Overt Distributivity in Algebraic Event Semantics*

Lucas Champollion

June 9, 2014

Abstract

Adnominal and adverbial distributive elements are known to differ across languages as to whether they can only distribute over individuals mentioned in the same sentence, or also over pragmatically salient occasions that have not been mentioned. Those that can also be used as distributive determiners, such as English *each*, never allow distribution over occasions (Zimmermann, 2002b). I explain this pattern, and the behavior of distance-distributivity more generally, in a surface-compositional theory that formally relates the differences in overt distributive elements to those between distributivity operators posited by Link (1983) and Schwarzschild (1996). I achieve surface compositionality by formulating the theory in algebraic event semantics (Krifka, 1998).

Keywords: distance-distributivity, crosslinguistic semantics, algebraic semantics, adnominal *each*, adverbial *each*, quantifier float, covers

1 Introduction

This paper presents a surface-compositional theory of distance distributivity that relates adnominal and adverbial distributive elements, atomic and cover-based distributivity, and distributive determiners to each other. This is one of two self-contained papers that can be read individually but that form a coherent whole. This paper focuses on overt distributivity. Its counterpart, Champollion (2014a), focuses on covert distributivity.

*For helpful discussions and comments, I am grateful to Chris Barker, Benjamin Bruening, Dylan Bumford, Seth Cable, Robert Henderson, Chris LaTerza, Anna Szabolcsi, Linmin Zhang, Malte Zimmermann and to the audiences of the Stuttgart 2011 workshop on distributivity, the Amsterdam Colloquium 2011, at the University of Potsdam, and to members of my 2013 NYU seminar on algebraic semantics, where earlier versions of this paper and its counterpart, Champollion (2014a), were presented. I thank Linmin Zhang for helping me collect English judgments via Amazon Mechanical Turk. For native speaker judgments on other languages, I thank my consultants Isaac Bleaman, Heather Burnett, Ivano Ciardelli, Tuğba Çolak-Champollion, Liz Coppock, Hana Filip, Floris Roelofsen, Gianina Iordăchioaia, Zack Jaggers, Songhee Kim, Sonia Kasyanenko, Jeremy Kuhn, Chigusa Kurumada, Roumyana Pancheva, Kjell Johan Sæbø, and Yohei Oseki. This paper and Champollion (2014a) build in part on some earlier work of mine, for which I am indebted to many people and institutions as indicated there (Champollion, 2012, 2013).

While this paper builds on results obtained in Champollion (2014a), the two papers can be read in either order.

Overt and covert distributivity are illustrated in the following examples:

(1) a. The girls each wore a black dress.
b. The girls wore a black dress.

In sentence (1a), the adverbial distributive element *each* distributes the predicate *wear a black dress* over the individual girls and leads to the entailment that each of the girls in question wears a black dress. Sentence (1b) is interpreted in the same way, even though there is no *each*. The ability of verb phrases to distribute in the absence of an overt distributive element has been attributed to what is traditionally called the D operator, a silent counterpart of adverbial *each* (Link, 1987; Roberts, 1987).

The purpose of this paper and of Champollion (2014a) is to bring together several strands of research on phenomena related to the semantics and pragmatics of distributivity in natural language. One of these strands deals with overt distributivity, which is is crosslinguistically often expressed via adverbials and adnominals, such as English *each* and German *jeweils*. Such elements differ with respect to whether they are restricted to distribution over individuals mentioned in the same sentence, or whether they can also distribute over pragmatically salient occasions that need not have been explicitly mentioned. It has been motivated by the properties of adverbial *each* and its adnominal and determiner counterparts both in English and other languages (Zimmermann, 2002b). As we will see, the meanings of these elements 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 *jeweils*, as illustrated in the following example. The focus of this paper is on this strand.

(2) Hans hat jeweils zwei Affen gesehen. *(German)*
Hans has DIST two monkeys seen.
'Hans has seen two monkeys on each occasion.'

The other strand concerns the properties of silent distributivity operators such as the one arguably present in (1b). Covert phrasal distributivity has been at the center of a long debate as to whether it always involves distribution over atoms – singular individuals – or whether it can also involve distribution over nonatomic entities (Lasersohn, 1989; Gillon, 1990). This is the focus of Champollion (2014a). I reserve the term *D operator* for distributivity operators that always distribute over atoms. As for the nonatomic version of the operator, whose meaning may be paraphrased as *each salient part of*, I will refer to it as the *Part operator*, following Schwarzschild (1996).

As this sketch already suggests, overt and covert distributivity share many similarities. In both cases, some elements can only distribute to atoms (*each, D*) while others can distribute to salient nonatomic entities (*jeweils, Part*). And as we will see, in both cases, the former elements can only distribute over pluralities that have been explicitly mentioned while the latter elements can also distribute over salient domains that have not been explicitly mentioned, such as temporal occasions. These similarities give rise to analogous questions in the overt and in the covert case. Can a given distributive element (be it a covert operator or an overt lexical item) 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 element to element? 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 theory should capture the semantic variation across distributivity-related elements should be captured . The resulting system should be fully formalized and explicit.

This paper, together with Champollion (2014a), contributes towards these goals. By combining ideas from algebraic semantics and event semantics, the two papers provide a framework in which the split in overt distance-distributive elements can be related to the debate in the literature on covert distributivity. In this framework, the various uses of *each* in English are all lexically related to the distributivity operator, either in its semantic, atomic form as defined by Link (1987) or in its pragmatic, salience-related form as defined by Schwarzschild (1996). As we will see, these various uses of *each* and these silent operators share some part of their meanings with each other and with their counterparts across languages. This fact is captured by deriving them from related distributivity operators which differ only in possible settings of two parameters and the ranges of values they allow for them. One parameter indicates the dimension in which distributivity takes place. This can be a thematic role in some semantic instances of distributivity, or a spatial or temporal dimension in other instances. The other parameter indicates the size of the entities over which distributivity takes place, such as atoms or salient amounts of space or 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 either nonatomic or in any case to not make its atoms available to the semantics of natural language. As a result, when the first parameter is set to time, the second 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 superficially unrelated facts observed in various places in the literature. It is situated within a broader framework that connects it to aspect and measurement under the name of strata theory (Champollion, 2010a).

In this paper, I provide a view on the seemingly noncompositional behavior of distance-distributive elements on which they look much more well-behaved than might be expected. The analysis is surface-compositional, reuses many independently motivated assumptions, and avoids unusual semantic mechanisms such as index-driven and crosswise lambda abstraction (Zimmermann, 2002b). I also avoid nonstandard and otherwise unmotivated concepts such as distributive polarity items (Oh, 2001, 2006). The analysis is placed in the context of algebraic event semantics. This allows us to formally model the relations between distribution over individuals and over events, as well as those between distribution over atoms and over nonatomic parts. The paper provides a formal framework for algebraic event semantics and makes an explicit proposal for how the compositional process can be modeled.

When we move to algebraic event semantics, the theory of distributivity operators developed by Link (1987) and extended by Schwarzschild (1996) turns out to require adjustments for a number of reasons, as discussed in Champollion (2014a). As shown by the following examples, the Neo-Davidsonian event semantic setting gives us the ability to think of the D and Part operators as being coindexable with different thematic roles. This allows us to capture through a simple change in coindexation the kinds of configurations that have otherwise been taken to require type-shifting-based reformulations of these operators (Lasersohn, 1998):

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

The reformulation of the distributivity operators in Champollion (2014a) provides the groundwork on which I build in this paper in order to formally relate this ambiguity to the one observable in examples like the following:

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

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

To capture this and other parallels between covert and overt distributivity, I will propose that distance-distributive elements across languages are in essence overt versions of the D and Part operators.

The theoretical picture that has been sketched so far, and that is developed below and in Champollion (2014a), provides us with a way to formulate commonalities and differences across instances of distributivity in natural language. Individual elements 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.

To develop this picture, Section 2 starts by describing relevant facts and generalizations about overt instances of distributivity across languages, drawing largely on the crosslinguistic discussion in Zimmermann (2002b). The two strands – overt and covert distributivity – are brought together in Section 3, which develops a surface-compositional account of overt distance distributivity. The typological facts about the range of meanings of overt distributive elements described in Section 2 are explained in Section 4. Section 5 compares the present analysis with Zimmermann (2002b). Section 6 shows how to extend the analysis to the case when English *each* modifies a prepositional phrase, which previously lacked a surface-compositional analysis and which bears special relevance to previous theories of plurals and events (Schein, 1993). Section 7 concludes.

2 Overt Distributivity Across Languages

This section presents the empirical basis for this paper. Distributive elements have different syntactic uses and different meanings across languages. In English, the distributive quantifier *each* can be used in at least three ways, which I will refer to as adnominal, adverbial, and determiner *each* respectively:

(6) a. **Adnominal each:** Two men have carried three suitcases *each*.
    b. **Adverbial each:** Two men have *each* carried three suitcases.
    c. **Determiner each:** *Each* man has carried three suitcases.

There are many terms for these three uses. Adnominal *each* has also been called shifted (Postal, 1974), an anti-quantifier (Choe, 1987), binominal (Safir and Stowell, 1988), or ditransitive (Roberts, 1987). Adverbial *each* has also been called floated (Choe, 1987). Determiner *each* is also called prenominal (Safir and Stowell, 1988). For the purpose of this paper, I set aside the use of *each* in other constructions, notably the reciprocal *each other* and the partitive *each of the men*. On the connection between reciprocals and *each*, see for example LaTerza (2014b). (I have not been able to access LaTerza (2014a) in the preparation of this paper, but I understand from its author that it is likely to be relevant to this discussion.) I will refer to the noun phrase *two men* in (6a) and (6b) as the antecedent of *each*, and to the noun phrase *three suitcases* in (6a) as the host of adnominal *each*.

I will refer to adnominal and adverbial *each* and to similar elements across languages as distance-distributive elements. That term is taken from Zimmermann (2002b). There is a slight difference in terminology: Zimmermann reserves the term *distance distributivity* for adnominal elements while I use it both for adnominal and for adverbial elements. This seems appropriate because adverbial *each* can be separated from its antecedent, for example by an auxiliary as shown in (6b).

Adnominal *each* can be shown by movement tests to form a constituent with its host noun phrase (Burzio, 1986; Safir and Stowell, 1988). Distance-distributive elements like it are sometimes thought of as a challenge for compositional semantics, because their interpretations are similar to those of distributive determiners even though their surface syntactic structure appears to be fundamentally different (Oh, 2001, 2006). For example, adnominal *each* in the object of sentence (6b) is contained in the constituent over which it seems at first sight to distribute, namely the verb phrase *carried three suitcases each*. This is of course similar to the challenge represented by quantifiers in object position (*carried every suitcase*), and the standard solutions to that challenge are available in both cases. For example, one can lift the type of the quantifier or the verb in order to give the quantifier scope over the verb phrase (Hendriks, 1993; Barker, 2002). I will follow the same general strategy in the formal analysis below. In fact, as we will see, the scope of adnominal *each* is even more restricted than that of object quantifiers: it takes scope only over its host noun phrase. So there is no special challenge for compositionality.

As we have seen in the pair of examples in (4) and (5) above, adnominal *each* can target different antecedents. This dependency is generally regarded as a case of ambiguity rather than underspecification, and I will follow this view. There are syntactic constraints on the distribution of adnominal *each* with respect to its antecedent, and the dependency has been variously argued to be similar to that of reflexive pronouns with respect to their antecedents (Burzio, 1986; Safir and Stowell, 1988) or to that of traces of noun phrases that undergo raising with respect to these noun phrases (Sportiche, 1988). Similarly, adverbial *each* has been variously claimed to be related to its antecedent by movement, in the sense that it modifies the trace of its antecedent, or to be base-generated, in which case its relation to its antecedent can be taken to be anaphoric. For an overview of these conflicting claims and their implications, see Bobaljik (2001). Since the nature of the dependency between adnominal and adverbial *each* and their antecedents is not the focus of this paper, I will not take a strong position on it. In the formal theory to be developed below, I will represent it via coindexing of thematic relations, which is more in line with anaphora-based accounts than with movement-based ones. In this, I follow previous semantic analyses that interpret adnominal *each* without any movement, such as Zimmermann (2002b). Should the movement-based view turn out to be the correct one, the syntax-semantics interface may need to be modified accordingly, for example by incorporating elements from the theory of Cable (2014).

