' \leq t[\exists e P(e) \ \wedge \ \tau(e) = t']$

I have replaced $\subseteq$ by $\leq$ because I assume a mereological rather than a set-theoretic structure for time (see Section 2.4.4).

After these modifications are carried out, Dowty's analysis produces representations like the following:

(10) ⟦John ran for an hour⟧
$= \exists t[\text{hours}(t) = 1 \ \wedge \ \forall t'[t' \leq t \rightarrow \exists e \ \text{run}(e, \text{john}) \ \wedge \ \tau(e) = t']]$
(There is a time interval t which lasts an hour and John runs at each of its parts.)

(11) ⟦John ran for a mile⟧
$= \exists t[\text{miles}(t) = 1 \ \wedge \ \forall t'[t' \leq t \rightarrow \exists e \ \text{run}(e, \text{john}) \ \wedge \ \sigma(e) = t']]$
(There is a spatial interval t which spans a mile and John runs at each of its parts.)

Dowty's analysis can be classified as an implementation of the strata-based approach because he lets *for*-adverbials introduce a test that applies to subintervals. This feature is also implemented in my own proposal, as presented in Chapters 4 and 5. My translation for the temporal adverbial *for an hour* is shown in (12).

(12) ⟦for an hour⟧ (my proposal, based on Section 4.7)
$= \lambda P_{\langle vt \rangle} \lambda e : \mathrm{SR}_{\tau, \varepsilon(\lambda t[\mathrm{hours}(t)=1])}(P) \cdot P(e) \wedge \mathrm{hours}(\tau(e)) = 1$
where $\varepsilon(\lambda t[\mathrm{hours}(t) = 1])$ is a predicate that holds of temporal intervals
which are very small compared to an hour

This entry results in logical representations like the one in (13).

(13) ⟦John ran for an hour⟧
Presupposition: $\mathrm{SR}_{\tau, \varepsilon(\lambda t[\mathrm{hours}(t)=1])}(\text{John run})$
Assertion: $\exists e[\llbracket \text{john run} \rrbracket(e) \wedge \mathrm{hours}(\tau(e)) = 1]$

Given the definition of stratified reference in Section 4.6, the presupposition of
(13) can be expanded as in (14).

(14) $\mathrm{SR}_{\tau, \varepsilon(\lambda t[\mathrm{hours}(t)=1])}(\llbracket \text{John run} \rrbracket) \Leftrightarrow$
$\forall e[\llbracket \text{John run} \rrbracket(e) \rightarrow e \in {}^*(\lambda e'.\varepsilon(\lambda t[\mathrm{hours}(t) = 1])(\tau(e')) \wedge \llbracket \text{John run} \rrbracket(e'))]$
(Every event e in ⟦John run⟧ can be divided into one or more parts each of
which is in ⟦John run⟧ and has a very small temporal extent compared to
an hour.)

The corresponding spatial translations are exactly parallel:

(15) ⟦for a mile⟧ (my proposal)
$= \lambda P_{\langle vt \rangle} \lambda e : \mathrm{SR}_{\sigma, \varepsilon(\lambda t[\mathrm{miles}(t)=1])}(P) \cdot P(e) \wedge \mathrm{miles}(\sigma(e)) = 1$
where $\varepsilon(\lambda t[\mathrm{miles}(t) = 1])$ is a predicate that holds of spatial intervals which
are very small compared to a mile

(16) ⟦John ran for a mile⟧
Presupposition: $\mathrm{SR}_{\sigma, \varepsilon(\lambda t[\mathrm{miles}(t)=1])}(\text{John run})$
Assertion: $\exists e[\llbracket \text{john run} \rrbracket(e) \wedge \sigma(e) = 1]$

(17) $\mathrm{SR}_{\sigma, \varepsilon(\lambda t[\mathrm{miles}(t)=1])}(\llbracket \text{John run} \rrbracket) \Leftrightarrow$
$\forall e[\llbracket \text{John run} \rrbracket(e) \rightarrow e \in {}^*(\lambda e'.\varepsilon(\lambda t[\mathrm{miles}(t) = 1])(\sigma(e')) \wedge \llbracket \text{John run} \rrbracket(e'))]$
(Every event e in ⟦John run⟧ can be divided into one or more parts each of
which is in ⟦John run⟧ and has a very small spatial extent compared to a
mile.)

6.3.2 Krifka (1998)

Krifka (1998) translates *for*-adverbials using a parametrized version of divisive
reference, and therefore instantiates the subregion-based approach. Krifka's posi-
tion has evolved over the years: in Krifka (1986), Chapter 3, *for*-adverbials impose
the requirement that predicates be divisive and strictly cumulative (that is, they

are cumulative and they have at least two distinct entities in their denotation), and the predicate that is output by the *for*-adverbial must not be cumulative. In Krifka (1989), atelic predicates are considered to be "strictly cumulative or at least non-quantized". Krifka (1998) contains the most recent entry for *for*-adverbials published by the author.

The translation given in Krifka (1998) for *for an hour* and its use in *John ran for an hour* are shown below. I present Krifka's entry here in its original form to show that the reason it fails is not due to a simplification of mine. I explain Krifka's notation below.

(18)   [[for an hour]] (Krifka)
       $= \lambda R.\lambda x.\lambda e.[R(x,e) \ \wedge \ H'(e) = 1 \ \wedge$
       $\partial \exists e'[e' <_\tau e \ \wedge \ \forall e''[e'' \leq_\tau e \rightarrow R(x,e'')]]]$

(19)   [[John ran for an hour]] (Krifka)
       $= \exists e.[\text{run}(\text{john}, e) \ \wedge \ H'(e) = 1 \ \wedge$
       $\partial \exists e'[e' <_\tau e \ \wedge \ \forall e''[e'' \leq_\tau e \rightarrow \text{run}(\text{john}, e'')]]]$

The translation in (18) consists of an assertion and of a presupposition, which is indicated by ∂. The translation takes a predicate of individuals and events R, an individual x, and an event e. In (19), the predicate R is assumed to have been contributed by the verb phrase, x by the subject, and e by existential closure. The entry states that the number of hours of the runtime of e is one. This is expressed by the function H' that maps entities to their runtimes in hours (with some special provisions for noncontinuous events that are irrelevant for the present point).

The presupposition of (18) makes use of a parametrized part-of relation $\leq_\tau$. Krifka defines that $e' \leq_\tau e$, read "e' is a temporal subevent of e", holds if and only if $e' \leq e$ and there is another part e'' of e whose runtime $\tau(e'')$ does not overlap with $\tau(e')$.[28] This is the case if the runtime of e' is a proper part of the runtime of e. Thus, Krifka relativizes the part-of relation to time. Note that nothing in the definition requires e and e' to coincide in their spatial extent or in any other way. In particular, e' could also have a smaller spatial extent than e.

The presupposition of (18) states that the predicate R denoted by the verb phrase must relate the subject to all the temporal subevents of the event e. I will refer to this part of the presupposition as the *divisiveness clause*, because it essentially requires the predicate to have divisive reference, except that $\leq$ is replaced by $\leq_\tau$. The presupposition also requires that the event has temporally shorter subevents. I call this the *existence clause*. The purpose of this clause is to exclude quantized telic

---

[28]Krifka actually writes $\leq_{H'}$ rather than $\leq_\tau$. However, the function H' is not used in his definition of $\leq_\tau$. I write Krifka's relation as $\leq_\tau$ because I find this notation more intuitive when it is used in parallel with its spatial counterpart, which I write $\leq_\sigma$.

predicates like *eat an apple*. These predicates vacuously satisfy the divisiveness clause, because an event to which *eat an apple* applies has no subevents – and, in particular, no temporally shorter subevents – that would also be in the denotation of *eat an apple*.

Although Krifka does not deal with spatial *for*-adverbials, his theory can be straightforwardly extended to them by stating that spatial *for*-adverbials check that each spatial (as opposed to temporal) subevent of the sum event is in the denotation of the main clause predicate. This is shown in (20), which uses a relation $\leq_\sigma$ or "is a spatial subevent of", to be understood in a parallel way to $\leq_\tau$.

(20) $\llbracket$for a mile$\rrbracket$ (based on Krifka)
$= \lambda R.\lambda x.\lambda e.[R(x,e) \ \wedge \ M'(e) = 1 \ \wedge$
$\partial \exists e'[e' <_\sigma e \ \wedge \ \forall e''[e'' \leq_\sigma e \to R(x,e'')]]]$

6.4   Comparing the approaches

In this section, I describe a case for which the proposal in Krifka (1998) makes the wrong predictions, and which the proposal in Dowty (1979) and my own proposal handle correctly. Consider again sentence (5a), repeated here:

(21)   John pushed carts to the store for fifty minutes.

The events of which this sentence holds are sums of events in which John goes back and forth to the store many times. For the present purpose, I assume that the predicate *John push carts to the store* is interpreted as in (22) (see also Section 2.9 for my assumptions concerning the semantics of *to the store*). Krifka (1998) assumes a similar representation, although I abstract away from some irrelevant details here.

(22) $\lambda e[{}^*\text{push}(e) \ \wedge \ {}^*\text{ag}(e) = j \ \wedge \ \text{end}(\sigma(e)) = \text{the.store}]$
(True of any potentially pushing event whose agent is John and whose spatial extent is an interval whose end is the store.)

I first illustrate the different behavior of the two approaches intuitively with figures, and then I discuss more formally what goes wrong with Krifka's account. I will overlay the constraints that the two approaches impose onto a schematic representation of an arbitrarily picked sum event of which (21) holds. This schematic representation is shown in Figure 6.1. It uses the visual metaphor described in Chapter 1 and above. This figure is a space-time diagram with time on the vertical axis, and with the path or spatial interval to which *to the store* refers along the horizontal axis. I have taken certain liberties here for expository purposes. Strictly speaking, in a space-time diagram, a stationary object like the store should be

represented as a vertical line, and John and his carts should be represented as zigzagging spacetime "worms". I have used more recognizable representations. One can imagine Figure 6.1 as taken from a vertical filmstrip that has been recorded from a perpendicular angle to John's path. The men and carts can be thought of as individual frames, each of which represents an event in which John pushes carts to the store.

Figure 6.1: An event in the denotation of (21).
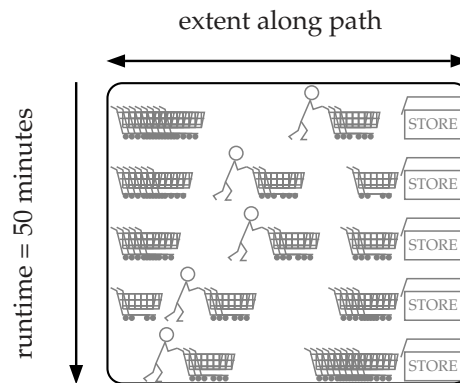
Figure 6.2 contrasts two ways of breaking down this sum event into subevents, corresponding to the two theories. Figure 6.2a illustrates the strata-based approach: the event is divided into subevents along the temporal axis. Each of these events is tested for whether it qualifies, that is, whether it is in the denotation of the predicate *John push carts to the store*. A checkmark indicates that an event qualifies, a cross shows that it does not. A picture in which there are only checkmarks and no crosses represents a case in which the theory predicts the sentence to be acceptable. Since this is the case in Figure 6.2a, this is a point for the strata-based approach.

The subregion-based approach is illustrated in Figure 6.2b. As shown by the abundance of crosses in this figure, this approach breaks down the sum event too much by insisting on considering all subevents, or at least all temporal subevents. Most of these subevents have locations that are remote from the location of the store, and are therefore not in the denotation of the predicate *John push carts to the store*. The restriction to temporally shorter subevents does not prevent this. The subregion-based approach in general, and Krifka's theory in particular, wrongly rules out sentence (21).

Figure 6.2: Comparison of the strata-based and subregion-based approaches

(a) Strata-based approach

(b) Subregion-based approach
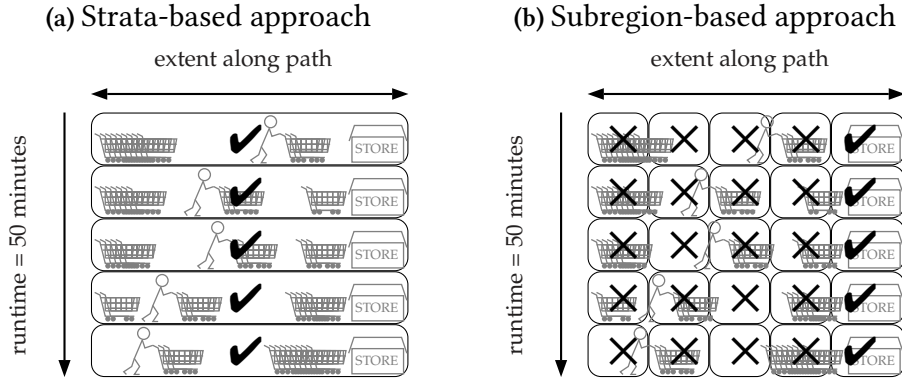
6.4.1 The subregion-based approach

Following this intuitive presentation, let me now show in detail how Krifka (1998)'s account wrongly rules out sentence (21) even though it is more sophisticated than a naïve implementation of the subregion-based approach. I will argue that the presupposition predicted by Krifka's entry in (18) is violated by the predicate denoted by (21). As a result, Krifka's account wrongly predicts that (21) should be unacceptable due to presupposition failure. His prediction is that (21) presupposes that any event e in its denotation satisfies the following condition:

(23) $\exists e'[e' <_\tau e] \ \wedge \ \forall e''[e'' \leq_\tau e \rightarrow e'' \in [\![\text{John push carts to the store}]\!]]$

As a reminder, I call the first conjunct of this presupposition the existence clause and the second conjunct the divisiveness clause. To argue that this presupposition is violated, I will proceed in four steps. First, I will argue that sentence (21) applies to a proper sum event, that is, an event that has subevents. This means that the existence clause of the presupposition is satisfied, which means that the problem must lie in the divisiveness clause and ultimately in Krifka's reliance on divisive reference. Second, I will argue that some of these shorter subevents are not in the denotation of the predicate *John push carts to the store*. Third, I will argue that some of these subevents stand in the relation $\leq_\tau$ and therefore $\leq$ to the sum event (they are parts of the sum event and their runtime is shorter than that of the sum event). Finally, I will argue that these subevents cannot be excluded on pragmatic grounds from the range of the universal quantifier in (23). These four steps together have the consequence that the presupposition in (23) does not hold.

I start by arguing that (21) applies to a sum event that has subevents. I have assumed that a verbal predicate that has been modified by *to the store* only applies to events whose spatial extent ends at the store (see Section 2.9). This assumption is intuitively plausible and corresponds to the account of spatial prepositional phrases in Krifka (1998). Suppose that (21) is uttered in a scenario in which John repeatedly pushed some carts from location A to the store. (In a pragmatically plausible scenario, he pushes a different set of carts on each of his back-and-forth trips, but nothing depends on this.) Call B the point halfway between location A and the store. Standard assumptions of mereological event semantics commit us to the (intuitively plausible) claim that (21) entails the existence of a sum event e which can be divided at least into two complex subevents: an event e_1 in which John pushed carts (iteratively) from A to B and an event e_2 in which he pushed carts (iteratively) from B to the store. I will leave e_2 aside and concentrate on the event e_1. This event does not itself qualify as an event of John pushing carts to the store, because its spatial extent does not include the store. To see that e entails the existence of e_1, note that (21) entails (24), and moreover (24) entails (25):

(24)   John pushed carts to the store.

(25)   John pushed carts halfway towards the store.

These sentences are of course less informative than (21) because they describe parts of the scenario evoked by it, so it is not immediately obvious that the entailment relations that I have mentioned hold. For example, one might take (25) to be false in such a scenario because it conveys that the carts in question did not reach the store. However, that the carts did not reach the store is not a part of the literal meaning of but an implicature. This is clear in contexts such as questions, where implicatures are usually not computed. For example, when we turn (25) into a question, it is possible to answer it affirmatively with (24), but the converse is not possible:

(26)   Did John push carts halfway towards the store?

(27)   Yes/*No – in fact, he pushed carts to the store.

(28)   Did John push carts to the store?

(29)   No/*Yes, he pushed carts halfway towards the store.

I conclude that the entailment relations do indeed hold. Each of these sentences entails the existence of an event, and we can plausibly model the fact that (21) entails (25) by assuming that any event denoted by (21) has an event denoted by (25) as one of its parts. This means that (21) indeed denotes a sum event, as claimed above.

The second step is to argue that the predicate denoted by (24), which applies to the sum event of (21), call it e, does not apply to the sum event of (25), call it e_1. This is easy to see, given that we know that (24) entails (25) but not vice versa. If *John pushed carts to the store* applied to the event in (25), then (25) should entail (24). The intuition behind this reasoning is that an event of pushing carts halfway towards the store neither is nor entails an event of pushing carts to the store.

Third, I argue that $e_1 \leq_\tau e$, that is, e_1 is a part of e and there is another part of e whose runtime does not overlap with the runtime of e_1. This is how Krifka expresses that e_1 has a shorter runtime than e. As mentioned, that $e_1 \leq e$ is intuitively plausible. Any assumption to the contrary would make it at best very difficult to explain why (21) entails (25). Moreover, observe that e_1 is a proper part of e, that is, it is not identical to e. This is so because (25) does not entail (24). By the Unique Separation axiom (see Section 2.3.1), there is an event e_2 that is a proper part of e and does not overlap with e_1, so that the sum of e_1 and e_2 is e. Intuitively, this models the following fact: e is a sum event in which John pushed carts from A to the store. e_1 is a part of e in which John pushed carts from A halfway towards the store, that is, from A to B; and e_2 is an event in which John pushed carts from B to the store, that is, the result of subtracting e_1 from e. Now, given the fact that John cannot be in two different locations at the same time, the runtimes of e_1 and e_2 do not overlap. This entails $e_1 \leq_\tau e$, which I set out to show.[29]

Summing up, we have the following situation. A *for*-adverbial can modify a sentence, namely (21), whose sum event e has a proper part e_1 that is not in the denotation of the sentence. This is in contradiction to divisive reference. Moreover, e_1 has a shorter runtime than the sum event e, therefore not only $e_1 < e$ holds but also $e_1 <_\tau e$. This is in contradiction to Krifka (1998)'s divisiveness clause, by which he implements the subregion-based approach as seen in (18). I will refer to e_1 as the "offending event" since it violates the divisiveness clause. We will see later that the strata-based approach fares better because e_1 does not lead to a violation of stratified reference.

Finally, let us consider what additional assumptions would be necessary to rescue the subregion-based approach. Although this is not always explicitly stated, universal quantifiers in semantic representations are generally assumed to be restricted to "relevant" entities. The following classical example, taken from Kratzer (1989), illustrates the point. We have an orchard whose trees are all laden with wonderful apples. A man who wants to buy the orchard asks us whether all its trees are apple trees. We answer: "Yes, and every tree is laden with wonderful

---

[29]The runtimes of each of these events will be discontinuous in scenarios where John goes back and forth between A and the store and pushes the carts to the store one by one or at least little by little. This does not affect the main point.

apples." It is clear from context that the universal quantifier supplied by *every tree* is implicitly restricted to the trees in the orchard. Otherwise, the sentence would be false because there are many trees in the world which are not apple trees and cannot have any apples. The relevance issue is a bit of a wildcard, because no commonly accepted semantic theory provides clear criteria for deciding whether a given entity is relevant. However, a common intuition is that all irrelevant entities must be "outside the situation" in which the sentence is understood. In the orchard example, most if not all of the trees inside the orchard are relevant from an intuitive point of view since the truth of the sentence depends on any one of them having apples, while the trees outside the orchard are irrelevant. This intuition is formally implemented in situation semantics (Kratzer 1989).

In our event-based framework, the situation in which a sentence is understood can be thought of as the sum event over which a sentence existentially quantifies (Dekker 1997; Kratzer 2010). According to this view, the relevant events in the example discussed above are the subevents of the sum event e. Event e_1, the "offending event" whose location does not contain the location of the store, is indeed a subevent of e and is therefore relevant. As shown in Figure 6.2b, there are many other such offending subevents. It would be difficult to explain why none of these subevents should be seen as relevant.

However, let us grant that an argument to the contrary can be made, and that the only relevant events turn out to be those subevents of e whose location contains the store. In Figure 6.2b, these subevents are the ones whose right-hand boundary is also the right-hand boundary of the sum event. These subevents all have checkmarks. This indicates that they all qualify as *push carts to the store*, which is the case because their location contains the store. It would then follow that all relevant subevents of e would qualify as *push carts to the store*, and the presupposition of (21) would indeed be satisfied. This would appear to rescue the subregion approach.

I see three problems with this line of thinking. First, there is no reason to assume that these subevents are indeed irrelevant for the felicity and truth conditions of the sentence, and that so many of them should be. If they were, we should expect most trees inside our orchard to be irrelevant for the truth conditions of *Every tree is laden with apples* as well. Second, there is no explanation why (21) requires an iterative interpretation. Suppose John starts pushing a set of carts towards the store all at once, and stops doing anything once he has arrived there. Although the locations of most subevents of such an event do not contain the store, those events whose location does contain the store all qualify as events in which John pushes carts to the store. If these are the only relevant subevents, the sentence should be modifiable by any temporal *for*-adverbial. Finally, there is no explanation why spatial *for*-adverbials cannot modify (21), since an analogous reasoning shows

that they should be acceptable just like temporal *for*-adverbials. I conclude that an appeal to relevance is unlikely to save the subregion-based approach.

6.4.2 The strata-based approach

I will now argue that the strata-based approach provides a handle on the observations presented in Section 6.2: that the sentence *John pushed carts to the store* is compatible with a temporal *for*-adverbial; that this adverbial enforces an iterative interpretation; and that the sentence is not compatible with a spatial *for*-adverbial. It will also be necessary to exclude irrelevant events, but this appeal to relevance does not undermine the approach because none of the parts of the sum event need to be considered irrelevant. The strata-based approach was illustrated in Figure 6.2a. The facts just mentioned are illustrated by example (5), repeated here:

(30) a. John pushed carts to the store for fifty minutes.
      b. *John pushed carts to the store for fifty meters.

   I start with the first observation: the sentence is compatible with a temporal *for*-adverbial. Take any event in the denotation of (30) and call it e. As implemented in Dowty's proposal, the strata-based approach checks if every subinterval of the runtime of e is the runtime of an event in the denotation of *John push carts to the store*. This condition is of course problematic, because there are no infinitely short events that qualify as pushing events, let alone events of pushing carts to the store. However, Chapter 5 shows that the strata-based approach can be maintained in spite of this minimal-parts problem. My implementation of the strata-based approach checks if there is a way to divide up the sum event in (21) along the temporal axis into very short strata that are all in the denotation of the main clause predicate. What counts as very short depends on the length of the temporal interval associated with the measure phrase of the *for*-adverbial. In the case of (21), "very short" means whatever counts as very short relative to fifty minutes – say, five minutes.

(31) $\text{SR}_{\tau, \varepsilon(\lambda t[\text{minutes}(t)=50])}(\llbracket \text{John push carts to the store} \rrbracket) \Leftrightarrow$
      $\forall e[e \in \llbracket \text{John push carts to the store} \rrbracket \rightarrow$
      $e \in {}^*(\lambda e'.\varepsilon(\lambda t[\text{minutes}(t) = 50])(\tau(e')) \wedge e' \in \llbracket \text{John push carts to the}$
      $\text{store} \rrbracket)]$
      (Every event e in $\llbracket$John push carts to the store$\rrbracket$ can be divided into one or more parts, each of which is also in the denotation of $\llbracket$John push carts to the store$\rrbracket$ and has a very small runtime compared with fifty minutes.)

Because the strata-based approach does not quantify over all temporally shorter subevents, it places a weaker requirement on the main clause predicate than Krifka's account does. We have seen that the sum event of (21), e, can be divided into two parts e_1 and e_2 such that e_1 does not qualify as an event in which John pushed carts to the store. More specifically, e_1 represented the sum of all subevents in which John pushed carts halfway towards the store. In other words, the existence of the parts e_1 and e_2 represents a way of dividing up e along the spatial axis. The existence of e_1 is without consequence for my proposal. All that needs to be checked is whether there exists a way of dividing e into subevents along the temporal axis, such that these subevents are each in the denotation of the main clause predicate. Once it has established that such a division exists, it does not matter whether or not there are also other ways of dividing e into parts to which the main clause predicate does not apply.

To check whether e can be divided up in the specified way, we need to check the scenarios with which (24) is compatible. Example (24) by itself does not specify whether John pushed carts to the store little by little (iteratively, that is, with many trips back and forth) or all at once, but as mentioned, adding a temporal *for*-adverbial as in (21) forces an iterative interpretation. On my proposal, this contrast is expected. A scenario in which John pushed the carts to the store little by little entails that the main clause predicate in (24) is true within many consecutive time intervals – whatever time it takes John to push a set of carts to the store and to go back to the origin. A scenario in which John pushed the carts to the store all at once is not compatible with such an entailment. This entailment is exactly the one that is also required by *for*-adverbials on the strata-based approach.

The strata-based approach not only accounts for the acceptability and licensed entailments of (21). It also explains the contrast between the acceptability of the temporal interval in (21) and its spatial counterpart. This minimal pair is repeated here from (5):

(32)   a.   John pushed carts to the store for fifty minutes.
       b.   *John pushed carts to the store for fifty meters.

The reason for the contrast between these two sentences is easily explained on the strata-based approach, given any background theory that makes the uncontroversial prediction that the PP *to the store* is only true of events whose spatial extent ends at the store. The idea is this: While the temporal *for*-adverbial in (32a) tests whether the main-clause predicate is true within every very short temporal interval that is a part of the fifty minutes in question, the spatial *for*-adverbial in (32b) tests whether the main-clause predicate is true within every very short spatial subinterval of the fifty meters in question (see Figure 6.3). Some of these

subintervals will not spatially contain the location of the store, that is, they will not contain the endpoint of the spatial interval at which the sum event is located. Any event that is located in such a subinterval will not qualify as *push carts to the store*, that is, it will not be in the denotation of the main-clause predicate. This means that there is no way of dividing up a sum event denoted by *John pushed carts to the store* along the dimension of the spatial *for*-adverbial in (32b) that would be conform to the requirements of the *for*-adverbial. This reasoning applies both to Krifka (1998) and to the strata-based approach. In other words, both theories rule out (32b). However, only the strata-based approach can explain the contrast of (32b) with (32a). As we have seen, Krifka (1998) would rule out both sentences.

Figure 6.3: Ruling out **John pushed carts to the store for fifty meters* on the strata-based approach

extent along path = 50 meters

runtime
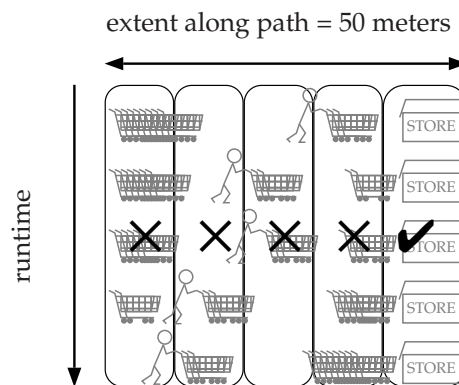
6.5   Summary

Temporal and spatial *for*-adverbials impose analogous constraints on the predicates with which they combine. Any theory of telicity should therefore account for temporal as well as spatial *for*-adverbials. Since the two kinds of adverbials differ in their distribution, any theory of telicity should be able to classify a given predicate as temporally telic but spatially atelic or vice versa. This situation supports a parametrized notion of the telic-atelic opposition, where the parameter is set either to time or to a spatial dimension.

The parametrized nature of aspect is expected within the general picture of this work. As argued in Chapter 4, the fact that a predicate can be distributive on its agent "dimension" without being distributive on its theme "dimension" or

vice versa supports a parametrized view of distributivity. In connection with measurement, I have argued in Chapter 4 that the singular-plural and count-mass distinctions that are operational in the pseudopartitive construction need to be parametrized to a "dimension" like width or diameter. The concept of stratified reference subsumes this picture. The dimension parameter of stratified reference gives us a handle on the distinction between spatial and temporal aspect.

I have argued that subinterval-property based or "strata-based" approaches to atelicity, modified as in Chapter 5, generalize to the spatial case in a better way than approaches which are based on divisive reference do. On strata-based approaches, an atelic predicate is one that holds of subevents which are constrained in size along the dimension with respect to which the predicate is atelic, but not along other dimensions. Approaches based on divisive reference impose too strong conditions because they also require the predicate to hold of subevents whose size is small along all dimensions. Once again, the picture is subsumed by the concept of stratified reference. According to this concept as applied in Chapter 4, a predicate that is distributive on its agent "dimension" is not required to apply to subevents that are constrained on their theme dimensions – that is, it is not required for an agent-distributive predicate to be distributive on its theme as well.

Chapter 7

Measure functions

7.1 Introduction

This chapter considers the formal properties of measure functions and the way in which distributive constructions constrain these properties. I start by describing the relevant constraint in relation with the extensive-intensive opposition known from physics. In line with the unified picture presented in this work, I show that this opposition is not only operative in pseudopartitives, but also in *for*-adverbials. As in Chapters 4 and 6, I draw attention to circumstances involving several dimensions. In this case, the dimensions correspond to different measure functions. I show that the constraint on measure functions only applies along one dimension at a time. I suggest that this behavior has the same source as the atelicity constraint in *for*-adverbials and the distributivity constraint in adverbial-*each* constructions.

My background assumptions on pseudopartitives and measure functions are laid out in Sections 2.5.3 and 3.2. Section 2.5.2 has also introduced trace functions, and Section 3.2 has assumed htat event pseudopartitives involve trace functions. Measure functions relate individuals and events to their degrees on various scales, while trace functions relate events to their runtimes and locations. The rest of this work distinguishes them from measure functions mainly for technical reasons, for example because they have different types. Trace functions behave analogously to measure functions with respect to the facts I will discuss here. Therefore, this chapter does not distinguish between measure functions and trace functions.

7.2 Intensive and extensive measure functions

Pseudopartitives can be used to describe the volume of a quantity of water, but not to describe its temperature:

(1) a. thirty liters of water   *volume*
     b. *thirty degrees Celsius of water   **temperature*

As discussed by Krifka (1998) and Schwarzschild (2006), this example is representative of a general constraint that corresponds to a distinction commonly made in measurement theory and in physics, namely the one between extensive and intensive measure functions. Following Schwarzschild (2006), I refer to this phenomenon as the monotonicity constraint on measure functions. (Monotonicity is Schwarzschild's term for the property of being an extensive measure function.)

In physics, measure functions are not defined to be intensive or extensive per se, but only with respect to a given system. The physical notion of system can be thought of as a given amount of substance or spacetime. An extensive measure function is one whose magnitude is additive for subsystems; an intensive measure function is one whose magnitude is independent of the extent of the system (Krantz et al. 1971; Mills et al. 2007). For example, when one considers the system consisting of the water in a tank, volume is an extensive measure function because the volume of the water as a whole is greater than the volume of any of its proper parts. But temperature is intensive with respect to this system because the temperature of the water as a whole is no different from the temperature of its proper parts.

I take it as a fact about the physical world that certain measure functions are intensive or extensive on certain systems. For example, it is a physical and not a linguistic fact that temperature is intensive on the set of all water amounts. The pseudopartitive construction appears to be sensitive to this physical fact. To model this sensitivity, I assume that the notions of extensive and intensive measure functions have counterparts in the mereological model. For this purpose, I formalize the linguistic counterpart of these terms as follows. Alternative definitions due to Krifka (1998) and Schwarzschild (2006) are discussed in Section 7.5.

I call a function intensive on an entity if and only if it returns the same value for every part of this entity. Intuitively, this entity stands for the system under consideration in the physical sense. In practice, it can be thought of as a substance or an event.

(2) **Definition: Intensive function on an entity**
     For any measure function f and for any entity $x \in \mathrm{dom}(f)$:
     $\mathrm{Intensive}(f, x) \stackrel{\mathrm{def}}{=} \forall y[y < x \ \wedge \ y \in \mathrm{dom}(f) \rightarrow f(y) = f(x)]$
     (A function is intensive on an entity if and only if it returns the same value for every part of this entity.)

The following definition generalizes this concept to sets of entities. In the analysis of pseudopartitives presented below, these sets will be supplied by the

denotation of the measure noun.

(3)   **Definition: Intensive function on a set**
For any measure function f and for any set $X \subseteq \mathrm{dom}(f)$:
Intensive$(f, X) \overset{\text{def}}{=} \exists x[x \in X \wedge \forall y[y < x \wedge y \in \mathrm{dom}(f) \rightarrow f(y) = f(x)]]$
(A function is intensive on a set of entities if and only if it is intensive on at least one of the entities in the set.)

According to this definition, a function that is intensive on a set does not need to be intensive on every entity in the set. The purpose behind this is the following. We want to be able to say that temperature is intensive on water because it is intensive on any typical amount of water. But there may be some amounts of water whose parts do not have the same temperature. The definition is not affected by these outliers. For example, if we heat up a body of water, the heat does not propagate instantly throughout the water and, at least for a while, this body of water will not have a constant temperature. The temperature of the whole body of water may be the average of the temperatures of every part of this water, but it cannot be equal to each of these temperatures. Nevertheless, temperature is still considered intensive despite the existence of these heterogeneous bodies of water. Another possibility would be to treat the temperature of a heterogeneous body of water as undefined. I hesitate to do so because it seems to me that this would be too restrictive compared with the natural language use of the term *temperature*. We can speak of the temperature of an object even if we do not know whether every part of the object has the same temperature. Definition (3) sidesteps this issue because it allows us to disregard bodies of water with heterogeneous temperatures.

