Covert Distributivity in Algebraic Event Semantics*

Lucas Champollion

July 22, 2015

Abstract

This is the first in a pair of papers that aim to provide a comprehensive analysis of the semantic phenomenon of distributivity in natural language. This paper investigates and formalizes different sources of covert distributivity. Apart from lexical distributivity effects, which are modeled by meaning postulates, phrasal distributivity is captured via two covert operators: (i) a D-operator distributing over atoms only (Link, 1987), and (ii) a cover-based Part-operator, which can also distribute over non-atomic pluralities under contextual licensing (Schwarzschild, 1996). The resulting theory surpasses accounts in which nonatomic distributivity is freely available, or not available at all; furthermore, it correctly predicts differences between lexical and phrasal nonatomic distributivity. D and Part are reformulated in Neo-Davidsonian algebraic event semantics, so that they apply to event predicates and make the sum event available for further modification by arguments and adjuncts. This paves the way for an account of the context-dependency of distributivity phenomena under *for*-adverbials, which improves on theories that predict indefinites to either always or never covary with *for*-adverbials.

Keywords: distributivity, algebraic semantics, covers, leakage, *for*-adverbials

1 Introduction

This paper presents a theory of covert distributivity that focuses on the distinction between lexical and phrasal distributivity and advocates a middle road in the debate on whether distributivity is atomic or nonatomic. I suggest a reformulation of the classical definition of distributivity operators and propose to expand their use into the temporal domain. This is one of two self-contained papers that can be read individually but that form a coherent whole. The purpose of these papers is to bring together several strands

*Thanks to Chris Barker, Benjamin Bruening, Seth Cable, Robert Henderson, Chris LaTerza, Anna Szabolcsi, Linmin Zhang, and Malte Zimmermann for helpful comments and discussion. I am grateful to an audience at the University of Potsdam and to members of my 2013 NYU seminar on algebraic semantics, where earlier versions of this paper were presented. This paper and its counterpart, Champollion (2014), build in part on some earlier work of mine, for which I am indebted to many people and institutions as indicated there (Champollion, 2010b, 2012, 2013).

of research on phenomena related to the semantics and pragmatics of distributivity in natural language. This paper focuses on covert distributivity. The counterpart of this paper focuses on overt distributivity. I will refer to it as Champollion (2014). The papers can be read in any order.

Covert distributivity as I will talk about it in this paper is diagnosed by the ability of indefinites or numerals in object position to covary in a way that is generally attributed to a covert verb-phrase-level modifier called the D operator (Link, 1987; Roberts, 1987). The meaning of this kind of operator is either similar to the adverbial modifier *each* or it corresponds to something like *each salient part of*, where salience is a context-dependent notion (Schwarzschild, 1996).

The ambiguity between distributive and scopeless readings in English can be modeled by assuming that the D operator is optionally present as a silent verb phrase modifier whose syntax and meaning correspond to that of adverbial *each*. For example, (1a) below represents a scopeless reading and (1b) a distributive reading. I use the term "scopeless" to refer both to collective and cumulative readings. The distinction between these two readings does not matter for this paper, since in these readings there is no scopal dependency of any noun phrase on anything else. See Landman (2000) for discussion.

(1) a. The boys saw two monkeys.
≈ The boys between them saw two monkeys. *scopeless*
b. The boys [D [saw two monkeys]].
≈ The boys each saw two monkeys. *distributive*

This paper makes both technical and empirical contributions to semantic theory. The main technical contribution of this paper is a reformulation of distributivity operators that makes them compatible with Neo-Davidsonian algebraic event semantics (Davidson, 1967; Parsons, 1990; Krifka, 1989, 1998). This reformulation makes it possible to draw on the resources of event semantics in order to formally model the relations between distributivity over individuals and over events and times in a parallel way. And it makes it possible to draw on the resources of algebraic semantics and mereology in order to formally model the relations between distributivity over atoms and over nonatomic entities like time intervals in a parallel way.

The main empirical contribution of this paper is a unified theory of covert atomic and nonatomic distributivity, over individuals and over temporal intervals, at the lexical and at the phrasal level. Distributivity in the nominal domain has not previously been related to distributivity in the temporal domain. Here I understand distributivity as involving the application of a predicate to the members or subsets of a set, or to the parts of an entity (individual, event, or interval). This application is diagnosed by the presence of certain entailments I call distributive entailments. An example of distributivity in the nominal domain was already given in (1b); here is another one:

(2) The girls are wearing a dress. (Winter, 2001)

This example involves distributivity because on its most plausible interpretation, it entails that each of the girls in question wears a (different) dress. Thus the predicate *wear a dress* is applied to the members or parts of the set or collective denoted by *the*

girls. The fact that the predicate includes an indefinite, *a dress*, makes it especially easy to diagnose the presence of distributivity, since in the absence of a universal quantifier, a distributivity operator needs to be postulated in order to explain why this indefinite can covary, that is, can be interpreted as involving reference to more than one dress in total.

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, to the extent that we admit strange models in which several girls can wear the same dress, the predicate *wear a dress* does not have to be interpreted distributively. I will set these two points aside in the following discussion.

There has been a long-standing debate about whether and to what extent distributivity over nonatomic entities, or "genuine plural quantification" (Link, 1987), ever occurs (Link, 1987; Gillon, 1987, 1990; Lasersohn, 1989; Schwarzschild, 1996; Winter, 2001; Kratzer, 2007). Here is a preview of the argument developed in Schwarzschild (1996). Schwarzschild argues that the distributivity operator should be modified to allow for "nonatomic distributive" interpretations in a limited set of circumstances, essentially whenever there is a particularly salient way to divide a plural individual into parts other than its atoms. Here is an example. Shoes typically come in pairs, so a sentence like (3) can be interpreted as saying that each pair of shoes costs fifty dollars, as opposed to each shoe or all the shoes together.

(3) **(Context: 3 pairs of shoes are on display, each with a $50 tag:)**
The shoes cost fifty dollars. (Lasersohn, 1998)

Since the numeral *fifty dollars* covaries with pairs of shoes rather than with shoes, this example is generally taken to involve nonatomic distributivity in the count domain. The presence of this kind of interpretation depends on contextual factors. For example, it is not part of the meaning of sentence (3) itself that shoes typically come in pairs. Verb phrases can only be interpreted as distributing over nonatomic entities if there is supporting context or world knowledge that makes these nonatomic entities pragmatically salient. In the absence of this support, verb phrases must distribute over atoms or not at all, as the following example shows.

(4) John, Mary, Bill, and Sue were paid fifty dollars. (based on Lasersohn, 1989)

This example can be interpreted as saying that the four people in question were each paid fifty dollars, or that they were paid fifty dollars together. Out of the blue, nonatomic interpretations are not available. For example, in a scenario where each of the four people in question were paid twenty-five dollars, the sentence is false, even though there are ways to group the four people into pairs such that each pair was paid a total of fifty dollars.

It has been suggested that the difference between (3), where a nonatomic distributive reading is available, and (4), where it is not, is due to the lack of a contextually salient partition or cover in the latter case (Schwarzschild, 1996). This can be modeled by making the distributivity operator anaphoric on such a partition or cover. Schwarzschild modifies the D operator accordingly and calls it the Part operator. In

this paper, I adopt this strategy and I extend it to the temporal domain. This makes it possible to create a parallel between the discussion of nonatomic distributivity in the nominal domain and in the temporal domain. The parallel can be illustrated by the following pair of examples:

(5) a. ??John found a flea on his dog for a month.
b. The patient took two pills for a month and then went back to one pill.

Example (5a) is from Zucchi and White (2001). Example (5b) is based on observations in Moltmann (1991). Out of the blue, examples like (5b) are odd just like (5a) because they suggest that the same flea is found repeatedly, the same pills are taken repeatedly, and so on. But context can improve such examples by making covariation of the indefinite or numeral possible. Thus example (5b) is acceptable in a context where the patient's daily intake is salient (in a hospital, for example). It does not require any pill to be taken more than once, so we have covariation.

I will argue that this kind of covariation, when it occurs, diagnoses verb-phrase-level distributivity, and that it is caused by the presence of a temporal analogue of the Part operator discussed above. The contribution of this operator in example (5b) can be paraphrased as *every day*. The operator takes scope over the verb phrase *take two pills* but under the *for*-adverbial. The month-long interval introduced by the *for*-adverbial plays the same role vis-à-vis the distributivity operator as the collective individual denoted by *the shoes* does in example (3).

I will contrast this theory with an alternative view, on which the covariation in (5b) is due to the *for*-adverbial itself, and no distributivity operator is present in examples like (5b). On that alternative view, the *for*-adverbial is interpreted as a universal quantifier meaning something like *at each relevant point during a month*, as has been suggested at various times in the literature (Dowty, 1979; Moltmann, 1991; Deo and Piñango, 2011). I will argue that *for*-adverbials cannot be interpreted as universal quantifiers, since out of the blue they do not induce covariation in indefinites they outscope (5a) except when an overt distributive quantifier intervenes (6). This is shown in the following example from Zucchi and White (2001):

(6) John found a flea on his dog every day for a month.

I will account for the contrast between (5a) and (6) by claiming that the temporal distributivity operator cannot occur in the former, and that *every day* plays its role in the latter. This requires an explanation of why the operator is not able to occur in (5a) out of the blue. I will suggest that its ability to occur or not is due to a contextual factor because it is anaphoric on a salient set of stretches of spacetime, in the same way as the nonatomic distributivity operator in (3) is anaphoric on a salient set of shoes.

While the topic of this paper is covert distributivity, the topic of Champollion (2014) is overt distributivity, as manifested in adverbial *each* and its adnominal and determiner counterparts both in English and other languages (Zimmermann, 2002). The meanings of these items varies in ways that sometimes require them to distribute over individuals (such as in the case of *each*) and in other cases allow them to distribute over salient parts of spacetime (such as in the case of German *jeweils*, for example). The main claim of Champollion (2014) is that the D operator relates to the Part operator in

the same way as *each* relates to *jeweils*. Thus, overt and covert distributivity share many similarities. This gives rise to similar questions in the two cases. Can a distributivity operator only distribute down to singular entities or also to plural entities? Do these entities need to be of a certain size or "granularity", and can this size vary from operator to operator? Must these entities have been overtly mentioned in the sentence and thereby contributed by semantic means, or can they also be supplied by the context via pragmatic means?

A unified semantic analysis of distributivity should make it apparent which aspects of the meanings of various distributivity operators are always the same, and along which dimensions these meanings can differ. The semantic variation across distributivity related items should be captured by parameters. I will explain the fact that the various overt and covert distributivity operators share some part of their meanings. To do so, I will view distributivity as the property of a predicate which, whenever if holds of a certain entity or event, also holds of its parts along a certain dimension and down to a certain granularity. I have previously developed and defended this view at length under the name of *strata theory* (Champollion, 2010b). Going beyond distributivity, strata theory has applications in the domains of aspect and measurement. Here, I focus on distributivity. Strata theory conceptualizes dimension and granularity as parameters which can be set to different values for different instances of distributivity. The dimension parameter specifies the domain in which the predicate in question is distributed. It can be set to a thematic role or to time, for example. For the purpose of this paper, different settings of this parameter will allow us to capture the commonalities and differences between distributivity over count domains and time. The granularity parameter can be used to specify that the parts in question must be atomic, or that they can be nonatomic but must be very small as measured along the dimension specified by the other parameter. This parameter accounts for the differences between distributive constructions over discrete (count) domains, and those over domains involving continuous dimensions, such as time. These parameters interact with each other against the background of assumptions about the metaphysics of natural language. For example, time is assumed to be nonatomic, or in any case to not make its atoms available to the semantics of natural language. As a result, when the dimension parameter is set to time, the granularity parameter cannot be set to anything involving atoms, because time does not provide any atoms to distribute over. This simple idea turns out to explain and connect a range of facts observed in various places in the literature.

As will become clear below, this understanding of distributivity provides several theoretical advantages. First, by understanding distributivity 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 distributivity operator of Link (1987), Roberts (1987), and Winter (2001) corresponds to one setting of the granularity parameter, and the nonatomic distributivity operator of Schwarzschild (1996) corresponds to another setting. Following Schwarzschild, I will assume that there is a distributivity operator whose granularity parameter is anaphoric on its context and can only be set to a nonatomic value when context supports a salient granularity level.

Second, by understanding distributivity as parametrized for dimension, we gain

the technical ability to distinguish agent-based from theme-based distributivity and the like (Lasersohn, 1998). Not only thematic roles like agent and theme can be considered dimensions, but also trace functions like runtime and location. We can therefore instantiate the dimension parameter of distributivity with time, or more specifically, with the temporal trace function τ that maps events to their runtimes (Krifka, 1998). Given the assumption that time and space are nonatomic, we expect that this should only be possible when the granularity parameter of the distributivity 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 parametrized distributivity, the asymmetry between the atomic domain of individuals and the nonatomic domain of time allows us to explain the scopal behavior of *for*-adverbials. If the distributivity 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 Neo-Davidsonian event semantic setting also gives us the ability to think of overt and covert distributivity operators as being (co-)indexable with different thematic roles. This allows us to capture through a simple change in indexation the kinds of configurations that have otherwise been taken to require type-shifting-based reformulations of the D operator (Lasersohn, 1998):

(7) a. The first-year students D(took an exam). *Target: agent*
b. John D(gave a pumpkin pie) to two girls. *Target: recipient*

As I discuss in Champollion (2014), this phenomenon has a direct counterpart in examples like the following, which involve adnominal *each* (Zimmermann, 2002; Blaheta, 2003):

(8) The boys told the girls two stories each. *Target: agent*
(two stories per boy)

(9) The boys told the girls two stories each. *Target: recipient*
(two stories per girl)

The theoretical picture that is developed here provides a way to formalize such parallels across instances of distributivity in natural language. Individual items can be analyzed as being hard-wired for certain parameter values, so that, for example, the difference between Link's and Schwarzschild's operators, as well as that between *each* and *jeweils*, can be described in terms of whether the value of the granularity parameter is prespecified to *Atom* or can be filled in by context. In this way, overt and covert instances of distributivity fit together and into distributivity theory more generally.