Turning now to other languages, adnominal and adverbial *each* are translated in German by one word, *jeweils* (Moltmann, 1997; Zimmermann, 2002b). Determiner *each*, however, is translated by another word, *jed-*. I gloss distance-distributive elements as Dist since, as we will see, they have a wider range of readings than *each*. Example (7a) is adnominal, example (7b) is adverbial, and example (7c) contains a determiner. Though adverbial and adnominal *jeweils* take the same surface position in (7a) and (7b), they can be teased apart syntactically, as discussed in Zimmermann (2002b).

(7) a. Die Jungen haben [*jeweils* [drei Koffer]] getragen.
The boys have Dist three suitcases carried.
b. Die Jungen haben [*jeweils* [drei Koffer getragen]].
The boys have Dist three suitcases carried.
c. *Jeder/*Jeweils* Junge hat drei Koffer getragen.
Each.sg.n/Dist boy has three suitcases carried.

As we will see, *each* and *jeweils* generalize to two classes of distance-distributive elements across languages. *Each*-type distance-distributive elements can also be used as determiners. *Jeweils*-type distance-distributive elements cannot double as determiners. Some languages have distance-distributive elements which can also function as distributive determiners, as in English, and others are like German in that they have no such elements (Zimmermann, 2002b). Across these languages, Zimmermann observes that distance-distributive elements which can also be used as determiners (e.g. *each*) always distribute over individuals, as determiners do. In contrast, many of those distance-distributive elements which are formally distinct from determiners (e.g. *jeweils*) can also distribute over salient occasions, that is, over chunks of time or space.

Let me illustrate this observation by using German *jeweils*, a distance-distributive element which cannot double as a distributive determiner. *Jeweils* can distribute over individuals like English *each*, but also over spatial or temporal occasions, as long as context provides a salient set of such occasions. I call this the occasion reading. It corresponds to what is also called the *spatial key reading* and the *temporal key reading* (Balusu, 2005; Balusu and Jayaseelan, 2013). I leave open the question of whether or not the spatial and temporal cases should be distinguished as two separate readings. The framework I will present can accommodate both possibilities. Another, less theory-neutral term for the occasion reading is *event-distributive reading* (Oh, 2001, 2006). Zimmermann (2002b) uses the term *adverbial reading* for it. This term is potentially misleading, because it suggests that only the adverbial use of *jeweils* can give rise to this reading. But adnominal *jeweils* can give rise to it as well (Zimmermann, 2002b, ch. 5). For example, in (8), *jeweils* is part of the subject noun phrase (we know this because German as a V2 language allows only one constituent before the tensed verb *standen*) and is therefore adnominal. However, as shown by the paraphrase, this instance of *jeweils* distributes over occasions, not over individuals.

(8) Jeweils zwei Jungen standen Wache.
DIST two boys stood watch.
'Each time, two boys kept watch.' (German)

The following examples illustrate the occasion reading. Sentence (9) is ambiguous between a reading that distributes over individuals – the ones of which their plural subject consists, (9a) – and one that distributes over occasions (9b).

(9) Die Jungen haben jeweils zwei Affen gesehen.
The boys have DIST two monkeys seen.
a. 'Each of the boys has seen two monkeys.'
b. 'The boys have seen two monkeys each time.' (German)

While the former reading is always available, the latter requires a supporting context. That is, when (9) is uttered out of the blue, it only has the reading (9a). The reading (9b), by contrast, is only available in contexts where there is a previously mentioned or otherwise salient set of occasions, such as contexts in which the boys have been to the zoo on several previous occasions.

Unlike *each*, *jeweils* can also occur with a singular subject, as in (10), repeated here from (2), which only has an occasion reading.

(10) Hans hat jeweils zwei Affen gesehen.
Hans has DIST two monkeys seen.
'Hans has seen two monkeys on each occasion.' (German)

This sentence is odd out of the blue, and it requires supporting context in the same way as reading (9b) does. Its other potential reading would involve vacuous distribution over only one individual, Hans. This is presumably blocked through the Gricean maxim of manner "Be brief" or a non-vacuity presupposition or implicature or whatever else prevents vacuous distributivity (Roberts, 1987, p. 219). For more discussion of this point, and for a fuller discussion of the kinds of noun phrases that can license adnominal *each* in English, see Champollion (2014b).

While *jeweils* allows distribution both over individuals and over salient occasions, this is not the case for all distance-distributive elements (Zimmermann, 2002b). Across languages, many adnominal distance-distributive elements can only distribute over individuals. For example, English adnominal *each* lacks the occasion reading:

(11) The boys saw two monkeys each.
  a. *Available:* 'Each of the boys saw two monkeys.'
  b. *Unavailable:* 'The boys saw two monkeys on each occasion.'

When adnominal *each* is used in a sentence whose subject is singular, distribution over individuals is not possible, again presumably for pragmatic reasons:

(12) *John saw two monkeys each.

Unlike (10), this sentence lacks an occasion reading, even with supporting context. To make the occasion reading surface, one needs to add an overt noun like *time*:

(13) John saw two monkeys each time.

We have seen that English *each* also differs from German *jeweils* in that only the former can also be used as a determiner. Coming back to what I mentioned at the beginning of this section, Zimmermann reports the following generalization (Zimmermann, 2002b):

(14) **Zimmermann's generalization:** All *each*-type distance-distributive elements (i.e. those that can also be used as determiners) can only distribute over individuals. This contrasts with *jeweils*-type distance-distributive elements, many of which can also distribute over salient spatial or temporal occasions.

This generalization is based in part on the following examples, which show that distance-distributive elements in Dutch, French, Italian, Japanese, Norwegian, and Russian all behave like English *each* in two ways: They can also be used as distributive determiners, and they lack the occasion reading except when an extra noun *time* is added. Many of the following examples are taken from Zimmermann (2002b).

(15) De jongens hebben *elk* twee boeken gelezen.
the boys have DIST two books read
'The boys have read two books each.' (Dutch)[1]

(16) *Elk* jonge heeft twee boeken gelezen.
DIST boy has two books read
'Each boy has read two books' (Dutch)[2]

(17) Hans heeft *elke* *(keer) twee boeken gelezen..
Hans has DIST time two books read.
'Hans has read two books each time.' (Dutch)[3]

(18) Les professeurs ont lu deux livres *chacun/chaque*.
the professors have read two books DIST
'The professors have read two books each.' (French)[4]

[1] Zimmermann (2002b, p. 40)
[2] Zimmermann (2002b, p. 44)
[3] Floris Roelofsen, p.c. to the author
[4] Tellier and Valois (1993, p. 574, ex. 1a) quoted in Zimmermann (2002b, p. 41). *Chaque* is colloquial as

(19) *Chaque* professeur a lu deux livres.
DIST professor has read two books
'Each professor has read two books.' (French)[5]

(20) Pierre a lu deux livres *chaque* *(fois) / **chac-un(e)*(fois).
Pierre has read two books DIST time
'Pierre read two books each time.' (French)[6]

(21) I ragazzi comprarono un libro *ciascuno*.
the boys bought a book DIST
'The boys bought one book each.' (Italian)[7]

(22) *Ciascun* ragazzo ha comprato due salsicce.
DIST boy has bought two sausages
'Each boy has bought two sausages.' (Italian)[8]

(23) *Peter ha comprato due salsicce *ciascun/-o/-e*.
Peter has bought two sausages DIST
Intended: 'Peter has bought two sausages.' (Italian)[9]

(24) Otoko=tati-ga *sorezore* huta=ri-no zyosei-o aisi teiru koto.
men=PL-NOM DIST two=CL-GEN women-ACC love-ASP FACT
'The men love two women each.' (Japanese)[10]

(25) *Sorezore*-no gakusei-ga iti=dai-no piano-o motiage-ta.
DIST-GEN student-NOM one=CL-GEN piano-ACC lift-PAST
'Each student lifted one piano.' (Japanese)[11]

(26) Taroo-wa sorezore-?(de) iti-dai-no piano-o motiage-ta.
Taroo-TOP DIST(-LOC) one-CL-GEN piano-ACC lift-PAST
'Taroo lifted one piano on each occasion.' (Japanese)[12]

(27) Guttene har kjøpt to pølser *hver*.
boys-the have bought two sausages DIST
'The boys bought two sausages each.' (Norwegian)[13]

(28) *Hver* gutt har kjøpt to pølser.
DIST boy has bought two sausages

---

an adnominal. While French adnominal *chacun* and determiner *chaque* are not exactly identical, they are historically related and can still be considered formally identical (Grevisse, 1980; Junker, 1995; Zimmermann, 2002b, p. 44, fn. 30).

[5] Zimmermann (2002b, p. 44)

[6] Author's judgment, adapted from Zimmermann (2002b, p. 47)

[7] Burzio (1986, p. 198, ex. 50b) quoted in Zimmermann (2002b, p. 41)

[8] Zimmermann (2002b, p. 44)

[9] Ivano Ciardelli, p.c. to the author

[10] Sakaguchi (1998, p. 115, ex. 1) quoted in Zimmermann (2002b, p. 41)

[11] Sakaguchi (1998, p. 4, ex. 7) The syntactic status of *sorezore* in this example, and therefore the import of Japanese on Zimmermann's generalization, is debatable since Japanese is usually assumed to lack overt determiners.

[12] Chigusa Kurumada, p.c. to the author. Kurumada comments that the sentence without *de* feels like an elliptical version of the sentence with *de*.

[13] Øystein Vangsnes, p.c. to Zimmermann (2002b, p. 40)

'Each boy has bought two sausages.' (Norwegian)[14]

(29) Jon har kjøpt to pølser *hver* *(gang).
Jon have bought two sausages Dist
Intended: 'Jon has bought two sausages each time.' (Norwegian)[15]

(30) Mal'chiki kupili (po) dve sosiski *kazhdyj*.
boys.nom bought two sausage.gen.sg Dist
'The boys bought two sausages each.' (Russian)[16]

(31) *Kazhdyj* mal'chik kupil dve sosiski.
Dist boy bought two sausages
'Each boy bought two sausages' (Russian)[17]

(32) Petja pokupal dve sosiski *kazhdyj* *(raz).
Petja buy.perf two sausage Dist (time)
'Peter bought two sausages each time.' (Russian)[18]

Zimmermann's generalization states that every distance-distributive element that can be used as a determiner lacks the occasion reading. The opposite is not the case, however. For example, the Romanian distance-distributive element *cîte* lacks the occasion reading, but it cannot be used as a determiner. This is illustrated in the following:

(33) Doi oameni au cărat *cîte* trei valize.
two men have carried Dist three suitcases
'Two men have carried three suitcases each.' (Romanian)[19]

(34) **Cît(e)* student a plecat.
Dist student has left.
Intended: 'Each student has left.' (Romanian)[20]

(35) *Petru a cîștigat *cîtă/cîte* (dată).
Peter has won Dist.f/Dist.adv time.
Intended: 'Peter won each time.' (Romanian)[21]

Zimmermann considers Japanese as an example of the same phenomenon that I have illustrated with Romanian, but whether this is correct is not very clear. Zimmermann bases his view on the fact that the Japanese distance-distributive item *sorezore* differs formally from what he calls the Japanese distributive determiner-quantifier *wh. . . + mo*, which is illustrated in (36).

(36) *Dono* gakusei-mo sooseezi-o hutatu katta.
which student-mo sausage-acc two-cl bought

[14] Zimmermann (2002b, p. 44)
[15] Kjell Johan Sæbø, p.c. to the author
[16] Olga Borik, p.c. to Zimmermann (2002b, p. 41)
[17] Olga Borik, p.c. to Zimmermann (2002b, p. 44)
[18] Sonia Kasyanenko, p.c. to the author
[19] Gil (1982, p. 19, ex. 1f), Zimmermann (2002b, p. 41)
[20] Brasoveanu and Farkas (2011, p. 10), Gianina Iordăchioaia, p.c.
[21] Gianina Iordăchioaia, p.c. There is a related expression *cîte-o-dată* that means *from time to time*.