Extensive measure functions are defined in the same way as intensive measure functions, except for two changes described below.

(4)   **Definition: Extensive function on an entity**
For any measure function f and for any entity $x \in \mathrm{dom}(f)$:
Extensive$(f, x) \overset{\text{def}}{=} \forall y[y < x \wedge y \in \mathrm{dom}(f) \rightarrow f(y) < f(x)]$
(A function is extensive on an entity if and only if it returns a smaller value for every part of this entity.)

(5)   **Definition: Extensive function on a set**
For any measure function f and for any set $X \subseteq \mathrm{dom}(f)$:
Extensive$(f, X) \overset{\text{def}}{=} \forall x[x \in X \rightarrow \forall y[y < x \wedge y \in \mathrm{dom}(f) \rightarrow f(y) < f(x)]]$
(A function is extensive on a set of entities if and only if it is extensive on at least one of these entities.)

The first change in these definitions compared to definitions (2) and (3) is that $=$ is replaced by $<$. This change captures the physical fact that extensive measure functions are additive, that is, they return larger values on larger entities. The second change is that we do not worry about outliers. That is, we define an extensive measure function on a set as one that is extensive on all its members rather than on some of its members. The purpose of this change is to avoid the counterintuitive possibility that a measure function can be both extensive and intensive. However, the precise definition of extensive measure function is not very important. As we will see, the linguistically relevant distinction is not between intensive and extensive measure functions but between intensive and nonintensive measure functions. The next section gives an example of a function that is neither intensive nor extensive, namely height, and shows that it is just as compatible with pseudopartitives as extensive measure functions are.

I assume that the physical facts about intensive and extensive measure functions are represented as meaning postulates that use these definitions. As described in Section 2.6.5, I assume that mass nouns like *water* denote sets. Given this assumption, the fact that volume is extensive on water but temperature is intensive on water is represented by the following meaning postulates:

(6)   a.   Extensive(volume, water)
       b.   Intensive(temperature, water)

The next section explains how these meaning postulates interact with the Distributivity Constraint to rule out constructions like (1b).

7.3   Why *thirty degrees of water* is unacceptable

This section shows that the account presented and motivated in the previous chapters can be used to explain the constraint against intensive measure functions in pseudopartitives without any additional assumptions. Chapter 4, especially Section 4.3.1, has motivated the assumption that pseudopartitives are distributive constructions and that as such they obey the following constraint:

(7)   **Distributivity Constraint**
     A distributive construction whose Key is K, whose Share is S and whose Map is M is acceptable only if S has stratified reference with respect to M (formally: $\text{SR}_{M,\varepsilon(K)}(S)$).

Let us apply this constraint to the pseudopartitives in (1a) and (1b). The background assumptions described in Section 3.2 assign them the following LFs and

logical representations (I have marked Key, Share, and Map for convenience):

(8) a. $[[_{Key}$ thirty liters] $[[_{Map}$ *volume*] [of $[_{Share}$ water]]]]]
     b. $[[_{Key}$ thirty degrees] $[[_{Map}$ *temperature*] [of $[_{Share}$ water]]]]]

(9) a. $\lambda x[\text{water}(x) \wedge \text{liters}(\text{volume}(x)) = 30]$
     b. $\lambda x[\text{water}(x) \wedge \text{degrees.celsius}(\text{temperature}(x)) = 30]$

I now show how the Distributivity Constraint in (7) accepts the representation (9a) but rules out the representation (9b), given that with respect to water, volume is extensive, but temperature is intensive. The abbreviation SR in (7) refers to stratified reference, the higher-order property that was introduced in Chapter 4 and further motivated in Chapters 5 and 6 on independent grounds. Its definition is repeated here from Section 4.6:

(10) **Definition: Stratified reference**

$$\text{SR}_{f,\varepsilon(K)}(P) \stackrel{\text{def}}{=} \forall x[P(x) \rightarrow x \in {}^*\lambda y \left(\begin{array}{c} P(y) \wedge \\ \varepsilon(K)(f(y)) \end{array} \right)]$$

(A predicate P has stratified reference with respect to a function f and a threshold $\varepsilon(K)$ if and only if there is a way of dividing every entity in its denotation exhaustively into parts ("strata") which are each in P and which have a very small f-value. Very small f-values are those that satisfy $\varepsilon(K)$.)

When the Distributivity Constraint is applied to examples (1a) and (1b), stratified reference expands into the following conditions:

(11) a. $\forall x[\text{water}(x) \rightarrow$
         $x \in {}^*\lambda y(\text{water}(y) \wedge \varepsilon(\lambda d[\text{liters}(d) = 30])(\text{volume}(y)))]$
         (Every amount of water x can be divided into one or more parts, each
         of which is itself water and whose volume is very small relative to
         thirty liters.)
     b. $\forall x[\text{water}(x) \rightarrow$
         $x \in {}^*\lambda y(\text{water}(y) \wedge \varepsilon(\lambda d[\text{degrees.celsius}(d) = 30])(\text{temperature}(y)))]$
         (Every amount of water x can be divided into one or more parts, each
         of which is itself water and whose temperature is very small relative
         to thirty degrees Celsius.)

As discussed in Section 2.6.5, given that *water* is a mass noun, it has approximate divisive reference with respect to a threshold, which I assume to be equal to or below $\varepsilon(\lambda d[\text{liters}(d) = 30])$. This means that every amount of water x above this threshold can be divided into one or more parts, each of which is itself water. Volume is extensive on water, as specified by the meaning postulate (6a). This

meaning postulate states that every amount of water x is such that each of its parts has a smaller volume than x. Condition (11a) is therefore fulfilled, and (1a) is predicted to be acceptable.

By contrast, temperature is intensive on water, as specified by the meaning postulate (6b). This postulate entails that there are water entities whose parts all have the same temperature. On the plausible assumption that some of these entities have a temperature that does not qualify as very low with respect to thirty degrees Celsius, condition (11b) is false, and (1b) is predicted to be unacceptable. This explanation makes use of the Distributivity Constraint, which was motivated extensively in Chapters 5 and 6.

7.4 Generalizing to other distributive constructions

The present work sees the Distributivity Constraint as applying to all distributive constructions. This leads to the prediction that all distributive constructions, and not just pseudopartitives, reject measure functions in the same way. In particular, both pseudopartitives and *for*-adverbials are predicted to be unacceptable when they relate events to measure functions that are intensive on the sets denoted by their substance nominals and by their verbal predicates respectively. The following examples confirm this prediction:

(12) a. thirty hours of driving   *runtime*
      b. thirty miles of driving   *location*
      c. *thirty miles an hour of driving   **speed*

(13) a. drive for thirty hours   *runtime*
      b. drive for thirty miles   *location*
      c. *drive for thirty miles an hour   **speed*

While this observation is novel as far as I know, the general idea that there is a connection between measure adverbials and measurement is not new. Krifka (1998), Section 3.4, notes that *for*-adverbials are like nominal measure phrases "insofar as they introduce a quantitative criterion of application." In his setup, the noun *hour* introduces an extensive measure function for events. Schwarzschild (2006), Section 3.2, similarly points out that a formalization of telic-atelic opposition in the line of Dowty (1979) can be couched in terms of monotonicity. For example, *in*-adverbials can only combine with telic predicates because, as he puts it, runtime "is nonmonotonic on the relevant part-whole relation in the domain given by" that predicate. For Schwarzschild, runtime is a dimension that is monotonic on the part-whole relation that relates events to their subevents. While Schwarzschild

does not go into any formal detail, the intuition is the same as the present one. I will discuss Krifka's and Schwarzschild's proposals in Section 7.5.

I discussed the fact that English has not only temporal but also spatial *for*-adverbials in Chapter 6. In that chapter, I argued for including spatial aspect in the study of aspect. The question arises, which categories of aspect there might be besides temporal and spatial aspect. It turns out that not every measure function is associated with an aspect in the way runtime and location are:

(14) a. drive ∅/for thirty hours   *runtime*
      b. drive ∅/for thirty miles   *location*
      c. *drive ∅/for thirty kilograms   **weight*
      d. *drive ∅/for thirty degrees Celsius   **temperature*
      e. drive ∅/***for** thirty miles an hour   **speed*

To some extent, these gaps have independent explanations. Unlike typical pseudopartitives, *for*-adverbials measure events rather than substances. Common sense suggests that driving events do not have weights or temperatures, and we can assume that examples like (14c-d) are category mistakes. More formally, I assume that the measure functions *weight* and *temperature* are partial functions (see Section 2.5.3) and do not have any driving events in their domains. What is surprising, however, is that (14e) is unacceptable with a *for*-adverbial but acceptable when the word *for* is left out. This contrast shows that even though we can talk in principle about the speed of an event, *for*-adverbials reject speed as a Map. The example from Quine (1985), mentioned in Section 2.4.4, of a sphere rotating slowly and heating up quickly at the same time also suggests that speed is among the possible properties of events. We can avoid the undesirable conclusion that the sphere is both quick and slow by assuming that the rotating and the heating up are two separate events, and that each one has a different speed.

One might think that it is just an idiosyncratic fact about *for*-adverbials that they reject speed as a measure function. However, we have seen above that speed is rejected as a Map by pseudopartitives as well. What is more, the following example shows that *for*-adverbials are productive in the way they combine with measure functions:

(15)   I'm interested in fashion for five pages, not for eighteen pages!

This sentence was used by Tuğba Çolak in conversation with me on April 25, 2010. Its intended interpretation was roughly *I'm willing to read five pages but not eighteen pages of articles on fashion.* Here, the *for*-adverbial is used in connection with a measure function that can best be described as page length. Given that *for*-adverbials can be put to use spontaneously even with unusual measure functions

such as page length, it is remarkable that they resist measure functions such as temperature in examples such as the ones above.

I now propose an explanation of these facts in terms of strata theory. The measure functions *runtime* and *location* are acceptable both in pseudopartitives and *for*-aderbials. They can both be regarded as nonintensive measure functions on events, because the runtime and location of a part of an event can be smaller than the one of the whole event. In contrast, there are driving events on which speed is intensive, that is, every part of the event has the same speed as the whole event. I assume that speed is intensive on driving events:

(16)   Intensive(speed, drive)

Based on the background assumptions in Chapter 3, the constructions in (12) and (13) involve different measure functions: runtime, location, and speed. The Distributivity Constraint therefore requires stratified reference to hold of the predicate *drive* with respect to different Maps (measure functions). The specific conditions imposed on (12) and (13) are the following (respectively for the a, b, and c examples):

(17)   a.   Every driving event e can be divided into one or more parts, each of which is a driving event whose runtime is very small.   *(true)*
        b.   Every driving event e can be divided into one or more parts, each of which is a driving event whose location is very small.   *(true)*
        c.   Every driving event e can be divided into one or more parts, each of which is a driving event of very short speed.   *(false)*

Condition (17c) fails because speed is intensive on driving events. This explains why (12c) and (13c) are unacceptable. In general, all measure functions that are intensive in the systems denoted by the substance nominal of the pseudopartitive are predicted to be unacceptable.

Since the Distributivity Constraint is relativized to the Share (the substance nominal of the pseudopartitive and the verbal predicate with which a *for*-adverbial combines), we expect different Shares to be compatible with different measure functions. Consider the nouns *fever* and *warming*: A sum of two consecutive one-degree fever bouts is a two-degree fever bout, and a sum of two consecutive one-degree warmings is a two-degree warming. Constrast this with the case of water: a sum of two water entities whose temperature is one degree does not equal a water entity whose temperature is two degrees. Generalizing from these examples, it is plausible to assume that the measure function *temperature* is intensive on the set ⟦water⟧ and extensive on the sets ⟦fever⟧ and ⟦warming⟧. We can express this distinction by assuming the following meaning postulates:

(18) a. Extensive(temperature, fever)
      b. Extensive(temperature, warming)
      c. Intensive(temperature, water)

The postulate (18c) is repeated from (6b) above. The point here is that it is not in logical contradiction with the meaning postulates (18a) and (18b).

Based on these meaning postulates, we expect that measure functions like temperature should be acceptable in pseudopartitives and in *for*-adverbials as long as their Share satisfies stratified reference. This prediction is confirmed by the following examples (emphasis mine):

(19) a. Emilia was lying on her bed, *with 41 degrees Celsius of fever.*[30]
      b. The scientists from Princeton and Harvard universities say just *two degrees Celsius of global warming*, which is widely expected to occur in coming decades, could be enough to inundate the planet.[31]
      c. The sample continued to *cool for several degrees* to point N and then suddenly increased to a temperature between the transition points of Form I and Form I1 with no indication of the presence of Form 111.[32]

These examples show that *temperature* is an acceptable measure function in pseudopartitives whose substance noun is *fever* or *warming*. We have seen above that temperature is not acceptable in pseudopartitives whose substance noun is *water*. This provides additional motivation for the point made above: We cannot simply categorize measure functions as intensive or extensive per se. What matters is whether they are intensive or extensive on the set denoted by the substance noun of the pseudopartitive in which they appear.

7.5   Comparison with previous work

The observation that pseudopartitives are sensitive to the distinction between intensive and extensive measure functions goes back to Krifka (1989, 1998) and is at the center of the work of Schwarzschild (2002, 2006). In this section, I compare strata theory with these theories. Since the two authors propose very similar constraints, I discuss them together. As elsewhere, I have attempted to be representative but not exhaustive. A comparison of my approach with the

---

[30] Attested example (http://www.fanfiction.net/s/3616691/). Accessed November 15, 2010.

[31] Attested example (Calgary Herald, December 17, 2009, article: *Two degrees is all it takes – Warming may trigger floods*).

[32] Attested example, from Daubert and Clarke (1944).

proposals in Nakanishi (2004), Heycock and Zamparelli (2005), Chierchia (2008), and Brasoveanu (2009) will have to wait for another occasion.

Krifka (1998) defines the notion of extensive measure function with respect to nonoverlapping entities:

(20)   **Definition: Extensive measure function (Krifka)**
        A measure function μ is *extensive* if and only if for any a, b that do not overlap, $\mu(a) + \mu(b) = \mu(a \oplus b)$; and for any c, d, if $c \leq d$ and $\mu(d) > 0$ then $\mu(c) > 0$.

Unlike Lønning (1987), Schwarzschild (2006) and myself, Krifka's ontology does not have a layer of degrees that distinguishes between measure functions like *height* and unit functions like *meters*. Instead, he assumes that measure phrases involve functions that relate entities directly to numbers in a way that maps the ordering relation between entities 'be smaller than' to the relation 'be less than' between numbers. For example, a function we might write *height-in-meters* relates entities directly to numbers that represent their height in meters, and similarly for functions like *volume-in-liters* and *temperature-in-degrees-Celsius*. I use these longer names only to distinguish the functions they denote from my measure and unit functions. Krifka (1998) refers to the functions he uses as measure functions and uses short names like *liters*. This analysis is advocated in Quine (1960, p. 244-5) and has been adopted in various other places, such as Chierchia (1998a, p. 74). As Schwarzschild (2002) points out, this Quinean analysis makes it difficult to analyze the semantics of comparatives like *six ounces heavier*. I adopt it only temporarily for the purpose of discussing the analysis of Krifka (1998). I will refer to the functions Krifka adopts as *Quinean measure functions*.

A Quinean measure function like *volume-in-liters* is extensive in the sense of Krifka's definition (20) because the volume of the sum of any two nonoverlapping entities is the sum of their volumes, and no entity with nonzero volume has a part with zero volume. (Such a part would have to be an empty part or bottom element. Krifka assumes a system equivalent to classical extensional mereology, in which there is no bottom element. See Section 2.3.1 for discussion.) A Quinean measure function like *temperature-in-degrees-Celsius* is not extensive because the temperature of the sum of two nonoverlapping entities is not the sum of their temperatures. Depending on the ontological setup, it might be undefined or their average.

Krifka places a constraint into the semantic translations of measure phrases like *thirty liters* that makes them compatible only with extensive Quinean measure functions in the sense of his definition. Schwarzschild (2006) notes that placing this constraint into the measure phrase causes problems because not all uses of

measure phrases occur in constructions that reject intensive measure functions. Of course, Krifka's constraint could equally well be placed into the lexical entry of *of* in order to tie it to pseudopartitives specifically.

Schwarzschild (2006) uses the term *monotonic measure function* for the same concept. While he does not provide a formal definition of monotonicity, the definition in (21) is implicit in the discussion in his paper. Roger Schwarzschild (p.c.) informs me that this definition is indeed what he had in mind.

(21)   **Definition: monotonic measure function (Schwarzschild)**
        A measure function μ is *monotonic* if and only if for any a, b, if a is a proper part of b, then $\mu(a) < \mu(b)$.

This definition is very similar to the one in (20) above. This becomes clear when we apply it to the same example: volume is monotonic in the sense of this definition because any proper part of an entity has a smaller volume than that entity. But temperature is not monotonic because proper parts of an entity are generally not colder than that entity.

The definitions by Krifka and Schwarzschild do not fit into the general picture pursued in this work. Through the Distributivity Constraint, adverbial-*each* constructions impose identical constraints on their Maps as pseudopartitives do (see Chapter 4). These thematic roles are generally not extensive or monotonic. Two nonoverlapping events, such as John's running from his house halfway to the store and his subsequent running to the store, can have the same agent, and their sum (John's running from his house to the store) again has the same agent.

In fairness, neither Krifka nor Schwarzschild attempted to relate pseudopartitives to adverbial-*each* constructions. But we find the same problem even if we only look at pseudopartitives. Height is acceptable in pseudopartitives like (22), so if Krifka's and Schwarzschild's accounts are correct, it should be extensive and monotonic.

(22)   Five feet of snow covered Berlin.

However, height is neither monotonic nor extensive according to Krifka's, Schwarzschild's, and my own definitions. Here is why: Imagine that it snows on Berlin for two days in a row, Monday and Tuesday. Imagine that the snow does not melt, so that after these two days there are two layers of snow on top of each other. Imagine that the height of the total snow cover is five feet. The lower layer $s_\downarrow$ has fallen on Monday, and the upper layer $s_\uparrow$ has fallen on Tuesday. Call s the sum of $s_\downarrow$ and $s_\uparrow$, the snow that fell on Berlin as a whole. There are of course different ways of dividing up s. We can look at its horizontal layers $s_\downarrow$ and $s_\uparrow$, but we can also separate it vertically, according to the different regions on which it has fallen.

For example, among the proper parts of s are the snow that fell on West Berlin, call it s_w, and the snow that fell on East Berlin, call it s_e. Then s is also the sum of s_w and s_e.

The problem is that s_w and s_e all have the same height. So height is, according to the formal properties defined above, not extensive and not monotonic. If height was extensive in the sense of Krifka (1998), the sum of the heights of s_w and s_e should be the same as the height of their sum $s_w \oplus s_e$. In Krifka's system, the fact that s_w, s_e and s have the same height, namely five feet, is described as follows:

(23)   height-in-feet$(s_w) = 5 \; \wedge \;$ height-in-feet$(s_e) = 5 \; \wedge \;$ height-in-feet$(s) = 5$

But we can refer to the snow in question with a pseudopartitive, for example by using sentence (22). This means that the formal properties *extensive measure function* and *monotonic measure function* do not correctly characterize the class of admissible measure functions in pseudopartitives.

Schwarzschild (2006) is aware of this problem. From similar examples, he concludes that pseudopartitives do not test for monotonicity with respect to the mereological part-whole relation, but with respect to a different part-whole relation which he sees as contextually supplied. In our example, his assumption would be that context provides a relation according to which the snow that fell on West Berlin, s_w, may well not be a part of the snow that fell on the entire city, s. Schwarzschild may well accept that s_w is a mereological part of s since snow is a mass noun, but this fact does not enter the picture.

I see two problems with this suggestion. First, Schwarzschild (2006) does not impose any formal constraints on the contextually supplied part-whole relation he assumes. We have already faced a similar situation in connection with the discussion of Moltmann's contextually determined part-whole relation in Section 5.3.3. At that time, I mentioned the objection by Zucchi and White (2001) to her proposal: "Since Moltmann does not tell us much about what *relevant* parts are, it is unclear to what extent her formulation actually solves the minimal parts problem." The same objection holds for Schwarzschild's relation. There is no way to know whether two entities stand in Schwarzschild's contextual part relation, so it is unclear how to test the predictions of his account. Second, many measure functions like *temperature* are already correctly ruled out even without replacing the mereological part-whole relation by a contextually supplied relation, so the two relations must coincide to a large extent.

In contrast to Schwarzschild (2006) but in keeping with Krifka (1998), my own account is based on the mereological parthood relation. This relation is assumed to be independent of context (see Section 2.3.1). Context does enter the picture of my account in two ways: to determine the threshold parameter $\varepsilon(K)$ that allows

stratified reference to avoid the minimal-parts problem (Chapter 5), and to exclude (some) entities from consideration that are not part of the entity that the sentence or phrase is about (Chapter 6). However, it is not necessary to appeal to context in order to account for the Berlin example, because stratified reference can already accommodate it. The issue and its solution are exactly the same as the solution discussed in Chapter 6 for the sentence *John pushed carts to the store.*

In contrast to both Krifka (1998) and Schwarzschild (2006), my account does not explicitly disallow nonextensive measure functions. Being an extensive measure function is not a necessary condition for appearing in a pseudopartitive. This is an advantage of the present account, since height is an admissible measure function in pseudopartitives like (22), and since we have seen above that height is not extensive. The Distributivity Constraint imposes the following condition on (22):

(24) $\forall x[\text{snow}(x) \rightarrow x \in {}^*\lambda y(\text{snow}(y) \ \wedge \ \varepsilon(\lambda t[\text{feet}(t) = 5])(\text{height}(y)))]$
     (Every amount of snow x can be divided into one or more parts (strata), each of which is itself snow and has a very small height.)

This condition is satisfied in the scenario described above despite the existence of the entities $s_\uparrow$ and $s_\downarrow$. The argumentation is analogous to Chapter 6. In the terms of that chapter, the present account is strata-based since it allows the snow parts in question to extend arbitrarily in all dimensions except height. The strata are very thin horizontal layers of snow (Figure 7.1). Their horizontal orientation is a result of the fact that they are only constrained along the height dimension. The account by Schwarzschild (2006) is also strata-based, but it assumes that context supplies the decomposition of the snow into horizontally oriented strata. The account by Krifka (1998) is subregion-based because it applies the relevant test to subregions of snow such as s_w. In this case, the test requires subregions of snow to have a smaller height than the entire snow does, and s_w fails this test.

7.6   Summary

The starting point for this chapter was the old observation that intensive measure functions like *temperature* may not occur in pseudopartitives. We have seen that the same constraint is also operative in *for*-adverbials. Both facts are predicted by the present framework.

The constraint against intensive measure functions is also observed by Krifka (1998) and Schwarzschild (2006). While they also discuss some parallels between pseudopartitives and aspect, the present account is the first to explore the connections between different distributive constructions systematically. As we have

Figure 7.1: Accepting *five feet of snow* on the strata-based approach
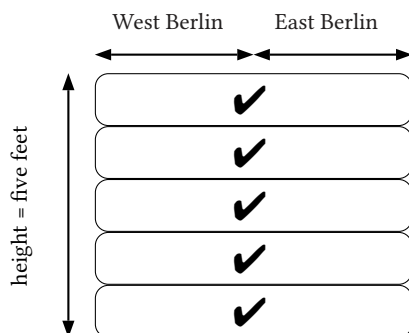
already seen in Chapter 3, an event pseudopartitive like *three hours of running* is given the same analysis as a *for*-adverbial like *run for three hours.* This has provided us with the basis for explaining that the two constructions also license the same measure functions.

The present account subsumes the insights of Krifka and Schwarzschild and makes sense of their observations within the larger picture of strata theory. Unlike previous accounts, the constraint against intensive measure functions does not have to be stipulated, because it is a consequence of the Distributivity Constraint.

In previous chapters, I have exploited one of the defining features of strata theory, namely that it pushes us towards thinking of distributivity as relativized to a certain dimension, thematic role, or measure function. For example, Chapter 6 has argued that we should understand temporal atelicity as something more akin to the subinterval property than to divisive reference, because only the subinterval property is relativized to time. Metaphorically speaking, we should consider only the strata of a given event and not all its subregions. In this chapter, I have transferred this insight to pseudopartitives. While the entities involved are substances rather than events and while the dimensions are measure functions rather than thematic roles, the distinction between strata-based and subregion-based approaches is still operative. In this chapter, the example *five feet of snow* has played the same role as the example *push carts to the store* in Chapter 6. Both examples force us to consider two dimensions at once: height and width in the former case, time and space in the latter case. The insight from Schwarzschild (2006) that the pseudopartitive construction must be checked on horizontal layers of snow rather than on every subregion of snow finds a natural explanation here.

Chapter 8

Distributivity and scope

8.1 Introduction

In the previous chapters, I have shown how the properties of distributivity, atelicity, and noncount reference can be seen as expressions of a parametrized higher-order property, stratified reference. We have seen that this property allows us to get a handle on a number of problems that occur in the domains of aspect and measurement by manipulating its dimension and granularity parameters. Throughout these chapters, I have not made a distinction between lexical and derived constituents, and I have not talked much about how predicates obtain stratified reference. I have assumed that distributive predicates such as *walk* and *smile* have stratified reference as a matter of world knowledge, which can be formally enforced through meaning postulates. A meaning-postulate approach to distributivity is fine for lexical predictes, but as we will see, it leads to well-known problems with phrasal constituents such as *wear a dress* in which the presence of an indefinite quantifier allows us to locate the source of distributivity at the level of the verb phrase or higher. Meaning postulates are a plausible option when it comes to explaining why certain lexical items are distributive, but when distributivity occurs at the phrasal level, this approach is no longer plausible.

In this chapter, I follow and extend the standard way the literature answers this problem by adopting an operator-based account of phrasal distributivity. A nondistributive verb phrase can be "shifted" to a distributive interpretation through the application of a variant of the D operator known from (Link 1987b). Since my Neo-Davidsonian setup represents verb phrases as event predicates, I formulate the D operator in a somewhat different way than Link does, namely as an event predicate modifier. The view developed in the previous chapters, according to which predicative distributivity is stratified reference, proves beneficial here.

Applying the D operator to a verb phrase can be seen as a means to locally ensure that the stratified reference presupposition of a distributive item like *each* is satisfied, a repair strategy not unlike presupposition accommodation. Since stratified reference is parametrized for dimension and granularity, we expect the D operator to be parametrized in the same way. I will explore the consequences of this view in two directions, corresponding to the two parameters.

- By understanding the D operator as parametrized for granularity, we gain a new perspective on the debate between proponents of atomic and cover-based formulations of this operator. The atomic D operator of Link (1987b), Roberts (1987), and Winter (2001) corresponds to one setting of the granularity parameter, and the nonatomic D operator of Schwarzschild (1996) corresponds to another setting. Following Schwarzschild, I will argue that the granularity parameter of the D operator is anaphoric on its context, and can only be set to a nonatomic value when context supports a salient granularity level.

- By understanding the D operator as parametrized for dimension, we gain the technical ability to distinguish agent-based from theme-based distributivity and the like (Lasersohn 1998). Perhaps more interestingly, in the present framework, not only thematic roles like agent and theme are considered dimensions, but also trace functions like runtime and location. We can therefore ask whether the dimension parameter of the D operator can also be instantiated as a trace function. Given the assumption that time and space are nonatomic, we expect that this should only be possible when the granularity parameter of the D operator is set to a nonatomic value, which in turn should require context to provide a salient granularity. I will argue that such contexts indeed exist although they are rare, and I show that the corresponding phenomenon was already noticed in the literature on aspect. Through the parametrized D operator, the asymmetry between the atomic domain of individuals and the nonatomic domain of time allows us to explain the different scopal behavior of *for*-adverbials and distributive items like *each*. If the D operator is easily available only when its granularity is atomic, then it is expected not to be easily available in the temporal domain of *for*-adverbials.

The following introductory discussion, as well as the discussion in Section 8.3, are inspired from Winter (2001), Section 6.2, who in turn builds on earlier contributions by Roberts (1987) and others. For other introductions to the same topic, see also Schwarzschild (1996), Chapter 6, and Link (1997), Section 7.4.

8.2   Lexical and phrasal distributivity

Section 4.2.3 introduced predicative distributivity as a property of predicates. This property was illustrated both with lexical predicates like *smile* and with phrasal predicates like *wear a dress*. To highlight this distinction, I speak of lexical and phrasal distributivity, respectively (and likewise for collectivity):

(1)   **Lexical distributivity/collectivity**

    a.   The children smiled.   *distributive*

    b.   The children were numerous.   *collective*

(2)   **Phrasal distributivity/collectivity**

    a.   The girls are wearing a dress.   *distributive*

    b.   The girls are sharing a pizza.   *collective*

Sentence (1a) entails that each child smiled, while sentence (1b) does not entail that each child was numerous. Similarly, sentence (2a) entails that each girl wears a different dress, but sentence (2b) does not entail that the girls ate different pizzas. At this point, two caveats need to be made. First, distributive readings with definite plurals taking scope over singular indefinites are somewhat marked, and not always easily available. See Dotlačil (2010) for extensive discussion. Second, the predicate *share a pizza* is actually not collective but mixed, because sentence (2b) does admit of a distributive interpretation, however dispreferred. Similarly, to the extent that we admit strange models in which several girls can wear the same dress, the predicate *wear a dress* is also mixed. These interpretations can be roughly paraphrased as follows:

(3)   a.   There is a dress that the girls are wearing.   *collective*

    b.   The girls are each sharing a pizza (with someone else).   *distributive*

However, let us ignore this fact and continue to use *wear a dress* and *share a pizza* as prototypical distributive and collective phrasal predicates. They are as close as one can get to the ideal.

The distinction between lexical and phrasal distributivity is closely related to the P/Q-distributivity distinction proposed by Winter (1997, 2001). Winter uses the term P-distributivity (where P stands for *predicate*) to refer to those cases of distributivity which can, in principle, be derived from some property of the lexical item involved. Q-distributivity (Q for *quantificational*) refers to cases where this approach is not possible because the distributive predicate contains an overt quantifier, as in (2a). In order for (2a) to entail that each girl wears a different dress, the entire verb phrase, including its object, must be distributed over the girls. This

means that the entire verb phrase *wear a dress* and not just the verb *wear* must be regarded as distributive. Since only phrasal constituents can contain quantifiers, Q-distributivity is by necessity always phrasal.

The difference between lexical and phrasal distributivity corresponds to the difference between what can and what cannot be described using meaning postulates. Assuming that *the children* refers to the sum of all children (see Section 2.3.1), it is possible to ascribe the difference between (1a) and (1b) to the meaning of *smile* and *be numerous*. The difference between these two verbs can be described as a meaning postulate to the effect that whenever *smile* applies to a plural event whose plural agent is a proper sum, it also applies to events whose agents are the atomic parts of that sum, while there is no such meaning postulate for *be numerous* (Scha 1981). Stratified reference allows us to formulate this meaning postulate concisely as follows:

(4)   **Meaning postulate: *smile* is distributive**
$SR_{ag,Atom}(\llbracket smile \rrbracket)$
$\Leftrightarrow SR_{ag,Atom}(\lambda e[smile(e)])$
$\Leftrightarrow \forall e[^*smile(e) \rightarrow e \in {}^*\lambda e'(^*smile(e') \wedge Atom(^*ag(e')))]$
(Every smiling event consists of one or more smiling events whose agents are atomic.)

Equivalently, *smile* can be assumed to apply to events whose agents are atomic individuals and (via lexical cumulativity) to events whose agents are the sums of these individuals, while *be numerous* also applies to events whose agents are sums of people even when it does not apply to events whose agents are the individual people in these sums (Hoeksema 1983).

The difference between the distributive interpretation in (2a) and the collective interpretation in (3a) is of a different kind, since it involves a scopal ambiguity. Neither an implementation based on meaning postulates nor one that is based on lexical cumulativity can handle Q-distributivity. The lexical cumulativity assumption, which entails that *wear* applies to singular and plural wearing events alike, cannot model phrasal distributivity because lexical cumulativity concerns only the verb level and not to the verb phrase level (see Section 2.7.2). For this reason, it cannot create a scopal dependency to between the definite subject and the indefinite object (Kratzer 2007). An implementation based on meaning postulates fares even worse. On the assumption that *the girls* denotes a proper sum, we would have to formulate a meaning postulate such as (5) to make sure that (2a) entails that each of the atomic individuals in this sum wears a dress.

(5)   **Problematic meaning postulate: *wear a dress* is distributive**
$SR_{ag,Atom}(\llbracket wear\ a\ dress \rrbracket)$

$\Leftrightarrow \text{SR}_{ag,Atom}(\lambda e[^*\text{wear}(e) \wedge \text{dress}(^*\text{th}(e))])$

$\Leftrightarrow \forall e[^*\text{wear}(e) \wedge \text{dress}(^*\text{th}(e)) \rightarrow$

$e \in {}^*\lambda e'(^*\text{wear}(e') \wedge \text{dress}(^*\text{th}(e')) \wedge \text{Atom}(^*\text{ag}(e')))]$

(Every event in which a dress is worn consists of one or more events in which a dress is worn and whose agents are atomic.)