The rest of this paper is organized as follows. In Section 2, I review the literature on covert distributivity in the nominal domain, focusing on the atomic distributivity operator introduced in Link (1987) and Roberts (1987). Nonatomic distributivity is discussed in Section 3, where I present the nonatomic distributivity operator introduced in Schwarzschild (1996), summarize the literature on the topic, and describe my own view. The question of how to adapt distributivity operators into event semantic frameworks, such as the one adopted here, is discussed in Section 4. In that section, I

argue for a specific way to do so, which I argue in Section 5 to be superior to a previous proposal by Lasersohn (1998). The specific way in which I formulate the D and Part operators involves outfitting them with dimension and granularity parameters. When the dimension parameter of the reformulated Part operator is set to time, the result induces covariation of indefinites over salient stretches of time. Section 6 builds on this result and provides an account of the limited ability of covariation by indefinites in the scope of *for*-adverbials. Section 7 tentatively discusses extensions to other temporal modifiers such as *every day* and to generic sentences. Section 8 concludes.

2 Atomic Distributivity

In this section and the next, I present the theory of atomic distributivity initiated by Link (1987) and Roberts (1987), and further developed by subsequent authors. This section focuses on the atomic distributivity operator D as originally defined by Link and Roberts, and the motivation that led to it. The next section focuses on the nonatomic distributivity operator Part as originally defined by Schwarzschild (1996). The discussion in this and in the next section takes inspiration from Winter (2001), Chapter 6, 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.

Distributivity can be understood, among other things, as a property of predicates, in opposition to collectivity. Lexical predicates like *smile* and phrasal predicates like *wear a dress* are distributive because, whenever several people smile or wear a dress, this entails that each of them smiles or wears a dress. The distinction between lexical and phrasal predicates that have distributive interpretations will be important throughout this paper. To highlight this distinction, from now on I will speak of lexical and phrasal distributivity. As the following examples show, we find distributive and collective interpretations in both classes of predicates:

(10) **Lexical distributivity/collectivity**

a. The children smiled. *distributive*
b. The children were numerous. *collective*

(11) **Phrasal distributivity/collectivity**

a. The girls are wearing a dress. *distributive*
b. The girls are sharing a pizza. *collective*

Distributivity and collectivity can be given operational definitions in terms of entailments or lack thereof (for details, see Champollion, 2010b). Sentence (10a) entails that each child smiled, while sentence (10b) does not entail that each child was numerous. Similarly, sentence (11a) entails that each girl wears a different dress, but sentence (11b) does not entail that the girls ate different pizzas.

The distinction between lexical and phrasal distributivity is related to the P/Q-distributivity distinction introduced in 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 indefinite or numeral quantifier, as in (11a). In order for (11a) 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 ascribed to the lexical semantics of the verb. It is possible to ascribe the difference between (10a) and (10b) to the meaning of *smile* and *be numerous*. The difference between the distributive interpretation in (11a) and the collective interpretation in (11b) is of a different kind, since it involves a scopal ambiguity. Accounts that are based only on lexical semantics cannot model Q-distributivity, and therefore cannot model phrasal distributivity, because they cannot create a scopal dependency between the definite subject and the indefinite object. For more details, see Winter (2001), Sections 3.2 and 6.1.

The usual way to model Q-distributivity, which I will 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 is the purpose of the D operator postulated by Link (1987) and Roberts (1987). It shifts a verb phrase to a distributive interpretation, more specifically, one that holds of any individual whose atomic parts each satisfy the unshifted verb phrase. (The operator goes back to Link (1991), which was written before Link (1987). See Roberts (1987, p. 157) for discussion.) This D operator is usually defined as follows.

(12) **Link's D operator:**
$[\![\text{D}]\!] = \lambda P_{et} \lambda x \forall y [y \leq x \wedge \text{Atom}(y) \rightarrow P(y)]$
(Takes a predicate P over individuals and returns a predicate that applies to any individual whose atomic parts each satisfy P.)

Here, the variable x is resolved to a plural entity, typically provided by the subject, and y ranges over its atomic parts, that is, the singular individuals of which it consists. The operator introduces a universal quantifier, and it is the scopal interaction of this quantifier with the indefinite inside a Q-distributive predicate that accounts for the covariation effects. For example, if the verb phrase *wear a dress* is represented as $\lambda x \exists z [\text{dress}(z) \wedge \text{wear}(x, z)]$, the meaning of (11a) can be represented in a way that places it in the scope of the universal quantifier introduced by the D operator. Here and below, I write $\bigoplus \text{girl}$ for the sum of all girls, the plural individual that corresponds to *the girls*. I will come back to mereological notation in Section 4.

(13) $\forall y [y \leq_{Atom} \bigoplus \text{girl} \rightarrow \exists z [\text{dress}(z) \wedge \text{wear}(y, z)]]$
(Every atomic part of the sum of all girls wears a dress.)

The D operator is able to apply to entire verb phrases and not just to lexical predicates. It is this property that allows the D operator to account for phrasal distributivity (Dowty, 1987; Roberts, 1987; Lasersohn, 1995). Moreover, at least Roberts (1987) allows the D operator to apply to any predicate, whether it is a verb phrase or not. For

example, it may apply to a predicate that has been derived by lambda abstraction over a nonsubject predicate in order to derive an interpretation of (14) where each of two girls received a pumpkin pie:

(14) John gave a pumpkin pie to two girls.
(two girls) D[λx. John gave a pumpkin pie to x] (Roberts, 1987)

This approach involves an otherwise unmotivated structure or perhaps an application of quantifier raising, and is criticized for this reason by Lasersohn (1998), whose own proposal is the topic of Section 5 in this paper. The need for the D operator to be able to target noun phrases other than the subject is an important point, and I return to it below. As will be discussed below, my own implementation deals with nonsubject predicates by parametrizing the D operator for different thematic roles. But first, I turn to a review of nonatomic distributivity.

3 Nonatomic Distributivity

In the previous section, I have presented the atomic view on distributivity. 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 (1996), Verkuyl and van der Does (1991), Schwarzschild (1996, ch. 5), Brisson (1998, 2003), and Malamud (2006a,b). This section presents and motivates the nonatomic view.

Section 1 has discussed examples like (3) (*The shoes cost fifty dollars*), where context can make nonatomic distributive interpretations available. This kind of example is discussed in Schwarzschild (1996). There, Link's distributivity operator is modified so that it no longer distributes over atoms but is instead anaphoric to a salient cover (a partition of a plural individual that allows overlap). The nonatomic view is based on sentences like the following, which is adapted from Gillon (1987):

(15) 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.

An early argument for the nonatomic view was given as follows (Gillon, 1987, 1990): in order to generate the reading on which (15) 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. In a set-based representation of plural individuals, covers are like partitions of a set except that their cells can overlap. Formally, a set 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.

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 (Schwarzschild, 1996). 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 (16). On this view, the denotation of sentence (15) can be represented as (17). This formula is verified in the actual world by the existence of the cover $\{rodgers \oplus hammerstein, rodgers \oplus hart\}$.

(16) **Nonatomic distributivity operator, existentially bound cover**
$[\![\mathrm{D}_\exists]\!] = \lambda P_{\langle et\rangle}\lambda x \exists C[\mathrm{Cov}(C, x) \wedge \forall y[\mathrm{C}(y) \wedge y \leq x \rightarrow P(y)]]$

(17) $\exists C[\mathrm{Cov}(C, \mathrm{rodgers} \oplus \mathrm{hammerstein} \oplus \mathrm{hart}) \wedge$
$\forall y[\mathrm{C}(y) \wedge y \leq x \rightarrow y \in [\![\mathrm{wrote\ musicals}]\!]]]$

Existentially bound covers are not a good way to model phrasal distributivity, because they overgenerate readings. These can be described as halfway between collective (or cumulative) and distributive readings, and they are sometimes called intermediate readings. I call them nonatomic distributive readings. For example, in a situation where John, Mary, and Bill are the teaching assistants and each of them was paid exactly $7,000 last year, sentences (18a) and (18b) are both true (Lasersohn, 1989). This is as is expected on the atomic approach. Sentence (18a) is true on its distributive reading, and Sentence (18b) is true on its collective or cumulative reading. But sentence (18c) is false, even though the cover $\{j \oplus m, m \oplus b\}$ would verify it if it was represented using the $\mathrm{D}_\exists$ operator in (16). That is, sentence (18c) does not have a nonatomic distributive reading.

(18) a. The TAs were paid exactly $7,000 last year. *atomic distributive*
b. The TAs were paid exactly $21,000 last year. *collective*
c. The TAs were paid exactly $14,000 last year. **nonatomic distributive*

Giving up the existential cover-based operator $\mathrm{D}_\exists$ in (16) explains why sentence (18c) is false, because without this operator, there is no way to generate a nonatomic distributive reading for this sentence. However, sentence (15) above does have a nonatomic distributive reading, so giving up $\mathrm{D}_\exists$ requires an alternative account of this reading. Lasersohn (1989) proposes to do so through the use of lexical meaning postulates that stipulate implications like the one in (19).

(19) $\forall w, x, y, z[\mathrm{write}(w, x) \wedge \mathrm{write}(y, z) \rightarrow \mathrm{write}(w \oplus x, y \oplus z)]$

This is actually a special case of an assumption called lexical cumulativity, to which I come back below. (Sternefeld (1998) and Beck and Sauerland (2000) generalize this approach to arbitrary constituents, but I will not do so in this paper because their

generalization makes it difficult to model the difference between lexical and phrasal distributivity.)

Further support for adopting Lasersohn's assumption while rejecting existentially bound covers comes from the difference between *write musicals* and *write a musical*. The following sentence is false in the actual world, that is, it does not have the nonatomic distributive construal that (15) has (Link, 1997).

(20) Rodgers, Hammerstein and Hart wrote a musical.

Given what is sometimes called the inclusive view of the plural, *write musicals* literally applies to entities who wrote one or more musicals (Krifka, 1986; Sauerland, Andersen, and Yatsushiro, 2005; Spector, 2007). Such entities include the sum $rodgers \oplus hammerstein$ and the sum $rodgers \oplus hart$, and via (19), the sum individual $rodgers \oplus hammerstein \oplus hart$, of which (19) 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 (15) and (20). (These assumptions do not explain why sentence (15) conveys that more than one musical in total is written. This fact can be explained, for example, by modeling this information as a scalar implicature (Spector, 2007; Zweig, 2008, 2009).)

As we have seen, it is important that a cover-based operator like (16) is not available in the grammar, because that operator would predict (20) to be true in the actual world. Lasersohn and many others conclude from this and similar examples that the atomic approach to phrasal distributivity is superior to covers (e.g. Winter, 2001).

However, there are cases in which a cover-based operator does seem necessary (Gillon, 1990; Schwarzschild, 1996). These cases typically involve special contexts in which discourse makes a specific cover pragmatically salient. This is where the shoe example in (3) comes in, repeated below as (21a). Shoes typically come in pairs, so a sentence like (21a) can be interpreted with respect to a cover whose cells each contain a matching pair of shoes. The relevant reading is a nonatomic distributive 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 (21b), which can only mean that each suitcase weighs fifty pounds or all of them together do so.

(21) a. The shoes cost fifty dollars. (Lasersohn, 1998)
b. The suitcases weigh fifty pounds.

In the nonatomic reading of (21a), 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 (15). Schwarzschild (1996) models the context dependency of this kind of intermediate reading by assuming that the distributivity operator (which he renames Part, to set it apart from Link's D operator) contains a cover variable whose value is supplied anaphorically by context. I will represent this variable as a subscripted C.

(22) **Schwarzschild's nonatomic distributivity operator, free cover**
$[\![\text{Part}_C]\!] = \lambda P_{\langle et \rangle} \lambda x \forall y [\text{C}(y) \wedge y \leq x \rightarrow P(y)]$

The difference between D and Part amounts to a division of labor between semantics and pragmatics. Semantics accounts for atomic phrasal distributivity and pragmatic accounts for nonatomic phrasal distributivity. Schwarzschild assumes that C is restricted through a pragmatic mechanism to be a cover over x, but he prefers not to write this condition into his operator.

The introduction of a pragmatic component into the analysis of what had previously be treated as a purely semantic phenomenon is discussed and justified at length in Schwarzschild (1996). Relevant evidence comes from sentences like the following:

(23) The young animals and the old animals were separated.
(Schwarzschild, 1996, p. 44)

This sentence typically entails that the young animals were separated from the old animals, but is compatible with each of these two groups staying together. This suggests that the verb phrase *be separated* is distributed down to the level of these two groups and not all the way down to individual animals. At this point, a proponent of atomic distributivity might argue that the reason that the verb phrase is able to apply at this intermediate level is that the two groups of animals are in fact atoms. The two conjuncts might then be analyzed as involving group-forming operators that map each of the two pluralities of animals to an atom. This route is taken by Landman (1989). Schwarzschild rejects this approach and argues for the essentially pragmatic nature of nonatomic distributivity by pointing out that the inference down to groups is cancelable. The following sentence leaves it open exactly how the animals were separated, a fact that is unexpected on the group-based analysis:

(24) The young animals and the old animals were separated, but not necessarily by age.