'Every student bought two sausages.' (Japanese)[22]

However, *sorezore* can also be used in the position of a determiner, as example (25) above shows. The syntactic status of *sorezore* in this example, and therefore the import of Japanese on Zimmermann's generalization, is debatable since Japanese is usually assumed to lack overt determiners.

While Romanian and possibly Japanese are counterexamples to the inverse of Zimmermann's generalization, that inverse direction still describes a tendency. That is, in many languages where adnominal distance-distributive elements have occasion readings, they cannot be used as determiners. In addition to German *jeweils*, adnominal distance-distributive elements in Bulgarian, Czech, and Korean have occasion readings cannot be used as determiners, as is shown below. Most of these observations are due to Zimmermann (2002b). The Korean case is also discussed in depth by Choe (1987) and by Oh (2001, 2006).

(37) John i Mary kupiha *po* edna tetradka.
John and Mary bought Dist one notebook
'John and Mary bought one notebook each.' (Bulgarian)[23]

(38) *Vsjako/*Po* momche kupi dve jabulki.
Dist boy bought two sausages
'Each boy bought two sausages.' (Bulgarian)[24]

(39) Mary byaga *po* 5 mili predi zakuska.
Mary runs Dist 5 miles before breakfast
'Mary runs 5 miles before breakfast (every morning).' (Bulgarian)[25]

(40) Chlapci koupili *po* dvou párcích/párkách.
boys bought Dist two sausages.loc
'The boys bought two sausages each.' (Czech)[26]

(41) *Každý/*Po* chlapec koupil dva párky.
Dist boy bought two sausages
'Each boy bought two sausages.' (Czech)[27]

(42) *Po* třech ženách vstupovalo do místnosti.
Dist three.loc women.loc entered.3sg into room
'Three women entered the room [i.e., one triplet after another]' (Czech)[28]

(43) ai-tul-i phwungsen-hana-*ssik*-ul sa-ess-ta.
child-pl-nom balloon-one-Dist-acc bought
'The children bought one balloon each.' (Korean)[29]

[22] Satoshi Tomioka, p.c. to Zimmermann (2002b, p. 45)
[23] Petrova (2000) quoted in Zimmermann (2002b, p. 41)
[24] Milena Petrova, p.c. to Zimmermann (2002b, p. 45); Roumyana Pancheva, p.c. to the author
[25] Petrova (2000, ex. 3b) quoted in Zimmermann (2002b, p. 41)
[26] Hana Filip, p.c. to Zimmermann (2002b, p. 41)
[27] Hana Filip, p.c. to Zimmermann (2002b, p. 45) and to the author
[28] Hana Filip, p.c. to Zimmermann (2002b, p. 47) and to the author
[29] Choe (1987, p. 49, ex. 13) quoted in Zimmermann (2002b, p. 41)

(44) Sonyen-mata chayk-ul twu kwen-*ssik* sa-(a)t-ta.
boy-Dist book-acc two cl-dist buy-past-dec
'Every boy bought two books.' (Korean)[30]

(45) na-nun [phwungsen- hana-*ssik*-ul] sa-ess-ta.
I-top balloon one-Dist-acc bought
'I bought a balloon (each time / each day / at each store).' (Korean)[31]

Many languages express adnominal distance distributivity by a bound morpheme (either an affix or a reduplicative morpheme) that attaches to a numeral (Gil, 1982). In this category, on the one hand, we find cases where this process does not give rise to occasion readings, such as Hungarian (Farkas, 1997; Szabolcsi, 2010) and Turkish (Tuğba Çolak-Champollion, p.c.). On the other hand, we find cases where it does, such as Telugu (Balusu, 2005; Balusu and Jayaseelan, 2013) and Tlingit (Cable, 2014). All these cases are illustrated below. The import of these facts on Zimmermann's generalization is unclear, since bound morphemes are not expected to be able to act as determiners. I mention them here for completeness and because the compositional analysis given below extends to them.

(46) pilla-lu renDu renDu kootu-lu-ni cuus-ee-ru.
kid-pl two two monkey-pl-acc see-past-3pl
   a. 'The kids each saw two monkeys.'
   b. 'The kids saw two monkeys each time.'
   c. 'The kids saw monkeys in groups of two.' (Telugu)[32]

(47) A gyerekek *két-két* majmot láttak.
The children two-two monkey.acc saw.3pl
   a. *Available:* 'Each of the children saw two monkeys.'
   b. *Unavailable:* 'The children saw two monkeys each time.' (Hungarian)[33]

(48) Çocuklar iki-ş-*er* sosis aldı.
The children two.Dist sausage bought.
'The children bought two sausages each.' (Turkish)[34]

(49) *Can iki-ş-*er* sosis aldı.
Can two.Dist sausage bought.
Intended: 'Can bought two sausages each time.' (Turkish)[35]

(50) *Nás'gigáa* xáat has aawasháat.
three.Dist fish 3pls.3o.caught
   a. 'They caught three fish each.'

[30] Kim, p.c. to Zimmermann (2002b, p. 45)
[31] Choe (1987, p. 49, ex. 13) quoted in Zimmermann (2002b, p. 47)
[32] Balusu and Jayaseelan (2013, p. 67, ex. 15a)
[33] Szabolcsi (2010, p. 138, ex. 99)
[34] Tuğba Çolak-Champollion, p.c. to the author.
[35] Tuğba Çolak-Champollion, p.c. to the author. This sentence is unacceptable on the intended reading. According to my consultant, it may be thatit is acceptable under the reading *Can bought two sausages of each kind.*

b. 'They caught three fish each time.' (Tlingit)[36]

The facts discussed so far suggest the following requirements for a semantic analysis of distance-distributivity. First, the synonymy of the determiner, adnominal and adverbial uses of *each* in English should be captured, ideally by essentially identical lexical entries. Second, the fact that distance-distributive elements across languages share some part of their meanings (namely their individual-distributive readings) should be represented, as well as the fact that some of them can also have occasion readings. Third, the analysis should clarify the connections between distance-distributive elements and distributivity theory more generally, and the semantic variation across distance-distributive elements should be captured. Finally, an explanation for Zimmermann's generalization should be readily available. The rest of the paper develops an analysis that fulfills these requirements.

3 Relating Overt and Covert Distributivity

The connection between the D operator from Link (1987) and adverbial *each* that was illustrated in (1) has been noted many times. I take adverbial and adnominal *each* and related distance-distributive elements in Dutch, French, Italian, Japanese, Norwegian, and Russian to be essentially D operators. These are the languages mentioned in Section 2 as being English-type. As for *jeweils* and its relatives in German-type languages like Bulgarian, Czech and Korean, we have seen that they can distribute over spatial and temporal intervals – arguably nonatomic entities. Link's D operator always distributes down to individual atoms and can therefore not be extended to these cases. So I will connect them to the nonatomic distributivity operator Part from Schwarzschild (1996).

I now present a formally explicit, surface-compositional account of the overt distributive elements described in Section 2 in terms of the covert distributive operators D and Part discussed in Champollion (2014a) and in the literature on covert distributivity as summarized there. For the sake of brevity, I will only execute the analysis for English *each* and German *jeweils*, but it should be clear how to extend it to other distributive elements depending on which one of these two they pattern with. I will only show how to model the individual-distributive and the temporal occasion readings. The extension from the temporal to the spatial occasion reading is straightforward.

The guiding idea of the analysis is that overt distributive elements include two versions of the distributivity operator. *Each* includes the atomic distributivity operator D, which can only distribute over count domains because only those domains have atoms. *Jeweils* includes the nonatomic distributivity operator Part. I argue in Champollion (2014a) that the latter operator can also distribute over noncount domains like time. I adopt the strata-theoretic perspective from Champollion (2010a). According to this theory, distributivity is a property with two parameters: dimension and granularity. I will suggest that *each*, just like the D operator, comes prespecified for "granularity=atom". This blocks the setting "dimension=time", so distributivity over occasions is unavailable. By contrast, *jeweils* does not come prespecified for anything

[36] Cable (2014, ex. 3b)

but is anaphoric on context, so it can distribute over salient covers, or salient stretches of time, just like the Part operator.

In claiming that *each* is an overt form of the D operator, I loosely follow proposals made for German *jeweils* and its short form *je* by Link (1986, 1987) and for English *each* by Roberts (1987). While Link and Roberts did not give explicit compositional implementations and did not or not fully consider the crosslinguistic picture, this paper can be seen, in a way, as an update to these early ideas which benefits from later work on algebraic semantics, nonatomic distributivity, and compositional implementations. In particular, Link and Roberts did not have access to the theory that has been built around the commonalities and differences between the D and Part operators (Champollion, 2014a, and references therein).

Arriving at the meaning of *each* from the meaning of distributivity operators is somewhat of the reverse of the process by which Schwarzschild arrived at his Part operator, which "was based on a generalization of Dowty and Brodie's (1984) account of floated quantifiers as verb phrase modifiers" (Schwarzschild, 1996, p. 137). Schwarzschild himself notes that the history of Part should not be taken for an endorsement that floated quantifiers are related to it, and argues that floated quantifiers should be distinguished from Part operators because he takes reciprocals to be licensed by distributivity operators, but not by adverbial *each*. He gives the following examples to support this claim. These kinds of examples, as well as the idea that reciprocals are licensed by distributivity operators of some kind or other, go back to Heim, Lasnik, and May (1991).

(51) a. They$_j$ Part$_i$ [saw each other$_{j,i}$].
  b. *They$_j$ each$_i$ saw each other$_{j,i}$.

Whatever the merits of the particular analysis of reciprocals underlying example (51a), Schwarzschild's argument can be doubted on empirical grounds. While some speakers of English reject (51b), many naturally-occurring examples analogous to this example can be found. The ones below are taken from the Corpus of Contemporary American English (Davies, 2008):

(52) You each had each other and yet the teasing was relentless and he could be very naughty with you.

(53) One of the leaders is Rich Borow, Stephanie's husband and nemesis on the bike; they each push each other hard.

(54) Communicate with other teachers, coaches, and paraprofessionals to see how you can each support each other.

I conclude that whatever fundamental differences there may be between overt and covert adverbial distributivity, the ability to license reciprocals is not one, at least not for those speakers that produced the examples above.

Before proceeding, let me specify a few background assumptions. The following discussion is partly taken from Champollion (2014a). 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 $\langle ve \rangle$. 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 lowecase 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 invidiuation 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, 2010a).

I use the star operator originally defined in Link (1983), which is defined as follows. Here, $\bigoplus C$ stands for the mereological sum of all entities to which C applies. Some authors write σ instead of $\bigoplus$.

(55) $x \in {}^*P \overset{\text{def}}{=} \exists C[x = \bigoplus C \ \wedge \ C \subseteq P]$
(x is the sum of all the elements of a subset C 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. For an overview of the axioms and linguistic applications of mereology, including the star operator, see Champollion and Krifka (2014).

I assume that verbs and their projections denote sets of events and are therefore of type vt (Parsons, 1990). Thematic roles, events, and verbs are each assumed to be closed under sum formation (Krifka, 1989; Champollion, 2010a). I discuss these assumptions in turn. The cumulativity assumption for thematic roles can be stated as follows:

(56) **Cumulativity assumption for thematic roles**
For any thematic role θ it holds that
$\forall e, e', x, y[\theta(e) = x \wedge \theta(e') = y \ \rightarrow \ \theta(e \oplus e') = x \oplus y]$

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, then there is an event $e \oplus e'$ whose agent is the sum of John and Mary. Thus, $e \oplus e'$ has a unique entity as its agent, even though this entity is a proper sum. A consequence of this assumption is that thematic roles are homomorphisms, or structure-preserving maps, with respect to the $\oplus$ operation (pace Kratzer 2003):

(57) **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.)

This property makes thematic roles very similar to the "trace functions" τ (runtime) and σ (spatial location) which are typically assumed to map events to those regions of spacetime that represent their temporal and spatial locations, respectively (Krifka, 1998). The location of the sum of two events is generally taken to be the sum of their locations, and similarly for their runtimes. For an algebraic semantic treatment of paths in spacetime, see Zwarts (2005). In the following, I will exploit this similarity between thematic roles and trace functions in the analysis of *jeweils* and other German-type distributive operators by allowing the dimension parameter on *jeweils* to be instantiated either with a thematic role or with τ. These two possibilities will lead to distribution over individuals and distribution over salient occasions respectively. In Champollion (2010b), I exploit this similarity a second time to account for examples like *The shoes cost $50*, where I claim that the Part operator is parametrized on the thematic role of the subject and distributes over salient pairs of shoes. Other examples, like *The patient took two pills (i.e. per day) for a month* are analyzed in an analogous way, except that the Part operator is parametrized on τ and therefore distributes over salient stretches of time.