There are two problems with this meaning postulate. First, in order to allow a sum event whose theme is one dress to consist of several events whose themes are different dresses, we would have to give up the assumption that thematic roles are sum homomorphisms (see Section 2.5.1), because one dress cannot be the sum of different dresses. Second, the postulate is stated with respect to a verb phrase, but meaning postulates are generally taken to be applicable only to individual lexical entries and not to larger constituents such as verb phrases. It may be possible to generate infinitely many such larger constituents for a given grammar, and it would be impractical to specify meaning postulates for them in advance.

The classical way out of this problem, which I adopt here too, is to introduce a covert distributive operator in the logical representation so that the indefinite can take scope at two different places with respect to it. This operator is usually called D, and it adjusts the meaning of a verb phrase like *wear a dress* into a form that satisfies a condition analogous to (5). The D operator was originally defined by Link (1987b) in a setting that translated intransitive verbs and verb phrases as predicates over individuals (type $\langle et \rangle$) rather than predicates over events (type $\langle vt \rangle$). Link's operator is specified as follows:

(6) $\llbracket D \rrbracket (\text{Link}) = \lambda P_{\langle et \rangle} \lambda x \forall y[y \leq_{Atom} x \rightarrow P(y)]$
(Takes a predicate P over individuals and returns a predicate that applies to any individual whose atomic parts each satisfy P.)

This operator introduces a universal quantifier, whose scopal interaction with the indefinite inside a Q-distributive predicate accounts for the covariation effects. For example, in a classical framework where the verb phrase *wear a dress* is translated as $\lambda x \exists z[\text{dress}(z) \wedge \text{wear}(x, z)]$, the meaning of (2a) can be represented in a way that places it in the scope of the universal quantifier introduced by the D operator.

(7) $\forall y[y \leq_{Atom} \bigoplus \text{girl} \rightarrow \exists z[\text{dress}(z) \wedge \text{wear}(y, z)]]$
(Every atomic part of the sum of all girls wears a dress.)

Unlike meaning postulates, the D operator is taken to be part of the lexicon, and therefore it is available to apply to entire verb phrases and not just to predicates. The significance of this fact in the context of phrasal distributivity is discussed in

Dowty (1987), Roberts (1987), and Lasersohn (1995) (Chapter 7), among others.

In the semantic framework adopted here, the formulation of the D operator must be modified for a number of reasons. First, I follow Landman (1989) in assuming that the definite plural *the girls* is ambiguous between a sum interpretation and a group interpretation, an assumption which is required to model the distinction between collective and cumulative readings (see Section 2.8).[33] Impure atoms are not intended to play a role in any account that uses a D operator, so I exclude them by letting the operator distribute over pure atoms rather than just over atoms. I do this with the help of the predicate *PureAtom* defined in Section 2.3.1. Second, I assume that verb phrases are predicates over events rather than individuals (see Section 2.7.1), so D must be of type $\langle vt, vt \rangle$ rather than $\langle et, et \rangle$. Third, I have argued in Section 4.2.2 that predicative distributivity must be specifiable with respect to a given thematic role, since the verb *kill*, for example, is distributive on its theme role but not on its agent role.

There are many possible ways to reformulate the D operator that fulfill these requirements. I find the following way insightful. We already have at our disposal a very similar notion, namely stratified distributive reference (SDR). Section 4.5 has developed this notion as a way to characterize predicative distributivity. The definition of SDR is repeated here:

(8)   **Definition: Stratified distributive reference**

$$\mathrm{SDR}_\theta(P) \overset{\text{def}}{=} \forall e[P(e) \rightarrow e \in {}^*\lambda e' \begin{pmatrix} P(e') \wedge \\ \mathrm{PureAtom}(\theta(e')) \end{pmatrix}]$$

(An event predicate P has stratified distributive reference with respect to a thematic role θ if and only if every event e to which P applies can be exhaustively divided into one or more subevents ("strata") to which P also applies and whose θ is a pure atom.)

As shown in Section 4.6, SDR is just a special case of stratified reference. We could equivalently use stratified reference for what follows, but the development

---

[33]Landman (1989), who does not assume the existence of a D operator, shows that the sum-group ambiguity of plural noun phrases together with a set of type-shifting operations can account for cases involving coordination of a distributive and a collective VP that were once thought to provide a knock-down argument for the D operator as a VP modifier (Dowty 1987; Roberts 1987; Lasersohn 1995). However, later authors such as Schwarzschild (1996) (p. 62) and Winter (2001) (Sections 3.3.2 and 6.3) show that the D operator is needed for reasons other than VP coordination. Schwarzschild discusses an example from Angelika Kratzer (p.c.) in which the D operator interacts scopally with a raising predicate, which I also mention below. Winter argues for the D operator on the basis of the observation, due to Eddy Ruys, that noun phrases can take existential and distributive scope at different places in the syntax. When the full range of phenomena is considered, including collective, cumulative, and distributive readings, it appears necessary to assume both the existence of the D operator and the sum-group ambiguity of noun phrases, as I do here.

would be less clear because stratified reference uses an additional granularity parameter. In the next section, I will introduce this parameter into the D operator to account for cases of nonatomic distributivity. The parallel between the D operator and stratified reference will then be clear. For now, I leave the granularity issue aside.

Since SDR tells us what it means for a predicate to be distributive, the purpose of the D operator can be understood as shifting arbitrary predicates to an interpretation that satisfies SDR. With this in mind, I define the D operator as follows:

(9) **Definition: Atomic event-based D operator**

$$\llbracket D_\theta \rrbracket \overset{\text{def}}{=} \lambda P_{\langle vt \rangle} \lambda e [e \in {}^* \lambda e' \left(\begin{array}{l} P(e') \wedge \\ \text{PureAtom}(\theta(e')) \end{array} \right)]$$

(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs are pure atoms.)

The subscript θ stands for a thematic role. I assume that the D operator is coindexed with the thematic role it targets. I provide justification for this assumption below.

As the following theorem shows, there is a close relationship between this D operator and SDR. I prove this theorem in the Appendix.

(10) **Theorem: D_θ is a repair strategy**
$$\forall P \forall \theta [\text{SDR}_\theta(D_\theta(P))]$$
(When the D operator coindexed with the thematic role θ is applied to any predicate, the result always has stratified distributive reference with respect to θ).

This theorem allows us to understand the D operator as a linguistic repair strategy: by applying D to a predicate, we change its denotation so that it satisfies the SDR requirement. We can now understand distributive markers like *each* as imposing pressure on a predicate to satisfy stratified reference, and this pressure can be resolved through the application of the D operator to the predicate. In Section 4.7, I implemented the SDR requirement of *each* as a presupposition that *each* imposes as a consequence of the distributivity constraint. Applying the D operator can then be seen as a strategy of locally satisfying this presupposition, somewhat similarly to presupposition accommodation. On this view, a sentence such as *The girls each wear a dress* involves the application of a D operator to *wear a dress* in order to satisfy the presupposition of *each*.

Let me illustrate how my D operator works with an example. In my framework, the verb phrase *wear a dress* is translated as $\lambda e[{}^*\text{wear}(e) \wedge \text{dress}({}^*\text{th}(e))]$, or the

set of all potentially plural wearing events whose theme is a dress. As such, this predicate cannot be used to derive the distributive reading of a sentence like (2a). As discussed in Section 2.8, I assume that this distributive reading involves the translation of *the girls* as $\bigoplus$ girl, and that composing the meaning of a subject and a verb phrase involves a thematic role head [ag] (also known as little v) and a referential type shifter. After existential closure applies, the result of this derivation is the following translation:

(11) $\exists e[{}^{*}\mathrm{ag}(e) = \bigoplus \mathrm{girl} \ \wedge \ {}^{*}\mathrm{wear}(e) \ \wedge \ \mathrm{dress}({}^{*}\mathrm{th}(e))]$
(There is a potentially plural wearing event whose agents sum up to the girls, and whose theme is a dress.)

Although this formula allows for the possibility that the event e is plural and consists of one wearing event per girl, it is only true in a model in which the theme of these events is one and the same dress. It is not possible for the dresses that the girls are wearing to be different. The formula requires the theme of e to be in the denotation of the predicate that is supplied by the singular count noun *dress*. As discussed in Section 2.6.1, I assume that singular count nouns apply only to mereological atoms, so ${}^{*}\mathrm{th}(e)$ must be an atom, and its only part is itself. Each part of e must also have this atom as its theme, which means that every girl must be wearing the same dress. This follows from the cumulativity assumption for thematic roles (see Section 2.5.1), which requires ${}^{*}\mathrm{th}(e)$ to be the sum of the themes of the parts of e.

However, consider now the result of applying the D operator in (9) to the verb phrase *wear a dress* before it is applied to the noun phrase in the same way as before. In that case, the result is the following:

(12) $\exists e[{}^{*}\mathrm{ag}(e) = \bigoplus \mathrm{girl} \ \wedge$
$e \in {}^{*}\lambda e'({}^{*}\mathrm{wear}(e') \ \wedge \ \mathrm{dress}({}^{*}\mathrm{th}(e')) \ \wedge \ \mathrm{PureAtom}({}^{*}\mathrm{ag}(e')))]$
(There is an event whose agents sum up to the girls, and this event consists of wearing events for each of which the agent is a pure atom and the theme is a dress.)

The star operator ${}^{*}\lambda e'$ is introduced through the D operator and takes scope over the predicate *dress* introduced by the theme. Unlike the previous formula, this representation does not require the theme of e to be a dress, though it requires e to consist of parts whose themes are dresses. This allows for the possibility that each girl wears a potentially different dress. The representation explicitly states that the dress-wearing events e' have pure atoms as agents, but not that these pure atoms are girls. However, this fact is entailed by cumulativity of thematic roles together with the assumption that the entities in the denotation of singular count nouns

are pure atoms. By cumulativity of thematic roles, any entity x which is the agent of one of the dress-wearing events e' is a part of the agent of e. This agent is the sum of all girls. By definition of sum (see Section 2.3.1), x overlaps with a part of this agent. Being atomic, x can only overlap with y if it is a part of y. This means that x is an atomic part of the girls. Since I assume that singular entities like girls are mereological atoms (Section 2.6.1), it follows that x is a girl. In this way, the distributive interpretation of (2a) is correctly captured.

Before moving on, let me compare my formulation to another way of reformulating the D operator, due to Lasersohn (1998). The focus of Lasersohn's paper is technical. He shows a way of generalizing Link's and other D operators so that they apply to other positions than the subject position both in both eventless and event-based frameworks. The following entry is a special case among these different combinations, namely the verb phrase-level version of an event-based version of Link's operator. It does not represent Lasersohn's entire proposal, but it is part of his proposal that is most closely related to mine.

(13)   **Distributivity operator over events (Lasersohn)**
       $[\![\text{D}]\!]$ (Lasersohn) $= \lambda P_{\langle e,vt \rangle} \lambda x \lambda e \forall y [y \leq_{Atom} x \rightarrow \exists e'[e' \leq e \ \wedge \ P(y)(e')]]$

This operator applies to a predicate of type $\langle e, vt \rangle$. It is based on the assumption that a verb phrase like *smile* that is about to combine with it is represented as something like $\lambda x \lambda e[\text{smile}(e) \ \wedge \ \text{ag}(e) = x]$. By combining with this type of predicate, the D operator compositionally acquires the information about which thematic role it modifies. In contrast, my proposal assumes that the operator is coindexed with the appropriate thematic role. My proposal therefore does not rely on the assumption that the D operator is immediately adjacent to its thematic role head, while Lasersohn's proposal does. Since there is no consensus on how thematic roles are introduced, both options appear viable, but I see a potential problem with Lasersohn's proposal in cases where the D operator appears to be remote from its target noun phrase. One such case is discussed in Schwarzschild (1996), p. 62, who attributes this observation to Angelika Kratzer. The observation involves a scopal interaction between the D operator and the modal predicate *likely*.

(14)   John and Mary are likely to win the lottery.

As Schwarzschild points out, (14) has the following two distributive readings:

(15)   a.   There is a good chance that John will win the lottery and that Mary
            will win the lottery.
       b.   John and Mary each have a good chance of winning the lottery.

Schwarzschild notes that this ambiguity can be explained by assuming that the scope of the D operator and of the modal predicate can alternate. This alternation is represented in the following two LFs, where [ag] stands for the little v thematic role head:

(16)   a.   [[John and Mary] [[ag] [D [are likely [PRO to win the lottery]]]]]
       b.   [[John and Mary] [[ag] [are likely [D [PRO to win the lottery]]]]]

No matter if one assumes that the thematic role *agent* (or whatever is the role that relates John and Mary to the lottery-winning event) is supplied high by [ag], or low by PRO, the D operator is separated from it by *be likely* in at least one of the two cases. For this reason, it seems advisable to assume that the relationship between D and the thematic role it modifies can be nonlocal. This is more compatible with my proposal than with Lasersohn's, because I represent this relationship through coindexation while Lasersohn represents it through function application.

Another difference between my proposal and Lasersohn's is the way in which the two operators access the events over which they distribute. My operator uses algebraic closure over events with atomic agents, while Lasersohn's operator uses universal quantification over individuals. The difference between the two formulations is apparent in the different translations that result from inserting a D operator into *The girls smiled* before existential closure applies.

(17)   a.   Lasersohn's representation:
            $\lambda e \forall y [y \leq_{Atom} \bigoplus \text{girl} \wedge \rightarrow \exists e'[e' \leq e \wedge \text{smile}(e') \wedge {}^*\text{ag}(e') = y]$
       b.   My representation:
            $\lambda e[{}^*\text{ag}(e) = \bigoplus \text{girl} \wedge e \in {}^*\lambda e'[\text{smile}(e') \wedge \text{PureAtom}({}^*\text{ag}(e'))]]$

Lasersohn's representation applies to all events that contain a smiling subevent for each girl, even if they also contain other subevents. My representation applies to all events that contain a smiling subevent for each girl and nothing else. To use a term from Bayer (1997), Lasersohn's translation suffers from leakage, as it does not prevent the events to which it applies from containing extraneous material.

Leakage causes problems in connection with adverbials such as *surprisingly* or *in slow procession*, provided that these adverbials are translated as event predicates. The following example illustrates the leakage problem (see also Schein (1993) for discussion). In a nutshell, this problem is that whenever Lasersohn's event predicate (17a) applies to an event e, it also applies to any event of which e is a part. Let L stand for (17a) and let M stand for my event predicate (17b). Imagine a scenario in which an event e_0 satisfies both Lasersohn's predicate L and my predicate M. That is, e_0 is an event in which the girls smile. Assume that e_1 is

an event in which the boys cry. Let e_2 be the sum of e_0 and e_1. Now e_2 does not satisfy my predicate M because it contains extraneous material. That is, its sum agent is not the girls, but the girls and the boys, and it is not a smiling event but the sum of a smiling and of a crying event. But e_2 does satisfy Lasersohn's predicate L, because by virtue of containing e_0, it contains a smiling event for every girl. Imagine that event e_0 is not surprising by itself, and that it is surprising that e_0 cooccurs with the event e_1 in which the boys cry. Imagine that the two events taken together are surprising, that is, e_2 is surprising. Since e_0 by itself is not surprising, sentence (18) is intuitively judged false. If the D operator is applied to *smile*, then on Lasersohn's account, this sentence is translated as (18a), while on my account it is translated as (18b). The problem is that e_2 satisfies both L (by leakage) and the predicate *surprising* (by assumption). Therefore, Lasersohn's D operator wrongly predicts that (18) is judged true.

(18)   Surprisingly, the girls smiled.

     a. $\exists e[\text{surprising}(e) \ \wedge \ L(e)]$
     b. $\exists e[\text{surprising}(e) \ \wedge \ M(e)]$

To sum up this section, we have seen that Lasersohn's implementation faces a leakage problem and requires that the D operator occur immediately adjacent to the thematic role with which it is associated. My implementation avoids these problems, and it makes the connection between the D operator and stratified reference clear: the operator is a way to minimally change the meaning of a predicate so that it satisfies stratified reference.

8.3   Atomic and nonatomic distributivity

Modeling Q-distributivity, and consequently modeling phrasal distributivity, requires quantifying over parts of a plural individual. Two variations on Link's operator have emerged in the literature, corresponding to two different views on distributivity. Winter (2001) labels these views atomic and non-atomic distributivity. What has been presented above is the atomic view. This view assumes that phrasal distributivity involves universal quantification over atomic parts of the plural individual, that is, over singular individuals. On this view, the distributive reading of a sentence like *The girls are wearing a dress* is equivalent to *The girls are each wearing a dress*. The indefinite *a dress* covaries with respect to a covert universal quantifier that ranges over individual girls. This view is defended in Lasersohn (1998, 1995), Link (1997), and Winter (2001), among others.

By contrast, the nonatomic view holds that phrasal distributivity involves universal quantification over certain parts of the plural individual, and that these

parts can be nonatomic. Variants of this view are defended in Gillon (1987, 1990), van der Does and Verkuyl (1995), Verkuyl and van der Does (1996), Schwarzschild (1996, ch. 5), Brisson (1998, 2003), and Malamud (2006a,b). This section presents and motivates the nonatomic view and shows how the implementation presented in the last section can be modified to take it into account.

The traditional argumentation for nonatomic view is based on sentences like the following, which is adapted from Gillon (1987):

(19)   Rodgers, Hammerstein, and Hart wrote musicals.

This sentence plays on a particular fact of American culture: neither did the three composers it mentions ever write any musical together, nor did any of them ever write one all by himself. However, Rodgers and Hammerstein wrote the musical *Oklahoma* together, and Rodgers and Hart wrote the musical *On your toes* together. On the basis of these facts, the sentence is judged as true in the actual world, although it is neither true on the collective interpretation nor on an "atomic distributive" interpretation.

As the traditional argument for the nonatomic view on phrasal distributivity goes, in order to generate the reading on which (19) is true, the predicate *wrote musicals* must be interpreted as applying to nonatomic parts of the sum individual to which the subject refers. This view is generally implemented with the concept of a cover (Gillon 1987), which was introduced in Chapter 5. In a set-based representation of plural individuals, covers are like partitions of a set except that their cells can overlap (20). The corresponding mereological notion is shown in (21).

(20)   **Definition: Cover (set-theoretic)**
        $\mathrm{Cov}(C, P) \stackrel{\text{def}}{=} \bigcup C = P \ \wedge \ \varnothing \notin C$
        (C is a cover of a set P if and only if C is a set of nonempty subsets of P whose union is P.)

(21)   **Definition: Cover (mereological)**
        $\mathrm{Cov}(C, x) \stackrel{\text{def}}{=} \bigoplus C = x$
        (C is a cover of a mereological object x is a set of parts of x whose sum is x.)

Following Schwarzschild (1996), cover-based approaches modify the distributivity operator by relaxing the "atomic part" condition and by quantifying over nonatomic parts of a cover of the plural individual. The first cover-based approaches assumed that the cover can be existentially quantified by the operator that introduces it. In an eventless setting, this assumption can be implemented by a distributivity operator such as (22). On this view, sentence (19) is translated as (23). This formula is verified in the actual world by the existence of the cover

$\{rodgers \oplus hammerstein, rodgers \oplus hart\}.$

(22) **Nonatomic distributivity operator, existentially bound cover**
$$[\![D_\exists]\!] = \lambda P_{\langle et \rangle} \lambda x \exists C[\mathrm{Cov}(C, x) \ \wedge \ \forall y[\mathrm{C}(y) \ \wedge \ y \leq x \rightarrow P(y)]]$$

(23) $\exists C[\mathrm{Cov}(C, rodgers \oplus hammerstein \oplus hart) \ \wedge$
$\forall y[\mathrm{C}(y) \ \wedge \ y \leq x \rightarrow y \in [\![\text{wrote musicals}]\!]]]$

Existentially bound covers are now generally considered untenable as a way of modeling phrasal distributivity because they overgenerate readings. These readings can be described as halfway between collective (or cumulative) and distributive readings, and they are commonly called intermediate readings. I will also sometimes call them nonatomic distributive readings. For example, Lasersohn (1989) points out that in a situation where John, Mary, and Bill are the teaching assistants and each of them was paid exactly $7,000 last year, sentences (24a) and (24b) are both true, as is expected on the atomic approach. The former is true on its distributive reading, and the latter on its collective or cumulative reading. But sentence (24c) is false, even though the cover $\{j \oplus m, m \oplus b\}$ would verify it if it was translated using the $D_\exists$ operator in (22). That is, sentence (24c) does not have an intermediate reading.

(24)   a.   The TAs were paid exactly $7,000 last year.   *distributive*
       b.   The TAs were paid exactly $21,000 last year.   *collective*
       c.   The TAs were paid exactly $14,000 last year.   **intermediate*

Giving up the existential cover-based operator $D_\exists$ in (22) explains why sentence (24c) is false, because without this operator, there is no way to generate an intermediate reading for this sentence. However, sentence (19) above does have an intermediate reading, so giving up $D_\exists$ requires an alternative account of this reading. Lasersohn (1989) proposes to do so by assuming the following meaning postulate:

(25) $\forall w, x, y, z[\mathrm{write}(w, x) \ \wedge \ \mathrm{write}(y, z) \rightarrow \mathrm{write}(w \oplus x, y \oplus z)]$

This meaning postulate is actually a special case of the lexical cumulativity assumption, for which there is ample independent support (see Section 2.7.2). Further support for adopting lexical cumulativity while rejecting existentially bound covers comes from the difference between *write musicals* and *write a musical*. As Link (1997) notes, the following sentence is false in the actual world, that is, it does not have the "intermediate" construal that (19) has.

(26)   Rodgers, Hammerstein and Hart wrote a musical.   (Link 1997)

Given an inclusive view of plurals (see Section 2.6.2), *write musicals* applies to entities who wrote one or more musicals. Such entities include the sum $rodgers \oplus hammerstein$ and the sum $rodgers \oplus hart$, and moreover (via (25)) the sum individual $rodgers \oplus hammerstein \oplus hart$, of which (25) entails that they wrote the sum individual $oklahoma \oplus on.your.toes$. This plural individual qualifies as *musicals*, but not as *a musical*, which explains the contrast between (19) and (26).[34]

For this explanation, it is crucial that a cover-based operator like (22) is not available in the grammar, because that operator would predict (26) to be true in the actual world. Lasersohn, as well as Winter (2001) and others, conclude from this and similar examples that the atomic approach to phrasal distributivity is superior to covers. However, Gillon (1990) and Schwarzschild (1996) identify a residue of cases in which a cover-based operator does seem necessary. These cases involve special contexts in which discourse makes a specific cover pragmatically salient. Lasersohn (1995) provides a particularly clear example. Shoes typically come in pairs, so a sentence like (27a) can be interpreted with respect to a cover whose cells each contain a matching pair of shoes. The relevant reading is an intermediate reading: it does not assert that each individual shoe costs fifty dollars, nor that all the shoes taken together cost that much, but that each pair of shoes does. By contrast, no such cover is salient for example (27b), which can only mean that each suitcase weighs fifty pounds or all of them together do so.

(27)   a.   The shoes cost fifty dollars.   (Lasersohn 1995)
        b.   The suitcases weigh fifty pounds.

In the intermediate reading of (27a), the quantifier introduced by the direct object takes scope under the distributivity operator. Therefore this reading cannot be modeled by lexical cumulativity alone, unlike the intermediate reading of sentence (19). Schwarzschild (1996) models the context dependency of this kind of intermediate reading by assuming that the D operator (which he renames Part) contains an anaphorically supplied cover, which I represent as a subscripted C:

(28)   **Schwarzschild's nonatomic distributivity operator, free cover**
        $[\![\text{Part}_C]\!] = \lambda P_{\langle et \rangle} \lambda x \forall y [\text{C}(y) \ \wedge \ y \leq x \rightarrow P(y)]$

Schwarzschild assumes that C is restricted through a pragmatic mechanism to be a cover over x ($\text{Cov}(C, x)$), but he prefers not to write this condition into his operator. Schwarzschild himself does not say much about the pragmatic mechanism that resolves C. See Malamud (2006a,b) for a proposal in which the D

---

[34]These assumptions do not explain why sentence (19) conveys that more than one musical in total is written. This fact can be explained by a theory of dependent plurals such as Zweig (2008, 2009), as discussed in Section 9.4.2.

operator is anaphoric on a decision problem in the sense of van Rooij (2003) rather than on an unanalyzed predicate as in Schwarzschild's account. These approaches make different predictions, but the important point that is common to them is that the D operator poses a stronger restriction on C than just existentially quantifying over it. This restriction rules out intermediate readings in sentences like (24c), (26), and (27b) but not in sentences like (27a). While I see no obstacles to using decision problems, I continue to use a Schwarzschild-style approach to keep the representation simple.

As in the case of Link's operator discussed above, Schwarzschild's operator applies to individuals rather than events and must be modified in order to be applicable in a Neo-Davidsonian framework. I discuss ways to do so below.

Sentence (27a) is structurally equivalent to sentences (26), (24b), and (27b), yet only (27a) has an intermediate or "cover-based" reading. As Heim (1994) and Schwarzschild (1996) argue, this fact provides strong evidence that models of (phrasal) distributivity need to contain a pragmatic factor. Note that the operator in (28) is more restricted than the existential cover-based operator $D_\exists$ in (22) because (28) presupposes the existence of a context through which the variable C can be resolved. The contrast between (27a), which has an intermediate reading, and (24c), (26) and (27b), which do not, is predicted on the assumption that a salient context is only available for (27a).

To summarize the empirical picture presented in this section, nonatomic distributivity is readily available at the level of the verb (lexical level), but at the level of the verb phrase (phrasal level) it only occurs when context supplies a pragmatically salient cover. Atomic distributivity is available both at the lexical level and at the phrasal level. Summarizing the insights of the previous literature, I assume that this pattern is explained as follows (see Table 8.1). The lexical cumulativity assumption accounts for the availability of atomic and nonatomic distributivity at the lexical level. Link's atomic D operator is always available at the level of the verb phrase. Schwarzschild's cover-based D operator is also available at the level of the verb phrase, but it is only available if context supplies a salient cover. When this cover contains only one atomic individual in every cell, Schwarzschild's D operator behaves like Link's operator.

Having distinct D operators might seem redundant. Instead, we can of course assume that there is only one D operator, namely Schwarzschild's, and that covers over atomic individuals are salient in every situation. This view amounts to the following idea: in an atomic domain, the atomic level always provides a salient cover in every context, and this explains the strong preference that speakers have for atomic-level distributivity. In effect, Link's atomic operator is made available again through the back door. As far as I can tell, both possibilities are empirically equivalent. However, the one-operator possibility makes it possible to draw an

Figure 8.1: V level versus verb phrase level distributivity in atomic domains

	(a) Empirical generalization			(b) Explanation	
	lexical (V level)	phrasal (VP level)		lexical (V level)	phrasal (VP level)
atomic	available	available		lexical cum.	Atomic D op.
nonatomic	available	only w. context		lexical cum.	Cover-based D op.

interesting parallel between the D operator and stratified reference. To model nonatomic distributivity, I make a simple change to the event-based atomic D operator defined in (9), repeated below as (29). The change consists in replacing the predicate *PureAtom* by a free predicate, which I call C as a reminder that it plays the same role as the C predicate in Schwarzschild's operator (28). The change is highlighted in (29) and (30).

(29)   **Definition: Atomic event-based D operator**

$$[\![\mathrm{D}_\theta]\!] \stackrel{\text{def}}{=} \lambda P_{\langle vt \rangle} \lambda e [e \in {}^*\lambda e' \left(\begin{array}{l} P(e') \wedge \\ \mathbf{PureAtom}(\theta(e')) \end{array} \right)] \qquad = (9)$$

(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs are **pure atoms**.)

(30)   **Definition: Generalized event-based D operator**

$$[\![\mathrm{D}_{\theta,\mathbf{C}}]\!] \stackrel{\text{def}}{=} \lambda P_{\langle vt \rangle} \lambda e [e \in {}^*\lambda e' \left(\begin{array}{l} P(e') \wedge \\ \mathbf{C}(\theta(e')) \end{array} \right)]$$

(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs **satisfy the predicate C**.)

The generalized D operator has two parameters: the dimension parameter θ, which specifies the thematic role, and the granularity parameter C, which specifies the size of the entities over which the predicate is distributed. This formulation is exactly parallel to the formulation of stratified reference in Chapter 4. Unlike Schwarzschild, I do not rely on pragmatics to ensure that C actually covers the θs of the event to which the output of the D operator is applied. The only task of pragmatics in my account is to supply an antecedent predicate for C. By virtue of appearing in the scope of a star operator, any predicate to which C is resolved is

guaranteed to cover the θs of this event. See Section 5.4.2 for discussion.

The move from atomic to nonatomic distributivity parallels the move from stratified distributive reference to stratified reference in Section 4.6. In both cases, the granularity is changed from a fixed value to a parameter. However, there is also a difference. Following Schwarzschild (1996), I assume that the C parameter of the D operator in (30) can only be set in one of two ways: either it is set to the predicate *PureAtom* or to an anaphorically salient level of granularity. This anaphoric dependence is not present in stratified reference, where I have assumed that the granularity parameter is instead provided by a function ε which is vague but *not* anaphoric. In other words, I follow Schwarzschild (1996) in assuming that the generalized D operator can only acquire a nonatomic granularity in contexts that support a salient value of C, but I do not assume that the ε function on which stratified reference relies requires a special context in order to return a value. On the view developed in the previous chapters, stratified reference is present in *for*-adverbials and pseudopartitives, and they routinely instantiate its granularity parameter with nonatomic values even if context provides no salient level of granularity.

In contexts that do provide a salient level of granularity, the generalized D operator can pick it up anaphorically. The following theorem, a generalization of Theorem (10), leads us to expect the generalized D operator should function as a repair strategy whenever a *for*-adverbial imposes a stratified reference presupposition. More specifically, we expect this repair strategy to succeed if and only if the salient level of granularity is at least as coarse as the granularity parameter of the *for*-adverbial. The proof of this theorem can be found in the Appendix.

(31)   **Theorem: $D_{\theta,C}$ is a repair strategy**
       $\forall P \forall \theta \forall C \forall C'[C \subseteq C' \rightarrow \text{SR}_{\theta,C'}(D_{\theta,C}(P))]$
       (When the D operator coindexed with thematic role θ and with granularity threshold C is applied to any predicate, the result always has stratified reference with respect to θ and C' for any threshold C' that is at least as coarse as C.)

The search for clear cases of nonatomic distributivity has been going on since at least Link (1987b), who called them "genuine plural quantification". One of the factors that have made it so hard to identify clear cases of nonatomic distributivity is the focus in the literature on predicates that modify subjects with count nouns. On the standard assumption that the denotations of count nouns are taken from an atomic domain (see Section 2.6.1), phrasal distributivity over atoms is naturally expected to be more salient than nonatomic distributivity in almost all contexts and will obscure the presence of nonatomic readings. The idea that atomic granularity is

more salient than nonatomic granularity is already present in Schwarzschild (1996). I now sketch an explanation of this idea in terms of Kennedy (2007)'s principle of Interpretive Economy, which can in turn be derived from first principles in an evolutionary game-theoretic setting (Potts 2008). The central idea is that whenever possible, speakers will converge on certain focal points because this maximizes successful communication. Interpretive Economy was originally proposed to explain why speakers converge on interpreting scalar items like *tall* and *full* as referring to endpoints of a scale whenever such endpoints exist, and resort to context-dependent values only when this is not the case. We can interpret mereological parthood as inducing such a scale. For example, in a *for*-adverbial like *for two hours*, this scale ranges from arbitrarily short amounts of time to the point described by two hours. Since CEM does not contain a bottom element, this scale does not contain a lowest point, that is, there is no time interval with zero length. In count domains, however, the scale induced by the parthood relation is lower closed because singular individuals are atomic. Interpretive Economy suggests that speakers who use the D operator and who need to agree on how to interpret its granularity parameter converge on atomicity as a focal point, except in contexts where another granularity value is salient. By looking at noncount domains such as time, we remove atomic granularity as a potential focal point, so any phrasal distributivity effects we find must be cases of nonatomic distributivity.

In the next section, I look at the domain of time through the lens of *for*-adverbials. In a nonatomic domain like time, there are necessarily no atomic covers, so the first row of Table 8.1 is not applicable. I will argue that the second row of Table 8.1 is mirrored precisely in the temporal domain. That is, the dimension parameter of the D operator in (30) can be instantiated to τ (runtime) and, in that case, its granularity parameter is dependent on an anaphorically salient level of granularity. To motivate this novel application of the D operator, it is necessary to consider the scopal behavior of *for*-adverbials with respect to Shares that contain an overt quantifier. This is the topic of the next section.

8.4 The scopal behavior of *for*-adverbials

For the purpose of this section only, assume counterfactually that a *for*-adverbial is translated as a universal quantifier over instants, as if it was the temporal counterpart of *every*. I call this the baseline analysis:

(32) ⟦for an hour⟧ (baseline) =
     $\lambda P_{\langle vt \rangle} \exists t[\text{hours}(t) = 1 \ \wedge \ \forall t'[t' <_{Atom} t \rightarrow \exists e[P(e) \ \wedge \ \tau(e) = t')]]$

This assumption allows us to see why the scopal behavior of *for*-adverbials is surprising when compared with the familiar scopal behavior of the universal quantifier *every*. In contrast to *every*, which can take scope anywhere in its clause, we will see that *for*-adverbials always seem to take narrow semantic scope with respect to quantifiers in their syntactic scope, except in a specific and limited set of cases. I present these cases in Section 8.5, where I argue that they represent nonatomic phrasal distributivity. A recent extensive discussion of this observation is found in Kratzer (2007), though the relevant observations are found in earlier work as well (e.g. Carlson 1977; Zucchi and White 2001; van Geenhoven 2004).