Schwarzschild assumes that the cancellation prevents the value of C from being set to the cover that is made pragmatically salient by the conjunction. Beyond this kind of case, Schwarzschild does not say much about the precise pragmatic mechanism that resolves C. For a proposal in which the Part operator is anaphoric on a decision problem in the sense of van Rooij (2003), see Malamud (2006a,b). These approaches differ in details, but the important point that is common to them is that the Part operator imposes a stronger restriction on the identity of the cover than would be achieved by just existentially quantifying over it. This restriction rules out intermediate readings in sentences like (18c), (20), and (21b) but not in sentences like (21a). While I see no obstacles to using decision problems, I continue to use a Schwarzschild-style approach to keep the representation simple.

Sentence (21a) is structurally equivalent to sentences (18c), (20), and (21b), yet only (21a) 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. The operator in (22) is more restricted than the existential cover-based operator $D_{\exists}$ in (16) because (22) presupposes the existence of a context through which the variable C can be resolved. The contrast between (21a), which has a nonatomic distributive reading, and (18c), (20) and (21b), which do not, is predicted, on the plausible assumption that a salient context is only available for (21a).

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 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 Part 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 Part operator behaves like Link's D operator.

Table 1: V level versus VP level distributivity in atomic domains

(a) Empirical generalization

	lexical (V level)	phrasal (VP level)
atomic	available	available
nonatomic	available	only w. context

(b) Explanation

lexical (V level)	phrasal (VP level)
lexical cum.	Atomic op.
lexical cum.	Cover-based op.

Even though the semantic effects of D can be subsumed under the workings of Part, I postulate two covert distributivity operators, D and Part, in grammar. Having distinct distributivity operators might seem redundant. For the purpose of this paper, it would be equally possible to assume that there is only one operator, namely Part, and that covers over atomic individuals are salient in every situation. Such a view would amount 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. When the granularity parameter of Part is set to atoms, it behaves equivalently to D. Therefore, on such a view, D would in effect be made available again through the back door, even though officially there would be only one operator in the grammar.

Although such a view might seem more economical, I do not adopt it here. Instead, I assume that both D and Part are present in the grammar. Thus, I split the operator that accounts for phrasal distributivity into two. In the companion paper, Champollion (2014), I motivate this split by arguing that D and Part are lexicalized as adverbial and adnominal distributivity operators in individual languages. This assumption will allow us to capture the distinction between English *each* and its German relative *jeweils*. I suggest that the former corresponds to D and the latter corresponds to Part, and I use this assumption to account for the fact that *jeweils* and its relatives across languages have a wider range of readings than *each* and its relatives do.

The search for clear cases of nonatomic distributivity has been going on since at least Link (1987). On the one hand, lexical nonatomic distributivity clearly occurs in examples like the following:

(25) a. All competing companies have common interests. (Link, 1987)
b. Five thousand people gathered near Amsterdam. (van der Does, 1993)

In example (25a), the predicate *have common interests* can be applied distributively (that is, it describes several instances of having common interests) to nonatomic entities, because it does not make sense to say of a single company that it has common interests with itself. In example (25b), the predicate *gather near Amsterdam* can be applied distributively (that is, it describes several gatherings) to nonatomic entities, because a single person cannot gather.

On the other hand, examples that involve *phrasal* nonatomic distributivity, such as the shoe example (21a), are harder to come by. I believe that one of the reasons why it has been so hard to identify clear cases of nonatomic distributivity is the focus in the literature on predicates that apply in count domains. On the standard assumption that the denotations of count nouns are taken from an atomic domain, 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. In count domains, the scale induced by the relation that orders singular and plural individuals according to their cardinalities is lower closed because singular individuals are atomic. Interpretive Economy suggests that speakers who use a covert distributivity 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. This suggests that by looking at noncount domains, we can remove atomic granularity as a potential focal point, so any phrasal distributivity effects we find must be cases of nonatomic distributivity. If we look at nonatomic domains, nonatomic distributivity should be easier to detect. A reviewer offers the following example as a case in point (I confirmed its grammaticality with two native speakers of English):

(26) At the garden party, they sell milk, lemonade and beer. Milk costs one dollar, lemonade costs two dollars, and beer costs four dollars.

Here, the predicates *cost 50 cent* etc. are distributed to a level that is made pragmatically salient by the context, namely the units in which the beverages are sold – presumably glasses, bottles or cans.

Another domain that is nonatomic is time. In Section 6 further below, I will identify cases of nonatomic phrasal distributivity involving time. I will look at this domain through the lens of *for*-adverbials, focusing on their scopal behavior with respect to verb phrases that contain an overt quantifier. In a nonatomic domain, there are necessarily no atomic covers, so the first row of Table 1 is not applicable. I will

argue that the second row of Table 1 is mirrored precisely in the temporal domain, as shown in Table 2. Remember that I assume that distributivity always involves a dimension and a granularity parameter. That is, the dimension parameter of the distributivity operator involved can be instantiated to τ (runtime) and, in that case, its granularity parameter is dependent on an anaphorically salient level of granularity. I have suggested that the distributivity operators contain a dimension and a granularity parameter. The granularity parameter can be understood as Schwarzschild's cover, but the dimension parameter does not yet figure in the implementations we have seen so far. Therefore, it is necessary at this point to provide a formal implementation of Link's and Schwarzschild's distributivity operators that supports the notion of dimension and granularity parameters. Section 4 provides this implementation.

Table 2: Distributivity in nonatomic domains

	(a) Empirical generalization		(b) Explanation	
	lexical (V level)	phrasal (VP level)	lexical (V level)	phrasal (VP level)
atomic	**n/a**	**n/a**	**n/a**	**n/a**
nonatomic	available	only w. context	lexical cum.	Cover-based op.

4 Reformulating the D and Part Operators

Once we move to event semantics, the distributivity operators developed and motivated in the last two sections have to be adjusted for a number of reasons. First, Link's and Schwarzschild's formulations of distributivity operators are based on the assumption that verb phrases denote sets of individuals, while event semantics typically assumes that verb phrases denote sets of events. Second, like other kinds of quantification in event semantics, distributivity over individuals involves the introduction of a variable that ranges over subevents of some large event (Taylor, 1985; Schein, 1993). Third, there is a technical flaw with the only existing proposal of how to represent the D operator in event semantics (Lasersohn, 1998). Roughly, Lasersohn's operator fails to prevent the large event from containing extraneous material in addition to the subevents over which the D operator ranges ("leakage" in the terms of Bayer (1997)). These problems also apply to the Part operator. Finally, since we are adopting a Neo-Davidsonian view, on which thematic roles are reified, it makes sense to think of the D and Part operators as being indexed with thematic roles. This was discussed in Section 1 in connection with examples like the following, repeated here from (7):

(27) a. The first-year students D(took an exam). *Target: agent*
b. John D(gave a pumpkin pie) to two girls. *Target: recipient*

This section shows a way to reformulate the D and Part operators that meets these requirements. The next section 5 compares it with a related previous proposal by Laser-

sohn (1998). I now introduce the notational conventions and background assumptions of algebraic event semantics (e.g. Krifka, 1998; Champollion, 2010b). I use the following typing conventions: t for propositions, e for ordinary objects, and v for events. I use the symbols x, y, z, x', y', z' and so on for variables that range over ordinary objects, and e, e', e'' for events. I use P for predicates of type $\langle et \rangle$, V for predicates of type $\langle vt \rangle$, θ and Θ for functions of type $\langle ve \rangle$. My notion of events is essentially the one of Krifka (1998) and Parsons (1995). In particular, events are individual rather than generic, they are mostly concrete rather than abstract, and they are located in spacetime. Like Parsons, I assume that event modifiers can provide clues about the identity of events. For example, if I both brush my teeth and sing, but if my singing is surprising while my brushing my teeth is not, then the singing is not the same event as the tooth-brushing (Parsons, 1990, p. 157). I make the standard assumption that different lexical predicates such as singing and tooth-brushing denote disjoint sets, that is, they do not apply to the same event.

I assume that ordinary objects and events are each closed under mereological sum formation (Link, 1998). 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.

I assume that singular individuals, by which I mean the entities in the denotation of singular count nouns, are mereological atoms. This means that the notion of mereological part must be distinguished from the intuitive notion of part, so that the leg of a table is not a mereological part of the table. My assumption is very convenient and widespread in the literature, but it has been disputed on various grounds, including the existence of nouns like *twig, rock*, and *sequence* (Zucchi and White, 2001). The account I am about to give could still be given without this assumption, as long as we replace it by another way to represent the distinguished level of individuation that is associated with singular individuals, and as long as we agree that the D operator is restricted to this distinguished level. One such other way is provided by the notion of a "natural unit" (Krifka, 1989). For more discussion on criteria for the individuation of nouns, see Rothstein (2010) and Barker (2010).

I adopt thematic uniqueness, that is, the assumption that each event has at most one agent, at most one theme, etc. (Carlson, 1984, 1998; Parsons, 1990). This is why I represent thematic roles as functions. Thematic uniqueness is useful when we extend distributivity to nonatomic domains like time and space, because it allows us to rely on the idea that thematic roles can be treated as being the same kinds of things as functions that map events to their locations in spacetime (Champollion, 2010b).

I adopt classical extensional mereology. I assume that the domains of individuals, of events, and of times are each partially ordered by a parthood relation $\leq$, and that within each of these domains, any nonempty set of entities has a unique sum. I write $x \circ y$ for "x overlaps y", meaning that x and y have a part in common:

(28) **Definition: Overlap**
$x \circ y \stackrel{\text{def}}{=} \exists z[z \leq x \wedge z \leq y]$

(Two things overlap if and only if they have a part in common.)

In the following, $\bigoplus P$ stands for the mereological sum of all entities to which P applies. Some authors write σ instead of $\bigoplus$. The following definition of sum comes from Tarski (1935).

(29) $x = \bigoplus P \stackrel{\text{def}}{=} \forall y[P(y) \rightarrow y \leq x] \wedge \forall z[z \leq x \rightarrow \exists z'[P(z') \wedge z \circ z']]$
(An entity x is the sum of a set of entities P iff everything in P is a part of x and every part of x overlaps with something in P.)

Given that $\leq$ is a partial order, the concept of sum provably coincides with that of least upper bound with respect to $\leq$, but in this paper I will use Tarski's definition in proofs. For an overview of the axioms and linguistic applications of mereology, see Champollion and Krifka (to appear).

I will use the star operator originally defined in Link (1983). The star operator applied to a predicate P gives us the algebraic closure of P, that is the closure of P under sum. It is defined as follows:

(30) $x \in {}^*P \stackrel{\text{def}}{=} \exists C[x = \bigoplus C \ \wedge \ C \neq \emptyset \ \wedge \ C \subseteq P]$
(x is the sum of all the elements of a nonempty subset of P)

According to this definition, $x \in {}^*P$ means that x consists of one or more parts such that P holds of each of these parts. For example, if x is a square, and P is the property of being a triangle, then P does not contain x, but ${}^*P$ does, since any square can be divided into triangles (along a diagonal, for example).

The star operator will play a crucial role in the technical development below. Let me emphasize, however, that if one wishes to avoid using this operator, it is possible to avoid it by relying on the following fact, which follows from the definition of sum.

(31) **Fact:** $x \in {}^*P \Leftrightarrow \forall y[y \leq x \rightarrow \exists z[y \circ z \wedge z \leq x \wedge P(z)]$
(x is in the algebraic closure of P iff every part of x overlaps with a part of x that is in P.)

I assume that verbs and their projections denote sets of events and are therefore of type vt. They are related to their arguments and adjuncts by thematic roles such as *agent*, *theme* and runtime, which I write τ. This is the Neo-Davidsonian view on event semantics (Carlson, 1984; Parsons, 1990; Krifka, 1992). The application of a predicative or referential noun phrase to a verbal projection amounts to intersecting two sets of events. 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. In a declarative sentence, this operator either existentially binds the event variable or resolves it to a salient event if the sentence refers specifically to the that event.

As is customary in event semantics, I will only deal with unmodified numerals like *three boys*. Modified numerals like *exactly three boys* and *at most three boys* can also be handled in principle but raise additional issues involving event maximality (Krifka, 1999; Brasoveanu, 2010).

There does not seem to be a best practice of how best to implement event semantics

and how to combine verbal projections with their arguments. Some authors adopt more traditional versions of event semantics in which verbs denote relations between events and their arguments (Landman, 2000; Krifka, 1998), or relations between events and their internal arguments (Kratzer, 1996; Harley, 2012). Let me briefly mention some empirical and theoretical motivations that lead me to adopt Neo-Davidsonian semantics over these approaches. In Neo-Davidsonian event semantics, thematic roles can be treated as of one and the same kind other event properties like duration and spatial extent. This makes it a natural fit for the parallel treatment of distribution over individuals and distribution over time intervals. Furthermore, Neo-Davidsonian semantics allows us to adopt a uniform semantic architecture that makes it easier to express the theories in this paper and in Champollion (2014). For example, it exposes thematic roles to the compositional semantics, as opposed to keeping some or all of them implicit within the lexical entry of the verb. As we will see in Champollion (2014), this makes it easier to access them via theta indexing or function application. It also allows us to give uniform types to all verbal projections, which makes it easier for the distributivity operator to apply to each projection. Finally, it allows us to treat verbal and nominal projections as predicates of the same arity, which is useful from the point of view of algebraic semantics as it allows us to formulate cross-domain generalizations more easily (Champollion, 2010b). As a reviewer notes, adopting an asymmetric approach to event semantics as in Kratzer (1996) would lead one to expect asymmetries between the agent role and other thematic roles, since those other roles would not be directly accessible to the semantic theory. I have argued elsewhere that the asymmetries Kratzer tries to model correlate with syntactic positions and not with thematic roles (Champollion, 2010a). For other concerns with Kratzer's asymmetries, see Williams (2009). That said, it may be possible to reformulate much of the content of this paper and Champollion (2014) in other approaches than Neo-Davidsonian event semantics if one wishes to do so. For a detailed study of what it takes to reformulate Neo-Davidsonian theories in classical Davidsonian and eventless frameworks, see Bayer (1997).