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. 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 (58) as plural count nouns on the inclusive view of the plural, represented in (59):

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

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

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 $\lambda e[\text{lift}(e)]$ for the meaning of the verb *lift*, I write $\lambda e[{}^*\text{lift}(e)]$. I do the same for thematic roles.

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

(60) a. John slept.
b. Mary slept.
c. $\Rightarrow$ John and Mary slept.

(61) a. John lifted box b.
b. Mary lifted table t.
c. $\Rightarrow$ John and Mary lifted box b and table t.

(62) a. John is a boy.
b. Bill is a boy.
c. $\Rightarrow$ John and Bill are boys.

I assume that noun phrases are interpreted in situ, because I do not consider quantifier raising in this paper. Silent theta role heads, which denote thematic roles of type ve (event to individual), are located between noun phrases and verbal projections. The one that denotes the agent role can be either thought of as a silent case-marker-like part of the noun phrase, or as little v or Voice, depending on whether it first combines with the subject noun phrase or with the verb phrase. The heads that denote the theme and recipient roles bear some conceptual similarity with applicative heads (Pylkkänen, 2008). I will occasionally omit or abbreviate these heads in my LFs but they should always be assumed to be there. The precise nature of the compositional process is not essential, but it affects the types of the lexical entries of distance-distributive elements so let me make it concrete. I assume that the following type shifters apply first to the theta role head, then to the noun phrase, and finally to the verbal projection. I will introduce a third type shifter in Section 6 below in order to accommodate the alternative theoretical assumption that the theta role head combines first with the verbal projection and then with the noun phrase it belongs to, as opposed to the other way round. The choice between these assumptions is immaterial for the theories in this paper, but it matters to theories of the syntax-semantics interface more generally.

(63) a. Type shifter for indefinites: $\lambda\theta_{ve}\lambda P_{et}\lambda e.P(\theta(e))$
b. Type shifter for definites: $\lambda\theta_{ve}\lambda x\lambda e.\theta(e) = x$

Each of these type shifters combines a noun phrase with its theta role head to build an event predicate of type $\langle vt\rangle$ which can combine with other predicates of the same type via intersection. For example, after the noun phrases *the boys* (definite) and *two* *monkeys* (indefinite) combine with the theta role heads [ag] and [th] respectively, their denotations are as follows.

(64) ⟦[[ag] [the boys]]⟧ = $\lambda e[{}^*\text{ag}(e) = \bigoplus \text{boy}]$

(65) ⟦[[th] [two monkeys]]⟧ = $\lambda e[|{}^*\text{th}(e)| = 2 \wedge {}^*\text{monkey}({}^*\text{th}(e))]$

After the verb has combined with all its arguments, the event variable is existentially bound if the sentence is uttered out of the blue. If the sentence is understood as referring to a specific event, the event variable is instead resolved to that event. If the noun phrases combine directly with the verb, we get a scopeless reading as in (66). Here and below, I write *2M* as a shorthand for $\lambda e[|{}^*\text{th}(e)| = 2 \wedge {}^*\text{monkey}({}^*\text{th}(e))]$.

(66) ⟦The boys saw two monkeys⟧ = $\exists e[{}^*\text{ag}(e) = \bigoplus \text{boy} \wedge {}^*\text{see}(e) \wedge \text{2M}(e)]$

To generate distributive readings, we use Link's D operator, reformulated in the counterpart of this paper and repeated here as (67). As in Champollion (2014a), I assume that the D operator is coindexed with the thematic role of its target.

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

As an example, the distributive reading of (66) is derived like this:

(68) ⟦[[ag] The boys] [D_{ag} [saw [[th] two monkeys]]]]]⟧
$= \exists e[{}^*\text{ag}(e) = \bigoplus \text{boy} \;\wedge\; e \in [[\![D_{ag}]\!](\lambda e'[{}^*\text{see}(e') \;\wedge\; \text{2M}(e')])]]$
$= \exists e[{}^*\text{ag}(e) = \bigoplus \text{boy} \;\wedge\; e \in {}^*\lambda e'[{}^*\text{see}(e') \;\wedge\; \text{2M}(e') \;\wedge\; \text{Atom}(\text{ag}(e'))]]$

This formula is true just in case there is an event e whose agent is the boys, and which consists of seeing-two-monkeys events whose agents are atomic. As discussed in Champollion (2014a), the background assumptions of algebraic semantics ensure that the seeing-two-monkeys events have boys as agents even though the formula does not explicitly state this. I come back to this point at the end of this section.

Here are the entries for adverbial, adnominal, and determiner *each*. An explanation follows below. An illustration of the derivation of a basic sentence like *The boys saw two monkeys each* is shown in Figure 1. Adverbial and determiner *each* work similarly.

(69) $[\![\text{each}_\theta]\!]_{adverbial}$ = $[\![\text{D}_\theta]\!]$ = (67)

(70) $[\![\text{each}_\theta]\!]_{adnominal} = \lambda P_{et}\lambda\Theta_{ve}\lambda e.e \in [\![D_\theta]\!](\lambda e'.P(\Theta(e')))$

(71) $[\![\text{each}]\!]_{determiner} = \lambda P_{et}\lambda\theta_{ve}\lambda V_{vt}\lambda e\;[\theta(e) = \bigoplus P \wedge [\![D_\theta]\!](V)(e)]$

I assume that adverbial *each*, as shown in (69), is a verb phrase modifier just like the D operator, and can therefore be given the same entry as that operator. I adopt for concreteness the assumption that adverbial *each* is an adverb adjoined to VP. This is similar what has been argued for floating quantifiers in general by Dowty and Brodie (1984), Bobaljik (1995) and Doetjes (1997). For another view that analyzes floating quantifiers as the remaining part of a noun phrase the rest of which has moved away from it, see for example Safir and Stowell (1988) and Sportiche (1988). The movement view makes a formal link between *each* and its antecedent available for independent

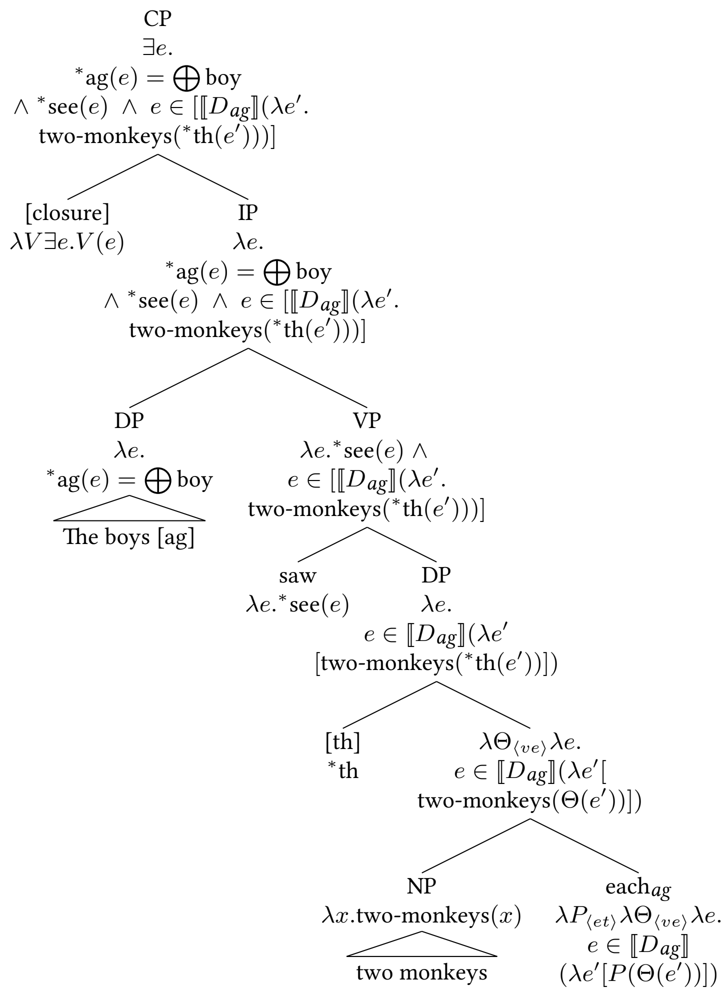

Figure 1: Deriving *The boys saw two monkeys each.*

reasons since there is a movement relation between them, while the adverbial view requires one to assume that the two are coindexed. Therefore the movement view, which I do not adopt here, would be likely to be at least as compatible with my approach than the adverbial view.

Adnominal and determiner *each*, as shown in (69) and (70), need to be type-shifted, but both are defined in terms of the D operator. This captures the fact that they are essentially synonymous. As shown in (70), adnominal *each* carries an index. I assume that it is coindexed with the theta role θ of its antecedent. It first combines with its host predicate P (e.g. *two monkeys*), and then with the theta head Θ of the host (which is not to be confused with the theta role of its antecedent). Afterwards, it combines intersectively with the verbal projection to which its host attaches (e.g. the verb *see*). This means that adnominal *each* does not take scope over the verbal projection but only over its own complement (Dotlačil, 2012). By contrast, adverbial *each* takes scope over the verbal projection. In this, I follow LaTerza (2014b), who evidence the following minimal pair as evidence for this contrast:

(72) a. John and Bill served [four meals each] to (exactly) three judges.
 b. John and Bill each [served four meals to (exactly) three judges].

As LaTerza reports, speakers judge (72a) true of situations where there are at most three judges, while (72b) is true in situations which allow up to six different judges. As he notes, this is predicted by giving adnominal and adverbial *each* different scopes, as indicated by the square brackets in these examples. For example, sentence (72b) can be derived as in Figure 2.

My entry for adnominal *each* combines with its host in two steps, in order to give it access to both the predicate and the theta head. This is not essential, but it allows us to ensure that the type of the predicate is $\langle et \rangle$. I do so to provide a hook on which to build future accounts of the "counting quantifier requirement" that prevents such phrases as **most men each* (Safir and Stowell, 1988; Sutton, 1993; Szabolcsi, 2010, §10.5). The theory in this paper does not to provide an account of this requirement, since it will not rule out bare plurals as in **They saw monkeys each*, as pointed out in Cable (2014). If an independent account of these kinds of mismatches can be provided that does not need to have access to the host predicate and its theta role separately, then it may not be necessary to place *each* between the host predicate and the theta head after all. The choice is immaterial for the rest of this paper.

Determiner *each*, as shown in (71), is not coindexed with anything because it is not a distance-distributive element. It combines first with a nominal and then with a theta head. The result is a phrase of VP modifier type $\langle vt, vt \rangle$, which is also the type of D_θ. Some intermediate steps of the derivations of sentence (6) are shown in (73) and (74).