The following examples, adapted from Kratzer (2007), illustrate the behavior of *for*-adverbials with respect to indefinites. Even though both scopal orders would be *a priori* plausible, the indefinite can only be interpreted with wide scope. For example, it would be plausible for (33b) to have an interpretation like *Over and over again over the course of five minutes, I dialed a different wrong phone number.* But this kind of interpretation is systematically absent from the sentences below.

(33)   a.   John pushed a cart for an hour. $\exists > \forall; {}^*\forall > \exists$
       b.   I dialed a wrong phone number for five minutes. $\exists > \forall; {}^*\forall > \exists$
       c.   She bounced a ball for 20 minutes. $\exists > \forall; {}^*\forall > \exists$
       d.   He kicked a wall for a couple of hours. $\exists > \forall; {}^*\forall > \exists$
       e.   She opened and closed a drawer for half an hour. $\exists > \forall; {}^*\forall > \exists$
       f.   I petted a rabbit for two hours. $\exists > \forall; {}^*\forall > \exists$

As Kratzer notes, the same phenomenon also holds in German, where the two following sentences have the same interpretation, namely one where the indefinite has wide scope with respect to the *for*-adverbial. This is remarkable since quantifier scope in German normally follows surface order. These two sentences are both translations of (33b), but in one case, the indefinite has been scrambled before the *for* adverbial.

(34)   a.   *Ich hab' fünf Minuten lang eine falsche Telefonnummer   gewählt.*
            I   have five minutes long a   wrong telephone.number dialed.
       b.   *Ich hab' eine falsche Telefonnummer   fünf Minuten lang gewählt.*
            I   have a   wrong telephone.number five minutes long dialed.

Even in cases where the wide scope interpretation of the indefinite is pragmatically odd and much less plausible than the narrow scope interpretation, it is still the only one available:

(35) ??John found a flea on his dog for a month.   (Zucchi and White 2001)

Sentence (35) is unambiguous: only one flea is found over and over again. If *for a month* was translated as a universal quantifier over instants, the indefinite *a flea* would have to be interpreted with wide scope over this quantifier in order to account for this behavior, even though the resulting interpretation is pragmatically odd because it is unusual to find the same flea repeatedly. The narrow scope interpretation, where the fleas can covary with the times, would be much more plausible, but it is not available out of the blue.

The examples so far have shown that singular indefinites must outscope *for*-adverbials. The same behavior can be observed if we replace the singular indefinite with certain other types of quantifiers, such as plural indefinites. For example, the following sentence cannot be interpreted as saying that over the course of three hours, John saw different sets of thirty zebras.[35] The only available interpretation is the one in which the plural indefinite *thirty zebras* takes wide scope.

(36)   John saw thirty zebras for three hours. $30 > \forall; {}^{*}\forall > 30$

Not only singular and plural indefinites must take wide scope over a *for*-adverbial (Zucchi and White 2001). In fact, the only items that seem to be able to take narrow scope with respect to a *for*-adverbial are bare plurals and bare mass nouns.

The scopal behavior of bare NPs with respect to *for*-adverbials mirrors their well-known tendency to take narrow scope in general (Carlson 1977). All VP-level and sentential predicates with bare NPs appear to be compatible with *for*-adverbials (Verkuyl 1972; Dowty 1979). The following sentences, taken from Dowty (1979), stand in marked contrast to the examples above, because there is no sense in which the bare NP has to be interpreted with wide scope. For example, (37a) is compatible with the plausible interpretation in which John finds different fleas on his dog and finds each of them only once.

(37)   a.   John discovered fleas on his dog for six weeks.
        b.   John discovered crabgrass in his yard for six weeks.

(38)   a.   Tourists discovered that quaint little village for years.
        b.   Water leaked through John's ceiling for six months.

We have seen that the ability of *for*-adverbials to give rise to quantifier scope ambiguities is much more limited than we would expect on the baseline analysis that translates *for*-adverbials as universal quantifiers over instants. Of course, there is no reason to actually assume that *for*-adverbials quantify over instants. As we

---

[35]A scopeless, cumulative reading which would correspond to the interpretation of *John saw thirty zebras in three hours* is not available either. I return to this point in Section 9.4.1.

will see in the next section, my own proposal as well as a number of previous proposals correctly explain the scopal behavior of *for*-adverbials with respect to indefinites and bare NPs. These proposals do not assume that *for*-adverbials quantify over instants.

Compare the baseline analysis, repeated here from (32), with my proposal in (40), repeated from Section 4.7:

(39) 〚for an hour〛 (baseline)
$= \lambda P_{\langle vt \rangle} \exists t[\text{hours}(t) = 1 \ \wedge \ \forall t'[t' <_{Atom} t \rightarrow \exists e[P(e) \ \wedge \ \tau(e) = t')]]]$

(40) 〚for an hour〛 (my proposal)
$= \lambda P_{\langle vt \rangle} \lambda e : \text{SR}_{\tau, \varepsilon(\lambda t[\text{hours}(t)=1])}(P).$
$P(e) \ \wedge \ \text{hours}(\tau(e)) = 1$

We have seen that indefinites take wide scope over *for*-adverbials. The baseline analysis of *for*-adverbials in (39) does not account for this behavior because it makes them universal quantifiers. The indefinite should be able to take narrow scope by remaining in situ. The translation in (40) states that the predicate P holds of the event e. (Ignore its presupposition for now.) Given our background assumptions, this immediately predicts that the indefinite in (33a), repeated here as (41), must take wide scope.

(41) John pushed a cart for an hour.   = (33a)

This prediction is obtained based on the following translation of the verb phrase *push a cart* (see Section 2.8):

(42) 〚push a cart〛 = $\lambda e[^{*}\text{push}(e) \ \wedge \ \text{cart}(^{*}\text{th}(e))]$
(True of any pushing event or sum of pushing events whose theme is one and the same cart.)

Importantly, even though the verbal denotation *push* is pluralized in this translation, the predicate *cart* is not pluralized. That is, the verb phrase only applies to events whose theme is exactly one cart, even if these events may be sums of events. This feature of Landman (2000)'s system is independently justified in his account of scopeless readings (see Section 2.8). In connection with the entry (40), it predicts that the entire event over which sentence (41) existentially quantifies must have a single cart as its theme:

(43) 〚John pushed a cart for an hour〛
$= \exists e : \text{SR}_{\tau, \varepsilon(\lambda t[\text{hours}(t)=1])}(\lambda e'[^{*}\text{push}(e') \ \wedge \ \text{cart}(^{*}\text{th}(e'))]).$
$[^{*}\text{ag}(e) = j \ \wedge \ ^{*}\text{push}(e) \ \wedge \ \text{cart}(^{*}\text{th}(e)) \ \wedge \ \text{hours}(\tau(e)) = 1]$
(There is a pushing event or sum of pushing events whose theme is one

cart, whose agent is John, and whose runtime measures one hour.)

As Kratzer (2007) points out, since we have adopted the background assumption of lexical cumulativity, we can also account for the behavior of achievement verbs like *find* even though these verbs are normally understood to have very short runtimes. For example, a sentence like (44) (repeated from (35) above) is now predicted to entail that there was a finding event e which lasted a month and whose theme is a flea. Lexical cumulativity allows this finding event to be plural. Since individual finding events have very short times, the finding event must indeed be plural (that is, repetitive) in order to be able to last a month. Importantly, the assumption of lexical cumulativity allows phrasal predicates like *find a flea* to involve reference to plural events only to the extent that the verb predicate (*find* in this case) already does so. The object *a flea* is not affected by pluralization and continues to involve reference to a singular flea. This means that sentence (44) requires a *single* flea to have been found repetitively over the course of a month.

(44)   [[John found a flea for a month]]
$= \exists e : \mathrm{SR}_{\tau, \varepsilon(\lambda t[\mathrm{months}(t)=1])}(\lambda e'[{}^*\mathrm{find}(e') \ \wedge \ \mathrm{flea}({}^*\mathrm{th}(e'))])$.
$[{}^*\mathrm{ag}(e) = j \ \wedge \ {}^*\mathrm{find}(e) \ \wedge \ \mathrm{flea}({}^*\mathrm{th}(e)) \ \wedge \ \mathrm{months}(\tau(e)) = 1]$
(There is a finding event or sum of finding events whose theme is one flea, whose agent is John, and whose runtime measures one month.)

As discussed in Sections 2.5.2 and 2.5.4, I do not require the runtimes of events to be continuous intervals, and I assume that the unit function *months* maps a discontinuous interval to the same number as the smallest continuous interval that contains it. This means that the representation in (44) does not require John to have been searching uninterruptedly at every moment of the month.

The present approach does not predict that all quantifiers take wide scope over *for*-adverbials. We have seen above that bare NPs do not they take wide scope with respect to *for*-adverbials, because they denote algebraically closed predicates. As described in Section 2.6.2, I assume that a bare plural like *fleas* has an inclusive meaning. That is, it is translated as *flea, the algebraic closure of its singular form, and its literal meaning is essentially *one or more fleas*. As described in Section 2.6.5, a mass noun like *crabgrass* is assumed to be its own algebraic closure, that is, *crabgrass = crabgrass. Lexical cumulativity has the effect that the bare plural in a predicate like *find fleas* stands in a cumulative-like relation to each of the subintervals over which the *for*-adverbial quantifies. For example, sentence (45) does not entail that any one flea has been found several times, but only that there is a plural hour-long interval over the course of which one or more fleas were found. This is entailed in the translation of (45), where *fleas* is translated in situ as a predicate that applies to the theme of the verb.

(45) ⟦John found fleas for a month⟧
$= \exists e : \mathrm{SR}_{\tau, \varepsilon(\lambda t[\mathrm{months}(t)=1])}(\lambda e'[^*\mathrm{find}(e') \ \wedge \ ^*\mathrm{flea}(^*\mathrm{th}(e'))])$.
$[^*\mathrm{ag}(e) = j \ \wedge \ ^*\mathrm{find}(e) \ \wedge \ ^*\mathrm{flea}(^*\mathrm{th}(e)) \ \wedge \ \mathrm{months}(\tau(e)) = 1]$
(There is a finding event or sum of finding events whose theme is a set of fleas, whose agent is John, and whose runtime measures one month.)

The stratified reference presupposition of the *for*-adverbial is fulfilled in this case. It requires that each plural finding event whose theme is a set of fleas and whose runtime is a month consists of finding events whose themes are also sets of fleas and whose runtime is very small. This is true for the following reason. *Find* is a punctual verb: any event in its denotation is a sum of finding events that have nearly instantaneous duration. By lexical cumulativity, an hour-long event e in the denotation of *find fleas* therefore consists of very short finding events. By cumulativity of thematic roles (Section 2.5.1), the themes of these events must be parts of the theme of e. These themes must themselves be individual fleas or sums of fleas, because *fleas* is the algebraic closure of the singular form of *flea*, and because the individual fleas in *flea* do not have any parts (Section 2.6.1).

8.5 Nonatomic distributivity in *for*-adverbials

Section 8.4 presented the generalization that indefinites must take wide scope over *for*-adverbials, and gave an account of this observation. However, based on the theoretical reasoning in Section 8.3, we expect there to be exceptions in special contexts that provide a salient level of granularity. This section presents such exceptions and proposes an explanation in terms of a nonatomic distributivity operator.

Example (46), repeated in adapted form from Section 8.4, illustrates again the generalization that an indefinite must "take scope" over the *for*-adverbial, even if the resulting interpretation is nonsensical. In this case, the interpretation is that there are two fleas that John found again and again on his dog for a month.

(46) ??John found two fleas on his dog for a month.

However, given the appropriate context (described below), examples (47) and (48) directly contradict this generalization. Example (47) is based on observations in Moltmann (1991), and example (48) is adapted from Landman and Rothstein (2009), where it is credited to Rothstein (2004).

(47) The patient took two pills for a month and then went back to one pill.

(48) This bicycle carried three children around Amsterdam for twenty years.

Example (47) sounds odd out of the blue, but as Moltmann (1991) observes, it is acceptable in a context where the patient's daily intake is discussed. Importantly, in this context, it does not require the same two pills to be taken again and again. For their example (48), Landman and Rothstein provide a supporting context as well: the bicycle is designed to carry around three children at a time, and over a period of twenty years it was used by different owners to carry different sets of three children around. In this context, the bicycle does not require the children to be the same across the twenty-year period.

Taken together, the three examples (46), (47), and (48) look puzzling, as there is no semantic criterion that distinguishes them from one another. In all three cases, the "wide scope" interpretation of the indefinite is pragmatically disfavored. It is highly unlikely to find the same two fleas repeatedly, or to take the same two pills repeatedly, and it is impossible for a set of children to do anything over twenty years because they would grow up and no longer be children.

However, there is a pragmatic criterion that distinguishes (47) and (48) from (46): the availability of a supporting context. We have observed the effect of context in Section 8.3, when we considered the following examples:

(49)   a.   The shoes cost fifty dollars.   = (27a)
       b.   The suitcases weigh fifty pounds.   = (27b)

Sentence (49a) has a distributive reading (each shoe costs fifty dollars), a collective reading (all the shoes together cost fifty dollars) and an intermediate reading (each pair of shoes costs fifty dollars). Sentence (49b), uttered out of the blue, lacks the intermediate reading. Following Schwarzschild (1996), this kind of contrast led us to the conclusion that atomic phrasal distributivity is always available, but nonatomic phrasal distributivity is only available when context provides a salient cover. We have furthermore adopted the insight from Lasersohn (1989) and Winter (2001) that lexical distributivity should be separated from phrasal distributivity.

The conceptual bridge between aspect and distributivity developed in this work allows us to borrow these insights for the present purpose. The idea is that the mechanism which allows context to rescue examples like (47) and (48) involves a distributivity operator over times which is anaphoric to a a salient level of granularity, analogously to the cover-based approach that Schwarzschild (1996) assumes for distributivity over count domains.

In Section 8.2, I developed a D operator which, like stratified reference, is relativized to two parameters: a thematic role, which was always set to *ag* for the examples we considered, and a level of granularity, which was assumed to be either atomic or provided by context. Given the parallel between thematic

roles and trace functions that this work pursues, we expect that the thematic role parameter can also be set to τ, or runtime. The idea is the following: in an atomic domain, the atomic level always provides a salient cover in every context, and this explains the strong preference that speakers have for atomic-level distributivity. In a domain like time, there are no atomic covers; or if there are atoms (instants of time), they are not directly accessible to natural language semantics (von Stechow 2009). If this is correct, we expect that setting the thematic role parameter to τ should be incompatible with setting the granularity parameter to *Atom*. Or, to put it differently, there is no atomic-level distributivity operator for time, just a nonatomic one which is subject to the same restrictions as its agent-based counterpart.

The following entry is repeated from (30), with the dimension parameter instantiated as τ:

(50) $[\![D_{\tau,C}]\!] = \lambda P_{\langle vt \rangle} \lambda e [e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ C(\tau(e')) \end{array} \right)]$

     (Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose runtimes satisfy the predicate C.)

Following Schwarzschild (1996), I assume that the predicate C in this operator is anaphoric on a contextually salient level of granularity.

The following fact is a special case of Theorem (31). It tells us that the temporal instantiation of the nonatomic D operator can be used as a repair strategy to satisfy the stratified reference presupposition a *for*-advebial:

(51)   **Fact: $D_{\tau,C}$ is a repair strategy**
     $\forall P \forall C \forall C' [C \subseteq C' \rightarrow \mathrm{SR}_{\tau,C'}(D_{\tau,C}(P))]$
     (When the granularity parameter of the temporal D operator is set to a given predicate C, then for any C' such that $C \subseteq C'$, the output of D satisfies stratified reference with respect to dimension τ and granularity C'.)

We are now ready to explain the difference between examples (46) and (47), repeated here:

(52) ??John found two fleas on his dog for a month.

(53)   The patient took two pills for a month.

Consider first example (53). The verb phrase *take two pills* is translated as $\lambda e [{}^*\text{take}(e) \wedge {}^*\text{pill}({}^*\text{th}(e)) \wedge |{}^*\text{th}(e)| = 2]$. This verb phrase applies to events in which a total of two pills are taken. It cannot combine directly with the *for-*

adverbial because in order to do so, it would have to satisfy the presupposition of *for a month*, namely $\mathrm{SR}_{\tau,\varepsilon(\lambda t[\mathrm{months}(t)=1])}$. Proof: In order for *take two pills* to have this property, any event in which two pills are taken should consist of one or more events in which two pills are taken and whose runtime is very short compared with one month. This is not the case. For example, a discontinuous sum event in which one pill is taken on Dec 1 and another pill is taken on Dec 31st has a runtime of a month (on the generous construal of the unit function *month*, see Section 2.5.4), but the only part of this event in which two pills are taken is the event itself. Its runtime is one month and therefore not very short compared with one month. End of proof. The same reasoning applies to example (52).

Given that *take two pills* and *find two fleas* do not satisfy the presupposition of *for a month*, we know from Fact (51) that the $\mathrm{D}_{\tau,C}$ operator can be used as a repair strategy, provided that the following two conditions hold: First, C can be resolved to a salient level of granularity. As mentioned, (53) is not acceptable out of the blue, but only if uttered in a context in which the level of granularity $\lambda t[\mathrm{days}(t) \leq 1]$ ("once a day") is salient, while there does not seem to be an easily available context that would provide a salient level of granularity for (52). Second, it must be the case that $\lambda t[\mathrm{days}(t) \leq 1] \subseteq \varepsilon(\lambda t[\mathrm{months}(t) = 1])$. That is, any interval that qualifies as "once a day" must also qualify as very short with respect to one month. While the definition of ε contains a vague component, I think this condition is plausibly fulfilled. It follows that (53) can be interpreted by applying $\mathrm{D}_{\tau,\mathrm{days}(t)\leq 1}$ to its verb phrase. The result of this operation is as follows:

(54) $\mathrm{D}_{\tau,\mathrm{days}(t)\leq 1}(\lambda e[^*\mathrm{take}(e) \ \wedge \ ^*\mathrm{pill}(^*\mathrm{th}(e)) \ \wedge \ |^*\mathrm{th}(e)| = 2])$

$= \lambda e[e \in {}^*\lambda e' \begin{pmatrix} {}^*\mathrm{take}(e') \ \wedge \ ^*\mathrm{pill}(^*\mathrm{th}(e')) \ \wedge \ |^*\mathrm{th}(e')| = 2] \ \wedge \\ \mathrm{days}(\tau(e')) \leq 1 \end{pmatrix}]$

(True of any plural event that consists of one or more events of taking two pills which each take place within a day.)

This predicate can now be combined with the *for*-adverbial *for a month* and with the agent noun phrase *the patient* as described in Sections 2.10 and 4.7. The result of this computation is this:

(55) $\exists e[^*\mathrm{ag}(e) = \mathrm{the.patient} \ \wedge \ \mathrm{months}(\tau(e)) = 1 \ \wedge$

$e \in {}^*\lambda e' \begin{pmatrix} {}^*\mathrm{take}(e') \ \wedge \ ^*\mathrm{pill}(^*\mathrm{th}(e')) \ \wedge \ |^*\mathrm{th}(e')| = 2] \ \wedge \\ \mathrm{days}(\tau(e')) \leq 1 \end{pmatrix}]$

(There is a plural event that consists of one or more events of taking two pills which each take place within a day. Its agent is the patient, and its runtime measures a month.)

Given the "generous" interpretation of *months*, this formula is verified by a

plural event with a discontinuous runtime, provided that there is a month between the beginning and the end of this event. Its truth conditions appear to correctly represent the way in which sentence (53) is interpreted.

8.6 Summary

This chapter has integrated the atomic D operator of Link (1987b) and the cover-based nonatomic D operator of Schwarzschild (1996) into the present framework. The formalization I chose makes the D operator complementary to stratified reference, in the sense that a predicate to which the D operator applies always has stratified reference. I have assumed that, like stratified reference, the D operator has two parameters: dimension and granularity. I have distinguished between lexical and phrasal distributivity. Following Lasersohn (1989) and Kratzer (2007), I have assumed that lexical distributivity is due to lexical cumulativity. Following Winter (2001), lexical distributivity must be distinguished from phrasal distributivity when it comes to finding cases of nonatomic distributivity. After reviewing the literature, I have concluded that there are some rare residual cases of nonatomic phrasal distributivity, and that these cases motivate the possibility of a nonatomic granularity parameter setting for the D operator. Following Schwarzschild (1996), I have assumed that this happens so rarely because the operator is anaphoric on its context with respect to this parameter. Examples like (56a) have a nonatomic distributive reading because context provides a salient level of granularity, while examples like (56b) do not have such a reading. I have extended this parallel to the temporal domain, where I have argued that a salient level of granularity provides a way for the indefinite (57a) to "take scope" under the *for*-adverbial, while such a reading is not present in (57b).

(56) a. The shoes cost fifty dollars.   = (27a)
      b. The suitcases weigh fifty pounds.   =(27b)

(57) a. The patient took two pills for a month and then went back to one pill.
                                                                          =(47)
      b. ??John found two fleas on his dog for a month.   = (46)

Chapter 9

For, *each*, and *all*

9.1   Introduction

The last chapter has addressed a problem involving *for*-adverbials by transfering an insight from distributivity to aspect. As we will see in this chapter, this flow of information can also be reversed: I will take an insight from the literature on aspect and apply it to a set of problems involving distributive quantification. The translation of *for*-adverbials presented in the previous chapters is shown to shed new light on a set of problems concerning the interaction of *all* with cumulative quantification and dependent plurals. To understand this interaction, *all* needs to be seen as a distributive item, which raises questions as to how and why it differs from the distributive item *each*.

9.2   *All* vs. *every/each*

The first set of problems addressed in this chapter concerns the semantic behavior of plural prenominal and adverbial *all*. The status of this word is a challenge for any theory of distributivity. *All* behaves similarly to distributive items like *each* in some respects, but in other respects it behaves decidedly differently from them. A theory of distributivity must be flexible enough to be applicable to *each* as well as to *all*.

In this chapter, I argue that *each* and *all* are both distributive items, and I account for their limited ability to take part in cumulative readings. The differences between *each* and *all* are explained by assuming that the former distributes over pure atoms while the latter distributes over both pure and impure atoms. The granularity parameter of stratified reference provides the means to specify this difference.

Historically, the first argument for *all* as a distributive quantifier came from the difference in interpretation between *all* and definite plurals. This difference has been called the "maximizing effect" of *all* (Dowty 1987) and, conversely, the "nonmaximality" of definite plurals (Brisson 1998). I will call it the argument from maximality. Consider for example the following minimal pair from Link (1983):

(1)   a.   The children built the raft.   *nonmaximal*
      b.   All the children built the raft.   *maximal*

As Link notes, "in [(1b)] it is claimed that every child took part in the action whereas in [(1a)] it is only said that the children somehow managed to build the raft collectively without presupposing an active role in the action for every single child." In other words, (1a) tolerates exceptions in a way that (1b) does not. Link proposes to account for this difference in meaning by giving *all* a translation as a universal quantifier which distributes the property of taking part in building the raft over all individual parts of the totality of children. The difference between (1a) and (1b) is accounted for by assuming that the translation of (1a) does not contain this universal quantifier. On this style of analysis, *all the children* but not *the children* involves distributivity. Link himself only discusses *all* as a side topic in this paper, but his idea underlies the influential analysis of Dowty (1987).

To avoid confusion in the following, let me point out right away that I do not consider the argument from maximality compelling. I present it here because it has been historically important, as it led Link (1983) and following him Dowty (1987) to treat *all* as a distributive quantifier. While I believe that *all* is a distributive item, I do not believe that maximality establishes this fact. The main reason I do not find the argument compelling lies in an observation by Brisson (1998): the maximality-nonmaximality opposition does not correlate with the distributive-collective opposition. Sentences (2a) and (2b) both have a collective reading, on which one raft was built in a coordinated action, and they also both have a distributive reading, on which each raft was built by a boy. Sentence (2a) tolerates exceptions on both readings, but sentence (2b) presupposes that every boy became involved on both the distributive and on the collective reading.

(2)   a.   The children built a raft.   ✓ *distributive nonmax.*, ✓ *collective nonmax.*
      b.   All the children built a raft.   ✓ *distributive max.*, ✓ *collective max.*

As this example shows, the "maximizing effect" of *all* is always present, even when the sentence in which it occurs is interpreted collectively. Likewise, the "nonmaximality" of definite plurals is always present, even when the sentence is interpreted distributively. Therefore, maximality is neither a sufficient nor a necessary condition for distributivity, and it seems doubtful to me to conclude on

187

this basis alone that *all* is a distributive item.

I will have nothing to say about the maximizing effect of *all the boys* and the nonmaximality of *the boys*. The translations I will give for these noun phrases ignore the effect. In this respect, my theory contrasts with accounts by Dowty (1987), and Brisson (2003), which use maximality as a way to explain other properties of *all*, such as its inability to license collective readings with numerous-type predicates (as discussed in the next section). These accounts face the problem that *all* shares these properties with other quantifiers in which a maximality effect does not obtain at first sight. According to Winter (2001), *all* patterns with *most of the*, *exactly four*, *at least twelve*, *many*, *few*, and other plural strong quantifiers in terms of its incompatibility with numerous-type predicates. It is difficult to see how an item like *few* or even *most* could be claimed to involve a maximality effect in any meaningful way.[36] For this reason, I do not use maximality to explain the properties of *all* which are of interest here. A full comparison of the present account with Dowty (1987) and Brisson (2003) will have to wait for another occasion, as will an extension to *most of the* and to other strong quantifiers. While each of the claims I make about *all* in the following seems to extend to *most of the*, other strong quantifiers such as *exactly three* pattern differently, for example because they can take part in cumulative readings (Brasoveanu 2010).

9.2.1 Numerous-type predicates

Having discarded maximality effects, I now discuss what I believe to be a convincing argument that *all* is a distributive item. This argument comes from a class of predicates with respect to which *all* behaves analogously to the uncontroversially distributive items *every* and *each*. This class was introduced in Section 4.2.3 under the name of *numerous-type predicates*. While these predicates can give rise to collective interpretations together with definite plurals and other noun phrases, they cannot be interpreted collectively when *all* or *every/each* are present.

This category has also been called *purely collective predicates*, *pure cardinality predicates* (Dowty 1987), and *genuine collective predicates* (Hackl 2002). It roughly corresponds to the *atom predicates* of Winter (2001, 2002), but that class is larger: as discussed in Section 4.2.3, Winter also includes run-of-the-mill distributive predicates like *smile*, which do not have collective interpretations to begin with. Following Dowty (1987), I do not include distributive predicates in the class of

---

[36]Sven Lauer (p.c.) suggests that items like *most*, as opposed to definite plurals, might involve a reduction of pragmatic slack in the sense that Lasersohn (1999) suggests for *all*. It is not clear to me how slack reduction can be diagnosed in *most*, or how it could be linked to the other properties of *all* under consideration in this chapter. Lasersohn (1999) himself does not provide such a link and refers to the theory of Dowty (1987) to account for these properties.

numerous-type predicates. Many of the facts I discuss in this section were first observed by Kroch (1974) and Dowty (1987) independently of each other.

In connection with definite plurals, numerous-type predicates easily give rise to collective interpretations. Indeed, the collective interpretation is often the only one available. For example, in (3a), the predicate *be numerous* can only be understood as applying collectively to the ants in the colony, because there is no sense in which an individual ant can be numerous. The sentence becomes ambiguous between a collective and distributive interpretation when its definite plural is headed by a group noun such as *committee* or *army*. For example, (3b) can mean either that each of the armies taken by itself was large in number of soldiers, or that the number of armies was large.

(3) a. The ants in the colony were numerous. *distributive, ✓ collective
     b. The enemy armies were numerous. ✓ distributive, ✓ collective

The distributive item *each* (and its relative *every*, which behaves analogously to it) only allows the distributive interpretation of a predicate of this type. When there is no such interpretation in the first place, the sentence becomes unacceptable altogether (4). Adverbial *each* behaves analogously to prenominal *each* in this respect (5), if one ignores the marked status of sentence (5b).

(4) a. *Each ant in the colony was numerous. *distributive, *collective
     b. Each enemy army was numerous. ✓ distributive, *collective

(5) a. *The ants in the colony were each numerous. *distributive, *collective
     b. ?The enemy armies were each numerous. ✓ distributive, *collective

The effect of *all* on this type of predicate is identical to the effect of *each*: if the sentence is acceptable at all, it only has a distributive interpretation. For example, sentences (6a) and (7a) are unacceptable, and sentences (6b) and (7b) can only be interpreted as saying that every enemy army had many members. These sentences are synonymous with sentence (4b), whose only interpretation is distributive.

(6) a. *All the ants in the colony were numerous. *distributive, *collective
     b. All the enemy armies were numerous. ✓ distributive, *collective

(7) a. *The ants in the colony were all numerous. *distributive, *collective
     b. The enemy armies were all numerous. ✓ distributive, *collective

The parallel between (4)-(5) and (6)-(7) on the other hand motivates treating *all* in analogous terms to the canonical distributive items *each* and *every*.

Other examples of the numerous-type class include *be politically homogeneous, be a motley crew, suffice to defeat the army* (Kroch 1974), *be a large group, be a group*

of four, be few in number, be a couple (Dowty 1987), *be denser in the middle of the forest* (Barbara Partee p.c. via Dowty 1987), *pass the pay raise, elect Bush, return a verdict of 'not guilty', decide unanimously to skip class, eat up the cake, finish building the boat* (Taub 1989), *be too heavy to carry* (Brisson 1998), *be a good team, form a pyramid, constitute a majority, outnumber* (Winter 2001).

It is important to remember that the items *all* and *every/each* do not block numerous-type predicates altogether. As (6b) and (7b) show, they only block the collective interpretation of these predicates. Their distributive interpretation becomes available when they are applied to entities in the denotation of group nouns such as *committee* and *army*.

Numerous-type predicates provide motivation for a parallel treatment of *each* and *all*. Before presenting such a treatment, I now turn to a class of predicates which seem to provide evidence that *all* is not a distributive item, thereby presenting a challenge to any parallel treatment of *each* and *all*.

9.2.2 Gather-type predicates

Gather-type predicates can lead to distributive and collective interpretations, even in the presence of the word *all*. They were introduced in Section 4.2.3 along with numerous-type predicates. I call this class after their prototypical member *gather*. As shown in Table 4.1, both the numerous-type and the gather-type predicates are subtypes of the collective predicates, in the sense that their distributive reading is only available if the subject argument involves reference to group individuals such as committees. The difference between gather-type and numerous-types of predicates concerns their interpretation. We have seen that the collective interpretation of numerous-type predicates is blocked both by *every/each* and by *all*, even if it is the only available interpretation. The collective interpretation of *gather*-type predicates is also blocked by *every/each*, but it is not blocked by *all*:

(8) a. All the students gathered in the hall.   *distributive, ✓ collective*
     b. *Each student gathered in the hall.   *distributive, *collective*

(9) a. All the committees gathered in the hall.   ✓ *distributive, ✓ collective*
     b. Each committee gathered in the hall.   ✓ *distributive, *collective*

The observation that some collective predicates (namely the gather-type ones) are compatible with *all* but not with *each* goes back at least to Vendler (1962). The first discussion of the two categories of collective predicates is found to Dowty (1987). The numerous-gather opposition has been subsequently discussed in Taub (1989), Brisson (1998, 2003), Winter (2001, 2002), and Hackl (2002). Gather-type predicates have also been called *essentially plural predicates* (Hackl 2002) and *set*

predicates (Winter 2001). Other examples of this type of predicate are *be similar, fit together* (Vendler 1957), *meet, disperse, scatter, be alike, disagree, surround the fort*, the object argument of *summarize* (Dowty 1987), and *form a big group* (Manfred Krifka p.c. via Brisson 2003).

Now that we have seen the lists of gather-type and numerous-type predicates, we can try and make generalizations about these classes. Taub (1989) hypothesizes that all gather-type predicates are activities and accomplishments, while all numerous-type predicates are states and achievements. Following her, Brisson (1998, 2003) proposes a syntactic account of the numerous-gather opposition that implements this in terms of a silent predicate DO. Brisson assumes this predicate is assumed to be present only on activities and on accomplishments. However, Taub's generalization seems more like a tendency than a hard and fast generalization. For example, the predicate *reach an agreement* is as good an achievement predicate as any other, but it is gather-type since it is compatible with *all* on a collective reading:

(10)   All the parties involved reached an agreement.

It is not easy to draw the boundary of the class of gather-type predicates. If one includes all collective predicates into this class as long as they are compatible with *all*, as does Winter (2001), one ends up with a heterogeneous class, including reciprocally interpreted predicates such as *admire each other*, and predicates formed with collectivizing adverbials such as *perform Hamlet together*. Following Dowty (1987) and Brisson (2003), I exclude these predicates from consideration here. Winter furthermore includes any predicate that is compatible both with *all* and with *each* as long as they bring about a difference in truth conditions. This difference cannot always be easily attested. For example, mixed predicates like *build a raft* and *perform Hamlet* belong to this class, so long as their collective reading remains available with *all* and can be distinguished truth-conditionally from their distributive reading (11). This is the case according to the judgment of Dowty (1987), but Dowty also reports that other speakers find these sentences completely synonymous. Following Winter (2001), I generalize this judgment to other mixed predicates like *build a raft* and I refer to these two classes of (possible) dialects of English as *Dowty's dialect* and *other dialects*.