I assume that thematic roles, events, and verbs are each closed under sum formation (Krifka, 1989; Champollion, 2010b). I discuss these assumptions in turn. The cumulativity assumption for thematic roles can be stated as follows:

(32) **Cumulativity assumption for thematic roles**
For any thematic role θ and any subset E of the domain of θ it holds that $\theta(\bigoplus E) = \bigoplus(\lambda x \exists e \in E.\ \theta(e) = x)$

This says that for any subset of the events on which a given thematic role θ is defined, we can compute the θ of their sum by summing up their θs. A consequence of this assumption is that thematic roles are homomorphisms, or structure-preserving maps, with respect to the $\oplus$ operation (pace Kratzer 2003 – for details see Champollion (2010b, p. 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.)

What this says is that, for example, if e is a talking event whose agent is John and e' is a talking event whose agent is Mary, $e \oplus e'$ is an event 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.

Turning now to events and verbs, I assume that the sum of any two events is itself an 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. Given that thematic roles are sum homomorphisms, the agent of the sum event $e_1 \oplus e_2$ is $j \oplus m$, the sum of their agents. I assume that whenever two events are in the denotation of a verb, then no matter whether they have the same thematic roles or not, and no matter if their runtimes are identical, adjacent, or otherwise, then the sum of these two events is also in the denotation of the verb. As was already mentioned in Section 3, this is a common and well-motivated assumption in event semantics (Scha, 1981; Schein, 1986, 1993; Lasersohn, 1989; Krifka, 1986, 1992; Landman, 1996, 2000). I refer to this assumption as *lexical cumulativity* (Kratzer, 2007). In the current 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 $e_1 \oplus e_2$. On the lexical cumulativity view, verbs can be said to have plural denotations, in the sense that their denotation obeys the same equation (34) as plural count nouns on the inclusive view of the plural, represented in (35):

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

(35) $[\![\mathrm{N}_{pl}]\!] = {}^*[\![\mathrm{N}_{sg}]\!]$

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

The lexical cumulativity assumption is motivated by the entailments in (36) and (37) (Krifka, 1989, 1992). Because of the parallelism between (34) and (35), the explanation of these entailments is completely analogous to the explanation of the entailment in (38), which motivated the treatment of plurality in Link (1983).

(36) a. John slept.
b. Mary slept.
c. ⇒ John and Mary slept.

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

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

Lexical cumulativity does not entail that all verb phrases are cumulative, only that all verbs are. For example, the sum of two events in the denotation of the verb phrase *be paid exactly $7,000* is not again in its denotation as a general rule, because it will

usually involve $14,000 rather than $7,000. Likewise, the sum of two events in the denotation of the verb phrase *find a flea* will only end up in its denotation if the two events happen to involve the same flea. This fact is important because of the distinction between lexical and phrasal distributivity. If all verb phrases were cumulative, then under the present assumptions there would be no way to explain why nonatomic distributivity is not readily available at the verb phrase level, as discussed in Section 3.

I have assumed above that the type of verbs and their projections is vt and that they denote predicates of events. I have argued elsewhere to the contrary that the type of verbs and verbal projections should be taken to be $\langle vt, t\rangle$ if we want to deal with the interaction of verbs with quantification and other scope-taking phenomena (Champollion, 2011, 2015b). Here I stick with the more standard vt assumption, in part in order to make sure the system remains compatible with the majority of existing theories and in part because the lower type is sufficent for present purposes. I will also assume that event predicates can combine with other event predicates by a generalized form of intersection, similar to the predicate modification rule in Heim and Kratzer (1998). This is similar to the event identification rule introduced by Kratzer (1996). The idea that verbs and their arguments are combined by intersection is also argued for in Carlson (1984) and is elevated to a general principle, called *conjunctivism*, in Pietroski (2005, 2006).

Section 2 has discussed the distinction between lexical and phrasal distributivity. As we have seen there, lexical distributivity is a property that a given verb may or may not have with respect to one of its thematic positions. I will represent lexical distributivity using meaning postulates, which put restrictions on admissible models (Hoeksema, 1983). Such postulates can be used to state that whenever an (agent-)distributive predicate applies to (an event whose agent is) a plurality of individuals, then it also applies to (events whose agents are) all the individuals in the plurality. For example, the verb *see* is lexically distributive on (at least) its agent and theme positions, since it follows both from *John and Mary saw Bill* and from *John saw Bill and Mary* that John saw Bill. As another example, *kill* is distributive on its theme role but not on its agent role (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, (39a) entails (39b) but does not entail (39c).

(39) a. The police officers killed Bonnie and Clyde.
b. $\Rightarrow$ The police officers killed Bonnie.
c. $\not\Rightarrow$ Sheriff Jordan killed Bonnie and Clyde.

As this example illustrates, whenever a group of people is killed then each of them is killed, while a group of people can kill a person without it being the case that each of them kills that person. To take another example, from *John and Mary lifted box b* it does not follow that *John lifted box b*.

I propose to capture lexical distributivity in event semantics by lexical meaning postulates such as the following:

(40) **Meaning postulate: *see* is distributive on its theme position**
$\forall e.{}^*\mathrm{see}(e) \rightarrow e \in {}^*\lambda e'({}^*\mathrm{see}(e') \wedge \mathrm{Atom}(\mathrm{theme}(e')))$

This meaning postulate can be read as follows. Whenever there is a seeing event e, it consists of one or more seeing events e' whose themes are atoms. For example, if e is an event in which John and Bill were seen, then that event consists of (in this case two) seeing events whose themes are atoms. From this and from the assumption that John and Bill are mereological atoms, one can conclude that John was seen and that Bill was seen. We can capture the difference between the agent and theme role of *kill* by adopting a meaning postulate analogous to (40) only for the theme position of that verb. So in effect, (40) states that only singular individuals can see.

It is important to note that the meaning postulate in (40) is not part of the lexical entry for *see*. That lexical entry is given in (41), which is closed under sum, in accordance with (34).

(41) **Lexical entry for *see***
$[\![\text{see}]\!] = \lambda e.{}^*\text{see}(e)$
(The meaning of "see" is the property that holds of any sum of one or more seeing events.)

A meaning postulate like (40) is not part of the meaning of the word *see*, but rather it describes something about "the way reality seems to be organized" (Hoeksema, 1983). It can be seen as a restriction on admissible models, with the consequence that the denotation of *see* always has a certain higher-order property, namely stratified reference (Champollion, 2010b, 2015a). Accordingly, it would be misguided to incorporate (40) into the lexical entry for *see*. This division of labor between lexical entries and higher is a hallmark of algebraic semantics. Formalizing distributivity in terms of the star operator might seem unusual because many treatments of distributivity use universal quantification instead. The advantage of using the star operator will become clearer in Section 4, where I will compare it with a universal-quantifier based approach.

We are now ready to reformulate the D operator. The main idea is that this operator shifts any predicate (typically a verb phrase) into a predicate that satisfies a condition that is analogous to the one captured in the meaning postulate in (40). Concretely, I propose to redefine the D operator as follows:

(42) **Definition: Event-based D operator**
$[\![\text{D}_\theta]\!] \stackrel{\text{def}}{=} \lambda V_{\langle vt\rangle}\lambda e[e \in {}^*\lambda e'(V(e') \wedge \text{Atom}(\theta(e')))]$

The D operator is indexed with a variable θ, which I assume is resolved to a theta role. It applies to an event predicate V, such as a verb phrase. It returns another event predicate, one which holds of any event e as long as it consists of one or more events that are in V and which are mapped by the function θ to atoms. The variable θ is instantiated by a thematic role – a function that maps events to their agents, themes, and so on. I adopt the assumption that each event has at most one agent, at most one theme, etc. (Carlson, 1984, 1998; Parsons, 1990). This is why I represent roles as functions. Modeling thematic roles as functions will be useful once we extend the D operator to nonatomic domains like time and space, because it will allow us to rely on the idea that thematic roles can be treated as being the same kinds of things as functions that map events to their locations in spacetime.

The following example illustrates how the D operator in (42) works. Sentence (43)

gives a baseline, a scopeless reading that does not use the D operator. Sentence (44) shows the D operator in action to model a distributive reading.

(43) The boys saw a monkey.
$\exists e[{}^*\text{agent}(e) = \bigoplus \text{boy} \ \wedge \ {}^*\text{see}(e) \ \wedge \ \text{monkey}(\text{theme}(e))]$
(There is a potentially plural seeing event whose agents sum up to the boys, and whose theme is one monkey. That is, only one monkey is seen.)

(44) The boys D(saw a monkey).
$\exists e[{}^*\text{agent}(e) = \bigoplus \text{boy} \wedge$
$e \in {}^*\lambda e'({}^*\text{see}(e') \ \wedge \ \text{monkey}(\text{theme}(e')) \ \wedge \ \text{Atom}(\text{agent}(e')))]$
(There is an event whose agents sum up to the boys, and this event consists of seeing events for each of which the agent is a boy and the theme is a monkey.)

The star operator ${}^*\lambda e'$ is introduced through the D operator and takes scope over the predicate *monkey* introduced by the theme. The representation (44) does not state explicitly that each boy sees a monkey so it might not be clear that it is an adequate way to capture what the sentence means. This is where the background assumptions introduced above come into play. The representation (44) explicitly states that the monkey-themed events e' have atoms as agents. The fact that these atoms are boys is entailed by the background assumption that the entities in the denotation of singular count nouns are atoms, together with the background assumption that thematic roles are cumulative, as discussed above. For details, see Champollion (2010b). The representation (44) also does not state explicitly that the monkey-themed events e' are seeing events. This is entailed by the meaning postulate in (40) together with the just-mentioned facts.

Coming now to the event-based reformulation of Schwarzschild's Part operator, this one can be seen as a generalization of Link's D operator. Instead of specifying the granularity parameter to be atomic, we leave it free. When we instantiate that parameter with *Atom*, we get Link's D operator as in (42). Accordingly, we obtain the reformulation by replacing *Atom* in (42) with a free variable C, which I will assume is anaphoric on an antecedent that can be provided by the context. This minimal change reflects the close connection between D and Part.

(45) **Definition: Event-based Part operator**
$[\![\text{Part}_{\theta,\text{C}}]\!] \stackrel{\text{def}}{=} \lambda P_{\langle vt\rangle}\lambda e[e \in {}^*\lambda e'(P(e') \ \wedge \ C(\theta(e')))]$

What this says is that Part takes an event predicate P and returns a predicate that holds of any event e which can be divided into events that are in P and whose θs satisfy the contextually salient predicate C. Note that whenever θ is a sum homomorphism, this will entail that the θs of these events sum up to the θ of e.

Definition (45) entails that C is what we may call a mereological cover of $\theta(e)$, that is, C is a set whose sum is $\theta(e)$. In this way, the notion of cover emerges naturally and does not need to be separately defined. This may be seen as a conceptual advantage over Schwarzschild (1996).

The following examples show how my reformulation of Schwarzschild's Part operator works. Again, I start by giving a sentence in which the operator is not used,

followed by a sentence in which the operator is used. The example is repeated from the shoe sentence (3). I assume for concreteness that the shoes play the theme role in this sentence. I assume that in these sentences, C is resolved to a contextual predicate that I call *pair* and that applies to a sum of two shoes just in case they are one of the pairs of shoes on display.

(46) The shoes cost fifty dollars. *collective*

a. $\exists e.{}^*\text{theme}(e) = \bigoplus \text{shoe} \wedge [\![\text{cost fifty dollars}]\!](e)$
(There is a costing-fifty-dollar event whose theme is the shoes.)

(47) The shoes $\text{Part}_{\textit{theme,pair}}$ cost fifty dollars. *distributive*

a. $\exists e.{}^*\text{theme}(e) = \bigoplus \text{shoe} \wedge \text{Part}_{\textit{theme,pair}}([\![\text{cost fifty dollars}]\!])(e)$

b. $\exists e.{}^*\text{theme}(e) = \bigoplus \text{shoe} \wedge e \in {}^*\lambda e' [e' \in [\![\text{cost fifty dollars}]\!] \wedge \text{pair}(\text{theme}(e'))]$
(There is a plural event whose themes sum up to the shoes and which consists of costing-fifty-dollars events with pairs as themes.)