(73) $[\![[[\text{two monkeys}]\ \text{each}_{ag}]\ [\text{th}]]\!]$
$= \lambda e[[\![D_{ag}]\!](\lambda e'[|{}^*\text{th}(e')| = 2 \wedge {}^*\text{monkey}({}^*\text{th}(e'))])(e)]$
$= \lambda e[e \in {}^*\lambda e'[{}^*\text{see}(e') \wedge |{}^*\text{th}(e')| = 2 \wedge {}^*\text{monkey}({}^*\text{th}(e')) \wedge \text{Atom}(\text{ag}(e'))]]$

(74) $[\![[[\text{Each boy}]\ [\text{ag}]]\!]$
$= \lambda V_{vt}\lambda e[{}^*\text{ag}(e) = \bigoplus \text{boy} \wedge [\![D_{ag}]\!](V)(e)]$
$= \lambda V_{vt}\lambda e[{}^*\text{ag}(e) = \bigoplus \text{boy} \wedge e \in {}^*\lambda e'[V(e') \wedge \text{Atom}(\text{ag}(e'))]]$

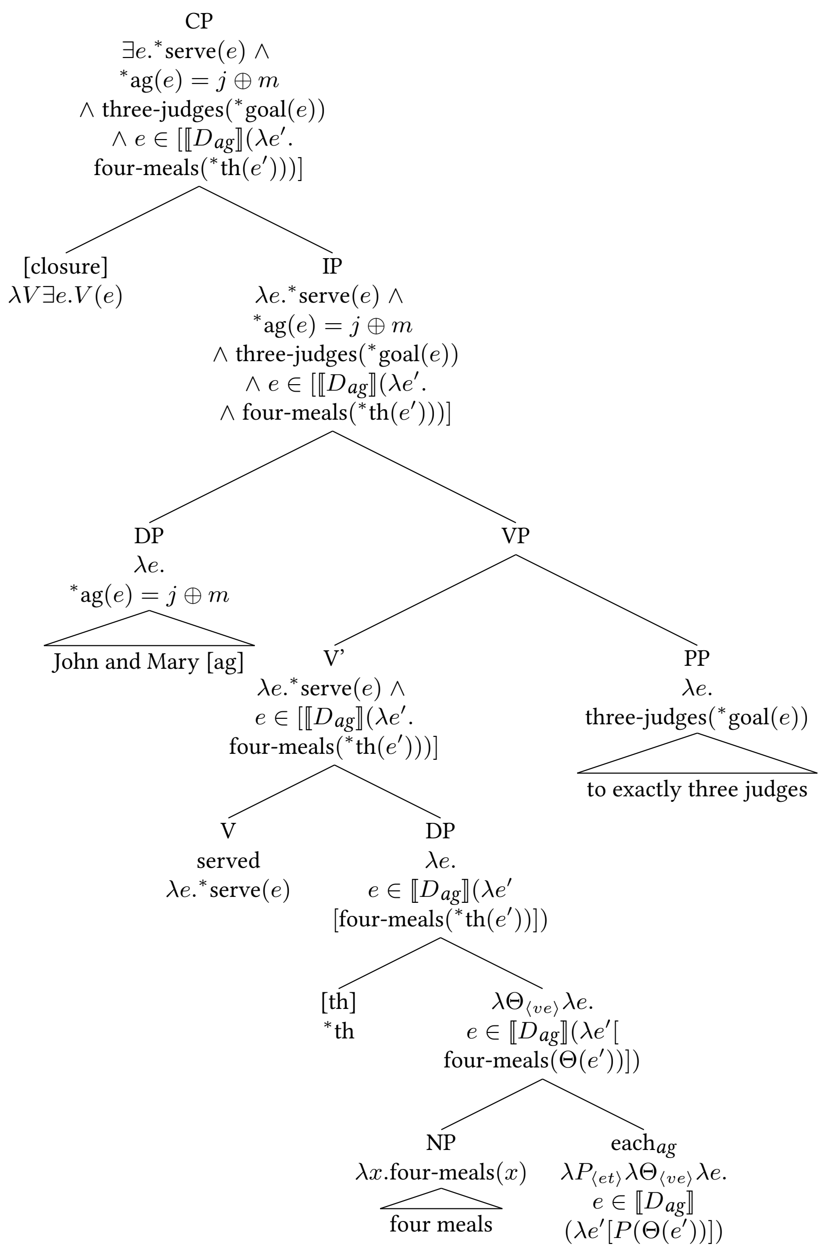

Figure 2: Deriving *John and Mary served four meals each to exactly three judges.*

Let us now turn to *jeweils*. My reformulation of the Part operator in Champollion (2014a), repeated here as (75), is also the lexical entry of adverbial *jeweils*, as shown in (76). The same type shift as in (70) brings us from (76) to adnominal *jeweils*, as shown in (77).

(75) **Definition: Event-based Part operator**
$[\![\text{Part}_{\theta,\text{C}}]\!] \stackrel{\text{def}}{=} \lambda P_{\langle vt \rangle} \lambda e[e \in {}^*\lambda e'(P(e') \ \wedge \ C(\theta(e')))]$
(Takes an event predicate P and returns a predicate that holds of any event e which consists entirely of events that are in P and whose θs satisfy the contextually salient 'cover predicate' C.)

(76) $[\![\text{jeweils}_{\theta,C}]\!]_{\textit{adverbial}} = [\![\text{Part}_{\theta,C}]\!] = (75)$

(77) $[\![\text{jeweils}_{\theta,C}]\!]_{\textit{adnominal}} = \lambda P \lambda \Theta \lambda e[[\![\text{Part}_{\theta,C}]\!](\lambda e'[P(\Theta(e'))])(e)]$

As in the case of the Part operator, the granularity parameter C of *jeweils* can be set to *Atom* so long as its dimension parameter θ is set to a function into a count domain, such as *ag*. In that case, *jeweils* distributes over individuals and is equivalent to *each*. The following example illustrates this with sentence (7a); sentence (7b) is equivalent.

(78) Die Jungen haben jeweils$_{ag,Atom}$ zwei Affen gesehen.
The boys have DIST two monkeys seen.
"The boys have each seen two monkeys."

If – and only if – there is a supporting context, the anaphoric predicate C can be set to a salient antecedent other than *Atom*, and in that case θ is free to adopt values such as τ (runtime). This leads to occasion readings. Suppose for example that it is in the common ground that the boys have been to the zoo to see animals last Monday, last Wednesday and last Friday, and that (7a) is uttered with reference to that state of affairs, or sum event. It is interpreted as follows.

(79) $[\![$Die Jungen haben jeweils$_{\tau,\text{zoovisit}}$ zwei Affen gesehen.$]\!]$ =
${}^*\text{ag}(e_0) = \bigoplus \text{boy} \wedge {}^*\text{see}(e') \wedge e_0 \in {}^*\lambda e'[|{}^*\text{th}(e')| = 2 \wedge {}^*\text{monkey}({}^*\text{th}(e')) \wedge \text{zoovisit}(\tau(e'))]$
"The boys have seen two monkeys on each occasion."

Since the sentence refers specifically to the sum e_0 of the three events in question, the event variable in (79) is resolved to e_0 rather than being existentially bound. The predicate that is true of any time interval at which a zoo visit takes place, call it *zoovisit*, is also salient in this context. So C can be resolved to *zoovisit* rather than to *Atom*. Since there are no atoms in time, it is only now that θ can be set to τ, rather than to *ag* as in (78). What (79) asserts is that e_0 has the boys as its agents; that it can be divided into subevents, each of whose runtimes is the time of a zoo visit; and that each of these subevents is an event whose theme are two monkeys. That these subevents are seeing events is entailed by the fact that *see* is lexically distributive on its theme argument, which in turn is formally represented as a meaning postulate, as discussed in Champollion (2014a). I assume that runtime is closed under sum just like other thematic roles ($\tau = {}^*\tau$), or in other words, it is a sum homomorphism. This means that any way of dividing e_0 must result in parts whose runtimes sum up to

$\tau(e_0)$. Assuming that $\tau(e_0)$ is the (discontinuous) sum of the times of the three zoo visits in question, this entails that each of these zoo visits is the runtime of one of the seeing-two-monkeys events. This is the occasion reading.

4 Zimmermann's Generalization Explained

This section discusses the question how to capture the correlation expressed in Zimmermann's generalization (14). That is to say, why does a distance-distributive element which can also be used as a distributive determiner lack the occasion reading? Zimmermann himself proposes a syntactic explanation: Determiners must agree with their complement; adnominal or adverbial *each* also has a complement, a proform that must acquire its agreement features from its antecedent, which is the antecedent of *each*; only overt antecedents have agreement features; so adnominal/adverbial *each* cannot have a covert antecedent; so it cannot refer to a contextually supplied but not overtly mentioned antecedent such as a salient set of occasions.

This explanation is compatible with the present framework, and it makes the right predictions given the assumption that covert antecedents cannot have agreement. However, this assumption is problematic. To mention a simple example, German has grammatical gender, so for example, the gender of *Tisch* 'table' is feminine. Knowing this, a German speaker can point to a table and say with reference to it:

(80) Den hab ich gebraucht gekauft.
This.m have I used bought.
'I have bought this used.' (German)

But I cannot point to it and say:

(81) *Die hab ich gebraucht gekauft.
'This.f. have I used bought.' (German)

As this example shows, a deictic pronoun in German has to agree in grammatical gender with the gender of the noun phrase that would most aptly describe this antecedent, even though the antecedent has not been mentioned explicitly.

A similar phenomenon was pointed in English by Tasmowski-De Ryck and Verluyten (1982). English pronouns agree with their antecedents based on syntactic rather semantic grounds, as is shown by *pluralia tantum* such as *pants* and *scissors* which are syntactically plural but semantically singular. Pronouns show syntactic agreement with their antecedents even when these antecedents are not overtly mentioned:

(82) a. (John wants his pants that are on a chair and he says to Mary:)
Could you hand them/*it to me, please?
b. (Same situation but with a shirt:)
Could you hand *them/it to me, please?

For a recent discussion of these facts, and an explanation in terms of covert syntactic antecedents that are included in the syntactic structure and c-command the entire sentence in question, see Collins and Postal (2012, ch. 4). In the following I will remain neutral on whether the covert syntactic antecedent should be thought of as being included in the syntactic structure or not.

Zimmermann's analysis assumes that *each* indeed has a silent complement and that the two agree. Since *each* does not show overt agreement in English, it is difficult to tell whether it in fact agrees with anything, just like English *all*. In general, adverbial distance-distributive elements do not have to agree with their antecedents. For example, German *all-* 'all' and *jed-* 'every, each' can both be used as determiners, as seen in (83) and (84), and as adverbial distance-distributive elements, as seen in (85) and (86). When used as determiners in count domains, the former is plural and requires plural agreement, and the latter is singular and requires singular agreement, just like their English counterparts. When used as adverbials, *alle* is plural, as shown in (85) but *jed-* must be singular, as shown in (86). The latter is the case even when the antecedent of *jed-* is plural.

(83)
Alle	Jungen	haben	geniest.
All.pl	boy.pl	have.pl	sneezed.

'All the boys have sneezed.'

(84)
Jeder	Junge	hat	geniest.
Dist.sg	boy.sg	has.sg	sneezed.

'Each boy has sneezed.'

(85)
Die	Jungen	haben	alle/*aller	geniest.
The.pl	boy.pl	have.pl	all.pl/all/sg.	sneezed.

'The boys have all sneezed.'

(86)
Die	Jungen	haben	jeder/*jede	geniest.
The.pl	boy.pl	have.pl	Dist.sg/Dist.pl	sneezed.

'The boys have each sneezed.'

This shows that at least some distance-distributive elements do not need to agree with their antecedents. To put it differently, on Zimmermann's assumptions, the proform associated with *jeder* in (86) is singular and so it cannot have acquired its agreement features from its antecedent *Jungen*, since that antecedent is plural and *alle* in (85) is also required to be plural.

While Zimmermann's account seems problematic, its difficulties can perhaps be overcome, and it is by itself not incompatible with the present framework. But in the context of the general theory adopted here, a more straightforward explanation suggests itself. Distributive determiners like English *each* are only compatible with count nouns, not with mass nouns. We know this since, for example, **each mud* is ungrammatical. Formally, this amounts to an atomicity requirement of the kind the Link's D operator provides, as discussed in Champollion (2014a). This requirement can be seen as independent evidence of the atomic distributivity hard-coded in the entry (71) via the D operator. In other words, the distance-distributive element inherits the atomicity requirement of the determiner. This explanation is in line with the notion of parameter settings imported from strata theory and described above. That is, in English, adnominal, adverbial and determiner *each* have identical meanings up to type-shifting. Determiner *each* is only compatible with count domains because its granularity parameter is hardwired to the value "Atom". Adnominal *each* is formally identical to determiner *each*, so it inherits this property.

5 Previous Work: Zimmermann (2002b)

This section discusses the semantic account of *jeweils* and *each* offered in Zimmermann (2002b). There are useful summarizations of the essence of Zimmermann's proposal in a number of places (Zimmermann, 2002a; Blaheta, 2003; Dotlačil, 2012). Here I provide my own and add some novel points to the discussion.