(11)   **Dowty's dialect**
    a.   All the students in my class performed Hamlet.   ✓ *dist,* ✓ *coll*
    b.   Each student in my class performed Hamlet.   ✓ *dist,* *coll

(12)   **Other dialects**
    a.   All the students in my class performed Hamlet.   ✓ *dist,* *coll

b.   Each student in my class performed Hamlet.   ✓ *dist, *coll*

Gather-type predicates bring up two questions: why do *all* and *every/each* behave differently with respect to gather-type predicates? And why do gather-type and numerous-type predicates behave differently with respect to *all*? For now, I leave these problems aside and I concentrate instead on a number of similarities between *for*-adverbials and *all*. Strata theory leads us to expect this kind of similarities because distributivity and aspect are seen as intimately related. I come back to the problems posed by numerous-type and gather-type predicates in Section 9.5.

9.3   Similarities between *for* and *all*

I have argued in Chapter 4 that *for*-adverbials are distributive items. If *all* is a distributive item too, then we expect that there should be similarities between *for*-adverbials and *all*. This section argues that the behavior of *all* with respect to quantifiers can be seen as parallel to the behavior of *for*-adverbials in two respects. First, they both fail to give rise to cumulative readings with quantifiers in their scope; second, they can both license dependent plurals.

9.3.1   *For* and *all* block cumulative readings

Cumulative readings were introduced in Section 2.8. As described there, cumulative readings are a special kind of scopeless readings in which the members of two pluralities A and B stand in a relation, such that each member of A stands in this relation to some member of B in such a way that B is exhaustively covered. Zweig (2008, 2009) notes that the presence of *all* can block otherwise available cumulative readings. The following contrast illustrates this behavior:

(13)   a.   All the safari participants saw thirty zebras.
              Unavailable cumulative reading: Each safari participant saw at least
              one zebra, and thirty zebras were seen overall.
        b.   Three safari participants saw thirty zebras.
              Available cumulative reading: Three safari participants each saw at
              least one zebra and thirty zebras were seen overall.

It is surprising that (13a) does not have the cumulative reading of (13b), because this reading cannot be ruled out in terms of lack of plausibility. For example, suppose that (13b) is uttered in a context where there are only three safari participants. In this context, the noun phrases *three safari participants* and *all the safari*

participants involve reference to the same plural individual. Yet, as (13a) shows, it is not possible to use them interchangeably. Only the former gives rise to a cumulative reading.

Compare this fact with the behavior of *for*-adverbials. It is an old observation that a *for*-adverbial normally cannot take scope over an indefinite (see Section 8.4). Moreover, a *for*-adverbial cannot enter a scopeless relation, in particular a cumulative relation, with an indefinite (14a). Cumulative readings are available in the temporal domain, as shown by the fact that *in*-adverbials do give rise to cumulative readings (14b). The *for*-adverbial in (14a) blocks this cumulative reading, just like *all* in (13a) blocks a cumulative reading.

(14) a. John saw thirty zebras for three hours.
          Unavailable reading: John saw a total of thirty zebras over the course of a three-hour timespan.
      b. John saw thirty zebras in three hours.
          Available reading: John saw a total of thirty zebras over the course of a three-hour timespan.

Similar facts are discussed in Krifka (1992) and Eberle (1998), among others. However, the connection between *for*-adverbials, cumulative readings and *all* has to my knowledge not been noted previously. Since it is not common to describe the interpretation of (14b) as a cumulative reading, let me describe it in more detail. First, note that the interpretation is scopeless because *thirty zebras* and *three hours* can both be understood as involving reference to a particular set of zebras and a particular time interval respectively; that is, they do not covary with each other. Second, it is cumulative because each of the thirty zebras is related to an instant of the three hours (leaving aside the minimal parts problem, see Chapter 5).

This cumulative interpretation is not available in (14a). If anything, this sentence can only be understood as asserting that there are thirty zebras, each of which was seen by John over the course of three hours. This reading is not a scopeless reading, because it does not involve reference to one specific three-hour long time interval but to potentially as many such intervals as there are zebras.

The two paradigms (13) and (14) are not completely parallel, because (13b) has other readings besides the cumulative reading, while (14b) only has a cumulative reading. For example, (13b) also has a distributive reading in which three safari participants each saw a potentially different set of thirty zebras, but (14b) does not have a reading in which John kept seeing potentially different sets of thirty zebras over the course of a three-hour timespan. I am not suggesting that the sentences (13b) and (14b) should receive a parallel treatment. The focus here is on sentences (13a) and (14a). The point of sentences (13b) and (14b) is to illustrate that a predicate

like *see thirty zebras* can give rise to a cumulative reading, so the fact that this reading is blocked by *for* and *all* in (13a) and (14a) must be explained.

9.3.2   *For* and *all* license dependent plurals

Dependent plurals were introduced in Section 2.6.2. To recapitulate briefly, the typical cases of dependent plurals are bare plurals which are c-commanded by a coargument whose head noun is also plural. It is useful to separate their semantic contribution conceptually into two parts, as shown in (15). The distributivity component relates the bare plural to each of the parts of its coargument, and the multiplicity component sets a minimum size for the cardinality of the bare plural.

(15)   Five boys flew kites.

    a.   *Distributivity component:* Each of these boys flew **at least one** kite.
    b.   *Multiplicity component:* **At least two** kites were flown in total.

A bare plural is called dependent when it is interpreted inclusively ("at least one") in the distributivity component. As the paraphrases show, this is the case in (15).

When the coargument of the bare plural is headed by a singular determiner, it is interpreted exclusively ("at least two"), as shown in (16). In such cases, we no longer speak of a dependent plural.

(16)   Each boy flew kites.

    a.   *Distributivity component:* Each of these boys flew **at least two** kites.
    b.   *Multiplicity component:* At least two kites were flown in total.

For-adverbials and *all* both give rise to dependent plurals. The observation that *all* can license dependent plurals is due to Zweig (2008) and is illustrated by (17). The observation that *for*-adverbials can license them too is novel, as far as I know, and is illustrated in (18).

(17)   All the boys flew kites.

    a.   *Distributivity component:* Each boy flew **at least one** kite.
    b.   *Multiplicity component:* At least two kites were flown in total.

(18)   John flew kites for five hours.

    a.   *Distributivity component:* At each time, John flew **at least one** kite.
    b.   *Multiplicity component:* John flew at least two kites in total.

In the distributivity component of (17) and (18), the bare plural *kites* has an inclusive interpretation. In other words, sentence (17) is compatible with scenarios

in which no boy flies more than one kite, and sentence (18) is compatible with scenarios in which John never flies more than one kite at a time.[37] Therefore, *all* and *for*-adverbials license dependent plurals.

Sentences (19) through (20) illustrate the same point in a different way, using a test from Zweig (2008). Since a person normally only wears one necktie at a time, interpreting the predicate *wear yellow neckties* with respect to a singular agent makes sense if *neckties* has the inclusive interpretation *at least one necktie.* But if *neckties* is interpreted exclusively, the result is a funny interpretation in which one person is said to wear two neckties at the same party. We can use this predicate to test whether an environment can license dependent plurals. For example, (19) does not commit us to the funny interpretation, so it shows that *five boys* licenses dependent plurals.

(19)   At the party, five boys wore yellow neckties.
         Distributivity component: Each boy wore **at least one** yellow necktie.

By contrast, (20) only has a funny interpretation, so it shows that *each boy* does not license dependent plurals.

(20)   At the party, each boy wore yellow neckties.
         Distributivity component: Each boy wore **at least two** yellow neckties.

These two baseline examples have shown how the test works. Examples (21) and (22) apply it to *all* and *for*-adverbials respectively.

(21)   At the party, all the boys wore yellow neckties.
         Distributivity component: Each boy wore **at least one** yellow necktie.

(22)   At the party, John wore yellow neckties for five hours.
         Distributivity component: At any given moment, John wore **at least one** yellow necktie.

As the paraphrases of the distributive component show, these examples do not give rise to the funny interpretation, so they show that *all* and *for* license dependent plurals.

The ability to license dependent plurals represents another similarity between *for* and *all*, and is another motivation for giving a unified explanation of their behavior.

---

[37]The paraphrase in (18a) in terms of "at each time" does not take the minimal-parts problem into account. This problem is addressed in Chapter 5. The point about the inclusive interpretation of the bare plural is independent of this problem. No matter how long or short the times in questions are, sentence (18a) does not require John to fly more than one kite at any given time.

9.4 Explaining the similarities between *for* and *all*

I now identify an insight into the treatment of *for*-adverbials from the literature on aspect, and apply it to *all* in order to account for the missing cumulative readings.

My analysis is based on the idea that *all* imposes a constraint on the verb phrase predicate which is analogous to the presupposition of *for*-adverbials, except that the "dimension" involved (the Map, in the terminology introduced in Chapter 4) is not runtime or location as in the case of *for*-adverbials, but the thematic role of the *all*-phrase. This constraint has the effect that when *all* modifies a verb phrase (either directly as an adverbial, or prenominally after combining with a bare plural or definite description), every event in the denotation of the verb phrase is required to consist of one or more subevents, or strata, which are in the denotation of the verb phrase and which the thematic role associated with *all* maps to an atomic value. This constraint blocks cumulative readings in *for* and *all* (Section 9.4.1), and makes it possible to adapt the theory of Zweig (2008, 2009) to explain why they both license dependent plurals (Section 9.4.2).

The constraint uses the concept of stratified reference, which was developed in Chapter 4 and whose definition is repeated here:

(23)   **Definition: Stratified reference**
$\mathrm{SR}_{f,\varepsilon(K)}(P) \overset{\text{def}}{=} \forall x[P(x) \rightarrow x \in {}^*\lambda y \left(\; P(y) \; \wedge \; \varepsilon(K)(f(y)) \; \right)]$
(A predicate P has stratified reference with respect to a function f and a threshold $\varepsilon(K)$ if and only if there is a way of dividing every entity in its denotation exhaustively into parts ("strata") which are each in P and which have a very small f-value. Very small f-values are those that satisfy $\varepsilon(K)$.)

We have seen in Chapter 4 that the threshold $\varepsilon(K)$ is not fully specified and that, in *for*-adverbials, it resolves to a value that is very small with respect to their key K. To represent the fact that *all* always distributes its verb phrase down to atoms, I assume that *all* lexically sets this threshold to the predicate *Atom*. This predicate holds of every entity that has only itself as a mereological part (see Section 2.3.1). Here is my lexical entry for prenominal *all* in agent position:

(24) $[\![\mathrm{all}_{ag}]\!] = \lambda x \lambda P_{\langle vt \rangle} \lambda e : \mathrm{SR}_{ag,Atom}(P).[P(e) \; \wedge \; {}^*ag(e) = x]$

I assume that it first combines with a definite description, and then with an event predicate.

As described in Section 2.8, I adopt the standard assumption that the definite plural *the boys* is translated as $\bigoplus$boy, the sum of all boys, or as $\uparrow (\bigoplus \text{boy})$, the impure atom derived from this sum. Suppose that prenominal *all* combines with *the boys* on its sum interpretation to form the noun phrase *all the boys*. The result

is the following:

(25) $[\![\text{all}_{ag} \text{ the boys}]\!] = \lambda P_{\langle vt \rangle} \lambda e : \text{SR}_{ag, Atom}(P).[P(e) \wedge {}^*\text{ag}(e) = \bigoplus \text{boy}]$

Adverbial *all* is translated similarly:

(26) $[\![\text{all}_{ag}]\!] = \lambda P_{\langle vt \rangle} \lambda e : \text{SR}_{ag, Atom}(P).P(e)$

As these translations show, I assume that *all* is coindexed with the theta role of the noun phrase with which it is associated. For adverbial *all*, this assumption is necessary when it is separated from the noun phrase it modifies (*The boys have all gathered*). For adnominal *all*, I adopt this assumption because in the system described in 2.10, the thematic role *agent* is introduced by a head that first combines with the verb phrase, so the role is no longer compositionally available to *all*. One could avoid some of these coindexations by letting the thematic role head *agent* combine first to the noun phrase instead. This does not pose any problems from my point of view, though it contradicts current syntactic assumptions about "little *v*" (the thematic role head for agents). By using coindexation, I do not make myself dependent on any assumption with respect to how the thematic role is introduced.

9.4.1 *For* and *all* block cumulative readings

The translations for *for* and *all* presented above, together with the background assumptions concerning cumulative readings presented in Section 2.8, explain why *for*-adverbials and *all* can both block cumulative readings. I illustrate the point first with *for*-adverbials and then with *all*. The reasoning is analogous in both cases. In the first case, the argumentation is familiar from theories of aspectual composition such as Krifka (1998). The relevant examples are repeated here from Section 9.3.1:

(27) a. John saw thirty zebras for three hours.
Unavailable reading: John saw a total of thirty zebras over the course of a three-hour timespan.
b. John saw thirty zebras in three hours.
Available reading: John saw a total of thirty zebras over the course of a three-hour timespan.

The following LF represents the unavailable reading of (27a):

(28) John saw thirty zebras for three hours.
$\exists e[{}^*\text{see}(e) \wedge {}^*\text{ag}(e) = j \wedge \text{hours}(\tau(e)) = 3 \wedge$
${}^*\text{zebra}({}^*\text{th}(e)) \wedge |{}^*\text{th}(e)| = 30]$
Presupposition: $\text{SR}_{\tau, \varepsilon(\lambda t[\text{hours}(t)=3])}(\lambda e[{}^*\text{see}(e) \wedge$

197

*zebra(*th(e)) $\wedge$ $|^*$th(e)$|$ = 30])
(Every event in which thirty zebras are seen consists of one or more seeing events whose runtimes are very short compared with three hours and whose themes are sums of thirty zebras.)

This reading is unavailable because its presupposition is not satisfied by the predicate *see thirty zebras*. To see this, note that the presupposition requires that any event in the denotation of this predicate can be decomposed into strata in each of which thirty zebras are seen. It requires *see thirty zebras* to be true at any of these strata as well as their sum. In other words, at each stratum of time it has to be the case that thirty zebras are seen. This is only possible if the thirty zebras are practically constantly on display, that is, if they are seen simultaneously. But given lexical cumulativity, there are (plural) events in the denotation of *see thirty zebras* in which the zebras are seen consecutively. Consider the following example:

(29)   Three safari participants saw thirty zebras.
$\exists e[^*$safari.participant(*ag(e)) $\wedge$ $|^*$ag(e)$|$ = 3 $\wedge$ *see(e) $\wedge$ *zebra(*th(e)) $\wedge$ $|^*$th(e)$|$ = 30]

In this example, event e is asserted to have a sum theme of thirty zebras. The cumulative reading of this sentence does not entail that the zebras were seen simultaneously. The reading is silent on this point. Section 2.10 shows how an analogous sentence is generated based on the predicate *see thirty zebras*. This establishes that *see thirty zebras* does not require simultaneity.

Now consider an event in the denotation of this predicate in which the zebras are indeed not seen simultaneously but consecutively. Any part of this event whose runtime is shorter than its own has a theme that consists of less than thirty zebras. The shorter the runtime, the less zebras are seen. It follows that *see thirty zebras* does not satisfy the presupposition of (28). There are only two ways to interpret that sentence, and neither of them leads to a cumulative reading because they both create scope dependencies. The first way is by applying distributive QR (what Landman (2000) calls SQI, or scopal quantifying-in, see Section 2.8) to the object, with the following result:

(30)   John saw thirty zebras for three hours.
$\exists x[^*$zebra(x) $\wedge$ $|x|$ = 30 $\wedge$ $\forall y \leq_{Atom} x$ $\exists e[^*$see(e) $\wedge$ *ag(e) = j $\wedge$ hours($\tau(e)$) = 3 $\wedge$ *th(e) = y]]
(There are thirty zebras each of which John saw for three hours.)
Presupposition: $\forall y \leq_{Atom} x[SR_{\tau,\lambda t[\text{hours}(t)=3]}(\lambda e[^*$see($e$) $\wedge$ *th(e) = y])]$
(For each y among the thirty zebras, every seeing event whose theme is y consists of one or more seeing events whose runtimes are very short with

respect to three hours, and whose theme is again y.)

In this representation, the presupposition applies to a predicate that contains a variable y left behind by the QRed object. It is satisfied because *see* is atelic.

The second way is to apply distributive QR to the object as before, and then to raise the temporal measure phrase above the object. The interpretation is the same except that the three hours are now fixed and do not covary with the zebras. That is, there are three hours and there are thirty zebras each of which John saw for these three hours. I do not show this reading here. The important point is not how these readings are derived, but that there is no way to derive a cumulative reading with a *for*-adverbial because its presupposition does not allow a numeral theme to stay in situ. The same argumentation can now be applied to explain why *all* does not license a cumulative reading. The following example illustrates what happens when *all the safari participants* combines directly with *see thirty zebras*.

(31) All the safari participants saw thirty zebras.
$\exists e[{}^*\text{see}(e) \wedge {}^*\text{ag}(e) = \bigoplus \text{safari.participant} \wedge {}^*\text{zebra}({}^*\text{th}(e)) \wedge |{}^*\text{th}(e)| = 30]$
Presupposition: $\text{SR}_{ag,Atom}(\lambda e[{}^*\text{see}(e) \wedge {}^*\text{zebra}({}^*\text{th}(e)) \wedge |{}^*\text{th}(e)| = 30])$
(Every event in which thirty zebras are seen consists of one or more seeing events whose agents are atomic and whose themes are sums of thirty zebras.)

In (31), the presupposition is introduced by *all*. The LF is ruled out because this presupposition is not fulfilled by the predicate *see thirty zebras*, for the following reason. Given lexical cumulativity, the predicate *see thirty zebras* can apply to a plural seeing event in which a sum of individuals sees a total of thirty zebras. In such an event, it is not necessarily the case that every individual in this sum sees thirty zebras by himself. For example, each zebra could have been seen by only one individual. Likewise, *see thirty zebras* fails to have stratified reference with respect to time, because in such an event, it is not necessarily the case that every zebra is seen throughout the entire runtime of the event. For example, each zebra could have been seen at a different moment. Again, this state of affairs is compatible with the cumulative reading of (27b), which uses the predicate *see thirty zebras*.

As in the case of *for*-adverbials, sentence (31) does have available interpretations, just not the cumulative one. For example, it has a distributive reading, on which the predicate *see thirty zebras* is applied to each of the participants. This reading can be derived by applying the D operator presented in Chapter 8 to the verb phrase *see thirty zebras* before it combines with *all* and its subject. As described in that chapter, my implementation of the D operator has a dimension and a granularity parameter. The dimension parameter needs to be set to *ag* and the

granularity parameter to *PureAtom*. Since the output of the D operator always satisfies stratified reference, the presupposition of *all* is satisfied. Since this operator creates a scopal dependency between the subject and the object of the sentence, its presence rules out a cumulative interpretation.

9.4.2   *For* and *all* license dependent plurals

Section 9.3.2 pointed out that *all* and *for*-adverbials are parallel in that they both license dependent plurals. In this section, I propose an explanation of this fact by applying an extended version of the theory of dependent plurals in Zweig (2008, 2009) to these two cases. Zweig's theory in its original formulation is only partly able to account for the fact that *all* and *for*-adverbials license dependent plurals. As we will see, the problem with his account is that it requires all items which license dependent plurals to be interpreted in situ. Zweig assumes that noun phrases which are interpreted in situ give rise to cumulative readings as a general rule. This is also the case in the present framework. We have seen that *all* and *for*-adverbials do not give rise to cumulative readings. The problem therefore consists in explaining why they can nevertheless give rise to dependent plurals. In order to extend his approach to noun phrases headed by *all*, Zweig assumes (as I do) that they are indeed interpreted in situ. As a consequence, he needs to introduce an additional assumption that explains why they are unable to give rise to cumulative readings. I have presented and motivated such an additional assumption in the previous section: constructions with *all* (along with *for*-adverbials) impose a presupposition on their predicate, which I have described in terms of stratified reference. Zweig tentatively introduces a different assumption for the same purpose: a subject noun phrase headed by *all* applies algebraic closure to its verb phrase. In this section, I compare the effect of stratified reference with the effect of Zweig's assumption in the context of his system. While his translation of *all* is based on an intuition similar to my own, we will see that it is not independently motivated and that it fails to assign the correct semantics to *all*. The present account does not face these problems. My criticism of Zweig's proposal is limited: while his assumptions concerning the meaning of *all* are problematic, they can be separated from the rest of his account and do not play a major role in it. A synthesis of Zweig's theory and strata theory makes the right predictions concerning the ability of *all* to license dependent plurals. Moreover, it extends straightforwardly to *for*-adverbials. This last fact in particular confirms the usefulness both of Zweig's general approach and of strata theory, since he himself did not discuss *for*-adverbials at all.

As I already mentioned in Section 2.6.2, Zweig assumes the framework of Landman (2000), from which I have also adopted many ideas. As a consequence, it shares a large number of background assumptions with strata theory, such as the

assumption that the domains of individuals and events are each closed under sum (Sections 2.4.1 and 2.4.3), lexical cumulativity (Section 2.7.2), as well as the Unique Role Requirement and cumulativity of thematic roles (Section 2.5.1).

Simplifying a bit, Zweig assumes that the meaning of a bare plural has two parts: a lexical entry, which is inclusive ("at least one"), and a scalar implicature, which is exclusive ("at least two"). This scalar implicature is treated along the lines of Chierchia (2006). This theory assumes that scalar implicatures can be inserted into the computation of the truth-conditional meaning at specific "implicature calculation points", provided that this insertion strengthens the truth conditions of the entire sentence. Although Chierchia's theory is formulated in semantic terms, I describe it here as if it were formulated in syntactic ones. Instead of speaking of implicature projection and implicature calculation points, I will speak of implicature movement and of landing sites. Put in these terms, Chierchia's theory assumes that a scalar implicature detaches itself from its source, moves up the tree, and is interpreted at whichever landing site yields the strongest truth conditions for the entire sentence. Zweig assumes that there is such a landing site just underneath the existential closure operator and that this is always the highest available landing site, meaning that QR above existential closure does not create any additional landing sites.

Together with Landman's theory of argument interpretation, these assumptions predict correctly that (32) has a dependent plural interpretation, while (33) does not.

(32)   Five boys flew kites.   = (15)

(33)   Each boy flew kites.   = (16)

These predictions are obtained in the following way. If the subject is headed by an indefinite, as in (32), Landman's theory predicts that it can be interpreted in situ (see Fig. 9.1). In this case, the highest landing site for the scalar implicature is just outside of the scope of the subject. The result is a dependent plural interpretation. If the subject is headed by *every* or *each*, as in (33), Landman's theory predicts that it must undergo quantifier raising (see Fig. 9.2). The existential operator, the highest landing site for the scalar implicature, is located inside the scope of the subject. The result is a non-dependent plural interpretation. In the figures, I do not represent the way in which thematic roles enter the derivation, because Zweig's account and my own differ in this respect. Zweig assumes Landman's system, in which the thematic roles are introduced in the lexical entry of the verb, while I assume that they are contributed by silent lexical items (see Section 2.10). This difference is immaterial for the way Zweig's theory works.

Zweig only considers adnominal *each*, which he assumes (plausibly for the

Figure 9.1: Simplified representation of *Five boys flew kites* in Zweig's theory.
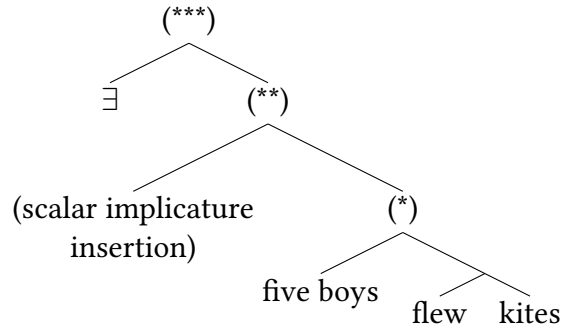
$[\![(*)]\!]$ $= \lambda e[*\mathrm{fly}(e) \wedge *\mathrm{boy}(*\mathrm{ag}(e)) \wedge |*\mathrm{ag}(e)| = 5 \wedge *\mathrm{kite}(*\mathrm{th}(e))]$

$[\![(**)]\!]$ $= \lambda e[*\mathrm{fly}(e) \wedge *\mathrm{boy}(*\mathrm{ag}(e)) \wedge |*\mathrm{ag}(e)| = 5 \wedge *\mathrm{kite}(*\mathrm{th}(e)) \wedge |*\mathrm{th}(e)| \geq 2]$

$[\![(***)]\!]$ $= \exists e[*\mathrm{fly}(e) \wedge *\mathrm{boy}(*\mathrm{ag}(e)) \wedge |*\mathrm{ag}(e)| = 5 \wedge *\mathrm{kite}(*\mathrm{th}(e)) \wedge |*\mathrm{th}(e)| \geq 2]$
(Five boys engaged in kite-flying and at least two kites were flown overall.)

Figure 9.2: Simplified representation of *Every boy flew kites* in Zweig's theory.
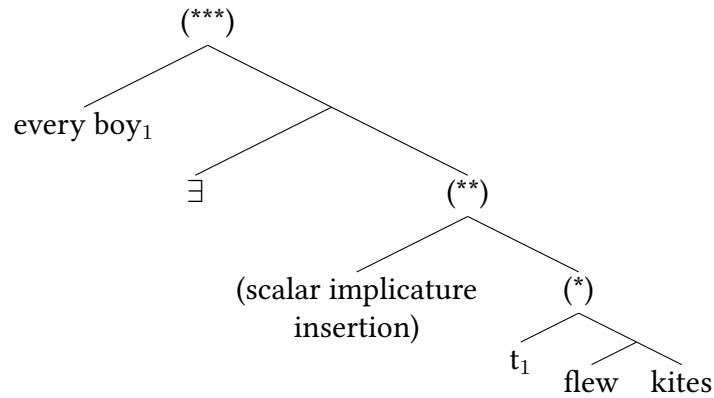
$[\![(*)]\!]$ $= \lambda e[*\mathrm{fly}(e) \wedge *\mathrm{ag}(e) = g(1) \wedge *\mathrm{kite}(*\mathrm{th}(e))]$

$[\![(**)]\!]$ $= \lambda e[*\mathrm{fly}(e) \wedge *\mathrm{ag}(e) = g(1) \wedge *\mathrm{kite}(*\mathrm{th}(e)) \wedge |*\mathrm{th}(e)| \geq 2]$

$[\![(***)]\!]$ $= \forall x[\mathrm{boy}(x) \rightarrow \exists e[*\mathrm{fly}(e) \wedge *\mathrm{ag}(e) = x \wedge *\mathrm{kite}(*\mathrm{th}(e)) \wedge |*\mathrm{th}(e)| \geq 2]]$
(Every boy flew at least two kites.)

present purpose) to have the same semantics as the determiner *every*. In particular, *each* must raise above existential closure, and the scalar implicature is therefore inserted underneath *each*. Adverbial *each*, which Zweig does not discuss, also fails to license dependent plurals, as the following example shows.

(34)   The boys each flew kites.

This sentence entails that each of the boys flew two or more kites. The bare plural *kites* in this example is therefore interpreted exclusively with respect to each boy. This raises the question of what is the best way to extend Zweig's account to adverbial *each*. The challenge is to explain why the scalar implicature can only be inserted in the scope of *each*. I sketch two possibilities here. One possibility would be to stipulate that the subject of a sentence with adverbial *each* must be given distributive wide scope, so that the account of adverbial *each* is reduced to Zweig's account of adnominal *each*. Another possibility would be to introduce barriers to implicature movement and to assume that adverbial *each* is such a barrier, so that the scalar implicature needs to be interpreted in its scope. The verb phrase *flew kites* would then be enriched to *flew two or more kites*. Since this predicate applies among other things to sums of events in which only one kite per event is flown, it does not satisfy stratified distributive reference with respect to agents (see Section 4.5). The D operator, with its parameters set to *agent* and *PureAtom*, is then applied to this enriched predicate to make sure that it satisfies stratified distributive reference (see Chapter 8). Then *each* is applied and checks that this is the case. Finally, [ag] and *the boys* apply, followed by existential closure. This yields the following result:

(35) $\exists e[{}^*\mathrm{ag}(e) = \bigoplus \mathrm{boy} \ \wedge \ e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\mathrm{fly}(e') \ \wedge \ {}^*\mathrm{kite}({}^*\mathrm{th}(e')) \ \wedge \\ \mathrm{PureAtom}({}^*\mathrm{ag}(e')) \ \wedge \ |{}^*\mathrm{th}(e')| \geq 2 \end{array} \right)]$
(There is a sum event whose agents are the boys, and which consists of flying events whose agents are pure atoms and whose themes are two or more kites.)

To summarize, on Zweig's account, the relative positions of the scalar implicature insertion point (the "landing site") and of the subject noun phrase determine whether a dependent plural interpretation is available. It is only available if the subject noun phrase does not QR above the landing site. The case of adverbial *each* poses problems because it is not clear how to make sure that the scalar implicature always lands below *each*, but I have sketched two potential solutions.

Consider now the case of *all the boys*. In order to generate a dependent plural interpretation of a sentence like *All the boys flew kites*, the subject noun phrase must be interpreted in situ. We have seen in Section 9.4.1 that the consequence is

an unattested cumulative reading if *all the boys* is translated analogously to *five boys*, as in (36) for example. Zweig is aware of this problem and proposes the translation in (37) as a remedy (Zweig 2008, Section 6.2.6).

(36) ⟦all the boys⟧ (naïve) $= \lambda P_{\langle et \rangle} \exists x[|x| = |\text{boy}| \ \wedge \ ^*\text{boy}(x) \ \wedge \ P(x)]$

(37) ⟦all the boys⟧ (Zweig) $= \lambda P_{\langle et \rangle} \exists x[|x| = |\text{boy}| \ \wedge \ ^*\text{boy}(x) \ \wedge \ ^*P(x)]$

To simplify the discussion, I rewrite these translations in an equivalent but more compact form:

(38) ⟦all the boys⟧ (naïve) $= \lambda P_{\langle et \rangle}[P(\bigoplus \text{boy})]$
      (True of predicates which hold of the sum of all boys.)

(39) ⟦all the boys⟧ (Zweig) $= \lambda P_{\langle et \rangle}[^*P(\bigoplus \text{boy})]$
      (True of predicates whose algebraic closure holds of the sum of all boys.)

Zweig claims that translating *all the boys* as in (39) blocks cumulative readings that would be available if the entry in (38) was used. However, this proposal does not have the desired effect, because any verb phrase that is compatible with (38) is also compatible with (39). The translation in (38) is the set containing every predicate that applies to the sum of all boys. We know from Theorem (18) in Section 2.3.1 that if a predicate P applies to a particular object, such as the sum of all boys, then $^*P$ also applies to this object. Therefore, the translation in (39) contains every predicate that is also contained in (38). In other words, (39) is a superset of (38), and any verb phrase whose denotation is contained in the former is also contained in the latter. In particular, a verb phrase like *see thirty zebras* will still be able to combine directly with a subject noun phrase like *all the boys* to produce a cumulative reading in the same way it would combine with a numeral subject like *three boys*.

While Zweig's implementation has formal problems, his own discussion of his proposal indicates that he intends *all the boys* to apply its verb phrase distributively to every boy, and to check whether the result is different from applying the verb phrase directly to the sum of all boys:

> [T]here is no difference between saying "there is a sum of boys, such that each of its atomic parts is the agent of an event of flying kites", and "there is a sum of boys that is the agent of an event of flying kites". This stands in contrast to the case with a numerical object; "there is a sum of boys, such that each of its atomic parts is the agent of an event of flying two kites" is not equivalent to "there is a sum of boys that is the agent of a sum of events of flying two kites". (Zweig 2008, p. 141)

Zweig notes that there may be a way to explain why *all* can stand in a cumulative-like relation with bare plurals (because it can license dependent plural interpretation) but cannot stand in a cumulative relation with overt numerals such as *two kites*. I have provided this explanation in Section 9.4.1 in terms of stratified reference. Although Zweig did not have the concept of stratified reference at his disposal, his description is remarkably accurate if one reads it as a description of the effect that stratified reference has on *all*.

I now show how a synthesis of Zweig's theory and my account can explain why *all* and *for*-adverbials license dependent plurals.

(40)   All the boys flew kites.   = (17)

(41)   John flew kites for five hours.   = (18)

The synthesis requires the assumption that *all* and *for*-adverbials can be interpreted in situ, so that they can take scope under the landing site of the scalar implicature of the bare plural object. This assumption is compatible with Zweig's framework. It is the null assumption because there are no type mismatches or other reasons that would force us to assume that these items must move.

In Section 4.6, stratified reference was implemented as a presupposition rather than as a part of the truth conditions of distributive items. In Zweig's account, presuppositions do not interact with the computation of the meaning of dependent plurals, so the presence of this kind of constraint does not interfere with his account.

Given these assumptions, Figures 9.3 and 9.4 show how the bare plural objects in (40) and (41) acquire their dependent plural reading. The computation is analogous to the case of the numeral *five boys* in 9.1 above. I leave out the presupposition of *for* and *all* in these figures.