The event-based reformulation of the D and Part operators allows us to think of distributivity as a parametrized operator. The D operator only has one parameter, θ, which specifies the thematic role over which it distributes. The Part operator has an additional granularity parameter, C, which corresponds to Schwarzschild's cover variable and which specifies the size of the things over which it distributes. These two parameters are at the core of strata theory, as discussed in Section 1 and in Champollion (2010b). The dimension parameter of strata theory corresponds to the θ parameter on our operators, and it indicates the domain that contains the entities over which the operator distributes. The granularity parameter of strata theory corresponds to the C parameter of the Part operator, and it indicates the size of the entities over which the operator distributes. I assume that the setting "granularity=atom" blocks the setting "dimension=time" because time is continuous and noncount – either there are no atoms to distribute over, or they are not accessible to natural language semantics (von Stechow, 2009). In Champollion (2014), I rely on this interaction in order to explain certain typological facts involving distance-distributive items.

5 Previous Work: Lasersohn (1998)

This section compares my reformulation with another proposal for reformulating the D operator in event semantics, due to Lasersohn (1998). The focus of Lasersohn's paper is technical. It shows a way of generalizing Link's and other D operators so that they apply to other positions than the subject position both in eventless and in event-based frameworks. The following entry is a special case among these different combinations, namely the VP-level version of an event-based version of Link's operator. It does not represent Lasersohn's entire proposal, but it is the part of his proposal that is most closely related to mine.

(48) **Distributivity operator over events (Lasersohn)**
$[\![\text{D}]\!]\ (\text{Lasersohn}) = \lambda P_{\langle e,vt\rangle} \lambda x \lambda e \forall y [y \leq_{Atom} x \rightarrow \exists e' [e' \leq e \wedge P(y)(e')]]$

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 $\lambda x \lambda e[\text{smile}(e) \wedge \text{agent}(e) = x]$. 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 indexed 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, unlike Lasersohn's. Since there is no consensus on how thematic roles are introduced, both options appear viable.

Another, more serious difference between my proposal and Lasersohn's consists in the way in which the two operators access the events over which they distribute. My operator in (42) 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 representations that result from inserting a D operator into *The girls smiled* before existential closure applies. (The subformula $y \leq_{Atom} \text{girl}$ is equivalent to $\text{girl}(x)$ on the assumption that girls are mereological atoms.)

(49) a. Lasersohn's representation:
$\lambda e \forall y[y \leq_{Atom} \bigoplus \text{girl} \rightarrow \exists e'[e' \leq e \wedge \text{smile}(e') \wedge \text{agent}(e') = y]$

b. My representation:
$\lambda e[^*\text{agent}(e) = \bigoplus \text{girl} \wedge e \in {}^*\lambda e'[\text{smile}(e') \wedge \text{Atom}(\text{agent}(e'))]]$

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. Thus, Lasersohn's solution suffers from leakage.

Leakage causes a problem in connection with subjects and other arguments and modifiers that take scope over the distributivity operator. This will be relevant in Section 6, where I will argue that the distributivity operator can also occur in the scope of *for*-adverbials. The problem can be illustrated with problems with adverbials such as *surprisingly* or *in slow procession*, provided that these adverbials are represented as event predicates, as has already been noted by Schein (1993). These predicates hold of an event even if they do not hold of its parts. The following example by Schein is a case in point:

(50) Unharmoniously, every organ student sustained a note on the Wurlitzer.

This example is true if the ensemble event was unharmonious even the same cannot be said of any one student's note.

The problem with Lasersohn's event predicate (49a) is that whenever it applies to an event e, it also applies to any event of which e is a part. Let L stand for (49a) and let M stand for my event predicate (49b). Imagine a scenario in which there are some girls and they smile, nobody else smiles, and there are some boys and they cry, and nobody else cries. In this scenario, consider the sum of all smiling events and call it e_{smile}. This itself counts as a smiling event (by lexical cumulativity), and its agent is the girls, so it satisfies both Lasersohn's predicate L and my predicate M. Consider the sum of

all crying events and call it e_{cry}. This is a crying event whose agent is the boys. Let $e_{smile \oplus cry}$ be the sum of e_{smile} and e_{cry}. Now $e_{smile \oplus cry}$ 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 $e_{smile \oplus cry}$ does satisfy Lasersohn's predicate L, because by virtue of containing e_{smile}, it contains a smiling event for every girl. Imagine that event e_{smile} is not surprising by itself, but that the two events taken together are surprising, that is, $e_{smile \oplus cry}$ is surprising. This might be the case, for example, because the girls and the boys normally always behave the same, so it is unusual that some of them smile if the other ones cry. Given that e_{smile} by itself is not surprising, sentence (51) is intuitively judged false in this scenario. If the D operator is applied to *smile*, then on Lasersohn's account, this sentence is represented as (51a), while on my account it is represented as (51b). The problem is that $e_{smile \oplus cry}$ satisfies both L (by leakage) and the predicate *surprising* (by assumption). Therefore, Lasersohn's D operator wrongly predicts that (51) is judged true.

(51) Surprisingly, the girls smiled.
   a. $\exists e[\text{surprising}(e) \wedge L(e)]$
   b. $\exists e[\text{surprising}(e) \wedge M(e)]$

We see 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. Similarly to what we have seen in connection with (40), a Lasersohn-style analysis could be rescued by means of event minimality or exemplification. Depending on the details, the result would likely come out as equivalent to my representation. For reasons of succintness, and in keeping with a longstanding tradition in the semantics of distributivity and plurality, I will continue to use star operator notation rather than relying on exemplification.

The nature of the problem with Lasersohn's implementation consists in the fact that it produces representations like (49a) that do not give special status to the sum of all of the subevents over which the D operator distributes. Instead, (49a) not only applies to that sum but also to any event that contains that sum. In other words, (49a) is persistent (Kratzer, 1989). By contrast, my representation (49b) only applies to the sum itself. This is due to the way the star operator works, as can be seen from the definition in (30). To expand on the geometrical example given there, if x is a square and P is the property of being a triangle, then x satisfies ${}^*P$ because x can be divided into triangles *without leaving anything out*. If x is a circle instead of a square, it does not satisfy ${}^*P$. This is because, even though every circle contains infinitely many triangles, the curvature of the circle makes it impossible to divide it into triangles without leaving anything out.

The star operator is not the only way to avoid the leakage problem. Instead, as a reviewer notes, one could combine Lasersohn's operator with the following minimization operator.

(52) $\min \stackrel{\text{def}}{=} \lambda V \lambda e.V(e) \wedge \forall e'[e' < e \rightarrow \neg V(e')]$
(An event is minimal with respect to V just in case V holds of it but not of

any of its proper parts.)

The combination of Lasersohn's operator (48) with the minimization operator would be as follows:

(53) $\lambda P_{\langle e,vt\rangle}\lambda x\lambda e.e \in \min(\lambda e'\forall y[y \leq_{Atom} x \rightarrow \exists e''[e'' \leq e' \wedge P(y)(e'')])$

When this amended version is inserted into *The girls smiled*, the resulting representation before existential closure applies is the following:

(54) $\lambda e.e \in \min(\lambda e'\forall y[y \leq_{Atom} \bigoplus \text{girl} \rightarrow$
$\exists e''[e'' \leq e' \wedge {}^*\text{agent}(e'') = y \wedge \text{smile}(e'') \wedge \text{agent}(e'') = y])$

This representation applies to all events that are minimal with respect to the property of containing a smiling subevent for each girl. In the example above, e_{smile} will be such an event, but $e_{smile \oplus cry}$ will not, because it is not minimal with respect to that property. Building minimality into the distributivity operator can lead to unexpected consequences, because they may not always be any minimal events. Take a predicate with the subinterval property, such as *sit*. If we assume that *sit* is distributive on its agent position and that time is nonatomic, applying (53) leads to the empty set because there are no minimal sitting events. One way to avoid this problem may be to replace minimality by exemplification. An entity e exemplifies a predicate V iff V holds of e and of either none or else all of the proper parts of e (Kratzer, 2007). Rather than pursuing this route further, I will adopt the operator in (42) because it is more simple and concise than (53).

As we have seen, to avoid leakage we need to reformulate the distributivity operator in a way that makes sure that no extraneous material can find its way into the sum event. This ensures that the output of the operator is the right kind of predicate to be passed on to the next argument or adjunct of the verb. In the previous examples, this was an event modifier like *surprisingly*. In the next section, I will exploit this property of the distributivity operator in another way: it will apply to a verb phrase as before, but its output will serve as the input to a *for*-adverbial, which in turn will pass it on to the subject of the sentence after making sure that the output of the distributivity operator is atelic.

6 The Scopal Behavior of *For*-Adverbials

This section identifies cases of nonatomic phrasal distributivity in the temporal domain. I take this to be a nonatomic domain. As discussed at the end of Section 3, this means there are necessarily no atomic covers.

In the following, τ stands for the "runtime" function that maps each event to the location in time at which it occurs. I assume that τ has the properties ascribed to it in Krifka (1998). Specifically, I assume that τ is a sum homomorphism. This means that runtimes can be discontinuous. Take for example an event e_1 whose runtime is the interval from 12:30 to 1pm, and an event e_2 whose runtime is the interval from 5pm to 6:15pm. The sum of these two events, $e_1 \oplus e_2$ will have a discontinuous runtime,

namely the sum of the interval from 12:30pm to 1pm and the interval and the interval from 5pm to 6:15pm.

As was already discussed in the preview in Section 1, as a general rule, *for*-adverbials by themselves do not cause indefinites in their scope to covary, but there are certain conditions in which they do: intervening modifiers like *every day*, and contexts where a temporal partition is salient (Zucchi and White, 2001; Deo and Piñango, 2011). This section presents these facts in more detail and provides an account of them. The main idea is that a distributivity operator can intervene between the verb phrase and the *for*-adverbial in the same way and with roughly the same kind of meaning as *every day* or *every hour* and so on, depending on the situation. This operator is anaphoric on context. In a nonatomic domain like time, atomic distributivity is not an option. So when the dimension parameter of the distributivity operator involved is instantiated to τ (runtime), its granularity parameter cannot be set to *Atom*. I assume that nonatomic levels of granularity are only available when they are salient, just as in the case of Schwarzschild's Part operator. On this view, then, indefinites that seem to covary with a *for*-adverbial actually covary with a covert Part operator.

Let me start by contrasting two simple theories of *for*-adverbials which I will call A and B. Theory A predicts that all indefinites in the scope of *for*-adverbials should covary while Theory B predicts that none of them should. Neither of these theories turns out to account for the facts by itself. I will propose to account for the limited covariation by adding a distributivity operator to Theory B. Roughly these two kinds of theories (but without the distributivity operator added to them) are also discussed by Zucchi and White (2001). On Theory A, the meaning of a *for*-adverbial can be represented as something like the following:

(55) $[\![\text{for an hour}]\!]$ (Theory A) =
$\lambda P_{\langle vt\rangle} \exists t[\text{hours}(t) = 1 \ \wedge \ \text{regular}(t) \wedge$
$\forall t'[t' \text{ is a very short part of } t \rightarrow \exists e[P(e) \ \wedge \ \tau(e) = t')]]$

On Theory B, the meaning of a *for*-adverbial can be represented as follows:

(56) $[\![\text{for an hour}]\!]$ (Theory B)
$= \lambda P_{\langle vt\rangle} \lambda e : \text{regular}(\tau(e)) \wedge \text{atelic}(P).P(e) \ \wedge \ \text{hours}(\tau(e)) = 1$

Theory A says that a *for*-adverbial is represented as a universal quantifier over very short subintervals of an approximately continuous interval, a bit as if it was the temporal counterpart of *every*. Theory B says that a *for*-adverbial is represented in a way that passes the denotation of the verb phrase up unchanged as long as it is atelic and approximately continuous. The predicate *atelic* and *regular* are shorthands whose precise elaboration is not at issue in this paper. However, let me comment briefly on each of them. The predicate *regular* is required because we allow temporal intervals to be discontinuous, as described above, and because *for*-adverbials are tolerant of gaps but only up to a point (Vlach, 1993). On how to spell out this requirement, see for example Piñón (1999) and Landman and Rothstein (2012a,b). This feature is not unique to *for*-adverbials but can be found in other temporal modifiers as well, albeit to varying degrees. For example, the sentences in (57) can all be imagined to be true even on the assumption that John doesn't work at night, or during mealtimes, or on

the weekends, but less likely so if he worked on the paper only on May 1st and May 31st. Between these extremes, the judgments get less clear.

(57) a. John worked on this paper for a month.
b. John worked on this paper from May 1 to May 31.
c. John wrote this paper in a month.

The amount of gaps that are tolerated in a true sentence can vary considerably from one case to another, even when we keep the verb and adverbial constant. This is illustrated in the following pair of examples (Barbara Partee p.c. to Vlach, 1993):

(58) a. Mary slept for a week.
b. Mary slept in the attic for a week.

As Vlach notes, the truth conditions of these sentences are very different. Sentence (58a) requires Mary to sleep almost continuously; sentence (58b) is compatible with an ordinary sleeping pattern of about eight hours a day. These variations are presumably due in part to pragmatic reasons. I will not try to account for them in this paper.

As for the atelicity requirement encoded here by the predicate *atelic*, it is the focus of Chapters 5 and 6 of Champollion (2010b), where I show that like the theory of distributivity presented here, it too can be formulated in terms of strata theory. Briefly, the requirement can be formulated in a way that is roughly parallel to the meaning postulate in (59), except that the dimension and granularity parameters involved are set to τ and to a very short part of the temporal interval introduced by the *for*-adverbial. When this is spelled out, in the concrete case of *for an hour*, it amounts to the following:

(59) **Atelicity requirement: *P* is distributive wrt. an hour**
$\forall e.P(e) \rightarrow [e \in {}^*\lambda e'(P(e') \ \wedge \ \tau(e') \text{ is very short wrt. an hour})]$

Theory A is somewhat similar to the influential theory of *for*-adverbials in Dowty (1979). However, there is an important difference because Dowty analyzes *for*-adverbials as universal quantifiers over subintervals rather than over instants. These do not have to be proper subintervals, so the predicate is required to hold at the entire interval described by the *for*-adverbial, in addition to all of its subintervals. A similar theory, in which *for*-adverbials quantify over "relevant" subintervals (whatever that may mean) is found in Moltmann (1991). Theory B is found in various forms in Krifka (1986, 1989, 1998). It is explicitly defended against Theory A in Kratzer (2007) and Champollion (2010b).