Zimmermann takes adnominal *each* and *jeweils* to be prepositional phrases that are only partially pronounced, but this aspect does not really influence the semantic composition. The meaning of *each*, or more precisely of the prepositional phrase that is supposed to contain it, is as follows (Zimmermann, 2002b, p. 210). While the relevant discussion is actually about adnominal *jeweils*, it carries over to adnominal *each* without changes, so I present it in terms of *each* for clarity.

(87) $[\![\text{each}]\!]$ (Zimmermann) = $\lambda P.\forall z[z \in Z_i \rightarrow \exists x[P(x) \wedge {}^*R_j(z,x)]]$

This meaning is a property of predicates that holds of a given predicate P iff every member of a certain plurality Z_i stands in the algebraic closure ${}^*R_j$ of a certain relation R_j to some entity of which P holds. In this definition, Z_i and R_j are free variables that are assumed to be coindexed, respectively, with the antecedent of *each* and with the relation that is denoted typically by the verb. Take for example sentence (11), repeated here as (88) with the coindexation added. Here P is the denotation of *two monkeys*, Z is coindexed with *the boys*, and R is coindexed with *saw*.

(88) [The boys]$_i$ saw$_j$ two monkeys each$_{i,j}$.

Here is how this sentence would be represented by Zimmermann (2002b). First, the entry for *each* is applied to *two monkeys*, which is taken to denote a predicate of sum individuals that I will represent here by the shorthand *two-monkeys*. This results in an open proposition with two free variables:

(89) $[\![\text{two monkeys each}_{i,j}]\!] = \forall z[z \in Z_i \rightarrow \exists x[\text{two-monkeys}(x) \wedge {}^*R_j(z,x)]]$

The next steps involve lambda-abstracting over the free variables, via a rule that Zimmermann calls "index-triggered lambda abstraction", a variant of a rule which is taken to apply when a type mismatch makes function application impossible (Bittner, 1994, p. 69).

(90) **Index-triggered λ-abstraction (Zimmermann, 2002b, p. 217)**:
If the semantic types of a proposition-denoting expression α and its syntactic sister β do not match, and if $[\![\alpha]\!]$ contains a free variable u_i that shares an index 'i' with β, λ-abstraction in $[\![\alpha]\!]$ over index 'i' is licensed, and $\lambda u_i.[\![\alpha]\!]$ is a value for α.

This rule allows a constituent with a free variable in it to combine with another constituent that is coindexed with that variable. For example, in (88), the constituent *two monkeys each*$_{i,j}$ has the free variable j in it, which carries the same index as the constituent *saw*$_j$. Since the two constituents are sisters, index-triggered λ-abstraction applies, with the result as shown in (91), as discussed in Zimmermann (2002b, p. 226).

(91) $\lambda R.\forall z[z \in Z_i \rightarrow \exists x[\text{two-monkeys}(x) \wedge {}^*R_j(z,x)]]$

Zimmermann takes the classical Davidsonian view on verb meaning, under which n-ary verbs denote $n+1$-ary relations between arguments and events, and he assumes that the event argument can be saturated at the level of the verb phrase by existential closure so that the verb *saw* can have the right type to combine with (91), as shown below:

(92) $[\![\text{saw}_j]\!] = \lambda y \lambda x.\exists e[\text{see}(x,y,e)]$

(93) $[\![(91)]\!]([\![(92)]\!]) = \forall z[z \in Z_i \rightarrow \exists x[\text{two-monkeys}(x) \wedge \exists e[{}^*\text{see}(z,x,e)]]]$

In this derivation, the existential quantifier over events has ended up outside the scope of the star operator. This is a bit mysterious if the star operator is introduced as a part of the meaning of *each*, but the problem would disappear if the star operator, or something to its effect such as lexical cumulativity was introduced as part of the meaning of the VP or through background assumptions. In any case, the result of the computation, as shown in (93), is another open proposition. The last step is to combine this with the antecedent, *the boys*$_i$, in another instance of index-triggered λ-abstraction. The result is as follows:

(94) $\forall z[(z \in \bigoplus \text{boy}) \rightarrow \exists x[\text{two-monkeys}(x) \wedge \exists e[{}^*\text{see}(z,x,e)]]]$

This formula says that for every boy there exists a sum of two monkeys such that the boy saw the monkeys. This is an accurate rendering of the truth conditions of sentence (88).

Zimmermann's system works as long as there are constituents that express the relation R and the sum Z that figure in the meaning of *each*. In the case of the sum, this is not a problem, since the antecedent of *each* is always a constituent. In the case of the relation, however, it is not always guaranteed that such a constituent exists (Dotlačil, 2011). The first two following examples are taken from Blaheta (2003), and the third one from Boeckx and Hornstein (2005).

(95)
a. Alex and Sasha visited the capitals of three states each.
b. Alex and Sasha dragged three bags through four puddles each.
c. Two women got the prize money and a silver medal each.

In these examples, the relation with which R in (87) would have to be coindexed is expressed by several words that do not form a constituent: "visited the capitals of", "dragged three bags through", "got the prize money and". Zimmermann's account is not flexible enough to model this. Similar problems also affect accounts that assume that adnominal *each* is licensed by, or otherwise requires, the presence of a D operator that takes scope over the verb phrase (e.g. Roberts, 1987; Oh, 2006). One solution is to give binominal *each* scope only over its noun phrase (Dotlačil, 2011; LaTerza, 2014b). This has been adopted here in the entry (70).

This solution predicts that binominal *each* should occur in sentences where its antecedent stands in a cumulative-quantification relation with another noun phrase. To test this prediction, a survey was run via the online marketplace Amazon Mechanical Turk (`www.mturk.com`) with 100 responses, of which 62 were from native speakers of English who had grown up in the US, the UK or Canada (the other responses were discarded). The text of the survey ran as follows:

(96) Imagine there are 20 boys and 10 girls in a room. Each of the boys gives one of the girls a kiss. Every girl receives at least one kiss. Your friend Bill saw what happened and describes it to you by saying: "20 boys gave 10 girls one kiss each."

The survey asked participants two questions: *Is he saying the truth?* and *How many kisses were given in total according to Bill's sentence?* The former question was implemented as a yes-no choice and the latter question required participants to type in a number. Of the 29 participants stated that Bill is saying the truth, 19 participants stated that according to Bill, 20 kisses were given in total and 7 stated that 200 kisses were given. Of the 33 participants stated that Bill is lying, 20 participants stated that according to Bill, 200 kisses were given in total and 8 stated that 20 kisses were given.

Of particular interest to us are the 19 survey participants who stated that Bill is saying the truth and who understood Bill as saying that 20 kisses were given. For at least these participants, the sentence has a cumulative reading in which *each* take the theta role of *twenty boys* as its antecedent. Thus, cumulative quantification is possible between the host phrase of *each* and a noun phrase that is not its antecedent – in this case, *10 girls*. (I suspect that those participants who understood Bill as saying that two hundred kisses were given understood *each* as taking the theta role of *ten girls* as its antecedent, and that they understood the entire vP *gave ten girls one kiss each* as being in the scope of a D operator that distributes it over the twenty boys.)

To confirm the prediction on a different sentence, another Mechanical Turk survey was run with 100 responses, of which this time 56 were from native speakers. Its text ran as follows:

(97) Imagine there is an election with two candidates. The day after, you read in the newspaper: "100 million voters gave two candidates one vote each." According to the newspaper, how many votes were cast in total?

Again, participants were asked to type in a number. By far the most frequent answers were "200 million" (32 participants) and "100 million" (19 participants). It is the latter group that is of interest to us, since they presumably took the sentence to describe a situation in which every one of the 100 million voters voted for only one of the candidates, with some of these votes going to the first candidates and the rest going to the second. This reading involves a cumulative relation between *two candidates* and *one vote each*, where the antecedent of *each* is the theta role of the subject. To see this, observe that *one vote each* has the same extension as *100 million votes* in the scenario described. Each of the 100 million votes went to one of the two candidates, and each of the two candidates got some of the 100 million votes. This is the hallmark of cumulative quantification (Scha, 1981; Krifka, 1999). (Those participants that answered

"200 million" presumably took the word *each* as taking the theta role of the indirect object *two candidates* as its antecedent, and interpreted the vP *gave two candidates one vote each* as being in the scope of a D operator that distributes it over the 100 million voters.)

The observations gathered in these two surveys confirm that cumulative quantification is possible between the host phrase of *each* and a noun phrase that is not its antecedent. This means that the universal quantifier or distributive operator that is contributed by *each* must not be given scope over the entire relation that holds between each member of the antecedent and each member of the host denotation. In Zimmermann's implementation, this is the relation R. As can be seen in (87), the universal quantifier introduced by *each* in Zimmermann's implementation actually takes scope over R. For this reason I see no way to represent this kind of cumulative quantification in Zimmermann's system. On my account, this is straightforward. Adnominal *each* does not take scope over the verbal predicate, just over the host noun phrase, as shown in the entry in (70). The result of analyzing sentence (97) is as follows:

(98) $\exists e.[{}^*\text{give}(e)$
$\wedge$ 120-million-voters$({}^*\text{ag}(e))$
$\wedge$ two-candidates$({}^*\text{ben}(e))$
$\wedge\ e \in {}^*\lambda e'[\text{vote}({}^*\text{th}(e'))$
$\wedge\ \text{Atom}(\text{ag}(e'))]]$

This representation states that there is a sum of giving events whose agents sum up to 100 million voters, whose beneficiaries sum up to two candidates, and which consist of subevents with atomic agents and individual votes as themes. The background assumptions of algebraic semantics ensure that this entails that there were 100 million such subevents.

Figure 3 shows the compositional interpretation of sentence (97). Since the current consensus is that the syntax of ditransitives differs somewhat from their surface structure, I have chosen to make this analysis harmonize with the syntax proposed in Marantz (1993) and recently defended by Bruening (2010). For this purpose, I have assumed that the *Voice* and *Appl* heads introduce the thematic roles *ag* and *ben* respectively, and I have type-shifted these heads accordingly. For reasons of clarity, I show the meanings of these heads only in their type-shifted form, but really they should be thought of as having the same meanings as above, that is, *ag and *ben. The type-shifter I have used here is the following:

(99) $\lambda\theta_{ve}\lambda V_{vt}\lambda P_{et}\lambda e.V(e) \wedge P({}^*\theta(e))$

An alternative, if one does not mind introducing additional compositional operations, would be to use the event identification operation from Kratzer (1996). For details on how this works, see Harley (2012).

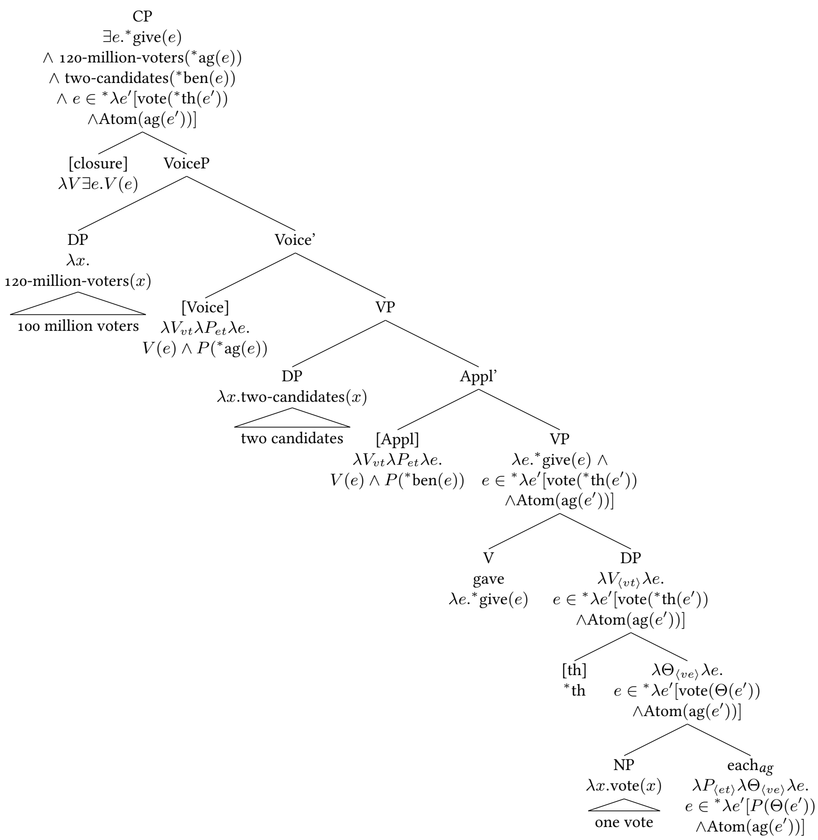

Figure 3: Deriving *100 million voters gave two candidates one vote each.*

6 *Each* as a PP Modifier

This section discusses the compositional interpretation of English *each* when it occurs as the modifier of a prepositional phrase. Two examples in which this is the case are the following. The first is from Maling (1976) and the second is from Schein (1993).