Figure 9.3: Accounting for the dependent plural in *All the boys flew kites*
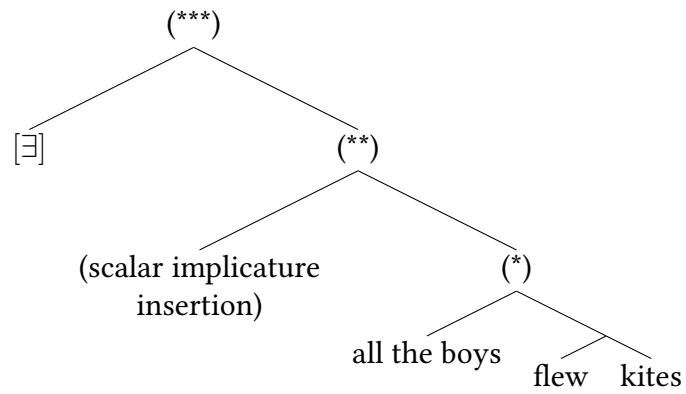
$\llbracket(*)\rrbracket = \lambda e[^*\text{fly}(e) \wedge {}^*\text{ag}(e) = \bigoplus \text{boy} \wedge {}^*\text{kite}(^*\text{th}(e))]$

$\llbracket(**)\rrbracket = \lambda e[^*\text{fly}(e) \wedge {}^*\text{ag}(e) = \bigoplus \text{boy} \wedge {}^*\text{kite}(^*\text{th}(e)) \wedge |^*\text{th}(e)| \geq 2]$

$\llbracket(***)\rrbracket = \exists e[^*\text{fly}(e) \wedge {}^*\text{ag}(e) = \bigoplus \text{boy} \wedge {}^*\text{kite}(^*\text{th}(e)) \wedge |^*\text{th}(e)| \geq 2]$
    (Every boy engaged in kite-flying and at least two kites were flown overall.)

Figure 9.4: Accounting for the dependent plural in *John flew kites for an hour*
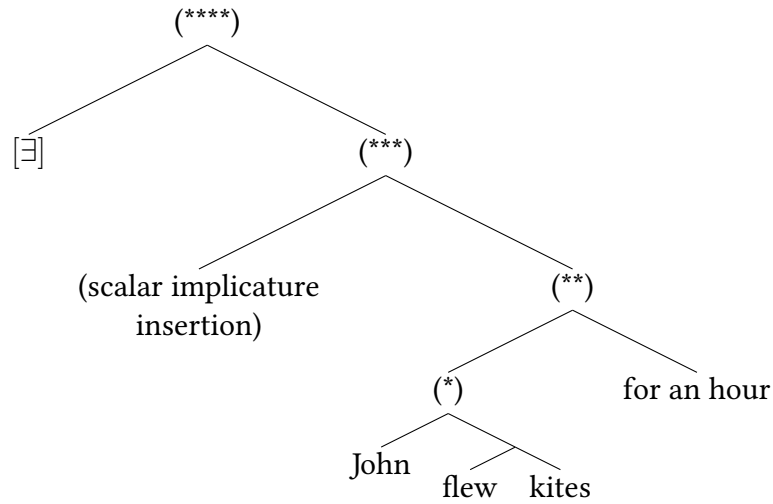
$\llbracket(*)\rrbracket = \lambda e[{}^*\text{fly}(e) \ \wedge \ {}^*\text{ag}(e) = j \ \wedge \ {}^*\text{kite}({}^*\text{th}(e))]$

$\llbracket(**)\rrbracket = \lambda e[{}^*\text{fly}(e) \ \wedge \ {}^*\text{ag}(e) = j \ \wedge \ {}^*\text{kite}({}^*\text{th}(e)) \ \wedge \ \text{hours}(\tau(e)) = 1]$

$\llbracket(***)\rrbracket = \lambda e[{}^*\text{fly}(e) \ \wedge \ {}^*\text{ag}(e) = j \ \wedge \ {}^*\text{kite}({}^*\text{th}(e)) \ \wedge \ \text{hours}(\tau(e)) = 1 \ \wedge \ |{}^*\text{th}(e)| \geq 2]$

$\llbracket(****)\rrbracket = \exists e[{}^*\text{fly}(e) \ \wedge \ {}^*\text{ag}(e) = j \ \wedge \ {}^*\text{kite}({}^*\text{th}(e)) \ \wedge \ \text{hours}(\tau(e)) = 1 \ \wedge \ |{}^*\text{th}(e)| \geq 2]$
(John flew two or more kites in total over the course of an hour.)

The constraint that rules out cumulative readings with *for*-adverbials and *all* operates analogously in both cases. It is a special case of the Distributivity Constraint, which has been motivated extensively throughout this work (see Section 4.6). The only additional assumption I have introduced is the granularity parameter of *all*, which I have assumed to be atomic. The following section motivates this choice and compares it with my treatment of the distributive item *each*.

9.5 Explaining the behavior of *numerous* and *gather*

Section 9.2.2 pointed out that the behavior of gather-type predicates with respect to distributive items differs from numerous-type predicates. In particular, *every*, *each*, and *all* behave analogously with respect to numerous-type predicates, while the distribution of gather-type predicates is more restricted with *every* and *each* than with *all*. In this section, I account for these differences in terms of a parameter setting: the granularity of *all* is at the level of both pure and impure atoms, and the granularity of *every* and *each* is at the level of pure atoms only. I argue that gather-type and numerous-type predicates instantiate two different notions that have both been called collective predication, and I propose to model this difference by assuming that only gather-type predicates can apply to events whose agents are impure atoms.

9.5.1 Collective predication

Collective predication generally involves the notion of a verbal predicate applying to a plural entity as a whole, as opposed to applying to the individuals that form this entity. Beyond this general idea, the criteria for what exactly constitutes collective predication in the semantic literature are often not clearly spelled out. Two similar but distinct views of collectivity can be distinguished. I distinguish between these two views with the terms *thematic* and *nonthematic collectivity*. The guiding idea behind this distinction is adapted from Verkuyl (1994).

On the first view, thematic collectivity, collective predication is defined positively, by the presence of certain kinds of entailments about a plural individual which cannot be induced from what we know about the parts of which this individual consists. For example, sentence (42) entails something about the Marines as an institution, an organized body which is able to take coordinated action and take responsibility, in this case for the action of invading Grenada (Roberts 1987, p. 147). The predicate *invade Grenada* instantiates thematic collective predication because it gives rise to the thematic entailment that the Marines as a whole were

responsible for invading Grenada.

(42)   The Marines invaded Grenada.   *collective*

The discussion of collectivity that most explicitly conforms to this view is found in Landman (2000). Landman calls the relevant entailments "thematic", because he sees them as analogous to the entailments which many theories associate with thematic roles. For example, such theories typically assume that the thematic role *ag* gives rise to the entailment that the agent of an event is responsible for this event. Landman makes the plausible assumption that the entailment about the collective responsibility of the Marines in (42) is of the same nature as the entailment of the individual responsibility of the agent in a sentence like *Saddam Hussein invaded Kuweit.*

As Landman acknowledges, it is difficult to identify or define thematic entailments exactly. Besides collective responsibility, he gives two other examples: collective body formation and collective action. Collective body formation is illustrated with examples such as *The boys touch the ceiling*, and collective action with examples such as *The boys carried the piano upstairs.* In both cases, the predicates license the same entailments about the boys as a whole that they do about individual boys in sentences like *The boy touched the ceiling* and *The boy carried the piano upstairs.* For example, in the case of *touch the ceiling*, one thematic entailment is that part of the agent is in surface contact with part of the ceiling, no matter if this agent is a boy or a group of boys. (This view is not uncontroversial. See Brisson (1998) for criticism.) Landman also notes that thematic entailments have a "non-inductive" character. Sentence (42) does not become true if two, ten, or even a very large number of members of the Marine Corps land on Granada in an unauthorized action. It requires that the Marines as an organization take responsibility for the invasion (Landman 2000, p. 171).

One of the consequences of lexical cumulativity is that predicates like *carry the piano upstairs* applies in principle not only to singular individuals but also to sums of individuals each of whom carry the piano upstairs. A sentence with a "mixed" predicate like *carry the piano upstairs* may be true both on its (thematic) collective reading and on its distributive reading at once, if the relevant individuals carried the piano upstairs together and also each of them did so separately. The non-inductive character of thematic entailments becomes visible here again: asserting (43a) does not entail the collective interpretation of (43b). More specifically, the distributive interpretation of (43b) and its collective interpretation are logically independent.

(43)   a.   The boys each carried the piano upstairs.
       b.   The boys carried the piano upstairs.

Following Landman (1989), I assume that this non-inductive character of the thematic implication is modeled by assuming that the collective interpretation of *carry the piano upstairs* applies to impure atoms like $\uparrow (a \oplus b)$. Schematically, if the boys are John and Bill, the meaning of (43a), as well as the distributive interpretation of (43b), is modeled as in (44a). I have marked lexical cumulativity explicitly with a star here. Its collective interpretation is modeled as in (44b). Lexical cumulativity is vacuous in this case. I have marked it by a star for consistency but it does not play any role.

(44) a. $\exists e[^*\text{carry.the.piano.upstairs}(e) \wedge {}^*\text{ag}(e) = j \oplus b]$
      b. $\exists e[^*\text{carry.the.piano.upstairs}(e) \wedge {}^*\text{ag}(e) = \uparrow (j \oplus b)]$

Since the sum $j \oplus b$ is different from the group $\uparrow (j \oplus b)$, the two readings are logically independent.

On the second view, nonthematic collectivity, collective predication is defined negatively. A collective predicate in this sense is defined as one that does not apply to the singular entities of which the individual to which it applies consists. This view is similar to what Verkuyl (1994) calls *kolkhoz collectivity*. As he explains, a kolkhoz is "a collective farm in the former Soviet Union owned by a group of individuals but none of these individuals . . . is its owner." Verkuyl traces a precursor of this view back to Jespersen (1913). Nonthematic collectivity allows the predicate to distribute down to subgroups but not down to the singular individual. For example, if a plurality of people is numerous (that is, if it has many members), some subpluralities of these people also have many members, but still *be numerous* exhibits nonthematic collectivity: it does not distribute down to individual people. It does not even make sense to apply the predicate *numerous* to a single person.

Instead of choosing between the two competing notions of collectivity, I adopt them both. The difference between the notions of thematic and nonthematic collectivity can help make sense of the empirically observable difference between gather-type and numerous-type predicates. I assume that being numerous is a property that a sum of individuals has *qua* sum, and that the notion of thematic entailments is not relevant in this case. Dowty (1987) seems to have a similar intuition when he calls numerous-type predicates *pure cardinality predicates*. My tentative conjecture is therefore the following:

(45) a. Thematic collectivity = gather-type predicates
      b. Nonthematic collectivity = numerous-type predicates

I propose to realize this distinction with the help of impure atoms. Landman (2000) only considers the thematic version of collectivity, and as we have seen, he models it by assuming that the agents of thematic collective predicates can be

impure atoms. I propose therefore that the agents of events in the denotations of nonthematic collective predicates can never be impure atoms. The agents related to a predicate like *be numerous* are typically sums. For example:

(46)   The boys are numerous.
        $\exists e[^*\text{numerous}(e) \ \wedge \ ^*\text{ag}(e) = \bigoplus \text{boy}]$

(47)   The boys gathered.
        $\exists e[^*\text{gather}(e) \ \wedge \ ^*\text{ag}(e) = \uparrow (\bigoplus \text{boy})]$

Let me point out an important difference between Landman's framework and mine. For reasons internal to his theory, Landman assumes that all basic predicates are distributive on all their argument positions. For example, he assumes that whenever the agent of an event to which a a basic predicate P applies is a sum individual, then each of the atomic parts of this sum individual is also the agent of an event to which P applies. In terms of this work, Landman assumes that $\text{SR}_{\theta, Atom}(P)$ holds for every basic verbal predicate P and for every argument theta role θ of P. This applies both to distributive predicates like *smile* and to thematic collective predicates like *gather*. In the latter case, collectivity effects are due to the possibility for impure atoms to be agents of gathering events.

Given Landman's assumptions, a representation like (46) would wrongly entail that each of the boys is the agent of an event to which *be numerous* applies. The corresponding entailment is blocked in (47) though the $\uparrow$ operator, because its output is not a sum individual, a welcome result.

To model the gather-type/numerous-type opposition, I adopt Landman's assumption for distributive predicates and for gather-type predicates, but not for numerous-type predicates. On my account, it is possible for *be numerous* to apply to an event whose agent is a sum individual even if none of the parts of this sum individual are agents of events to which *be numerous* applies.

Landman (1989) is aware that the framework of Link (1984), which contains sums and impure atoms, opens up the possibility for a three-way distinction between distributive predicates, and two types of collective predicates. Namely, collective predicates can be modeled as applying to sums without applying to their parts, or as applying to impure atoms without applying to the parts of their underlying sums. Like Link, Landman is only concerned about modeling the distributive-collective opposition, so he rejects Link's three-way distinction as unmotivated. However, neither Link nor Landman consider the empirical distinction between gather-type and numerous-type predicates. In a mereological setting, I do not see any way to model this empirical distinction without introducing some kind of formal distinction between the two types of predicates. Resurrecting the two types of collective predicates from Link (1984) seems like an obvious choice.

9.5.2 *All* distinguishes between *be numerous* and *gather*

Consider again the following contrast, repeated in adapted form from (6) and (8) above:

(48) a. *All the boys were numerous.
      b. All the boys gathered.

By the assumptions in the previous section, (48a) involves nonthematic collectivity and (48b) involves thematic collectivity. The subject in (48a) is interpreted as the sum of all boys, while the subject in (48b) is interpreted as the impure atom which is derived from this sum. In Section 9.4, I have assumed that *all* does not contribute to the truth conditions of a sentence, but only to its presuppositions. The truth conditions of (48a) and (48b) therefore can be determined by leaving out the word *all* and interpreting the resulting sentences:

(49) a. The boys were numerous.
          $\exists e[{}^*\text{numerous}(e) \ \wedge \ {}^*\text{ag}(e) = \bigoplus \text{boy}]$
      b. The boys gathered.
          $\exists e[{}^*\text{gather}(e) \ \wedge \ {}^*\text{ag}(e) = \uparrow (\bigoplus \text{boy})]$

These are not the only interpretations the system will generate. Given the assumption that *the boys* is ambiguous between a sum and a group interpretation, the grammar will also generate two more interpretations, namely the following:

(50) a. $\exists e[{}^*\text{numerous}(e) \ \wedge \ {}^*\text{ag}(e) = \uparrow (\bigoplus \text{boy})]$
      b. $\exists e[{}^*\text{gather}(e) \ \wedge \ {}^*\text{ag}(e) = \bigoplus \text{boy}]$

However, given our assumptions on nonthematic collectivity, we can disregard these interpretations. Since *be numerous* is a nonthematic collective predicate, it does not apply to any events whose agents are impure atoms, which is contradicted by the representation in (50a). As for *gather*, since it is a thematic collective predicate, it only applies to an event whose agent is a sum individual when it applies to events whose agents are the atomic parts of that sum individual. The representation in (50a) therefore entails that each boy is the agent of a gathering event. We can reject it as a category mistake.

This leaves us with the representations in (49). The task is to explain why the presence of *all* rules out (49a) but not (49b). I now show that the verb phrase in (49a) violates the stratified reference presupposition imposed by *all*. In (51), I repeat my translation for agent-coindexed *all the boys* from (25) above, where *the boys* has a sum interpretation.

(51)　⟦all$_{ag}$ the boys⟧ = $\lambda P_{\langle vt \rangle} \lambda e : \text{SR}_{ag, Atom}(P).[P(e) \wedge {}^{*}\text{ag}(e) = \bigoplus \text{boy}]$

After (51) combines with *be numerous*, its presupposition can be expanded as follows:

(52)　$\text{SR}_{ag,Atom}(\lambda e[\text{numerous}(e)]) \Leftrightarrow$
$\forall e[\text{numerous}(e) \rightarrow e \in {}^{*}(\lambda e'.\text{Atom}(\text{ag}(e')) \wedge \text{numerous}(e'))]$
(Every event e in the denotation of *be numerous* can be divided into one or more parts each of which is in the denotation of *be numerous* and has an atomic agent.)

This presupposition is violated because *be numerous* is nonthematic collective, and can therefore apply to an event whose agent is a sum individual even in cases where it does not apply to any events whose agents are the parts of this sum individual. In general, all nonthematic collective predicates violate the presupposition of *all*. This explains why sentences like (48a) are ruled out by the presence of *all*.

The account so far predicts that *be numerous* is never compatible with the presupposition of *all*, and that they should never cooccur. Section 9.2.1 has pointed out that this is not correct. The predicate *be numerous* is compatible with *all* when the subject of the sentence is headed by a group noun. This is illustrated by example (6b), repeated here:

(53)　All the enemy armies were numerous.　　　　✓ *distributive,* *collective

This sentence poses two challenges. First, we need to explain why it does not conform to the prediction that *all* and *be numerous* cannot cooccur. Second, we need to explain why it only has a distributive reading, in contrast to its counterpart without *all*, which also has a collective reading. This is illustrated by example (3b), repeated here:

(54)　The enemy armies were numerous.　　　　✓ *distributive,* ✓ *collective*

I propose to account for these cases as follows. Chapter 8 has presented the D operator as a strategy to change a predicate so that it satisfies the presupposition of a distributive item. I assume that the LF for sentences like (53) contains a D operator. This assumption is justified because it allows us to explain the collective-distributive ambiguity of sentences like (54). The presence of the D operator in (54) leads to a distributive reading and its absence leads to a collective reading.

My formulation of the D operator is repeated here from Section 8.3:

(55)　**Definition: Generalized event-based D operator**

$$\llbracket D_{\theta,C} \rrbracket \overset{\text{def}}{=} \lambda P_{\langle vt \rangle} \lambda e [e \in {}^{*}\lambda e' \begin{pmatrix} P(e') \wedge \\ C(\theta(e')) \end{pmatrix}]$$

(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs satisfy the predicate C.)

This operator has a dimension parameter θ and a granularity parameter C. As described in that section, I have assumed that the granularity parameter can always be resolved to the default value *PureAtom*. As for the dimension parameter, it can be instantiated with any thematic role. Assume that it is instantiated with *agent*. With these settings, the D operator looks as follows:

(56) $\llbracket D_{\text{ag,PureAtom}} \rrbracket = \lambda P_{\langle vt \rangle} \lambda e [e \in {}^{*}\lambda e' \begin{pmatrix} P(e') \wedge \\ \text{PureAtom}(\text{ag}(e')) \end{pmatrix}]$

(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose agent is a pure atom.)

When this operator is applied to *be numerous*, it returns a predicate that holds of any event e which consists of one or more events that are in *be numerous* and whose agent is a pure atom. This predicate applies to less events than *be numerous* does. In particular, it does not apply to any events whose agent is an impure atom. For example, imagine a model in which there are at least the following entities: one thousand boys, and the Golden Horde (a famous army). The boys have nothing to do with the Golden Horde. The members of the Golden Horde might also be in the model, but they are not of interest here. The word *army* is a group noun, and the Golden Horde is one of its referents, so the Horde represented as a pure atom, call it g (see Section 2.6.4). Assume that the sum of all boys in the model qualifies as numerous, and the Golden Horde does too. Then the predicate *be numerous* applies at least to two events e_1 and e_2 such that the agent of e_1 is $\bigoplus$ boy and the agent of e_2 is g. However, the predicate $D_{ag,PureAtom}(\llbracket$be numerous$\rrbracket)$ only applies to e_2 and not to e_1, because only e_2 consists entirely of events whose agent is an atom – in this case a pure atom.

In this way, the D operator acts as a filter on the predicate *be numerous*: its output only contains those events that have as their agents the referents of group nouns such as armies, committees, and so on, or which are built up from such events. In other words, every event in the denotation of the predicate $D_{ag,PureAtom}(\llbracket$be numerous$\rrbracket)$ can be divided into one or more parts each of which is in the denotation of $D_{ag,PureAtom}(\llbracket$be numerous$\rrbracket)$ and has an atomic agent. Therefore, $D_{ag,PureAtom}(\llbracket$be numerous$\rrbracket)$ satisfies the presupposition of *all* in (52), even though *be numerous* does not. This explains why sentences like (54) are

acceptable and why they only have a distributive interpretation: they contain the covert operator $D_{ag,PureAtom}$.

Consider now the case of *gather*. Here we need to explain why the presupposition of *all* is not violated by the predicate *gather*. Sentence (48b) involves a group agent, so the translation of *all the boys* is slightly different than in (51):

(57) $[\![\text{all}_{ag} \text{ the boys}]\!] = \lambda P_{\langle vt \rangle} \lambda e : SR_{ag,\ Atom}(P).[P(e) \ \wedge \ ^*ag(e) =\uparrow (\bigoplus \text{boy})]$

However, the presupposition of this translation is the same as (51). It expands as follows:

(58) $SR_{ag,Atom}(\lambda e[\text{gather}(e)]) \Leftrightarrow$
        $\forall e[\text{gather}(e) \rightarrow e \in \ ^*(\lambda e'.\text{Atom}(ag(e')) \ \wedge \ \text{gather}(e')])]$
        (Every event e in the denotation of *gather* can be divided into one or more parts each of which is in the denotation of *gather* and has an atomic agent.)

The predicate *gather* satisfies this presupposition because it is thematic collective, and can therefore only apply to an event whose agent is a sum individual as a result of lexical cumulativity, that is, it can only apply to this event if it also applies to events whose agents are atomic individuals. To put it differently, the agents of *gather* (and of any thematic collective predicate) form a complete Boolean algebra generated by a set of atoms. Many of these atoms are impure (the ones that are referenced by sentences such as *The boys gather*). For example, the event that verifies (50b) has as its agent the impure atom $\uparrow (\bigoplus boy)$. This event satisfies the presupposition (58) because it can be divided into one or more parts – in this case, into only one part, namely itself – which is in the denotation of *gather* and has an atomic agent.

In sum, the difference between numerous-type and gather-type predicates with respect to *all* can be explained in the following way. The intuition that thematic collective predication involves non-inductive attribution is modeled by the idea that only thematic collective predicates apply to impure atoms. The intuitive contrast between thematic and nonthematic collective predication is implemented by assuming that nonthematic collective predicates can apply to proper sums without applying to their parts. I have proposed that numerous-type predicates are nonthematic collective predicates and that gather-type predicates are thematic collective predicates. The prediction of these assumptions is that gather-type predicates always fulfill the presupposition of *all*. The difference between numerous-type and gather-type predicates boils down to whether they apply to sums or to groups.

Numerous-type predicates violate the presupposition of *all* at the lexical level but gather-type predicates never do. Only numerous-type predicates can apply to

a proper sum without applying to the parts of this sum, and this is prohibited by the presupposition of *all*. A repair strategy is available that applies the D operator to a numerous-type predicate. The output satisfies the presupposition of *all*, but it only applies to events whose agents are pure atoms such as committees and armies, or sums of such atoms. Gather-type predicates apply to groups, which are modeled as atoms. These atoms by definition do not have any proper parts. Gather-type predicates only apply to proper sums when they apply to each atomic part of that sum, by virtue of lexical cumulativity. The presupposition of *all* is therefore fulfilled by any gather-type predicate.

9.5.3 *Gather* distinguishes between *each* and *all*

The previous section has explained the behavior of *all* in connection with numerous-type predicates. This section concentrates on the distinctions between *all* and *each*. As discussed in Section 9.2.2, the former is compatible with gather-type predicates on a collective interpretation, but the latter is not. The following example is adapted from example (8):

(59) a. All the students gathered.
      b. *Each student gathered.

I see two possible strategies to explain this contrast. First, we can claim that *every/each* obligatorily undergoes quantifier raising and that *all* can be interpreted in situ. This strategy is applied in Zweig (2008, 2009) as presented and extended in Section 9.4.2 in order to account for the fact that *all* but not *every/each* is able to license dependent plurals. While such an account is plausible for noun phrases headed by the determienrs *every/each* and *all*, it is less plausible for the difference between adverbial *each* and *all*, which is analogous to (59):

(60) a. The students all gathered.
      b. *The students each gathered.

There do not seem to be any syntactic arguments that adverbial *each* undergoes movement (Zimmermann 2002). To make headway into explaining the difference between adverbial *all* and *each*, I therefore pursue a different strategy.

As Chapters 4 and 5 have argued, any theory of distributivity should include a granularity parameter, which indicates the things over which distributive items distribute. We can now use this granularity parameter to express that *every* and *each* distribute over pure atoms, while *all* distributes over atoms no matter whether they are pure or impure. On its group interpretation, the noun phrase *the students* refers to an impure atom, so it is compatible with *all* but not with *each*.

To implement this assumption, I assume that *all* and *each* have the following entries. Entry (61) is from Section 4.7, slightly simplified for ease of exposition.

(61) $[\![\text{each}_{ag}]\!] = \lambda P_{\langle vt\rangle}\lambda e : \mathrm{SR}_{ag,\,PureAtom}(P).P(e)$

(62) $[\![\text{all}_{ag}]\!] = \lambda P_{\langle vt\rangle}\lambda e : \mathrm{SR}_{ag,\,Atom}(P).P(e)$ $\qquad\qquad = (26)$

The presupposition of (61) expands as follows:

(63) $\mathrm{SR}_{ag,\,PureAtom}(\lambda e[P(e)])$

$\Leftrightarrow \forall e[^*\text{walk}(e) \to e \in {}^*\lambda e' \begin{pmatrix} P(e')\ \wedge \\ \mathrm{PureAtom}(\mathrm{ag}(e')) \end{pmatrix}]$

(Every event in P can be divided into parts ("strata") which are in P and whose agents are pure atoms.)

The presupposition of (62) expands as follows:

(64) $\mathrm{SR}_{ag,\,Atom}(\lambda e[P(e)])$

$\Leftrightarrow \forall e[^*\text{walk}(e) \to e \in {}^*\lambda e' \begin{pmatrix} P(e')\ \wedge \\ \mathrm{Atom}(\mathrm{ag}(e')) \end{pmatrix}]$

(Every event in P can be divided into parts ("strata") which are in P and whose agents are atoms.)

We can now explain the contrast in (60). The predicate *gather*, and likewise any other gather-type predicate, satisfies the presupposition of *all* but not the one of *each*. It satisfies the presupposition of *all* because it is thematic collective, and as such it only applies to events whose agents are atoms, apart from the effect of lexical cumulativity. It fails to satisfy the presupposition of *each* because these agents are not guaranteed to be pure atoms. For example, suppose that the sentence *The boys gathered* is true. Its representation (65) involves an event e whose agent is the group individual $\uparrow (\bigoplus \text{boy})$. This group is derived via the group formation operator $\uparrow$, so it is an impure atom (see Section 2.8).

(65) The boys gathered.
$\exists e[^*\text{gather}(e)\ \wedge\ {}^*\text{ag}(e) = \uparrow (\bigoplus \text{boy})]$ $\qquad\qquad = (47)$

Of course, *gather* does sometimes cooccur with *each*, namely when the subject noun phrase is a group noun:

(66) The committees each gathered.

In this case, *gather* can only be interpreted distributively: Sentence (66) entails for any one of the committee, that this committee gathered. Similarly, sentence (67)

is ambiguous between a collective reading, on which the committees come together, and a distributive interpretation, on which each committee gathers separately:

(67)   The committees all gathered.

Such cases are expected based on the result in Chapter 8. As shown there, the D operator is predicted to behave as a repair strategy that minimally changes a predicate so that it satisfies stratified reference. This prediction is a consequence of Theorem (31), repeated here:

(68)   **Theorem: $D_{\theta,C}$ is a repair strategy**
$\forall P \forall \theta \forall C \forall C'[C \subseteq C' \to \mathrm{SR}_{\theta,C'}(D_{\theta,C}(P))]$
(When the D operator coindexed with thematic role θ and with granularity threshold C is applied to any predicate, the result always has stratified reference with respect to θ and C' for any threshold C' that is at least as coarse as C.)

The following facts are a direct consequence of this theorem. Fact (70) follows given that *PureAtom $\subseteq$ Atom*.

(69)   **Fact: The output of D satisfies the presupposition of *each***
$\forall P[\mathrm{SR}_{\mathrm{ag,PureAtom}}(D_{\mathrm{ag,PureAtom}}(P))]$
(When the D operator coindexed with thematic role *ag* and with granularity threshold *PureAtom* is applied to any predicate, the result always has stratified reference with respect to *ag* and *Atom*.)

(70)   **Fact: The output of D satisfies the presupposition of *all***
$\forall P[\mathrm{SR}_{\mathrm{ag,Atom}}(D_{\mathrm{ag,PureAtom}}(P))]$
(When the D operator coindexed with thematic role *ag* and with granularity threshold *PureAtom* is applied to any predicate, the result always has stratified reference with respect to *ag* and *Atom*.)

These facts are relevant here because they have the consequence that the D operator can rescue sentences like (66) from presupposition failure by applying to *gather* before it combines with *each*. For this purpose, I assume that D is coindexed with the thematic role *ag* and that its granularity parameter is set to *PureAtom*. As discussed in Chapter 8, I assume that this value is always available even without supporting context. Applying the D operator with its parameters instantiated in this way to *gather* yields the following predicate:

(71) $[\![D_{\mathrm{ag,PureAtom}}(\mathrm{gather})]\!] = \lambda e[e \in {}^{*}\lambda e' \begin{pmatrix} {}^{*}\mathrm{gather}(e') \wedge \\ \mathrm{PureAtom}(\mathrm{ag}(e')) \end{pmatrix}]$
(Takes an event predicate P and returns a predicate that holds of any event

e which consists entirely of events that are in P and whose θs satisfy the predicate C.)

This predicate can now be used together with the sum interpretation of *the committees* to derive an interpretation for (66) and also to derive the distributive interpretation of (67):

(72)   The committees each gathered.

$$\exists e[{}^*\mathrm{ag}(e) = \bigoplus \mathrm{committee} \ \wedge \ e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\mathrm{gather}(e') \ \wedge \\ \mathrm{PureAtom}(\mathrm{ag}(e')) \end{array} \right)]$$

In this case, *gather* has a distributive reading: sentence (66) entails for any one of the committee, that this committee gathered. Similarly, sentence (67) is ambiguous between a collective reading, on which the committees come together, and a distributive interpretation, on which each committee gathers separately. The collective reading is derived by interpreting the subject on its group interpretation and by not applying the D operator to the verb phrase. The distributive reading is derived by interpreting the subject on its sum interpretation and by applying the D operator, instantiated as described above, to the verb phrase:

(73)   The committees all gathered.
   a.   **Collective reading:**
        $\exists e[{}^*\mathrm{ag}(e) = \uparrow (\bigoplus \mathrm{committee}) \ \wedge \ {}^*\mathrm{gather}(e)]$
        (There is a gathering event whose agent is the group of all committees.)
   b.   **Distributive reading:**
        $$\exists e[{}^*\mathrm{ag}(e) = \bigoplus \mathrm{committee} \ \wedge \ e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\mathrm{gather}(e') \ \wedge \\ \mathrm{PureAtom}(\mathrm{ag}(e')) \end{array} \right)]$$
        (There is a plural gathering event whose agent is the sum of all committees and which consists of gathering events whose agents are pure atoms.)

By cumulativity of thematic roles, the latter representation entails that each such pure atom is a committee.

9.6   Winter (2001) on *all*

Let me conclude by discussing one particularly significant piece of previous work on *all*. Winter (2001) develops a sophisticated framework in which *all* is the plural counterpart of *every*. This claim is made in the context of a larger type-driven theory of quantification, coordination, distributivity and predicative interpretations. Winter (2002) is a summary of the relevant parts of his analysis. Here, I briefly

review the parts that concern *all*. Winter's system is quite different from the present one because he assumes that verb phrases can have different types. For Winter, plural quantifiers like *all* are derived from singular quantifiers like *each* by a type shifter called *dfit*, and, as a result, plural quantifiers expect predicates of a different type as arguments than singular quantifiers do. Winter's translation for *all students* is given in (74). Like myself, Winter does not distinguish between different forms like *all students*, *all the students*, and *all of the students*.[38]

(74) ⟦all students⟧
= $\lambda S_{\langle et,t \rangle}.\forall x[\text{student}(x) \to \exists P[S(P) \ \wedge \ x \in P \ \wedge$
$\forall y[y \in P \to \text{student}(y)]]]$

To represent the fact that *all* and *every* are compatible with different types of predicates, Winter assumes that a verb phrase that is uninflected for number has one of two different types: type $\langle et \rangle$ (his atom predicates) and type $\langle et, t \rangle$ (his set predicates). *Each* expects an atom predicate and *all* expects a set predicate. I have introduced Winter's atom and set predicates in Section 4.2.3. Winter makes different assumptions about uninflected and inflected predicates. Uninflected atom predicates include all lexically distributive predicates as well as numerous-type predicates. Uninflected set predicates are the gather-type predicates. Number inflection shifts the types of verb phrases in different ways: Singular inflection maps atom predicates to themselves and restricts set predicates to their singleton sets, while plural inflection maps atom predicates to their powersets (minus the empty set) and set predicates to themselves. These assumptions are illustrated below in (75) and (76). Since powersets have the type of set predicates, this means that all plural verb phrases are set predicates. Winter notes that powerset formation (the effect of plural morphology on atom predicates) can be thought of as Link's D operator, but the equivalence is not perfect: Link's D operator can apply to any verb phrase predicate, but Winter's powerset formation can only apply to atom predicates.