Neither Theory A nor Theory B are able to account for the limited ability of indefinites to covary in the scope of *for*-adverbials, at least not without further modifications. This is because Theory A predicts that they should always covary and Theory B predicts that they should never covary. On Theory A, 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, *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 as we have seen in Section 1. I suggest that these limited cases 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. These sentences are supposed to be understood as uttered out of the blue, without any special context.

(60)	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 these examples show, indefinites are not able to covary with a *for*-adverbial when they occur in its syntactic scope in the way they can covary with a universal quantifier when they occur in its syntactic scope.

As I have mentioned in the introduction, 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. These examples should also be understood as uttered out of the blue. Example (61a) is from Zucchi and White (2001). Examples (61b) and (61c) are from Deo and Piñango (2011).

(61)	a.	??John found a flea on his dog for a month.	$??\exists > \forall$; $^*\forall > \exists$
	b.	??John noticed a discrepancy for a week.	$??\exists > \forall$; $^*\forall > \exists$
	c.	??John discovered a new proof for a week.	$??\exists > \forall$; $^*\forall > \exists$

The sentences in (60) and in (61) all involve reference to a single cart that is pushed, a single flea that is found repeatedly, and so on. It would be plausible for (60b) to have an interpretation like *Over and over again over the course of five minutes, I dialed a different wrong phone number.* Despite its plausibility, this kind of interpretation is systematically absent from such examples when they are understood as uttered out of the blue.

This lack of covariation is surprising if *for an hour* is interpreted as *at each very short part of a given hour*, as is the case in Theory A. On Theory A, 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. One might try to account for these facts by stipulating obligatory quantifier raising of the indefinite above the *for*-adverbial, as suggested by Krifka (1998). But the phenomenon can also be observed in German, a language that is otherwise known for having a preference for surface scope. Each of the following two sentences must involve reference to a single phone number (Kratzer, 2007).

(62)	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.

To account for this fact, an explanation based on obligatory quantifier raising would have to work against the German preference for surface scope, as Kratzer explains. And it is contradicted by the fact that covariation is possible in examples like (63), where a temporal universal quantifier intervenes. This example is repeated here from (6), and was pointed out by Zucchi and White (2001).

(63) John found a flea on his dog every day for a month.

This example does not require John to find the same flea multiple times, so covariation is possible here. If *for*-adverbials forced all indefinites in their syntactic scope to move and take semantic scope above them, then *every day* should not be able to prevent this from happening in this case, contrary to what we observe.

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, sentence (64) cannot be interpreted as saying that over the course of three hours, John saw different sets of thirty zebras. A scopeless, cumulative reading which would correspond to the interpretation of *John saw thirty zebras in three hours* is not available either. The only available interpretation is the one in which the plural indefinite *thirty zebras* does not covary.

(64) John saw thirty zebras for three hours.

The behavior just described does not hold across the board for all types of noun phrases. Bare plurals and mass nouns do not have to take distributive wide scope over *for*-adverbials (Carlson, 1977; Verkuyl, 1972; Dowty, 1979). The following sentences are taken from Dowty (1979):

(65) a. John found fleas on his dog for a month.
b. John discovered crabgrass in his yard for six weeks.

(66) a. Tourists discovered that quaint little village for years.
b. Water leaked through John's ceiling for six months.

The scopal behavior of bare noun phrases with respect to *for*-adverbials mirrors their well-known tendency to take narrow scope in general (Carlson, 1977). VP-level and sentential predicates with bare noun phrases appear to be generally compatible with *for*-adverbials (Verkuyl, 1972; Dowty, 1979). The sentences above stand in marked contrast to the examples with indefinites, because there is no sense in which the bare noun phrase in them has to be interpreted with wide scope. For example, (65a) is compatible with the plausible interpretation in which John finds different fleas on his dog over the course of a month and finds each of them only once.

Covariation is also possible when a salient level of granularity can be inferred from context. To repeat the example from Section 1, in a context where the daily pill intake of patients is salient such as a hospital, sentence (67) is licit despite the fact that it does not require any pill to be taken more than once.

(67) **Context: discussing daily pill intake**
The patient took two pills for a month and then went back to one pill.

Example (68) shows the same point. It is understood as involving reference to several snowmen. This is presumably because world knowledge makes the cycle of seasons salient here. This example is taken form Deo and Piñango (2011).

(68) We built a huge snowman in our front yard for several years.

Finally, example (69) is adapted from Landman and Rothstein (2009). These authors 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, which would of course be impossible.

(69) This bicycle carried three children around Amsterdam for twenty years.

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 Theory A, which analyzes *for*-adverbials as universal quantifiers over instants. At the same time, Theory B by itself is not a good fit either, because it would predict no scope ambiguities at all. So I will extend Theory B with a temporal version of the nonatomic distributivity operator Part that we have seen in Section 3. As we have seen, this operator introduces a contextual variable that is resolved to a cover. This is related to the notion of a contextually determined partition that originates independently in Moltmann (1991) and in Deo (2009). These authors place the anaphoricity on context into the *for*-adverbial itself. This approach is also advocated in Deo and Piñango (2011). For a detailed argument to the effect that it is the distributivity operator and not the *for*-adverbial that is anaphoric on context, see Champollion (2013).

In the following development, I first discuss cases where Theory B works without modifications, namely those in which the indefinite does not covary, and then extend it to those cases where it covaries by adding the distributivity operator. Given our background assumptions, Theory B immediately predicts that the indefinite in (60a), repeated here as (70), must take wide scope.

(70) John pushed a cart for an hour. = (60a)

This prediction is obtained based on the following representation of the denotation of the verb phrase *push a cart*, which assumes lexical cumulativity (34):

(71) $[\![\text{push a cart}]\!] = \lambda e[{}^*\text{push}(e) \ \wedge \ \text{cart}({}^*\text{theme}(e))]$
(True of any pushing event or sum of pushing events whose theme is one and the same cart.)

Even though the verbal denotation *push* is pluralized here, 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. In connection with the entry (56), it predicts that the entire event over which sentence (70) existentially quantifies must

have a single cart as its theme:

(72) ⟦John pushed a cart for an hour⟧
$= \exists e : \text{regular}(\tau(e)) \ \wedge \ \text{atelic}(\lambda e'[{}^*\text{push}(e') \ \wedge \ \text{cart}({}^*\text{theme}(e'))]).$
$[{}^*\text{agent}(e) = j \ \wedge \ {}^*\text{push}(e) \ \wedge \ \text{cart}({}^*\text{theme}(e)) \ \wedge \ \text{hours}(\tau(e)) = 1]$
(Defined iff *push a cart* is atelic. If defined, true iff 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.)

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 (Kratzer, 2007). For example, a sentence like (73) (repeated from (61a) 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. The lexical cumulativity assumption 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 (73) requires a single flea to have been found repeatedly over the course of a month.

(73) ⟦John found a flea for a month⟧
$= \exists e : \text{regular}(\tau(e)) \ \wedge \ \text{atelic}(\lambda e'[{}^*\text{find}(e') \ \wedge \ \text{flea}({}^*\text{theme}(e'))]).$
$[{}^*\text{agent}(e) = j \ \wedge \ {}^*\text{find}(e) \ \wedge \ \text{flea}({}^*\text{theme}(e)) \ \wedge \ \text{months}(\tau(e)) = 1]$
(Defined iff *find a flea* is atelic. If defined, true iff 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.)

The representation in (73) does not require John to have been searching uninterruptedly at every moment of the month, provided that the runtimes of events are not required to be continuous, and that the function *months* maps a discontinuous interval to the same number as the smallest continuous interval that contains it. I will therefore adopt these assumptions from now on.

So far we have seen that only Theory B is able to deal with all the facts. But as discussed, without further modifications it also predicts that singular indefinites and numerals in the scope of *for*-adverbials should never be able to show covarying behavior, consistent with what we have seen in the examples in (60) but contrary to what we have seen in examples (67), (68), and (69). The only criterion that distinguishes these two groups of examples is the availability of a supporting context. We have observed an analogous effect in Section 3, when we considered the following examples, repeated here from (21):

(74) a. The shoes cost fifty dollars.
b. The suitcases weigh fifty pounds.

I claim that the mechanism which allows context to rescue examples like (67), (68), and (69) involves a distributivity operator over times which is anaphoric to a salient

cover of the time interval in question, in the same way as examples like (74a) involve a distributivity operator over shoes which is anaphoric to a salient cover of the collection of shoes in question. In other words, the reason that predicates like *find a flea* in examples like (61a) are not able to distribute over days out of the blue is the same as the reason that predicates like *weigh fifty pounds* in examples like (74b) are not able to distribute over nonatomic sums of suitcases. Both would need a Part operator in order to distribute, but in these examples the Part operator is not licensed because there is no salient cover either over days or over sums of suitcases.

The reformulation of the distributivity operator developed in Section 4 was 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. I suggest that the thematic role parameter can also be set to τ, or runtime. As mentioned before, I assume that 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). This leads to the expectation 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 very same restrictions as its counterpart, the Part operator.

I will obtain this temporal nonatomic distributivity operator by instantiating my reformulation of the Part operator in (45) with suitable dimension and granularity parameters. Let me describe each of these parameters in term. I will instantiate the dimension parameter as τ.

(75) **Definition: Event-based temporal Part operator**
$[\![\text{Part}_{\tau,\text{C}}]\!] \stackrel{\text{def}}{=} \lambda P_{\langle vt\rangle}\lambda e[e \in {}^*\lambda e'(P(e') \ \wedge \ C(\tau(e')))]$
(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 contextually salient 'cover predicate' C.)

We can think of the insertion of this temporal instantiation of the nonatomic distributivity operator Part as a repair strategy that shifts a predicate into another one in order to satisfy the atelicity presupposition a *for*-adverbial. For example, the predicate *take a pill* by itself is punctual and telic, an achievement predicate in the terminology of Vendler (1957). Even the lexical cumulativity assumption does not change much about this fact since it widens this predicate only insofar as there exist pills that are eaten repeatedly. But after applying the operator in (75) to it, the predicate is now iterative and atelic. Depending on the value to which the variable C of this operator is resolved, the meaning of the shifted predicate can be paraphrased as *take a pill every day*, or *take a pill every hour*, etc. This kind of predicate can be empirically and formally shown to be atelic. Empirically, these predicates are atelic as shown by examples like (76), modeled on (63). Formally, under an appropriate definition of atelicity formulated in terms of strata theory, it can be shown that any predicate that results from the application of this operator is atelic (Champollion, 2010b).

(76) The patient took a pill every day for a month.

We are now ready to explain the difference between examples (61a) and (67), repeated here:

(77) **Uttered out of the blue:**
??John found a flea on his dog for a month.

(78) **Context: discussing daily pill intake**
The patient took two pills for a month and then went back to one pill.

In example (78), I assume that the verb phrase *take two pills* has the following denotation prior to the application of the distributivity operator:

(79) $[\![\text{take two pills}]\!] = \lambda e[{}^*\text{take}(e) \ \wedge \ {}^*\text{pill}({}^*\text{theme}(e)) \ \wedge \ |{}^*\text{theme}(e)| = 2]$

This predicate applies to events in which a total of two pills are taken. It cannot combine directly with the *for*-adverbial because it is not atelic. The operator in (75) can be used as a repair strategy, provided that it is available. In (78), this is the case, because there is a salient level of granularity in the context and C can be resolved to it. That is, (78) is uttered in a context in which the granularity *daily* is salient. This is as opposed to (77), which is uttered in a default context that does not contain anything that C can be resolved to.

Having explained the contrast between (77) and (78), I now turn to explaining the interpretation of the latter. (78) can be interpreted by applying $\text{Part}_{\tau,\lambda t.\text{days}(t)\leq 1}$ to its verb phrase. The result of this operation is as follows:

(80) $\text{Part}_{\tau,\lambda t.\text{days}(t)\leq 1}(\lambda e[{}^*\text{take}(e) \ \wedge \ {}^*\text{pill}({}^*\text{th}(e)) \ \wedge \ |{}^*\text{th}(e)| = 2])$
$= \lambda e[e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\text{take}(e') \ \wedge \ {}^*\text{pill}({}^*\text{th}(e')) \ \wedge \ |{}^*\text{th}(e')| = 2] \wedge \\ \text{days}(\tau(e')) \leq 1 \end{array} \right)]$
(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*. Under plausible assumptions about the compositional process, the result of this computation is the following (I am omitting the atelicity requirement for clarity):

(81) $\exists e : \text{regular}(\tau(e)).[{}^*\text{ag}(e) = \text{the.patient} \ \wedge \ \text{months}(\tau(e)) = 1 \wedge$
$e \in {}^*\lambda e' \left(\begin{array}{l} {}^*\text{take}(e') \ \wedge \ {}^*\text{pill}({}^*\text{th}(e')) \ \wedge \ |{}^*\text{th}(e')| = 2] \wedge \\ \text{days}(\tau(e')) \leq 1 \end{array} \right)]$
(There is an approximately continuous 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.)