(100) Mary put the books each back on the bookshelf.

(101) 300 quilt patches covered two workbenches each with two bedspreads.

Distance distributive *each* in sentences like (101) plays a role in an argument for Neo-Davidsonian event semantics and thematic roles offered by Schein (1993). I discuss this argument in Champollion (2009). Here, I take Neo-Davidsonian event semantics for granted and I focus on giving a compositional implementation, which Schein himself does not provide.

The word *each* appears to occupy the same syntactic position in these sentences as the word *both* in the following example from Janke and Neeleman (2012):

(102) He read the books both to someone famous.

While not all speakers accept such sentences, there is some indication that distance-distributive and other floated elements can modify prepositional phrases that follow them, as shown by the fronting test. The examples in (103) are from Bobaljik (1995, p. 214) and the one in (104) is due to Benjamin Bruening (p.c.).

(103) a. The magicians disappeared all at the same time.
b. All at the same time, the magicians began to appear.
c. *At the same time, the magicians began to appear all.

(104) All back at the farm, the boys sat down to lunch.

Aside from the parallel with *all*, two facts suggest that in sentences like (100), the word *each* does not modify the noun phrase to its left but the prepositional phrases to its right. First, definite noun phrases like *the books* cannot normally be modified by adnominal *each*. Second, (101) can be read with a prosodic break between *two workbenches* and *each* and in that case it has a reading where there is a total of two workbenches and a total of four bedspreads. This is the reading that Schein intends it to have, as evidenced by the scenario description he offers:

> "The tailor has taken 300 quilt patches from his stock to fabricate four bedspreads, which now cover over the workbenches. It cannot be assumed that any of the patches of irregular shape and size was sewn into any one bedspread. The tailor may have divided it among several. It may furthermore be the case that no group of fewer than the 300 quilt patches cover any workbench. All that can be said is that there was an event of two workbenches each being covered over with two bedspreads, and *there* 300 quilt patches made up the covers. The cross-reference to events is essential." (Schein, 1993, p. 9)

The word *each* in sentence (101) could also be read as an adnominal modifier that modifies *two workbenches*. In that case, the sentence requires there to be more than two workbenches, while Schein's scenario description suggests that there are only two workbenches in total. The use of *each* here, and in sentences (100) and (102), can be called adverbial or perhaps "ad-prepositional", depending on what syntactic structure one assumes: a right-branching structure like [V [DP PP]] or a left-branching structure like [[V DP] PP]. Both structures have been advocated in the literature. For a recent overview, see for example Janke and Neeleman (2012). On some incarnations of the right-branching analysis, *each* can modify the trace of a verb including the one advocated for sentence (102) by Janke and Neeleman (e.g. 2012). For example, Janke and Neeleman take *read* in (102) to originate between *both* and *to* and to move leftward. On the left-branching analysis, the most natural option is to regard these elements as "ad-prepositional" modifiers of a prepositional phrase, like the adverb *back* in *back at the farm*. I will adopt this option here.

The following formula captures this reading:

(105)
$$\begin{aligned}&\exists e.{}^*\text{cover}(e)\\ &\wedge {}^*\text{quilt-patch}({}^*\text{ag}(e)) \wedge |{}^*\text{ag}(e)| = 300\\ &\wedge {}^*\text{workbench}({}^*\text{th}(e)) \wedge |{}^*\text{th}(e)| = 2\\ &\wedge e \in {}^*\lambda e'[{}^*\text{bedspread}({}^*\text{instr}(e')) \wedge |{}^*\text{instr}(e')| = 2 \wedge \text{Atom}(\text{th}(e'))]\end{aligned}$$

Formula (105) is derived as in Figure 4, where I have used shortcuts like *300-quilt-patches* for better readability. It can be paraphrased as follows: There is a sum of covering events whose agents sum up to 300 quilt patches, whose themes sum up to two workbenches, and which can be divided into two smaller sum events, each of which involves two bedspreads and one of the workbenches. All this is true in the scenario Schein describes. The cross-reference to events, as Schein calls it, is achieved by having the predicate denoted by the subject noun phrase *three hundred quilt patches* modify the same event variable as the PP *each with two bedspreads*.

7 Summary and Discussion

At the end of Section 2, I had suggested that the facts reviewed in that section suggest the following requirements for a theory of distributivity. First, the synonymy of the determiner, adnominal and adverbial uses of *each* in English should be captured, ideally by essentially identical lexical entries. Second, the fact that distance-distributive elements across languages share some part of their meanings (namely their individual-distributive readings) should be represented, as well as the fact that some of them can also have occasion readings. Third, the analysis should clarify the connections between distance-distributive elements and distributivity theory more generally, and the semantic variation across distance-distributive elements should be captured. Finally, an explanation should be readily available for the crosslinguistic observation that distance-distributive elements that can also be used as determiners can only distribute over individuals (Zimmermann's generalization).

Let me briefly summarize how the theory presented in this paper addresses these issues. The synonymy of the determiner, adnominal and adverbial uses of *each* in

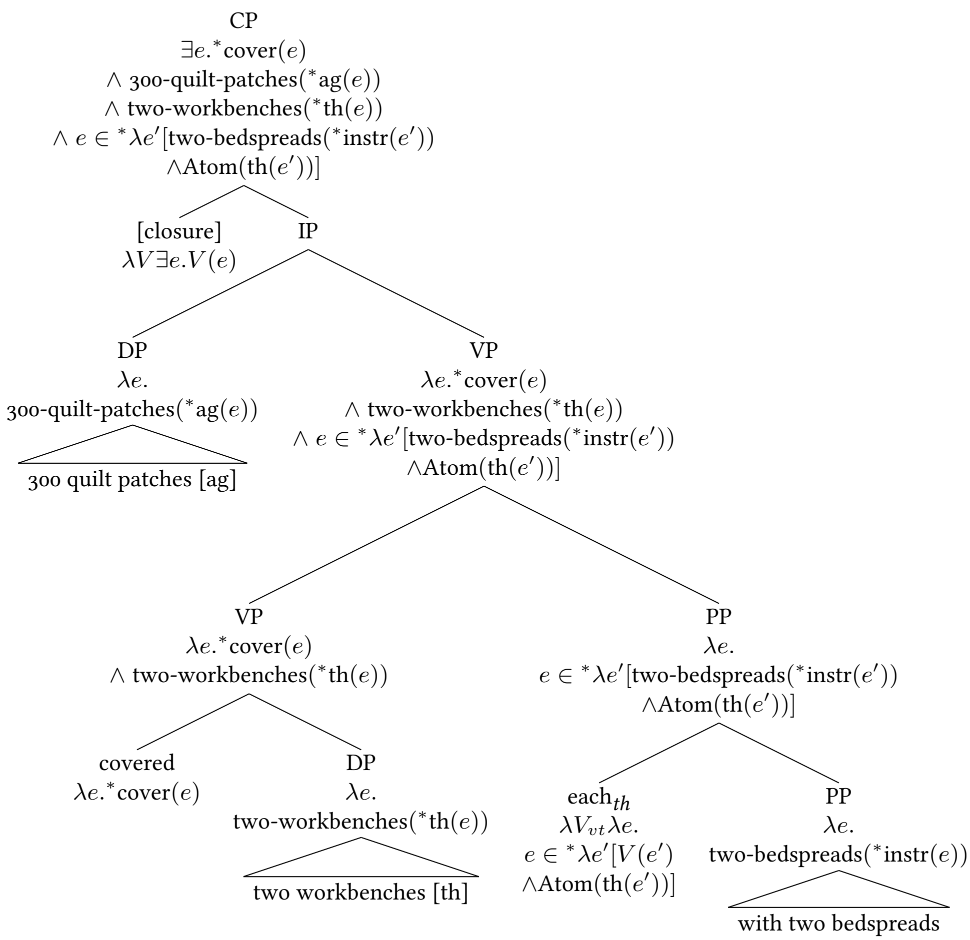

Figure 4: Deriving *300 quilt patches covered two workbenches, each with two bedspreads.*

English is captured by the fact that they are all derived from Link's D operator. I have represent the fact that distance-distributive elements across languages share some part of their meanings by derived them from related distributivity operators (Link's or Schwarzschild's) which differ from each other only in their parameter settings. On the theory presented here, distance-distributive items display the same parametric variation as covert distributivity operators do, not only insofar as nonatomic distributivity is concerned, but also insofar as the ability is concerned to target different thematic roles or time. The semantic variation among distance-distributive elements is captured by restriction on parameter settings. One type of element, exemplified by English *each*, is hard-wired for distribution over atoms and the other one, exemplified by German *jeweils* also allows distribution over nonatomic contexts. Zimmermann's generalization is explained by the natural assumption that distance-distributive elements are formally identical to distributive determiners and therefore inherit their inability to distribute over nonatomic domains, no matter if these domains are mass or temporal.

Taken together, this paper and its counterpart, Champollion (2014a), suggest the following general picture of distributivity. No matter whether distributivity is introduced by an overt or by a covert element, it always involves a certain domain that contains the individuals or the material to be distributed over, and a certain size or granularity that specifies how finely the relevant predicates are distributed. When the domain question is a count domain, for example when we distribute over people or objects, then it is always possible to distribute over these objects one by one. When the domain in question does not make such atomic units available, as in the case of time or space, two things can happen. Either the element in question does not allow distribution over such nonatomic domains, for example because it is incompatible with non count domains to begin with, or else it looks for a salient cover or set in the context, such as a salient set of temporal locations. Those distributive elements that can do this in principle can also do this in count domains even though atoms are available.

References

Balusu, Rahul (2005). "Distributive reduplication in Telugu". In: *Proceedings of the Thirty-Sixth Conference of the North Eastern Linguistics Society (NELS 36)*. Ed. by Christopher Davis, Amy Rose Deal, and Youri Zabbal. Graduate Linguistic Student Association (GLSA).

Balusu, Rahul and Karattuparambil A. Jayaseelan (2013). "Distributive quantification by reduplication in Dravidian". In: *Strategies of quantification*. Oxford, UK: Oxford University Press, pp. 60–86. DOI: 10.1093/acprof:oso/9780199692439.003.0004.

Barker, Chris (2002). "Continuations and the nature of quantification". In: *Natural Language Semantics* 10, pp. 211–242.

Barker, Chris (2010). "Nominals don't provide criteria of identity". In: *The Semantics of Nominalizations across Languages and Frameworks*. Ed. by Monika Rathert and Artemis Alexiadou. Mouton de Gruyter, pp. 9–24.