(75) VPs that start out as atom predicates: e.g. sleep$_{\langle et \rangle}$, be-numerous$_{\langle et \rangle}$

---

[38]Here is the complete derivation in Winter's framework:

(i) a. $dfit = \lambda D_{\langle et, \langle et,t \rangle \rangle} \lambda A_{\langle et,t \rangle} \lambda B_{\langle et,t \rangle}.D(\bigcup A)(\bigcup(A \cap B))$
   b. ⟦every⟧ = $\lambda P_{\langle et \rangle} \lambda Q_{\langle et \rangle}.P \subseteq Q$
   c. ⟦all⟧ = $dfit(⟦\text{every}⟧) = \lambda A_{\langle et,t \rangle} \lambda B_{\langle et,t \rangle}.\bigcup A \subseteq \bigcup(A \cap B)$
   d. ⟦*⟧ $= \lambda P_{\langle et \rangle} \lambda Q_{\langle et \rangle}.Q \subseteq P$
   e. ⟦students⟧ $= {}^*⟦\text{student}⟧ = \lambda Q.Q$ is a set of students
   f. ⟦all students⟧ $= \lambda B_{\langle et,t \rangle}.\bigcup {}^*⟦\text{student}⟧ \subseteq \bigcup({}^*⟦\text{student}⟧ \cap B)$
      = $\lambda S_{\langle et,t \rangle}.\forall x[\text{student}(x) \to \exists P[S(P) \ \wedge \ P(x) \ \wedge \ \forall y[P(y) \to \text{student}(y)]]]$

a. $[\![\text{sleep}_{sg}]\!] = \text{sleep}_{\langle et\rangle}$ *atom predicate*
b. $[\![\text{sleep}_{pl}]\!] = \lambda Q_{\langle et\rangle}[\exists x Q(x) \;\wedge\; \forall y[Q(y) \rightarrow \text{sleep}_{\langle et\rangle}(y)]]$ *set predicate*

(76) VPs that start out as set predicates: e.g. $\text{meet}_{\langle et,t\rangle}$, $\text{gather}_{\langle et,t\rangle}$

a. $[\![\text{meet}_{sg}]\!] = \lambda x.\text{meet}_{\langle et,t\rangle}(\{x\})$ *atom predicate*
b. $[\![\text{meet}_{pl}]\!] = \text{meet}_{\langle et,t\rangle}$ *set predicate*

Winter's translation of *all students* requires a verb phrase of type $\langle et, t\rangle$, that is, a set predicate. In case this argument is shifted from an uninflected atom predicate like *sleep*, the semantics of *all* undoes the effect of the shift, so that the sentence is equivalent to its counterpart with *each*.

(77) $[\![\text{all students slept}]\!]$
$= \forall x[\text{student}(x) \rightarrow \exists P[P(x) \;\wedge\; \forall y[P(y) \rightarrow \text{student}(y) \;\wedge\; \text{sleep}_{\langle et\rangle}(y)]]$
$= \forall x[\text{student}(x) \rightarrow \text{sleep}_{\langle et\rangle}(x)]$
(Every student is in a set of students each of whom slept, that is, every student slept.)

By contrast, when *all students* combines with a verb phrase that starts out as a set predicate, the effect is different from *each*:

(78) $[\![\text{all students met}]\!]$
$= \forall x[\text{student}(x) \rightarrow \exists P[\text{meet}_{\langle et,t\rangle}(P) \;\wedge\; P(x) \;\wedge\; \forall y[P(y) \rightarrow \text{student}(y)]]]$
(Every student is in a set which met and which consists only of students.)

Winter's system is technically ingenious and is much more ambitious than the small excerpt I have presented here. To be sure, it begs the question: why do predicates denoted by VPs come in different types? As I have discussed in Section 4.2.3, Winter offers a criterion based on compatibility with *all*, *each* and with similar quantifiers. Since Winter relates the atom-set opposition to compatibility with the very same quantifiers on which he bases his account, his criterion is circular: the different behavior of *be numerous* and *gather* with respect to quantifiers is derived from their uninflected forms having different types, and these types are in turn assigned as a result of a criterion that tests for compatibility with quantifiers. I have tentatively related the empirical distinction between *be numerous* and *gather* to the conceptual distinction between nonthematic and thematic collectivity. However, I believe that this difference between Winter's and my own account is not crucial, since he could equally appeal to this conceptual distinction.

Mixed predicates such as *build a raft* pose a problem for Winter's account. A sentence like (79) has both a distributive and a collective reading. As noted in Landman (2000), the difference between the distributive and the collective readings cannot be reduced to a scope difference (see also Section 2.8). The collective reading

of (79) does not entail that each girl built a raft. The distributive reading does have that entailment, no matter whether *a raft* takes wide scope (in which case the same raft is built and disassembled repeatedly) or narrow scope (in which case each girl builds a raft by herself).

(79)   All the girls built a raft.   ✓ *distributive*, ✓ *collective*

To derive this ambiguity, Winter has no choice but to assign different types to *build a raft* (and indeed he mentions similar predicates like *lift a piano* both in his list of atom predicates and in his list of set predicates). This is for the following reason:

- To derive the distributive reading, Winter needs to apply his equivalent of the D operator, that is, powerset formation. But unlike Link's D operator, Winter's powerset formation changes the semantic type of the predicate to which it applies, namely from atom predicate to set predicate. To be able to apply the D operator to *build a raft*, he therefore needs to assume that it is an atom predicate. It is a necessary feature of Winter's system that the D operator changes semantic types. Since he works in a setup where pluralities have a different type than singular individuals, it is not possible for him to formulate a D operator that does not change the semantic type of the predicate to which it applies.

- To derive the collective reading, Winter needs to generate *build a raft* as a set predicate. Since *all the girls* requires a set predicate, he cannot reuse the atom predicate version of *build a raft* directly, and he cannot lift it because the only atom-to-set lift available in Winter's system is powerset formation. Nor is it possible to enrich Winter's system by adding another atom-to-set lift that would not bring along distributivity: if such a lift was available, nothing would stop predicates like *be numerous* to shift from atom type to set type and then freely combine with *all*.

Against this background, consider what needs to be done in Winter's framework to model the ambiguity of sentence (79). Winter does not say much about the internal composition of verb phrases in his system, but it seems unavoidable for him to assume the existence of two logical constants $\text{build}_{\langle et,t \rangle}$ and $\text{build}_{\langle e, \langle et,t \rangle \rangle}$, as follows:

(80)   Distributive reading: starting from the predicate $\text{build}_{\langle et,t \rangle}$
   a. $[\![\text{build a raft}_{sg}]\!] = \lambda x. \exists y [\text{raft}(y) \wedge \text{build}_{\langle et,t \rangle}(y)(x)]$ *atom predicate*
   b. $[\![\text{build a raft}_{pl}]\!] = \lambda Q_{\langle et \rangle} [\exists x Q(x)] \wedge$
        $\forall x [Q(x) \rightarrow \exists y [\text{raft}(y) \wedge \text{build}_{\langle et,t \rangle}(y)(x)]]$ *set predicate*

(81)   Collective reading: starting from the predicate build$_{\langle e,\langle et,t\rangle\rangle}$
        a.   [[build a raft$_{sg}$]] = $\lambda x.\exists y[\text{raft}(y) \,\wedge\, \text{build}_{\langle e,\langle et,t\rangle\rangle}(y)(\{x\})]$ *atom pred.*
        b.   [[build a raft$_{pl}$]] = $\lambda Q_{\langle et\rangle}[\exists y[\text{raft}(y) \,\wedge\, \text{build}_{\langle e,\langle et,t\rangle\rangle}(y)(Q)]]$ *set pred.*

By contrast, on my account, the ambiguity of sentence (79) is simply a matter of whether or not the D operator applies to its verb phrase. The collective reading of (79) has clear non-inductive implications. For example, it entails that the girls coordinated their actions and were jointly responsible for the result. The predicate *build a raft* is therefore a thematic collective predicate like *gather*. As discussed in Section 9.5.3, thematic collective predicates satisfy the presupposition of *all* both with and without application of the D operator.

9.7   Summary

In this chapter, I have proposed that *all* is a distributive item. Like *each*, it requires the verbal predicate with which it combines to have stratified reference with respect to its own thematic role. Unlike *each*, the granularity parameter of *all* is set to *Atom* rather than *PureAtom*.

Modeling *all* as a distributive item provides us with an explanation of its puzzling scopal behavior. Like *for*-adverbials, it cannot give rise to cumulative readings, except when the other quantifier is a bare NP. Both *for* and *all* reject predicates like *see thirty zebras* that would give rise to cumulative readings. I have shown that these predicates do not fulfill the stratified reference presuppositions imposed by *all* and by *for*. Predicates with bare NPs, like *fly kites*, satisfy stratified reference and give rise to cumulative readings. I have shown that *for*-adverbials license dependent plurals, and that this fact can be accounted for by treating bare NPs as in Zweig (2008, 2009). While Zweig's theory, taken by itself, wrongly predicts that all dependent plural licensers can give rise to cumulative readings, stratified reference provides an independent factor that explains why *all* and *for*-adverbials block cumulative readings even though they license dependent plurals.

The facts concerning cumulative readings and dependent plurals have not often been discussed in previous work on *all* and *for*. As far as I know, the present account is the first to establish a formal connection between the scopal behavior of *all* and of *for*-adverbials. The literature on *all* has instead concentrated on why *all* is compatible with certain collective predicates like *gather*, but rejects others like *be numerous*. I have proposed an account of these two different types of collectivity in terms of the sum-group distinction known from Link (1984) and Landman (1989), but my assumptions differ from these authors. A numerous-type predicate can apply to an event whose agent is the sum $a \oplus b$ even if it does not apply to any

event whose agent is a. Neither distributive predicates nor gather-type predicates allow this state of affairs. However, a gather-type predicate can apply to an event whose agent is the group or "impure atom" $\uparrow (a \oplus b)$ even if it does not apply to any event whose agent is a. Distributive predicates do not apply to events with impure atoms as agents.

Numerous-type predicates violate the stratified reference presuppositions both of *each* and of *all*. This explains why *each* and *all* do not allow these predicates to have collective readings. Gather-type predicates violate the presupposition of *each* but not of *all*. This explains why *all*, but not *each*, allows them to have collective readings.

The D operator can rescue both numerous-type and gather-type predicates from presupposition failure. The result is a distributive reading, which is compatible both with *each* and with *all* but only if the subject of the sentence is a group noun. Following Barker (1992) and Winter (2001), I assume that group nouns like *committee* and *army* apply to pure atoms. Since a committee can be numerous and can gather, this means that even numerous-type and gather-type predicates have some events in their denotations whose agents are pure atoms.

Unlike accounts such as Dowty (1987) and Winter (2001), the present account does not build on the fact that the noun phrase *all the boys* is often understood as involving reference to more boys than the definite plural *the boys*. The behavior of *all* is not linked to these maximality effects. I have argued that this is a welcome fact because other quantifiers such as *most* do not involve maximality effects but pattern with *all* in other respects. It remains to be seen whether this is the right conclusion.

Unlike Winter (2001), the present account explains the difference between numerous-type and gather-type predicates without assuming that there is a type-theoretic distinction between the two. Winter assumes that gather-type collective predicates apply to sets, while numerous-type and distributive predicates apply to atoms. This makes it difficult for him to treat predicates like *build a raft*. These predicates have distributive as well as gather-type collective interpretations, and Winter must assign them two different types.

The gather-numerous opposition has resisted many attempts to establish reliable and independent criteria for membership in its two classes. While Winter (2001) uses a circular criterion to decide whether a given predicate is numerous-type or gather-type, I have tried to link this opposition to two different notions of collectivity. In order to justify the idea that numerous-type predicates apply to sums where gather-type predicates apply to groups, I have suggested that numerous-type predicates do not have any "non-inductive", thematic entailments in the sense of Landman (2000), while gather-type predicates do. Thematic entailments are a slippery and ill-defined concept, and it remains to be seen whether

my suggestion is ultimately tenable. Collective responsibility is one of several sufficient conditions that Landman mentions for group predication, but it is not a necessary condition. In fact, Landman does not give a set of necessary and sufficient conditions for identifying group predication. As he himself admits, this is a weakness in his account, and I inherit this weakness. The present account can be falsified by sentences in which a predicate that involves collective responsibility is rejected by *all* and thereby patterns with numerous-type predicates. It cannot be falsified, however, by sentences in which a predicate that does not involve collective responsibility is accepted by *all* on its collective interpretation. Such a predicate could always be assumed to involve group predication without contradicting Landman's theory or mine. This situation is not a very attractive, but I believe it is still preferable to giving neither sufficient nor necessary conditions for the relevant predicate classes.

Chapter 10

Conclusion

If this work is on the right track, then distributivity is ubiquitous. We just need to recognize it when it presents itself in unusual ways. I have made the case for this idea using *each*, *all*, *for*-adverbials and pseudopartitives. Now that we know what we are looking for, it should be easy to find more distributive constructions. Here are some possible places to look:

- With respect to distributive constructions in the traditional sense, I have only scratched the surface. I have completely ignored distance distributivity, in which a distributive item like *each* occurs remotely from the position in which it is compositionally interpreted (Zimmermann 2002). German and Japanese split quantifier constructions, in which a quantifier appears in adverbial position apart from the noun phrase over which it quantifies, are similar to adverbial-*each* distributive constructions in that they are incompatible with collective interpretations, and they are similar to pseudopartitive constructions in that their measure functions are subject to the same monotonicity constraint (Nakanishi 2004).

- I have only considered pseudopartitives like *three liters of water*. As discussed by Schwarzschild (2002, 2006), true partitives like *three liters of the water* and comparatives like *more water* are subject to the same constraint on measure functions as pseudopartitives. An extension of the present account to true partitives is straightforward if we assume that the constituent *of the water* has divisive reference, stratified reference, or whatever is the relevant property of the substance nominal of pseudopartitives. However, the assumption that the *of*-PP has divisive reference is not controversial: Ladusaw (1982) and many accounts that follow him adopt it, but Matthewson (2001) argues against it.

- *For*-adverbials are not the only examples of aspectually sensitive constructions. As argued in Hitzeman (1991, 1997), *until* is also sensitive to the atelic-telic distinction. The same appears to be true for *since*, though the situation is more complicated here. In English, *since* requires the Perfect, which is often analyzed as introducing an Extended Now interval (Dowty 1979; von Stechow 2002). This muddles the picture, but once we move to German, where the equivalent *seit* does not require the Perfect, we see the correlation emerge: "An Extended Now Perfect modified by *since* α may embed any aktionsart. German perfects modified by *seit* α may have these readings, though they are a bit marked. In contrast to English, *seit* α may combine with simple tenses as well, but then it behaves differently. The aktionsart modified *must be a state or an activity*" (von Stechow 2002, emphasis mine).

- The scopal behavior of generics appears to be analogous to that of *for*-adverbials, as shown by the following pair (example (1a) is from Rimell (2004)). This suggests that the generic quantifier might carry a stratified reference presupposition.

(1) a. He drinks beer/#a beer/#three beers/#a pint of beer.
    b. Last night, he drank beer/#a beer/#three beers/#a pint of beer for an hour.

Appendix

This Appendix contains the proof of the two following theorems from Sections 8.2 and 8.3:

(1) **Theorem: D_θ is a repair strategy**
$\forall P \forall \theta [\text{SDR}_\theta(D_\theta(P))]$
(When the D operator coindexed with the thematic role θ is applied to any predicate, the result always has stratified distributive reference with respect to θ).

(2) **Theorem: $D_{\theta,C}$ is a repair strategy**
$\forall P \forall \theta \forall C \forall C'[C \subseteq C' \rightarrow \text{SR}_{\theta,C'}(D_{\theta,C}(P))]$
(When the D operator coindexed with thematic role θ and with granularity threshold C is applied to any predicate, the result always has stratified reference with respect to θ and C' for any threshold C' that is at least as coarse as C.)

These theorems refer to the definitions repeated here from Chapters 4 and 8:

(3) **Definition: Atomic event-based D operator**
$[\![D_\theta]\!] = \lambda P_{\langle vt \rangle} \lambda e [e \in {}^*\lambda e' \begin{pmatrix} P(e') \wedge \\ \text{PureAtom}(\theta(e')) \end{pmatrix}]$
(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs are pure atoms.)

(4) **Definition: Generalized event-based D operator**
$[\![D_{\theta,C}]\!] = \lambda P_{\langle vt \rangle} \lambda e [e \in {}^*\lambda e' \begin{pmatrix} P(e') \wedge \\ \text{C}(\theta(e')) \end{pmatrix}]$
(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs satisfy the predicate C.)

(5) **Definition: Stratified distributive reference**
$\text{SDR}_\theta(P) \stackrel{\text{def}}{=} \forall e[P(e) \rightarrow e \in {}^*\lambda e' \begin{pmatrix} P(e') \wedge \\ \text{PureAtom}(\theta(e')) \end{pmatrix}]$

(An event predicate P has stratified distributive reference with respect to a thematic role θ if and only if every event e to which P applies can be exhaustively divided into one or more subevents ("strata") to which P also applies and whose θ is a pure atom.)

(6)   **Definition: Stratified reference**

$$\mathrm{SR}_{f,\varepsilon(K)}(P) \stackrel{\mathrm{def}}{=} \forall x[P(x) \to x \in {}^*\lambda y \left(\begin{array}{c} P(y) \wedge \\ \varepsilon(K)(f(y)) \end{array} \right)]$$

(A predicate P has stratified reference with respect to a function f and a threshold $\varepsilon(K)$ if and only if there is a way of dividing every entity in its denotation exhaustively into parts ("strata") which are each in P and which have a very small f-value. Very small f-values are those that satisfy $\varepsilon(K)$.)

Note that stratified distributive reference is the limiting case of stratified reference where $\varepsilon(K) = \textit{PureAtom}$, and the atomic D operator is the limiting case of the generalized D operator where $C = \textit{PureAtom}$. Theorem (1) is therefore a special case of Theorem (2), and strictly speaking, it would be sufficient to prove Theorem (2). Nevertheless, for clarity I give both proofs separately. They are essentially parallel.

Proof of Theorem (1)

To prove Theorem (1), we start with the following tautology:

(7) $\forall P \forall \theta \forall e[[e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \mathrm{PureAtom}(\theta(e')) \end{array} \right)]$

   $\to [e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \mathrm{PureAtom}(\theta(e')) \end{array} \right)]]$

We rewrite (7) as follows:

(8) $\forall P \forall \theta \forall e[[e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \mathrm{PureAtom}(\theta(e')) \end{array} \right)]$

   $\to [e \in {}^*\lambda e' \left([e' \in \lambda e'' \left(\begin{array}{c} P(e'') \wedge \\ \mathrm{PureAtom}(\theta(e'')) \end{array} \right)] \wedge \mathrm{PureAtom}(\theta(e')) \right)]]$

From Theorem (18) in Section 2.3.1, we know that $\forall e[P(e) \to {}^*P(e)]$. Using this fact, we rewrite (8) as follows:

(9) $\forall P \forall \theta \forall e[[e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ \mathrm{PureAtom}(\theta(e')) \end{array} \right)]$

$$\to [e \in {}^*\lambda e' \left(\ [e' \in {}^*\lambda e'' \left(\begin{array}{c} P(e'') \wedge \\ \mathrm{PureAtom}(\theta(e'')) \end{array} \right)] \ \wedge \ \mathrm{PureAtom}(\theta(e')) \ \right)]]$$

By two applications of the definition of D_θ, we rewrite (9) as follows:

(10) $\forall P \forall \theta \forall e [D_\theta(P)(e) \to e \in {}^*\lambda e' \left(\ D_\theta(P)(e') \ \wedge \ \mathrm{PureAtom}(\theta(e')) \ \right)]$

Theorem (1) follows from (10) by the definition of stratified distributive reference. End of proof.

Proof of Theorem (2)

To prove Theorem (2), we start with the following tautology:

(11) $\forall P \forall \theta \forall e \forall C \forall C'[[C \subseteq C' \ \wedge \ e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ C(\theta(e')) \end{array} \right)]$
$\to [e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ C(\theta(e')) \ \wedge \ C'(\theta(e')) \end{array} \right)]]$

We rewrite (11) as follows:

(12) $\forall P \forall \theta \forall e \forall C \forall C'[C \subseteq C' \to [[e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ C(\theta(e')) \end{array} \right)]$
$\to [e \in {}^*\lambda e' \left(\ [e' \in \lambda e'' \left(\begin{array}{c} P(e'') \wedge \\ C(\theta(e'')) \end{array} \right)] \ \wedge \ C'(\theta(e')) \ \right)]]]$

From Theorem (18) in Section 2.3.1, we know that $\forall e [P(e) \to {}^*P(e)]$. Using this fact, we rewrite (12) as follows:

(13) $\forall P \forall \theta \forall e \forall C \forall C'[C \subseteq C' \to [[e \in {}^*\lambda e' \left(\begin{array}{c} P(e') \wedge \\ C(\theta(e')) \end{array} \right)]$
$\to [e \in {}^*\lambda e' \left(\ [e' \in {}^*\lambda e'' \left(\begin{array}{c} P(e'') \wedge \\ C(\theta(e'')) \end{array} \right)] \ \wedge \ C'(\theta(e')) \ \right)]]]$

By two applications of the definition of $D_{\theta,C}$, we rewrite (13) as follows:

(14) $\forall P \forall \theta \forall e \forall C \forall C'[C \subseteq C' \to [D_{\theta,C}(P)(e) \to$
$e \in {}^*\lambda e' \left(\ D_{\theta,C}(P)(e') \ \wedge \ C'(\theta(e')) \ \right)]]$

Theorem (2) follows from (14) by the definition of stratified reference. End of proof.

Bibliography

Abusch, Dorit and Mats Rooth (1990). "Temporal adverbs and the English Perfect". In: *Proceedings of the Twentieth Conference of the North Eastern Linguistics Society (NELS 20)*. University of Massachusetts. Amherst, MA: GLSA (Graduate Linguistic Student Association), pp. 5–19.

Akmajian, Adrian and Adrienne Lehrer (1976). "NP-like quantifiers and the problem of determining the head of an NP". *Linguistic Analysis* 2.4, pp. 395–413.

Bach, Emmon (1986). "The algebra of events". *Linguistics and Philosophy* 15, pp. 5–16.

Bale, Alan (2009). "Yet more evidence for the emptiness of plurality". In: *Proceedings of the Thirty-Eighth Conference of the North Eastern Linguistics Society (NELS 38)*.

Barker, Chris (1992). "Group terms in English: representing groups as atoms". *Journal of Semantics* 9.1, pp. 69–93.

Barner, David and Jesse Snedeker (2005). "Quantity judgments and individuation: evidence that mass nouns count". *Cognition* 97, pp. 41–66.

Barwise, Jon and Robin Cooper (1981). "Generalized quantifiers and natural language". *Linguistics and Philosophy* 4, pp. 159–219.

Bayer, Samuel L. (1997). *Confessions of a Lapsed Neo-Davidsonian: Events and Arguments in Compositional Semantics*. New York, NY: Garland.

Bealer, George (1979). "Predication and Matter". In: *Mass terms: Some philosophical problems*. Ed. by Francis Jeffry Pelletier. Vol. 6. Studies in Linguistics and Philosophy. Berlin, Germany: Springer, pp. 279–294.

Beavers, John (2008). "Multiple incremental themes and Figure/Path relations". In: *Proceedings of the Eighteenth Conference on Semantics and Linguistic Theory (SALT XVIII)*.

Bennett, Michael R. (1974). "Some extensions of a Montague fragment of English". PhD thesis. Los Angeles, CA: University of California.

Bennett, Michael R. and Barbara H. Partee (1972). *Toward the logic of tense and aspect in English.* Indiana University Linguistics Club.

van Benthem, Johan (1983). *The logic of time.* Dordrecht, Netherlands: Reidel.

Binnick, Robert I. (1969). "Studies in the derivation of predicative structures". PhD thesis. Chicago, IL: University of Chicago.

Blaheta, Don (2003). "Binominal *each*: evidence for a modified type system". MA thesis. Brown University.

Brasoveanu, Adrian (2009). "Measure noun polysemy and monotonicity: Evidence from Romanian pseudopartitives". In: *Proceedings of the Thirty-Eighth Conference of the North Eastern Linguistics Society (NELS 38).* Amherst, MA: GLSA (Graduate Linguistic Student Association, University of Massachusetts).

Brasoveanu, Adrian (2010). "Modified Numerals as Post-Suppositions". In: *Logic, Language and Meaning.* Ed. by Maria Aloni et al. Vol. 6042. Lecture Notes in Computer Science. Berlin, Germany: Springer, pp. 203–212.

Breheny, Richard (2008). "A new look at the semantics and pragmatics of numerically quantified noun phrases". *Journal of Semantics* 25, pp. 93–139.

Brisson, Christine (1998). "Distributivity, maximality and floating quantifiers". PhD thesis. New Brunswick, NJ: Rutgers University.

Brisson, Christine (2003). "Plurals, *all*, and the nonuniformity of collective predication". *Linguistics and Philosophy* 26, pp. 129–184.

Bunt, Harry C. (1979). "Ensembles and the formal semantic properties of mass terms". In: *Mass terms: Some philosophical problems.* Ed. by Francis Jeffry Pelletier. Vol. 6. Studies in Linguistics and Philosophy. Berlin, Germany: Springer, pp. 249–277.

Bunt, Harry C. (1985). *Mass terms and model-theoretic semantics.* Cambridge, UK: Cambridge University Press.

Bunt, Harry C. (2006). "Mass expressions". In: *Encyclopedia of Language and Linguistics.* Ed. by Keith Brown. 2nd. Amsterdam, Netherlands: Elsevier.

Carlson, Gregory N. (1977). "Reference to Kinds in English". PhD thesis. Amherst, MA: University of Massachusetts.

Carlson, Gregory N. (1984). "Thematic roles and their role in semantic interpretation". *Linguistics* 22, pp. 259–279.

Carlson, Gregory N. (1989). "The semantic interpretation of generic sentences". In: *Semantic issues.* Ed. by Gennaro Chierchia, Barbara H. Partee, and Ray-

mond Turner. Vol. 2. Property theory, type theory, and semantics. Dordrecht, Netherlands: Kluwer, pp. 167–192.

Carlson, Gregory N. (1998). "Thematic roles and the individuation of events". In: *Events and grammar*. Ed. by Susan Rothstein. Dordrecht, Netherlands: Kluwer, pp. 35–51.

Carlson, Gregory N. (2006). "Generics, habituals and iteratives". In: *Encyclopedia of Language and Linguistics*. Ed. by Keith Brown. 2nd. Amsterdam, Netherlands: Elsevier.

Carnie, Andrew (2006). *Syntax: A generative introduction*. Oxford, UK: Blackwell Publishing.

Carston, Robyn (1998). "Informativeness, relevance and scalar implicature". In: *Relevance theory: Applications and implications*. Ed. by Robin Carston and Seiji Uchida. Amsterdam, Netherlands: John Benjamins, pp. 179–236.

Cartwright, Helen (1975). "Amounts and measures of amount". *Noûs* 9, pp. 143–164.

Casati, Roberto and Achille C. Varzi (1999). *Parts and places: the structures of spatial representation*. Cambridge, MA: MIT Press.

Champollion, Lucas (2009). "A unified account of distributivity, *for*-adverbials, and measure constructions". In: *Proceedings of the Fourteenth Sinn und Bedeutung conference*. Ed. by Martin Prinzhorn, Viola Schmitt, and Sarah Zobel. Universität Wien, pp. 84–100.

Champollion, Lucas (2010). "Cumulative readings of *every* do not provide evidence for events and thematic roles". In: *Logic, Language and Meaning: Proceedings of the Nineteenth Amsterdam Colloquium*. Ed. by Maria Aloni et al. Vol. 6042. Lecture Notes in Computer Science. Berlin, Germany: Springer, pp. 213–222.

Cheng, Chung-Ying (1973). "Comments on Moravcsik's paper". In: *Approaches to Natural Language*. Ed. by Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes. Dordrecht, Netherlands: Reidel, pp. 286–288.

Chierchia, Gennaro (1998a). "Plurality of mass nouns and the notion of 'Semantic Parameter'". In: *Events and Grammar*. Ed. by Susan Rothstein. Dordrecht, Netherlands: Kluwer, pp. 53–104.

Chierchia, Gennaro (1998b). "Reference to kinds across languages". *Natural Language Semantics* 6, pp. 339–405.

Chierchia, Gennaro (2006). "Broaden your views: Implications of domain widening and the "logicality" of language". *Linguistic Inquiry* 37, pp. 535–590.

Chierchia, Gennaro (2008). *On being 'one'*. Paper presented at the 2008 Workshop on Markedness and Underspecification in the Morphology and Semantics of Agreement (MUMSA).

Chierchia, Gennaro (2010). "Mass nouns, vagueness and semantic variation". *Synthese* 174, pp. 99–149.

Choe, Jae-Woong (1987). "Anti-quantifiers and a theory of distributivity". PhD thesis. Amherst, MA: University of Massachusetts.

Choe, Jae-Woong (1991). "Georgian Reduplication and a Relational Theory of Distributivity". In: *The Sixth Japanese-Korean Joint Conference on Formal Linguistics*. Ed. by Akira Ikeya. Tokyo, Japan: The Logico-Linguistic Society of Japan.

Chomsky, Noam (1981). *Lectures on government and binding*. Dordrecht, Netherlands: Foris.

Chomsky, Noam (1982). *The generative enterprise. A discussion with R. Huybregts and H. van Riemsdijk*. Dordrecht, Netherlands: Foris.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Cohen, Ariel (2005). "More than bare existence: An implicature of existential bare plurals". *Journal of Semantics* 22, pp. 389–400.

Comrie, Bernard (1976). *Aspect*. Cambridge, UK: Cambridge University Press.

Cresswell, Max J. (1976). "The semantics of degree". In: *Montague grammar*. Ed. by Barbara H. Partee. New York, NY: Academic Press, pp. 261–292.

Cresswell, Max J. (1985). "Review of Landman and Veltman (eds.), 1984, *Varieties of Formal Semantics*". *Linguistics* 23.

Daubert, B. F. and T. H. Clarke (1944). "Unsaturated Synthetic Glycerides. VI.1 Polymorphism of Symmetrical Monoöleyldisaturated Triglycerides2". *Journal of the American Chemical Society* 66.5, pp. 690–691.

Davidson, Donald (1967). "The logical form of action sentences". In: *The logic of decision and action*. Ed. by Nicholas Rescher. Pittsburgh, PA: University of Pittsburgh Press, pp. 81–95.

Dekker, Paul (1997). "Cases, adverbs, situations and events". In: *Proceedings of the workshop on context dependence*. Ed. by Hans Kamp and Barbara H. Partee. IMS, Stuttgart, Germany and UFAL, Prague, Czech Republic.

Delfitto, Denis (2005). "Bare plurals". In: *The Blackwell companion to syntax*. Ed. by Martin Everaert and Henk van Riemsdijk. Oxford, UK: Blackwell Publishing.

van der Does, Jaap and Henk J. Verkuyl (1995). "Quantification and predication". In: *Semantic ambiguity and underspecification.* Ed. by Kees van Deemter and Stanley Peters. Stanford, CA: CSLI Publications.

Doetjes, Jenny (1997). "Quantifiers and selection: on the distribution of quantifying expressions in French, Dutch and English". Volume 32 of HIL Dissertations. The Hague, Netherlands: Holland Institute of Generative Linguistics.

Dotlačil, Jakub (2010). "Anaphora and distributivity: A study of *same, different,* reciprocals and *others*". PhD thesis. Utrecht, Netherlands: Utrecht University.

Dowty, David R. (1972). "Studies in the logic of verb aspect and time reference in English". PhD thesis. Austin, TX: University of Texas.

Dowty, David R. (1979). *Word meaning and Montague grammar.* Dordrecht, Netherlands: Reidel.

Dowty, David R. (1987). "Collective predicates, distributive predicates, and *All*". In: *Proceedings of the Third Eastern States Conference on Linguistics (ESCOL 3).* Columbus, OH: The Ohio State University, pp. 97–115.

Dowty, David R. (1991). "Thematic proto-roles and argument selection". *Language* 67.3, pp. 547–619.

Eberle, Kurt (1998). "The influence of Plural NPs on the Aktionsart in DRT". In: *Plural quantification.* Ed. by Fritz Hamm and Erhard Hinrichs. Dordrecht, Netherlands: Kluwer.

Egg, Markus (1995). "Aspect and quantification - an iterative approach". In: *Proceedings of the Tenth Amsterdam Colloquium.* ILLC. Amsterdam, Netherlands.

Farkas, Donka F. and Henriëtte de Swart (2010). "The semantics and pragmatics of plurals". *Semantics and pragmatics* 3.6, pp. 1–54.

Fine, Kit (1999). "Things and their parts". *Midwest Studies in Philosophy* 23.1, pp. 61–74.

Fox, Danny and Martin Hackl (2006). "The universal density of measurement". *Linguistics and Philosophy* 29, pp. 537–586.

Gärdenfors, Peter, ed. (1987). *Generalized Quantifiers: Linguistic and Logical Approaches.* Dordrecht, Netherlands: Reidel.

Garey, Howard B. (1957). "Verbal aspect in French". *Language*, pp. 91–110.

Gawron, Jean Mark (2002). *Two kinds of determiners in DP.* Talk at LSA winter meeting.