Under suitable assumptions regarding the interpretation of *months*, this formula is verified by an approximately continuous plural event provided that there is a month between the beginning and the end of this event. The subformula $\text{days}(\tau(e')) \leq 1$ makes sure that each of the taking-two-pills events e' takes place within a day. I assume that the *regular* conjunct ensures that the taking-two-pills events happen on every day of the month. I come back to this point in Section 7.

These truth conditions appear to correctly represent the way in which sentence (78) is interpreted. In particular, the pairs of pills covary with the days even though there is no predicate *daily* or *every day* explicitly mentioned in the sentence.

The strategy I have adopted here – Theory B plus a distributivity operator anaphoric on context – is broadly compatible with many lexical entries given for *for*-adverbials in the literature, including those that rely on higher-order properties such as divisive reference and the subinterval property (Dowty, 1979; Krifka, 1986, 1989). The strategy does not predict that all quantifiers take wide scope over *for*-adverbials. We have seen above that bare noun phrases do not they take wide scope with respect to *for*-adverbials, because they denote algebraically closed predicates. This is illustrated in the following minimal pair, repeated here from (65a) and (77):

(82) a. John found fleas on his dog for a month. *OK*
b. ??John found a flea on his dog for a month. *odd out of the blue*

As discussed in Section 3, I assume that a bare plural like *fleas* has an inclusive meaning (Spector, 2007; Zweig, 2009). That is, its denotation is *flea, the algebraic closure of its singular form, essentially *one or more fleas*. 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. Sentence (82a) does not entail that any one flea has been found several times, but only that there is a plural month-long interval over the course of which one or more fleas were found. This is entailed in the following representation of the denotation of (82a), where *fleas* is interpreted in situ as a predicate that applies to the theme of the verb:

(83) $[\![$John found fleas (on his dog) for a month$]\!]$
$= \exists e : \text{regular}(\tau(e)) \wedge \text{atelic}(\lambda e'[{}^*\text{find}(e') \ \wedge \ {}^*\text{flea}({}^*\text{theme}(e'))]).$
$[{}^*\text{agent}(e) = j \ \wedge \ {}^*\text{find}(e) \ \wedge \ {}^*\text{flea}({}^*\text{theme}(e)) \ \wedge \ \text{months}(\tau(e)) = 1]$
(Defined iff *find fleas* is atelic. If defined, true iff 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 definedness condition of this sentence is fulfilled since *find fleas* is atelic by most definitions of atelicity (Dowty, 1979; Krifka, 1998; Champollion, 2010b). Given the lexical cumulativity meaning postulate, any event in virtue of which (82a) is true can consist of several individual findings. No distributivity operator is needed because this is a case of lexical and not phrasal distributivity. This observation is parallel to the argument about meaning postulates by Lasersohn (1989). I discussed this argument in Section 3 in connection with sentences like (15) and (20), repeated here as (84a) and (84b):

(84) a. Rodgers, Hammerstein, and Hart wrote musicals. *OK nonatm. dist.*
b. Rodgers, Hammerstein and Hart wrote a musical. **nonatm. dist.*

The system presented here explains the contrast in (84) in the same way as the contrast in (82). In both cases, lexical cumulativity causes the sentence with the bare plural object to exhibit nonatomic lexical distributivity. And in both cases, the sentence with the singular indefinite object does not exhibit nonatomic phrasal distributivity because

the lack of supporting context means that the nonatomic distributivity operator Part is not available.

The idea that covariation of singular indefinites is due to the insertion of a covert operator finds additional support in the observation that covarying singular indefinites take extra time to process compared with bare plurals, as discussed in Deo and Piñango (2011). Reading time increases at the *for*-adverbial in (85a) compared with (85b) in self-paced reading tests conducted by Todorova et al. (2000). The indefinite in (85a) is able to covary, that is, it involves reference to more than one large check.

(85) a. Even though Howard sent a large check to his daughter for many years, she refused to accept his money.
b. Even though Howard sent large checks to his daughter for many years, she refused to accept his money.

On the view presented here, example (85a) involves the covert presence of a Part operator but example (85b) does not. This kind of contrast can be explained if we assume that the retrieval of an antecedent for the anaphoric variable C in the Part operator leads to higher processing load. For another view on these and related facts, on which this anaphoricity is built into the *for*-adverbial itself, see Deo and Piñango (2011). I take examples like (85a) to show that anaphoricity is built into the distributivity operator rather than into the *for*-adverbial. For a detailed argument in favor of this view, and a critique of the alternative view, see Champollion (2013).

7 Extensions

The theory of *for*-adverbials presented here can be extended to other modifiers that do not or not easily induce covariation of indefinites in their scope. For example, habitual sentences show analogous scopal effects to *for*-adverbials. Example (86a) is from Anna Szabolcsi (p.c.) and example (86b) from Krifka et al. (1995, 39f.).

(86) a. Mary smokes cigarettes / *a cigarette.
b. Mary smokes cigarettes / a cigarette after dinner.

(87) a. Yesterday, Mary smoked cigarettes / *a cigarette for an hour.
b. Last month, Mary smoked a cigarette after dinner for a week.

The scopal behavior of generics appears to be analogous to that of *for*-adverbials, as shown by the following pair (example (88a) is from Rimell, 2004).

(88) 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.

This suggests that the generic operator can be modeled using an analogue of what I have called Theory B of *for*-adverbials – essentially as a modifier rather than as a quantifier (see also Rimell, 2004).

Coming back to *for*-adverbials, I have discussed the ability of temporal universal quantifiers in connection with examples like (6), repeated here (Zucchi and White, 2001):

(89) John found a flea on his dog every day for a month.

In addition to universal quantifiers, other interveners that allow indefinites to seemingly covary with *for*-adverbials include adverbials like *day after day*:

(90) John found a flea on his dog day after day for a month.

Such adverbials are similar to the pluractional adverbials *dog after dog* discussed for example by Beck and von Stechow (2007). This suggests that an extension of the present account to pluractionals might be promising.

Similarly, *day after day* and *every day* both hard-wire the granularity parameter to the value $\lambda t[\text{days}(t) \leq 1]$ ("daily") instead of retrieving it anaphorically from context:

(91) $[\![\text{day after day}]\!] = [\![\text{every day}]\!] \stackrel{\text{def}}{=} \lambda P_{\langle vt \rangle} \lambda e[e \in {}^*\lambda e'(P(e') \wedge \text{days}(\tau(e')) \leq 1)]$
(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 are at most one day long)

These entries can be refined in various ways. For example, we may wish to add a condition which ensures that every day in a consecutive span is actually represented. The underlined subformula performs this role:

(92) $[\![\text{every day}]\!] \stackrel{\text{def}}{=} \lambda P_{\langle vt \rangle} \lambda e[e \in {}^*\lambda e'(P(e') \ \wedge \ \text{days}(\tau(e')) \leq 1) \wedge$
$\underline{\exists t.\text{continuous}(t) \wedge \forall t'.[t' \leq t \wedge \text{days}(t') = 1 \rightarrow \exists e'.e' \leq e \wedge P(e') \wedge \tau(e') \leq t']}]$
(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 are at most one day long, in such a way that every day in a consecutive span of days is represented)

Alternatively, the underlined subformula could be replaced by a contextually supplied variable that specifies the domain of quantification (see e.g. Stanley and Szabó, 2000). This variable could then be made dependent on another operator on the sentence, such as a *for*-adverbial that takes scope over *every day*. One could then try to derive the requirement that days be consecutive from general properties of domain restriction.

8 Summary and Discussion

I have shown how to integrate the atomic distributivity operator D of Link (1987) and the cover-based nonatomic distributivity operator Part of Schwarzschild (1996) into a unified framework, in which distributivity has two parameters: dimension and granularity. I have empirically distinguished lexical and phrasal distributivity. Building on earlier work, I have suggested that lexical distributivity should be modeled by the lexical cumulativity assumption (Lasersohn, 1989; Kratzer, 2007). Building on the observation that nonatomic phrasal distributivity exists but requires supporting context to surface, I have suggested that this fact motivates the possibility of a nonatomic granularity parameter setting for the distributivity operator. I have assumed that this happens so rarely because the operator is anaphoric on its context with respect to this

parameter (Schwarzschild, 1996). Examples like (21a), repeated below as (93a), have a nonatomic distributive reading because context provides a salient level of granularity, while examples like (21b), repeated below as (93b), 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 in (78), repeated below as (94a), to involve reference to different sets of two pills, while such a reading is not present in (94b).

(93) a. The shoes cost fifty dollars.
b. The suitcases weigh fifty pounds.

(94) a. The patient took two pills for a month and then went back to one pill.
b. ??John found a flea on his dog for a month.

References

Barker, Chris (2010). Free choice permission as resource-sensitive reasoning. *Semantics and Pragmatics* 3(10), 1–38. DOI: 10.3765/sp.3.10.

Bayer, Samuel L. (1997). *Confessions of a lapsed neo-Davidsonian: Events and arguments in compositional semantics*. New York, NY: Garland. DOI: 10.4324/9780203055106.

Beck, Sigrid and Uli Sauerland (2000). Cumulation is needed: A reply to Winter 2000. *Natural Language Semantics* 8(4), 349–371. DOI: 10.1023/A:1011240827230.

Beck, Sigrid and Arnim von Stechow (2007). Pluractional adverbials. *Journal of Semantics* 24(3), 215–254. DOI: 10.1093/jos/ffm003.

Blaheta, Don (2003). Binominal *each*: Evidence for a modified type system. MA thesis. Providence, RI: Brown University. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.7455&rep=rep1&type=pdf.

Brasoveanu, Adrian (2010). Modified Numerals as Post-Suppositions. *Logic, Language and Meaning* 6042, 203–212. DOI: 10.1007/978-3-642-14287-1_21.

Brisson, Christine (1998). Distributivity, maximality and floating quantifiers. PhD thesis. New Brunswick, NJ: Rutgers University. URL: http://ling.rutgers.edu/images/10.1.1.121.4641.pdf.

Brisson, Christine (2003). Plurals, *all*, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26(2), 129–184. DOI: 10.1023/A:1022771705575.

Carlson, Gregory N. (1977). Reference to kinds in English. PhD thesis. Amherst, MA: University of Massachusetts. URL: semanticsarchive.net/Archive/jk3NzRlY/.

Carlson, Gregory N. (1984). Thematic roles and their role in semantic interpretation. *Linguistics* 22(3), 259–280. DOI: 10.1515/ling.1984.22.3.259.

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. DOI: 10.1007/978-94-011-3969-4_3.

Champollion, Lucas (2010a). Cumulative readings of *every* do not provide evidence for events and thematic roles. *Logic, Language and Meaning* 6042, 213–222. DOI: 10.1007/978-3-642-14287-1_22.

Champollion, Lucas (2010b). Parts of a whole: Distributivity as a bridge between aspect and measurement. PhD thesis. Philadelphia, PA: University of Pennsylvania. URL: http://repository.upenn.edu/edissertations/958/.

Champollion, Lucas (2011). Quantification and negation in event semantics. *The Baltic International Yearbook of Cognition, Logic and Communication* 6, 1–23. DOI: 10.4148/biyclc.v6i0.1563.

Champollion, Lucas (2012). Each vs. Jeweils: A cover-based view on distance distributivity. *Logic, Language and Meaning* 7218, 251–260. DOI: 10.1007/978-3-642-31482-7_26.

Champollion, Lucas (2013). Man and woman: The last obstacle to boolean coordination. In: *Proceedings of the 19th Amsterdam Colloquium*. Ed. by Maria Aloni, Michael Franke, and Floris Roelofsen. Amsterdam, Netherlands: ILLC Publications, pp. 83–90. DOI: 10.13140/2.1.4779.1045.

Champollion, Lucas (2014). Overt distributivity in algebraic event semantics. Ms. New York University. URL: http://ling.auf.net/lingbuzz/002098.

Champollion, Lucas (2015a). Stratified reference: the common core of distributivity, aspect, and measurement. To appear in *Theoretical Linguistics*. URL: http://ling.auf.net/lingbuzz/002366.

Champollion, Lucas (2015b). The interaction of compositional semantics and event semantics. *Linguistics and Philosophy* 38(1), 31–66. DOI: 10.1007/s10988-014-9162-8.

Champollion, Lucas and Manfred Krifka (to appear). Mereology. In: *Cambridge handbook of semantics*. Ed. by Paul Dekker and Maria Aloni. Cambridge, UK: Cambridge University Press. URL: http://ling.auf.net/lingbuzz/002099.

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. DOI: 10.1093/0199246270.003.0006.

Deo, Ashwini (2009). Unifying the imperfective and the progressive: Partitions as quantificational domains. *Linguistics and Philosophy* 32(5), 475–521. DOI: 10.1007/s10988-010-9068-z.

Deo, Ashwini and Maria Mercedes Piñango (2011). Quantification and context in measure adverbs. In: *Proceedings of the 21st Semantics and Linguistic Theory Conference (SALT 21)*. Ed. by Neil Ashton, Anca Chereches, and David Lutz. Ithaca, NY: Cornell University CLC Publications, pp. 295–312. URL: http://elanguage.net/journals/salt/article/view/21.295/2516.

van der Does, Jaap (1993). Sums and quantifiers. *Linguistics and Philosophy* 16(5), 509–550. DOI: 10.1007/978-94-017-2706-8_6.

van der Does, Jaap and Henk J. Verkuyl (1996). Quantification and predication. In: *Semantic ambiguity and underspecification*. Ed. by Kees van Deemter and Stanley Peters. Stanford, CA: CSLI Publications, pp. 27–54. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.37.8829&rep=rep1&type=pdf.