Bittner, Maria (1994). “Cross-linguistic semantics”. In: *Linguistics and Philosophy* 17, pp. 53–108.
Blaheta, Don (2003). “Binominal *each*: evidence for a modified type system”. MA thesis. Brown University.
Bobaljik, Jonathan D. (1995). “Morphosyntax: The syntax of verbal inflection”. PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.
Bobaljik, Jonathan D. (2001). “Floating quantifiers: Handle with care”. In: *The second state-of-the-article book*. Ed. by Lisa Cheng and Rynt Sybesma. 45-68. Berlin, Germany: Mouton de Gruyter.
Boeckx, Cedric and Norbert Hornstein (2005). “On eliminating D-structure: The case of binominal *each*”. In: *Syntax* 8, pp. 23–43.
Brasoveanu, Adrian and Donka F. Farkas (2011). “How indefinites choose their scope”. In: *Linguistics and Philosophy* 34.1, pp. 1–55.
Bruening, Benjamin (2010). “Ditransitive asymmetries and a theory of idiom formation”. In: *Linguistic Inquiry* 41.4, pp. 519–562.
Burzio, Luigi (1986). *Italian syntax: A government-binding approach*. Vol. 1. Springer.
Cable, Seth (2014). “Distributive numerals and distance distributivity in Tlingit (and beyond)”. In: *Language*.
Carlson, Gregory N. (1984). “Thematic roles and their role in semantic interpretation”. In: *Linguistics* 22, pp. 259–279.
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.
Champollion, Lucas (2009). “Cumulative readings of *every* do not provide evidence for events and thematic roles”. In: *Proceedings of the Nineteenth Amsterdam Colloquium*.
Champollion, Lucas (2010a). “Parts of a whole: Distributivity as a bridge between aspect and measurement”. PhD thesis. Philadelphia, PA: University of Pennsylvania.
Champollion, Lucas (2010b). “Quantification and negation in event semantics”. In: *The Baltic International Yearbook of Cognition, Logic and Communication* 6. DOI: 10.4148/biyclc.v6i0.1563.
Champollion, Lucas (2012). “Each vs. Jeweils: A cover-based view on distance distributivity”. In: *Logic, Language and Meaning*. Ed. by Maria Aloni et al. Vol. 7218. Lecture Notes in Computer Science. Springer Berlin Heidelberg, pp. 251–260. ISBN: 978-3-642-31481-0. DOI: 10.1007/978-3-642-31482-7_26.
Champollion, Lucas (2013). “The scope and processing of for-adverbials: A reply to Deo and Piñango”. In: *Proceedings of the 23rd Semantics and Linguistic Theory Conference (SALT 23)*. Ed. by Todd Snider, pp. 432–452.
Champollion, Lucas (2014a). “Covert distributivity in algebraic event semantics”. Submitted.
Champollion, Lucas (2014b). *Every boy bought two sausages each: Distributivity and dependent numerals*. Presented at WCCFL 32 (32nd West Coast Conference on Formal Linguistics).
Champollion, Lucas and Manfred Krifka (2014). “Mereology”. In: *Cambridge handbook of semantics*. Ed. by Paul Dekker and Maria Aloni. Cambridge University Press.
Choe, Jae-Woong (1987). “Anti-quantifiers and a theory of distributivity”. PhD thesis. Amherst, MA: University of Massachusetts.

Collins, Chris and Paul Martin Postal (2012). *Imposters: a study of pronominal agreement.* MIT Press.
Davies, Mark (2008). *The Corpus of Contemporary American English: 450 million words, 1990-present.* Available online at http://corpus.byu.edu/coca/.
Doetjes, Jenny (1997). "Quantifiers and selection: on the distribution of quantifying expressions in French, Dutch and English". Volume 32 of HIL Dissertations. The Hague, Netherlands: Holland Institute of Generative Linguistics.
Dotlačil, Jakub (2011). "Fastidious distributivity". In: *Proceedings of SALT 21*. Ed. by Neil Ashton, Anca Chereches, and David Lutz, pp. 313–332.
Dotlačil, Jakub (2012). "Binominal *each* as an anaphoric determiner: Compositional analysis". In: *Proceedings of Sinn und Bedeutung 16*. Ed. by Ana Aguilar, Anna Chernilovskaya, and Rick Nouwen. Cambridge, MA: MIT working papers in linguistics.
Dowty, David R. and Belinda Brodie (1984). "The semantics of "floated" quantifiers in a transformationless grammar". In: *Proceedings of the Third West Coast Conference on Formal Linguistics (WCCFL 3).* Ed. by Mark Cobler, Susannah McKaye, and Michael T. Westcoat. Stanford Linguistics Association. Stanford, CA, pp. 75–90.
Farkas, Donka F. (1997). "Dependent indefinites". In: *Empirical issues in formal syntax and semantics.* Ed. by Francis Corblin, Danièle Godard, and Jean-Marie Marandin. Bern: Peter Lang, pp. 243–268.
Gil, David (1982). "Distributive numerals". PhD thesis. Los Angeles, CA: UCLA.
Gillon, Brendan S. (1990). "Plural noun phrases and their readings: A reply to Lasersohn". In: *Linguistics and Philosophy* 13, pp. 477–485.
Grevisse, Maurice (1980). *Le bon usage: grammaire française, avec des remarques sur la langue française d'aujourd'hui.* Paris, France: Duculot.
Groenendijk, Jeroen, Theo Janssen, and Martin Stokhof, eds. (1984). *Truth, interpretation, information.* Dordrecht, Netherlands: Foris.
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.
Heim, Irene, Howard Lasnik, and Robert May (1991). "Reciprocity and plurality". In: *Linguistic Inquiry* 22.1, pp. 63–101.
Hendriks, Herman (1993). "Studied flexibility". PhD thesis. Amsterdam, Netherlands: University of Amsterdam.
Janke, Vikki and Ad Neeleman (2012). "Ascending and descending VPs in English". In: *Linguistic Inquiry* 43.1, pp. 151–190.
Junker, Marie-Odile (1995). *Syntaxe et sémantique des quantifieurs flottants* tous *et* chacun*: distributivité en sémantique conceptuelle.* Vol. 28. Geneva, Switzerland: Librairie Droz.
Kratzer, Angelika (1996). "Severing the external argument from its verb". In: *Phrase structure and the lexicon.* Ed. by J. Rooryck and L. Zaring. Dordrecht, Netherlands: Kluwer.
Kratzer, Angelika (2003). "The Event Argument and the Semantics of Verbs, chapter 3". Manuscript. Amherst: University of Massachusetts, Amherst, MA.

Kratzer, Angelika (2007). “On the plurality of verbs”. In: *Event structures in linguistic form and interpretation*. Ed. by Johannes Dölling, Tatjana Heyde-Zybatow, and Martin Schäfer. Berlin, Germany: de Gruyter, pp. 269–300.

Krifka, Manfred (1986). *Nominalreferenz und Zeitkonstitution. Zur Semantik von Massentermen, Pluraltermen und Aspektklassen*. Munich, Germany (published 1989): Fink.

Krifka, Manfred (1989). “Nominal reference, temporal constitution and quantification in event semantics”. In: *Semantics and contextual expression*. Ed. by Renate Bartsch, Johan van Benthem, and P. van Emde Boas. Dordrecht, Netherlands: Foris, pp. 75–115.

Krifka, Manfred (1992). “Thematic relations as links between nominal reference and temporal constitution”. In: *Lexical Matters*. Ed. by Ivan A. Sag and Anna Szabolcsi. Stanford, CA: CSLI Publications, pp. 29–53.

Krifka, Manfred (1998). “The origins of telicity”. In: *Events and grammar*. Ed. by Susan Rothstein. Dordrecht, Netherlands: Kluwer, pp. 197–235.

Krifka, Manfred (1999). “At Least Some Determiners Aren’t Determiners”. In: *The Semantics/Pragmatics Interface from Different Points of View*. Ed. by K. Turner. Amsterdam, Netherlands: Elsevier, pp. 257–291.

Landman, Fred (1996). “Plurality”. In: *Handbook of Contemporary Semantics*. Ed. by Shalom Lappin. Oxford, UK: Blackwell Publishing, pp. 425–457.

Landman, Fred (2000). *Events and plurality: The Jerusalem lectures*. Dordrecht, Netherlands: Kluwer.

Lasersohn, Peter (1989). “On the readings of plural noun phrases”. In: *Linguistic Inquiry* 20, pp. 130–134.

Lasersohn, Peter (1998). “Generalized distributivity operators”. In: *Linguistics and Philosophy* 21.1, pp. 83–93.

LaTerza, Chris (2014a). “Distributivity and plural anaphora”. PhD thesis. University of Maryland.

LaTerza, Chris (2014b). *Local plural anaphora as sub-event distributivity*. Presentation at the 32^{nd} West Coast Conference on Formal Linguistics (WCCFL 32).

Link, Godehard (1983). “The logical analysis of plurals and mass terms: A lattice-theoretical approach”. In: *Meaning, use and interpretation of language*. Ed. by Reiner Bäuerle, Christoph Schwarze, and Arnim von Stechow. Berlin, Germany: de Gruyter, pp. 303–323.

Link, Godehard (1986). “Je drei Äpfel - three apples each: Quantification and the German ’je’”. In: *Link, 1998*, pp. 117–132.

Link, Godehard (1987). “Generalized quantifiers and plurals”. In: *Generalized Quantifiers: Linguistic and Logical Approaches*. Ed. by Peter Gärdenfors. Reprinted in Link, 1998, Chapter 4. Dordrecht, Netherlands: Reidel, pp. 151–180.

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

Maling, Joan M. (1976). “Notes on quantifier-postposing”. In: *Linguistic Inquiry*, pp. 708–718.

Marantz, Alec (1993). “Implications of asymmetries in double object constructions”. In: *Theoretical aspects of Bantu grammar* 1, pp. 113–151.

Moltmann, Friederike (1997). *Parts and wholes in semantics*. Oxford, UK: Oxford University Press.

Oh, Sei-Rang (2001). “Distributivity in an event semantics”. In: *Proceedings of SALT XI*. Ed. by Rachel Hastings, Brendan Jackson, and Zsófia Zvolensky. Ithaca, NY: CLC Publications, pp. 326–345.

Oh, Sei-Rang (2006). “Plurality markers across languages”. PhD thesis. University of Massachusetts.

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

Petrova, Milena (2000). “The distributive marker *po* in Bulgarian”. Manuscript, USC.

Postal, Paul Martin (1974). *On raising*. Cambridge, MA: MIT Press.

Pylkkänen, Liina (2008). *Introducing arguments*. Cambridge, MA: MIT Press.

Roberts, Craige (1987). “Modal subordination, anaphora, and distributivity”. PhD thesis. Amherst, MA: University of Massachusetts.

Rothstein, Susan (2010). “Counting and the mass/count distinction”. In: *Journal of Semantics* 27.3, pp. 343–397.

Safir, Ken and Tim Stowell (1988). “Binominal each”. In: *Proceedings of the Eighteenth Conference of the North Eastern Linguistics Society (NELS 18)*. University of Massachusetts. Amherst, MA: GLSA (Graduate Linguistic Student Association), pp. 426–450.

Sakaguchi, Mari (1998). “Distributivity in English and Japanese”. PhD thesis. Los Angeles, CA: UCLA.

Scha, Remko (1981). “Distributive, collective and cumulative quantification”. In: *Formal methods in the study of language*. Ed. by Jeroen Groenendijk, Theo Janssen, and Martin Stokhof. Reprinted in Groenendijk, Janssen, and Stokhof, 1984. Amsterdam, Netherlands: Mathematical Center Tracts.

Schein, Barry (1986). “Event logic and the interpretation of plurals”. PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

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

Schwarzschild, Roger (1996). *Pluralities*. Dordrecht, Netherlands: Kluwer.

Sportiche, Dominique (1988). “A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure.” In: *Linguistic Inquiry* 19.3, p. 425. ISSN: 00243892.

Sutton, Melody (1993). “Binominal “each””. Master’s Thesis, UCLA.

Szabolcsi, Anna (2010). *Quantification*. Research Surveys in Linguistics. Cambridge University Press.

Tasmowski-De Ryck, Liliane and S. Paul Verluyten (1982). “Linguistic control of pronouns”. In: *Journal of Semantics* 1.4, pp. 323–346.

Tellier, Christine and Daniel Valois (1993). “Binominal *chacun* and pseudo-opacity”. In: *Linguistic Inquiry* 24.3, pp. 399–589.

Zimmermann, Malte (2002a). “A compositional analysis of anti-quantifiers as quantifiers”. In: *Proceedings of SALT XII*. Ed. by Brendan Jackson. Vol. 12, pp. 322–338.

Zimmermann, Malte (2002b). “Boys buying two sausages each: On the syntax and semantics of distance-distributivity”. PhD thesis. Amsterdam, Netherlands: University of Amsterdam.

Zucchi, Sandro and Michael White (2001). “Twigs, sequences and the temporal constitution of predicates”. In: *Linguistics and Philosophy* 24, pp. 187–222.

Zwarts, Joost (2005). “Prepositional Aspect and the Algebra of Paths”. In: *Linguistics and Philosophy* 28, pp. 699–740.