Gawron, Jean Mark (2004). *Motion, scalar paths, and lexical aspect.* Paper presented at *Workshop on Event Structures in Linguistic Form and Interpretation, Leipzig.*

Gawron, Jean Mark (2005). "Generalized paths". In: *Proceedings of the Fifteenth Conference on Semantics and Linguistic Theory (SALT XV).*

van Geenhoven, Veerle (2004). "For-adverbials, frequentative aspect, and pluractionality". *Natural Language Semantics* 12, pp. 135–190.

Geurts, Bart (2006). "Take 'five': The meaning and use of a number word". In: *Non-definiteness and plurality.* Ed. by Svetlana Vogeleer and Liliane Tasmowski. Amsterdam, Netherlands: John Benjamins, pp. 311–329.

Gil, David (1982). "Quantifier scope, linguistic variation, and natural language semantics". *Linguistics and Philosophy* 5, pp. 421–472.

Gil, David (1989). "Nouns, verbs, and quantification". Manuscript, The Porter Institute for Poetics and Semiotics, Tel Aviv University.

Gillon, Brendan S. (1987). "The readings of plural noun phrases in English". *Linguistics and Philosophy* 10, pp. 199–219.

Gillon, Brendan S. (1990). "Plural noun phrases and their readings: A reply to Lasersohn". *Linguistics and Philosophy* 13, pp. 477–485.

Gillon, Brendan S. (1992). "Towards a common semantics for English mass and count nouns". *Linguistics and Philosophy* 15, pp. 597–640.

Graham, Daniel W. (1980). "States and performances: Aristotle's test". *The Philosophical Quarterly* 30.119, pp. 117–130.

Groenendijk, Jeroen, Theo Janssen, and Martin Stokhof, eds. (1984). *Truth, interpretation, information.* Dordrecht, Netherlands: Foris.

Gruber, Jeffrey S. (1965). "Studies in lexical relations". PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

Guéron, Jacqueline (1979). "On the structure of "pseudo-partitive" NP's". *Bulletin of the ICU Summer Institute in Linguistics* 12, pp. 21–48.

Hackl, Martin (2001). "Comparative quantifiers". PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

Hackl, Martin (2002). "The Ingredients of Essentially Plural Predicates". In: *Proceedings of the Thirty-Second Conference of the North Eastern Linguistics Society (NELS 32).* University of Massachusetts. Amherst, MA: GLSA (Graduate Linguistic Student Association), pp. 171–182.

Hackl, Martin (2009). "On the grammar and processing of proportional quantifiers: *most* versus *more than half*". *Natural Language Semantics* 17, pp. 63–98.

Harnish, Robert M. (1976). "Logical Form and Implicature". In: *An Integrated Theory of Linguistic Ability*. Ed. by Thomas G. Bever, Jerrold J. Katz, and D. Terence Langendoen. New York, NY: Thomas Y. Crowell, pp. 313–392.

Hausser, Roland (1974). "Quantification in an extended Montague grammar". PhD thesis. Austin, TX: University of Texas.

Heim, Irene (1994). "Plurals". Lecture notes, MIT.

Heim, Irene and Angelika Kratzer (1998). *Semantics in Generative Grammar*. Oxford, UK: Blackwell Publishing.

Hellan, Lars (1981). *Towards an integrated analysis of comparatives*. Tübingen, Germany: Narr.

Heny, Frank (1982). "Tense, aspect and time adverbials, Part II". *Linguistics and Philosophy* 5, pp. 109–154.

Heycock, Caroline and Roberto Zamparelli (2005). "Friends and colleagues: Plurality, coordination, and the structure of DP". *Natural Language Semantics* 13.3, pp. 201–270.

Hinrichs, Erhard (1985). "A Compositional Semantics for Aktionsarten and NP Reference in English". PhD thesis. Columbus, OH: Ohio State University.

Hitzeman, Janet (1991). "Aspect and adverbials". In: *Proceedings of the First Conference on Semantics and Linguistic Theory (SALT I)*, pp. 107–127.

Hitzeman, Janet (1997). "Semantic partition and the ambiguity of sentences containing temporal adverbials". *Natural Language Semantics* 5.2, pp. 87–100.

Hoeksema, Jack (1983). "Plurality and Conjunction". In: *Studies in Modeltheoretic Semantics*. Ed. by Alice ter Meulen. Dordrecht, Netherlands: Foris, pp. 63–83.

Horn, Lawrence (1972). "On the semantics and pragmatics of logical operators in English". PhD thesis. Los Angeles, CA: University of California.

Horn, Lawrence (1989). *A natural history of negation*. Chicago, IL: University of Chicago Press.

Hovda, Peter (2009). "What is classical mereology?" *Journal of Philosophical Logic* 38.1, pp. 55–82.

Iatridou, Sabine, Elana Anagnostopoulou, and Roumyana Izvorski (2001). "Observations about the form and meaning of the Perfect". In: *Ken Hale: a life in language*. Ed. by Michael Kenstowicz. Cambridge, MA: MIT Press, pp. 189–238.

Ionin, Tania, Ora Matushansky, and Eddy G. Ruys (2006). "Parts of speech: Towards a unified semantics for partitives". In: *Proceedings of the Thirty-Sixth Conference of the North Eastern Linguistics Society (NELS 36)*. Ed. by Christopher Davis, Amy Rose Deal, and Youri Zabbal. Amherst, MA: Graduate Linguistic Student Association.

Jackendoff, Ray S. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

Jespersen, Otto (1913). *A modern English grammar on historical principles*. Vol. 1 Part II. Copenhagen, Denmark: Munksgaard.

Kamp, Hans (1979). "Events, instants and temporal reference". In: *Semantics from different points of view*. Ed. by Rainer Bäuerle, Urs Egli, and Arnim von Stechow. Berlin, Germany: Springer, pp. 376–417.

Kamp, Hans and Uwe Reyle (1993). *From discourse to logic*. Dordrecht, Netherlands: Kluwer.

Karttunen, Lauri and Stanley Peters (1979). "Conventional Implicature". In: *Presupposition*. Ed. by Choon-Kyu Oh and David A. Dinneen. Vol. 11. Syntax and Semantics. New York, NY: Academic Press, pp. 1–56.

Keefe, Rosanna (2000). *Theories of Vagueness*. Cambridge, UK: Cambridge University Press.

Keizer, Evelien (2007). *The English noun phrase: the nature of linguistic categorization*. Cambridge, UK: Cambridge University Press.

Kennedy, Christopher (2007). "Vagueness and grammar: the semantics of relative and absolute gradable adjectives". *Linguistics and Philosophy* 30, pp. 1–45.

Kennedy, Christopher (2009). *The number of meanings of English number words*. Manuscript.

Kennedy, Christopher (2010). "Ambiguity and Vagueness: An Overview". In: *Semantics: An International Handbook of Natural Language Meaning*. Ed. by Claudia Maienborn, Klaus von Heusinger, and Paul Portner. Berlin, Germany: de Gruyter.

Kipper-Schuler, Karin (2005). "VerbNet: A broad-coverage, comprehensive verb lexicon". PhD thesis. Philadelphia, PA: University of Pennsylvania.

Klein, Ewan (1991). "Comparatives". In: *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*. Ed. by Dieter Wunderlich and Arnim von Stechow. Berlin, Germany: de Gruyter, pp. 673–691.

Krantz, David H. et al. (1971). *Foundations of measurement.* New York, NY: Academic Press.

Krasikova, Sveta (2009). "Scales in the meaning of adjectives". In: *Russian in contrast.* University of Oslo.

Kratzer, Angelika (1989). "An investigation of the lumps of thought". *Linguistics and Philosophy* 12, pp. 607–653.

Kratzer, Angelika (1996). "Severing the external argument from its verb". In: *Phrase structure and the lexicon.* Ed. by J. Rooryck and L. Zaring. Dordrecht, Netherlands: Kluwer.

Kratzer, Angelika (2000). "The Event Argument and the Semantics of Verbs, chapter 2". Manuscript. Amherst: University of Massachusetts, Amherst, MA.

Kratzer, Angelika (2001). "The Event Argument and the Semantics of Verbs, chapter 4". Manuscript. Amherst: University of Massachusetts, Amherst, MA.

Kratzer, Angelika (2003). "The Event Argument and the Semantics of Verbs, chapter 3". Manuscript. Amherst: University of Massachusetts, Amherst, MA.

Kratzer, Angelika (2007). "On the plurality of verbs". In: *Event structures in linguistic form and interpretation.* Ed. by Johannes Dölling, Tatjana Heyde-Zybatow, and Martin Schäfer. Berlin, Germany: de Gruyter, pp. 269–300.

Kratzer, Angelika (2010). "Situations in Natural Language Semantics". In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Fall 2010. Stanford, CA. URL: http://plato.stanford.edu/archives/fall2010/entries/situations-semantics/.

Krifka, Manfred (1986). *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen.* Munich, Germany (published 1989): Fink.

Krifka, Manfred (1989). "Nominal reference, temporal constitution and quantification in event semantics". In: *Semantics and contextual expression.* Ed. by Renate Bartsch, Johan van Benthem, and P. van Emde Boas. Dordrecht, Netherlands: Foris, pp. 75–115.

Krifka, Manfred (1990). "Four thousand ships passed through the lock: Object induced measure functions on events". *Linguistics and Philosophy* 13, pp. 487–520.

Krifka, Manfred (1991). "Massennomina". In: *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung / Semantics: An international handbook of*

comtemporary research. Ed. by Arnim von Stechow and Dieter Wunderlich. Berlin, Germany: de Gruyter, pp. 399–417.

Krifka, Manfred (1992). "Thematic relations as links between nominal reference and temporal constitution". In: *Lexical Matters*. Ed. by Ivan A. Sag and Anna Szabolcsi. Stanford, CA: CSLI Publications, pp. 29–53.

Krifka, Manfred (1998). "The origins of telicity". In: *Events and grammar*. Ed. by Susan Rothstein. Dordrecht, Netherlands: Kluwer, pp. 197–235.

Krifka, Manfred (1999). "At Least Some Determiners Aren't Determiners". In: *The Semantics/Pragmatics Interface from Different Points of View*. Ed. by K. Turner. Amsterdam, Netherlands: Elsevier, pp. 257–291.

Krifka, Manfred (2004). "Bare NPs: Kind-referring, Indefinites, Both, or Neither?" In: *Proceedings of the Thirteenth Conference on Semantics and Linguistic Theory (SALT XIII)*. Ed. by Robert B. Young and Yuping Zhou. Cornell University. Ithaca, NY: CLC Publications, pp. 180–203.

Krifka, Manfred et al. (1995). "Genericity: An Introduction". In: *The Generic Book*. Ed. by Gregory N. Carlson and Francis Jeffry Pelletier. Chicago, IL: University of Chicago Press, pp. 1–124.

Kroch, Anthony S. (1974). "The semantics of scope in English". PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

Kuno, Susumu (1982). "The focus of the question and the focus of the answer". In: *Papers from the Parasession on Nondeclaratives*. Chicago Linguistics Society, pp. 134–157.

Ladusaw, William (1982). "Semantic constraints on the English partitive construction". In: *Proceedings of the First West Coast Conference on Formal Linguistics (WCCFL 1)*. Ed. by Daniel P. Flickinger, Marlys Macken, and Nancy Wiegand. Stanford, CA: CSLI Publications, pp. 231–242.

Landman, Fred (1989). "Groups, I". *Linguistics and Philosophy* 12, pp. 559–605.

Landman, Fred (1991). *Structures for semantics*. Dordrecht, Netherlands: Kluwer.

Landman, Fred (1996). "Plurality". In: *Handbook of Contemporary Semantics*. Ed. by Shalom Lappin. Oxford, UK: Blackwell Publishing, pp. 425–457.

Landman, Fred (2000). *Events and plurality: The Jerusalem lectures*. Dordrecht, Netherlands: Kluwer.

Landman, Fred (2004). *Indefinites and the type of sets*. Oxford, UK: Blackwell Publishing.

Landman, Fred and Susan Rothstein (2009). "Incremental homogeneity in the semantics of aspectual *for*-phrases". In: *Syntax, lexical semantics and event structure*. Ed. by Malka Rappaport Hovav, Ivi Sichel, and Edit Doron. Amsterdam, Netherlands: John Benjamins.

Langacker, Ronald (1986). *Foundations of cognitive grammar*. Stanford, CA: Stanford University Press.

Lasersohn, Peter (1988). "A semantics for groups and events". PhD thesis. Columbus, OH: Ohio State University.

Lasersohn, Peter (1989). "On the readings of plural noun phrases". *Linguistic Inquiry* 20, pp. 130–134.

Lasersohn, Peter (1995). *Plurality, conjunction and events*. Dordrecht, Netherlands: Kluwer.

Lasersohn, Peter (1998). "Generalized distributivity operators". *Linguistics and Philosophy* 21.1, pp. 83–93.

Lasersohn, Peter (1999). "Pragmatic Halos". *Language* 75.3, pp. 522–551.

Lasersohn, Peter (2010). "Mass Nouns and Plurals". In: *Semantics: An International Handbook of Natural Language Meaning*. Ed. by Claudia Maienborn, Klaus von Heusinger, and Paul Portner. Berlin, Germany: de Gruyter.

Lassiter, Daniel (2010a). "Explaining a restriction on the scope of the comparative operator". *University of Pennsylvania Working Papers in Linguistics* 16.1.

Lassiter, Daniel (2010b). "The algebraic structure of amounts: Evidence from comparatives". In: *Interfaces: Explorations in Logic, Language, and Computation (LNAI 6211)*. Ed. by Thomas Icard and Reinhard Muskens. Berlin, Germany: Springer.

Levin, Beth (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Levinson, Stephen C. (2000). *Presumptive meanings*. Cambridge, MA: MIT Press.

Lewis, David (1972). "General semantics". In: *Semantics of natural language*. Ed. by Donald Davidson and Gilbert Harman. Dordrecht, Netherlands: Reidel, pp. 169–218.

Link, Godehard (1983). "The logical analysis of plurals and mass terms: A lattice-theoretical approach". In: *Meaning, use and interpretation of language*. Ed. by Reiner Bäuerle, Christoph Schwarze, and Arnim von Stechow. Berlin, Germany: de Gruyter, pp. 303–323.

Link, Godehard (1984). "Hydras. On the logic of relative clause constructions with multiple heads". In: *Varieties of formal semantics*. Ed. by Fred Landman and Frank Veltman. GRASS 3. Dordrecht, Netherlands: Foris.

Link, Godehard (1986). "Je drei Äpfel - three apples each: Quantification and the German 'je'". In: *Link (1998)*, pp. 117–132.

Link, Godehard (1987a). "Algebraic semantics of event structures". In: *Proceedings of the Sixth Amsterdam Colloquium*. Ed. by Jeroen Groenendijk, Martin Stokhof, and Frank Veltman. Reprinted in Link (1998), Chapter 11. University of Amsterdam, Institute for Language, Logic and Information. Amsterdam, Netherlands, pp. 243–262.

Link, Godehard (1987b). "Generalized quantifiers and plurals". In: *Generalized Quantifiers: Linguistic and Logical Approaches*. Ed. by Peter Gärdenfors. Reprinted in Link (1998), Chapter 4. Dordrecht, Netherlands: Reidel, pp. 151–180.

Link, Godehard (1991). "Quantity and number". In: *Semantic universals and universal semantics*. Ed. by Dietmar Zaefferer. Dordrecht, Netherlands: Foris, pp. 133–149.

Link, Godehard (1997). "Ten Years of Research on Plurals — Where Do We Stand?" In: *Plurality and quantification*. Ed. by Fritz Hamm and Erhard Hinrichs. Dordrecht, Netherlands: Kluwer, pp. 19–54.

Link, Godehard (1998). *Algebraic semantics in language and philosophy*. Stanford, CA: CSLI Publications.

Lønning, Jan Tore (1987). "Collective readings of definite and indefinite noun phrases". In: *Generalized Quantifiers: Linguistic and Logical Approaches*. Ed. by Peter Gärdenfors. Dordrecht, Netherlands: Reidel, pp. 203–235.

Malamud, Sophia (2006a). "(Non-)Maximality and distributivity: a decision theory approach". In: *Proceedings of the Sixteenth Conference on Semantics and Linguistic Theory (SALT XVI)*. Tokyo, Japan.

Malamud, Sophia (2006b). "Semantics and pragmatics of arbitrariness". PhD thesis. Philadelphia, PA: University of Pennsylvania.

Marantz, Alec (1984). *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

Markosian, Ned (2009). "Time". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2009. Stanford, CA. URL: http://plato.stanford.edu/archives/fall2009/entries/time/.

Matthewson, Lisa (2001). "Quantification and the nature of crosslinguistic variation". *Natural Language Semantics* 9, pp. 145–189.

McCawley, James D. (1968). "Concerning the base component of a transformational grammar". *Foundations of Language* 4, pp. 243–269.

de Mey, Sjaap (1981). "The dependent plural and the analysis of tense". In: *Proceedings of the Eleventh Conference of the North Eastern Linguistics Society (NELS 11)*. Ed. by Victoria A. Burke and James Pustejovsky. University of Massachusetts. Amherst, MA: GLSA (Graduate Linguistic Student Association).

Mills, Ian et al., eds. (2007). *Quantities, units and symbols in physical chemistry (IUPAC Green Book)*. Oxford, UK: Blackwell Science.

Milsark, Gary (1974). "Existential sentences in English". PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

Mittwoch, Anita (1988). "Aspects of English aspect: On the interaction of perfect, progressive and durational phrases". *Linguistics and Philosophy* 11, pp. 203–254.

Mollá-Aliod, Diego (1997). "Aspectual composition and sentence interpretation: A formal approach". PhD thesis. Edinburgh, UK: University of Edinburgh.

Moltmann, Friederike (1989). "Measure adverbials as part quantifiers". In: *Proceedings of the Eighth West Coast Conference on Formal Linguistics (WCCFL 8)*. Ed. by E. Jane Fee and Katherine Hunt. Stanford, CA: Stanford Linguistics Association, pp. 247–261.

Moltmann, Friederike (1991). "Measure adverbials". *Linguistics and Philosophy* 14, pp. 629–660.

Moltmann, Friederike (1997). *Parts and wholes in semantics*. Oxford, UK: Oxford University Press.

Moltmann, Friederike (1998). "Part structures, integrity, and the mass-count distinction". *Synthese* 116.1, pp. 75–111.

Mourelatos, Alexander P. D. (1978). "Events, Processes, and States". *Linguistics and Philosophy* 2, pp. 415–434.

Muskens, Reinhard (1996). *Meaning and Partiality*. Stanford, CA: CSLI Publications.

Nakanishi, Kimiko (2004). "Domains of measurement: Formal properties of non-split/split quantifier constructions". PhD thesis. Philadelphia, PA: University of Pennsylvania.

Nouwen, Rick (2006). *Natural language quantifiers*. Lecture notes for ESSLLI 2006 course.

Ojeda, Almerindo E. (2003). "A Russellian interpretation of measure nouns". In: *From NP to DP: The syntax and semantics of noun phrases*. Amsterdam, Netherlands: John Benjamins, pp. 255–276.

Parsons, Terence (1970). "An analysis of mass and amount terms". *Foundations of Language* 6, pp. 362–388.

Parsons, Terence (1987). "Underlying states in the semantical analysis of English". In: *Proceedings of the Aristotelian Society 88*, pp. 1–30.

Parsons, Terence (1990). *Events in the semantics of English*. Cambridge, MA: MIT Press.

Partee, Barbara H. (1987). "Noun phrase interpretation and type-shifting principles". In: *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*. Ed. by Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof. Dordrecht, Netherlands: Foris, pp. 115–143.

Partee, Barbara H. and Mats Rooth (1983). "Generalized conjunction and type ambiguity". In: *Meaning, use and interpretation of language*. Berlin, Germany: de Gruyter, pp. 361–383.

Partee, Barbara Hall, Alice G. B. ter Meulen, and Robert Eugene Wall (1990). *Mathematical methods in linguistics*. Vol. 30. Studies in Linguistics and Philosophy. Dordrecht, Netherlands: Kluwer.

Pearson, Hazel (2009). "Towards a semantics for group nouns". Unpublished manuscript, Harvard University.

Pearson, Hazel (2010). "A new semantics for group nouns". In: *Proceedings of the Twenty-Eighth West Coast Conference on Formal Linguistics (WCCFL 28)*.

Pelletier, Francis Jeffry and Lenhart K. Schubert (2002). "Mass expressions". In: *Handbook of Philosophical Logic*. Ed. by Dov M. Gabbay and Franz Guenthner. Vol. 10. Dordrecht, Netherlands: Kluwer, pp. 1–87.

Pianesi, Fabio (2002). "Book review: Friederike Moltmann - Parts and Wholes in Semantics". *Linguistics and Philosophy* 25, pp. 97–120.

Pianesi, Fabio and Achille C. Varzi (1996). "Events, topology, and temporal relations". *The Monist* 78, pp. 89–116.

Piñón, Christopher (1995). "An ontology for event semantics". PhD thesis. Stanford, CA: Stanford University.

Piñón, Christopher (1999a). *A semantics for durative adverbials*. Talk at Sinn und Bedeutung 99, Düsseldorf.

Piñón, Christopher (1999b). "Durative adverbials for result states". In: *Proceedings of the Eighteenth West Coast Conference on Formal Linguistics (WCCFL 18)*, pp. 420–433.

Pontow, Carsten and Rainer Schubert (2006). "A mathematical analysis of theories of parthood". *Data & Knowledge Engineering* 59.1, pp. 107–138.

Postal, Paul (1974). *On raising*. Cambridge, MA: MIT Press.

Potts, Christopher (2008). *Interpretive economy, Schelling points, and evolutionary stability*. Manuscript, University of Massachusetts, Amherst, MA.

Quine, Willard van Orman (1960). *Word and Object*. Cambridge, MA: MIT Press.

Quine, Willard van Orman (1985). "Events and reification". In: *Actions and events: Perspectives on the philosophy of Donald Davidson*. Ed. by Ernest LePore and Brian P. McLaughlin. Oxford, UK: Blackwell Publishing, pp. 162–171.

Rathert, Monika (2004). *Textures of time*. Berlin, Germany: Akademie Verlag.

Richards, Barry (1982). "Tense, aspect and time adverbials, Part I". *Linguistics and Philosophy* 5, pp. 59–107.

Rimell, Laura (2004). "Habitual sentences and generic quantification". In: *Proceedings of the Twenty-Third West Coast Conference on Formal Linguistics (WCCFL 23)*. Ed. by G. Garding and M. Tsujimura. Somerville, MA: Cascadilla Press.

Robaldo, Livio (2010). "Distributivity, collectivity, and cumulativity in terms of (in)dependence and maximality". *Journal of Logic, Language and Information*, pp. 1–39.

Roberts, Craige (1987). "Modal subordination, anaphora, and distributivity". PhD thesis. Amherst, MA: University of Massachusetts.

van Rooij, Robert (2003). "Questioning to resolve decision problems". *Linguistics and Philosophy*, pp. 727–763.

van Rooij, Robert (2008). "Comparatives and quantifiers". In: *Empirical issues in syntax and semantics*. Vol. 7. Paris, France: Colloque de Syntaxe et Sémantique à Paris, pp. 393–414.

van Rooij, Robert and Katrin Schulz (2006). "Pragmatic meaning and non-monotonic reasoning: the case of exhaustive interpretation". *Linguistics and Philosophy* 29, pp. 205–250.

Rothstein, Susan (2004). *Structuring events*. Oxford, UK: Blackwell Publishing.

Rothstein, Susan (2009). "Individuating and measure readings of classifier constructions: Evidence from Modern Hebrew". *Brill's Annual of Afroasiatic Languages and Linguistics* 1.1, pp. 106–145.

Rothstein, Susan (2010). "Counting and the mass/count distinction". *Journal of Semantics* 27.3, pp. 343–397.

Sadock, Jerry (1984). "Whither radical pragmatics". In: *Meaning, Form and Use in Context: Linguistic Applications*. Ed. by Deborah Schiffrin. Georgetown University Press, pp. 139–149.

Safir, Ken and Tim Stowell (1988). "Binominal each". In: *Proceedings of the Eighteenth Conference of the North Eastern Linguistics Society (NELS 18)*. University of Massachusetts. Amherst, MA: GLSA (Graduate Linguistic Student Association), pp. 426–450.

Sauerland, Uli (2003). "A new semantics for number". In: *Proceedings of the Thirteenth Conference on Semantics and Linguistic Theory (SALT XIII)*. Cornell University. Ithaca, NY: CLC Publications, pp. 258–275.

Sauerland, Uli, Jan Anderssen, and Kazuko Yatsushiro (2005). "The plural is semantically unmarked". In: *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*. Ed. by Stephan Kepser and Marga Reis. Berlin, Germany: Mouton de Gruyter, pp. 413–434.

Scha, Remko (1981). "Distributive, collective and cumulative quantification". In: *Formal methods in the study of language*. Ed. by Jeroen Groenendijk, Theo Janssen, and Martin Stokhof. Reprinted in Groenendijk, Janssen, and Stokhof (1984). Amsterdam, Netherlands: Mathematical Center Tracts.

Schein, Barry (1986). "Event logic and the interpretation of plurals". PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

Schein, Barry (1993). *Plurals and events*. Cambridge, MA: MIT Press.

Schein, Barry (2006). "Plurals". In: *Oxford Handbook of Philosophy of Language*. Ed. by Ernie Lepore and Barry C. Smith. Oxford, UK: Oxford University Press.

Schwarzschild, Roger (1996). *Pluralities*. Dordrecht, Netherlands: Kluwer.

Schwarzschild, Roger (2002). "The grammar of measurement". In: *Proceedings of the Twelfth Conference on Semantics and Linguistic Theory (SALT XII)*. Ed. by Brendan Jackson. Cornell University. Ithaca, NY: CLC Publications.

Schwarzschild, Roger (2006). "The role of dimensions in the syntax of noun phrases". *Syntax* 9.1, pp. 67–110.

Selkirk, Lisa (1977). "Some remarks on noun phrase structure". In: *Formal Syntax*. Ed. by Peter Culicover, Tom Wasow, and Adrian Akmajian. New York, NY: Academic Press, pp. 285–316.

Sharvy, Richard (1980). "A more general theory of definite descriptions". *The Philosophical Review* 89, pp. 607–624.

Simons, P. (1987). *Parts: a study in ontology*. Oxford, UK: Clarendon Press.

Singh, Raj (2007). "Presupposition and selectional restrictions". Unpublished manuscript, online at `http://web.mit.edu/singhr/www/`.

Smith, Carlota (1997). *The parameter of aspect (second edition).* Dordrecht, Netherlands: Reidel.

Smollett, Rebecca (2001). "Quantized direct objects don't delimit after all". In: *Perspectives on Aspect.* Ed. by Henk J. Verkuyl, Henriëtte de Swart, and Angeliek van Hout. Berlin, Germany: Springer, pp. 41–59.

Spelke, Elizabeth S. (1990). "Principles of object perception". *Cognitive science* 14.1, pp. 29–56.

von Stechow, Arnim (1984). "My reaction to Cresswell's, Hellan's, Hoeksema's, and Seuren's comments". *Journal of Semantics* 3, pp. 183–199.

von Stechow, Arnim (2002). "German *seit* 'since' and the ambiguity of the German Perfect". In: *More than words. A Festschrift for Dieter Wunderlich.* Ed. by Ingrid Kaufmann and Barbara Stiebels. Berlin, Germany: Akademie Verlag, pp. 393–432.

von Stechow, Arnim (2009). "Tenses in compositional semantics". In: *The expression of time.* Ed. by Wolfgang Klein and Ping Li. Berlin, Germany: Mouton de Gruyter, pp. 129–166.

Stickney, Helen (2008). "The emergence of DP in the partitive structure". PhD thesis. Amherst, MA: University of Massachusetts.

Stump, Gregory T. (1981). "The interpretation of frequency adjectives". *Linguistics and Philosophy* 4, pp. 221–257.

Szabolcsi, Anna and Frans Zwarts (1993). "Weak islands and an algebraic semantics for scope taking". *Natural Language Semantics* 1.3, pp. 235–284.

Talmy, Len (1996). "Fictive motion and 'ception'". In: *Language and space.* Ed. by Paul Bloom et al. Cambridge, MA: MIT Press.

Tarski, Alfred (1929). "Les fondements de la géométrie des corps (Foundations of the geometry of solids)". *Księga Pamiątkowa Pierwszego Polskiego Zjazdu Matematycznego, suppl. to Annales de la Société Polonaise de Mathématique* 7, pp. 29–33.

Tarski, Alfred (1935). "Zur Grundlegung der Booleschen Algebra". *Fundamenta Mathematicae* 24. Translated to English as Tarski (1956), pp. 177–198.

Tarski, Alfred (1956). "On the Foundations of the Boolean Algebra". In: *Logics, Semantics, Metamathematics. Papers from 1923 to 1938.* Ed. by Joseph Henry Woodger. Translation of Tarski (1935). Oxford, UK: Clarendon, pp. 320–341.

Tatevosov, Sergej (2002). "The parameter of actionality". *Linguistic Typology* 6, pp. 317–401.

Taub, Alison (1989). "Collective predicates, aktionsarten and *all*". In: *Papers on quantification.* Ed. by Emmon Bach, Angelika Kratzer, and Barbara H. Partee. Amherst, MA: University of Massachusetts.

Taylor, Barry (1977). "Tense and continuity". *Linguistics and Philosophy* 1, pp. 199–220.

Tsoulas, George and Eytan Zweig (2009). *Distributivity across languages and logics.* Course at ESSLLI. URL: http://distshare.blogspot.com.

Ursini, Francesco-Alessio (2006). *Measure phrases and directional PPs: measuring out movement.* M. Phil. thesis, Utrecht University, Netherlands. URL: http://igitur-archive.library.uu.nl/student-theses/2006-1113-200125/Thesis.pdf.

Vaillette, Nathan (2001). "Flexible summativity: A type-logical approach to plural semantics". In: *OSU Working Papers in Linguistics 56.* Columbus, OH: Ohio State University, pp. 135–157.

Varzi, Achille C. (2006). "A note on the transitivity of parthood". *Applied ontology* 1.2, pp. 141–146.

Varzi, Achille C. (2010). "Mereology". In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Spring 2010. Stanford, CA. URL: http://plato.stanford.edu/archives/spr2010/entries/mereology/.

Vendler, Zeno (1957). "Verbs and times". *The Philosophical Review* 66, pp. 143–160.

Vendler, Zeno (1962). "Each and every, any and all". *Mind* 71.282, pp. 145–160.

Verkuyl, Henk J. (1972). *On the compositional nature of the aspects.* Dordrecht, Netherlands: Reidel.

Verkuyl, Henk J. (1989). "Aspectual classes and aspectual composition". *Linguistics and Philosophy* 12.1, pp. 39–94.

Verkuyl, Henk J. (1993). *A theory of aspectuality: The interaction between temporal and atemporal structure.* Cambridge, UK: Cambridge University Press.

Verkuyl, Henk J. (1994). "Distributivity and collectivity: a couple at odds". In: *Dynamics, polarity, and quantification.* Ed. by Makoto Kanazawa and Christopher Piñón. Stanford, CA: CSLI Publications, pp. 49–80.

Verkuyl, Henk J. (2005). "Aspectual composition: Surveying the ingredients". In: *Perspectives on Aspect.* Ed. by Henk J. Verkuyl, Henriëtte de Swart, and Angeliek van Hout. Berlin, Germany: Springer, pp. 19–39.

Verkuyl, Henk J. and Jaap van der Does (1996). "The semantics of plural noun phrases". In: *Quantifiers, logic, and language*. Ed. by Jaap van der Does and Jan van Eijck. Stanford, CA: CSLI Publications.

Vlach, Frank (1993). "Temporal adverbials, tenses and the Perfect". *Linguistics and Philosophy* 16, pp. 231–283.

Williams, Alexander (2009). "Themes, cumulativity, and resultatives: Comments on Kratzer 2003". *Linguistic Inquiry* 40.4, pp. 686–700.

Williamson, Timothy (1994). *Vagueness*. London, UK: Routledge.

Winter, Yoad (1997). "Choice Functions and the Scopal Semantics of Indefinites". *Linguistics and Philosophy* 20, pp. 399–467.

Winter, Yoad (2001). *Flexibility principles in Boolean semantics*. Cambridge, MA: MIT Press.

Winter, Yoad (2002). "Atoms and sets: A characterization of semantic number". *Linguistic Inquiry* 33.3, pp. 493–505.

Zimmermann, Malte (2002). "Boys buying two sausages each: On the syntax and semantics of distance-distributivity". PhD thesis. Amsterdam, Netherlands: University of Amsterdam.

Zucchi, Sandro and Michael White (2001). "Twigs, sequences and the temporal constitution of predicates". *Linguistics and Philosophy* 24, pp. 187–222.

Zwarts, Joost (1997). "Vectors as relative positions: A compositional semantics of modified PPs". *Journal of Semantics* 14, pp. 57–86.

Zwarts, Joost (2005). "Prepositional Aspect and the Algebra of Paths". *Linguistics and Philosophy* 28, pp. 699–740.

Zwarts, Joost (2006). "Event shape: Paths in the semantics of verbs". Manuscript. Radboud University Nijmegen and Utrecht University. URL: http://www.let.uu.nl/users/Joost.Zwarts/personal/EventShape.pdf.

Zweig, Eytan (2008). "Dependent plurals and plural meaning". PhD thesis. New York, NY: New York University.

Zweig, Eytan (2009). "Number-neutral bare plurals and the multiplicity implicature". *Linguistics and Philosophy* 32, pp. 353–407.