Dotlačil, Jakub (2010). Anaphora and distributivity: A study of *same*, *different*, reciprocals and *others*. PhD thesis. Utrecht, Netherlands: Utrecht University. URL: http://www.jakubdotlacil.com/proefschrift.pdf.

Dowty, David R. (1979). *Word meaning and Montague Grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht, Netherlands: Reidel. DOI: 10.1007/978-94-009-9473-7.

Dowty, David R. (1987). Collective predicates, distributive predicates, and *All*. In: *Proceedings of the 3rd Eastern States Conference on Linguistics (ESCOL)*. Ed. by Ann Miller and Zheng-sheng Zhang. Columbus, OH: Ohio State University, pp. 97–115.

van Geenhoven, Veerle (2004). For-adverbials, frequentative aspect, and pluractionality. *Natural Language Semantics* 12(2), 135–190. DOI: 10.1023/b:nals.0000031389.36427.af.

Gillon, Brendan S. (1987). The readings of plural noun phrases in English. *Linguistics and Philosophy* 10(2), 199–219. DOI: 10.1007/bf00584318.

Gillon, Brendan S. (1990). Plural noun phrases and their readings: A reply to Lasersohn. *Linguistics and Philosophy* 13(4), 477–485. DOI: 10.1007/bf00630751.

Groenendijk, Jeroen, Theo Janssen, and Martin Stokhof, eds. (1984). *Truth, interpretation and information: Selected papers from the Third Amsterdam Colloquium*. Dordrecht, Netherlands: Foris. DOI: 10.1515/9783110867602.

Harley, Heidi (2012). Semantics in distributed morphology. In: *Semantics: An International Handbook of Natural Language Meaning*. Ed. by Claudia Maienborn, Klaus von Heusinger, and Paul Portner. Vol. 3. de Gruyter, pp. 2151–2172. DOI: 10.1515/9783110253382.2151.

Heim, Irene (1994). Plurals. Ms. Massachusetts Institute of Technology.

Heim, Irene and Angelika Kratzer (1998). *Semantics in Generative Grammar*. Oxford, UK: Blackwell Publishing.

Hoeksema, Jack (1983). Plurality and conjunction. In: *Studies in modeltheoretic semantics*. Ed. by Alice ter Meulen. Dordrecht, Netherlands: Foris, pp. 63–83.

Kennedy, Christopher (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1), 1–45. DOI: 10.1007/s10988-006-9008-0.

Kratzer, Angelika (1989). An investigation of the lumps of thought. *Linguistics and Philosophy* 12(5), 607–653. DOI: 10.1007/bf00627775.

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, pp. 109–137. DOI: 10.1007/978-94-015-8617-7_5.

Kratzer, Angelika (2003). The event argument and the semantics of verbs, chapter 3. Ms. University of Massachusetts, Amherst. URL: http://semanticsarchive.net/Archive/GU1NWM4Z/.

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. DOI: 10.1515/9783110925449.269.

Krifka, Manfred (1986). *Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. Munich, Germany: Wilhelm Fink. URL: http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka_1989_Nominalreferenz_Zeitkonstitution.PDF.

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 Peter van Emde Boas. Dordrecht, Netherlands: Foris, pp. 75–115. URL: http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka%201989%20Nominal%20Reference.pdf.

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. URL: http://amor.cms.hu-berlin.de/~h2816i3x/Publications/Krifka%201992%20Thematic%20Relations.PDF.

Krifka, Manfred (1998). The origins of telicity. In: *Events and grammar*. Ed. by Susan Rothstein. Dordrecht, Netherlands: Kluwer, pp. 197–235. DOI: 10.1007/978-94-011-3969-4_9.

Krifka, Manfred (1999). At least some determiners aren't determiners. In: *The semantics/pragmatics interface from different points of view*. Ed. by Ken Turner. Amsterdam, Netherlands: Elsevier, pp. 257–291. URL: http://philpapers.org/rec/KRIALS.

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.

Landman, Fred (1989). Groups, I. *Linguistics and Philosophy* 12(5), 559–605. DOI: 10.1007/BF00627774.

Landman, Fred (1996). Plurality. In: *Handbook of contemporary semantic theory*. Ed. by Shalom Lappin. Oxford, UK: Blackwell Publishing, pp. 425–457. DOI: 10.1111/b.9780631207498.1997.00019.x.

Landman, Fred (2000). *Events and plurality: The Jerusalem lectures*. Dordrecht, Netherlands: Kluwer. DOI: 10.1007/978-94-011-4359-2.

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, pp. 229–251. DOI: 10.1093/acprof:oso/9780199544325.003.0011.

Landman, Fred and Susan Rothstein (2012a). The Felicity of Aspectual For-Phrases–Part 1: Homogeneity. *Language and Linguistics Compass* 6(2), 85–96. DOI: 10.1002/lnc3.324.

Landman, Fred and Susan Rothstein (2012b). The Felicity of Aspectual For-Phrases–Part 2: Incremental Homogeneity. *Language and Linguistics Compass* 6(2), 97–112. DOI: 10.1002/lnc3.323.

Lasersohn, Peter (1988). A semantics for groups and events. PhD thesis. Columbus, OH: Ohio State University. URL: http://semanticsarchive.net/Archive/zQ2ODQxN/.

Lasersohn, Peter (1989). On the readings of plural noun phrases. *Linguistic Inquiry* 20(1), 130–134. URL: http://www.jstor.org/stable/4178619.

Lasersohn, Peter (1995). *Plurality, conjunction and events*. Dordrecht, Netherlands: Kluwer. DOI: 10.1007/978-94-015-8581-1.

Lasersohn, Peter (1998). Generalized distributivity operators. *Linguistics and Philosophy* 21(1), 83–93. DOI: 10.1023/A:1005317815339.

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. Reprinted in Link, 1998, Chapter 1. Berlin, Germany: de Gruyter, pp. 303–323. DOI: 10.1515/9783110852820.302.

Link, Godehard (1987). Generalized quantifiers and plurals. In: *Generalized quantifiers*. Ed. by Peter Gärdenfors. Reprinted in Link, 1998, Chapter 4. Berlin, Germany: Springer, pp. 151–180. DOI: 10.1007/978-94-009-3381-1_6.

Link, Godehard (1991). Plural. In: *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung / Semantics: An international handbook of contemporary research*. Ed. by Arnim von Stechow and Dieter Wunderlich. Reprinted in Link, 1998, Chapter 2, translated by the author. Berlin, Germany: de Gruyter, pp. 418–440. DOI: 10.1515/9783110203363.

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. DOI: 10.1007/978-94-017-2706-8_2.

Link, Godehard (1998). *Algebraic semantics in language and philosophy*. Stanford, CA: CSLI Publications.

Malamud, Sophia (2006a). (Non-)Maximality and distributivity: A decision theory approach. In: *Proceedings of the 16th Semantics and Linguistic Theory Conference (SALT 16)*. Ed. by Masayuki Gibson and Jonathan Howell. Ithaca, NY: Cornell University CLC Publications, pp. 120–137. URL: http://research.nii.ac.jp/salt16/proceedings/malamud_salt16.pdf.

Malamud, Sophia (2006b). Semantics and pragmatics of arbitrariness. PhD thesis. Philadelphia, PA: University of Pennsylvania. URL: http://people.brandeis.edu/~smalamud/Malamud-diss-single.pdf.

Moltmann, Friederike (1991). Measure adverbials. *Linguistics and Philosophy* 14(6), 629–660. DOI: 10.1007/bf00631962.

Parsons, Terence (1990). *Events in the semantics of English*. Cambridge, MA: MIT Press.

Parsons, Terence (1995). Thematic relations and arguments. *Linguistic Inquiry* 26(4), 635–662. URL: http://www.jstor.org/stable/4178917.

Pietroski, Paul M. (2005). *Events and semantic architecture*. Oxford, UK: Oxford University Press. DOI: 10.1093/acprof:oso/9780199244300.001.0001.

Pietroski, Paul M. (2006). Interpreting concatenation and concatenates. *Philosophical Issues* 16(1), 221–245. DOI: 10.1111/j.1533-6077.2006.00111.x.

Piñón, Christopher (1999). A semantics for durative adverbials. In: *Proceedings of Sinn und Bedeutung 4*. Düsseldorf, Germany: University of Düsseldorf.

Potts, Christopher (2008). Interpretive economy, Schelling points, and evolutionary stability. Ms. University of Massachusetts, Amherst. URL: http://semanticsarchive.net/Archive/jExYWZlN/.

Rimell, Laura (2004). Habitual sentences and generic quantification. In: *Proceedings of the 23rd West Coast Conference on Formal Linguistics (WCCFL 23)*. Ed. by Vineeta Chand et al. Somerville, MA: Cascadilla Proceedings Project, pp. 663–676. URL: http://www.nyu.edu/projects/szabolcsi/Rimell_WCCFL23_Habitual.pdf.

Roberts, Craige (1987). Modal subordination, anaphora, and distributivity. PhD thesis. Amherst, MA: University of Massachusetts. URL: http://scholarworks.umass.edu/dissertations/AAI8710499/.

van Rooij, Robert (2003). Questioning to resolve decision problems. *Linguistics and Philosophy* 26(6), 727–763. DOI: 10.1023/B:LING.0000004548.98658.8f.

Rothstein, Susan (2010). Counting and the mass/count distinction. *Journal of Semantics* 27(3), 343–397. DOI: 10.1093/jos/ffq007.

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: de Gruyter, pp. 413–434. DOI: 10.1515/97831101975549.413.

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, pp. 483–512. DOI: 10.1515/9783110867602.131.

Schein, Barry (1986). Event logic and the interpretation of plurals. PhD thesis. Cambridge, MA: Massachusetts Institute of Technology. URL: http://hdl.handle.net/1721.1/15107.

Schein, Barry (1993). *Plurals and events*. Cambridge, MA: MIT Press.

Schwarzschild, Roger (1996). *Pluralities*. Dordrecht, Netherlands: Kluwer. DOI: 10.1007/978-94-017-2704-4.

Spector, Benjamin (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In: *Presuppositions and implicature in compositional semantics*. Ed. by Uli Sauerland and Penka Stateva. London, UK: Palgrave, pp. 243–281. DOI: 10.1057/9780230210752.

Stanley, Jason and Zoltán Gendler Szabó (2000). On Quantifier Domain Restriction. *Mind and Language* 15(2&3), 219–261. DOI: 10.1111/1468-0017.00130.

von Stechow, Arnim (2009). Tenses in compositional semantics. In: *The expression of time*. Ed. by Wolfgang Klein and Ping Li. Berlin, Germany: de Gruyter, pp. 129–166. DOI: 10.1515/97831101990031.129.

Sternefeld, Wolfgang (1998). Reciprocity and cumulative predication. *Natural Language Semantics* 6(3), 303–337. DOI: 10.1023/A:1008352502939.

Tarski, Alfred (1935). Zur Grundlegung der Booleschen Algebra. *Fundamenta Mathematicae* 24. Translated to English as Tarski, 1956, 177–198. URL: https://eudml.org/doc/212745.

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.

Taylor, Barry (1985). *Modes of occurrence*. Oxford, UK: Basil Blackwell.

Todorova, Marina et al. (2000). Aspectual coercion and the online computation of sentential aspect. In: *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*. Ed. by Lila R. Gleitman and Aravind K. Joshi, pp. 3–8. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.230.9686&rep=rep1&type=pdf#page=545.

Vendler, Zeno (1957). Verbs and times. *The Philosophical Review* 66(2), 143–160. DOI: 10.2307/2182371.

Verkuyl, Henk J. (1972). *On the compositional nature of the aspects*. Dordrecht, Netherlands: Reidel. DOI: 10.1007/978-94-017-2478-4.

Verkuyl, Henk J. and Jaap van der Does (1991). *The semantics of plural noun phrases*. Amsterdam, Netherlands: University of Amsterdam. URL: https://www.illc.uva.nl/Research/Publications/Reports/LP-1991-07.text.pdf.

Vlach, Frank (1993). Temporal adverbials, tenses and the Perfect. *Linguistics and Philosophy* 16(3), 231–283. DOI: 10.1007/bf00985970.

Williams, Alexander (2009). Themes, cumulativity, and resultatives: Comments on Kratzer 2003. *Linguistic Inquiry* 40(4), 686–700. DOI: 10.1162/ling.2009.40.4.686.
Winter, Yoad (1997). Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20(4), 399–467. DOI: 10.1023/A:1005354323136.
Winter, Yoad (2001). *Flexibility principles in Boolean semantics: The interpretation of coordination, plurality, and scope in natural language*. Cambridge, MA: MIT Press.
Zimmermann, Malte (2002). Boys buying two sausages each: On the syntax and semantics of distance-distributivity. PhD thesis. Amsterdam, Netherlands: University of Amsterdam. URL: http://www.lotpublications.nl/Documents/62_fulltext.pdf.
Zucchi, Sandro and Michael White (2001). Twigs, sequences and the temporal constitution of predicates. *Linguistics and Philosophy* 24(2), 187–222. DOI: 10.1023/a:1005690022190.
Zweig, Eytan (2008). Dependent plurals and plural meaning. PhD thesis. New York, NY: New York University. URL: http://semanticsarchive.net/Archive/GJhMzhhZ/.
Zweig, Eytan (2009). Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy* 32(4), 353–407. DOI: 10.1007/s10988-009-9064-3.