Dependent plurals and three levels of multiplicity

Serge Minor[1]

Abstract

The paper focuses on the semantics of distributivity, grammatical number, and cardinality predicates (numerals and modifiers like *several*). I argue that constructions involving so-called 'dependent plurals', i.e. plurals lacking cardinality predicates occurring in the scope of certain quantificational items such as *all* and *most* (e.g. *All the girls were wearing hats*), pose a challenge to familiar semantic frameworks that distinguish between two sources of multiplicity: mereological plurality and distributive quantification. I argue that dependent plural readings should be analysed as distinct both from cumulative readings and distributive readings, in the classical sense. I demonstrate how this can be accomplished in a semantic framework where expressions are evaluated relative to sets of assignments, or plural info states (van den Berg, in Stokhof and Torenvliet (eds) Proceedings of the 7th Amsterdam Colloquium, ILLC, University of Amsterdam, Amsterdam, 1990, in Dekker and Stokhof (eds) Proceedings of the 9th Amsterdam Colloquium, ILLC, University of Amsterdam, Amsterdam, 1994, Some aspects of the Internal Structure of Discourse. The Dynamics of Nominal Anaphora. PhD thesis, University of Amsterdam, 1996). The specific formal implementation is based on a modified version of Brasoveanu's (Structured nominal and modal reference. PhD thesis, Rutgers, The State University of New Jersey, 2007, Linguist Philos 31(2):129–209. https://doi.org/10.1007/s10988-008-9035-0, 2008) Plural Compositional DRT. In this framework we are able to distinguish between two types of distributivity: weak distributivity across the assignments in a single plural info state and strong distributivity across multiple info states. I argue that both of these types of distributivity play a role in the semantics of natural language, accounting for the contrasting properties of 'singular quantifiers', such as *each* and *every*, and 'plural quantifiers', such as *all* and *most*. The contrasting properties of bare plurals and plurals involving cardinality modifiers are analysed in terms of the distinction between state-level and assignment-level (mereological) plurality.

✉ Serge Minor
  sergey.minor@uit.no

1 Department of Language and Culture, UiT – The Arctic University of Norway, Tromsø, Norway

1 Introduction

Bare plurals in the scope of some quantificational noun phrases allow an interpretation which at first glance appears to be similar to that of singular indefinites:

(1) a. All the girls were wearing hats.
     b. All the girls were wearing a hat.

In a neutral context, sentence (1a) is interpreted as stating that each girl was wearing *a single hat*, i.e. its truth conditions appear to be very close to those of (1b). Crucially, sentence (1a) does not entail that each girl was wearing *more than one hat*. De Mey (1981) introduced the term *dependent plurals* for plural DPs that are used 'in what would appear to be a singular meaning' (see also Partee 1985; Roberts 1990; Zweig 2008, 2009; Ivlieva 2013, a.o.). I will adopt the term *licensor* to refer to the other member of the dependency, e.g. the DP *all the girls* in (1a), and *co-distributivity* as a pre-theoretic umbrella term for all readings that are compatible with a one-to-one correspondence between the set of individuals referred to (or quantified over) by the licensor DP and the set of individuals referred to by the dependent (cf. Sauerland 1994).[1] I will use the term *dependent plural* to refer specifically to plurals that occur in the scope of quantificational items (including floating quantifiers), and allow a co-distributive interpretation with their licensor.

This paper centres around three contrasts characteristic of dependent plural constructions. First, only a subset of quantificational DPs is able to license dependent plurals:

(2) a. Each girl was wearing hats.
     b. Each girl was wearing a hat.

Sentence (2a), in contrast to (1a), implies that each girl was wearing *more than one hat*, and thus differs sharply in its truth conditions from (2b) (cf. De Mey 1981; Zweig 2008, 2009; Kamp and Reyle 1993; Champollion 2010b for similar observations and discussion).

Second, only a subset of plural DPs can function as dependent plurals. Thus, in (3) the object DP contains the modifier *several*, and the sentence again entails that each girl was wearing *more than one hat* (cf. Zweig 2008, 2009):

(3) All the girls were wearing several hats.

Finally, sentence (3) contrasts with (4), where the subject is a definite plural:

(4) The girls were wearing several hats.

---

[1] I will reserve the more familiar term *cumulative interpretation* for cases like (4) below which involve two non-quantificational DPs.

Table 1 Availability of co-distributive readings (preliminary)

	Definite plurals	*all*-DPs	*each*-DPs
Bare plurals	✓	✓	*
Several-plurals	✓	*	*

Unlike (3), this sentence does not necessarily entail that each girl was wearing more than one hat. Instead, it allows for a cumulative interpretation, on which there exists some kind of correspondence between the girls and the hats with each girl wearing *at least one hat* (and each hat being worn by at least one girl).

Thus, on the one hand we have a three-way contrast between definite plurals vs *all*-DPs vs *each*-DPs. On the other hand we have a two-way contrast between bare plurals and indefinites with *several*. This is summarized in Table 1, where the columns represent the three types of DPs in the subject position and the rows represent the two types of DPs in the object position. The cells indicate the availability of a co-distributive interpretation in each configuration.

In the following we will see that each label in this table actually stands for a whole class of items which pattern together with respect to their semantic behaviour in the types of contexts discussed above. The aim of this paper is to provide a unified account of all the contrasts represented in Table 1. As we will see, the main challenge lies in accounting for the center column in Table 1, i.e. for the distinct semantic properties of plural quantificational DPs involving *all* (as well as *most*, *both*, *few* etc.) and their interaction with different types of plurals in their scope.[2]

The paper is structured as follows. Section 2 reviews the core empirical generalizations related to dependent plurals. Section 3 provides an overview of existing accounts of dependent plurals, and discusses some of the challenges they face. Section 4 introduces the core features of the semantic framework that I will use to couch my analysis. Sections 5 and 6 present a detailed analysis of the semantics of number features, numerals, distributivity operators and quantificational determiners, accounting for the core generalizations that govern the availability of co-distributive readings. Section 7 concludes the paper.

---

[2] Apart from nominal licensors dependent plurals can be licensed by various types of pluractional (quantificational, frequentitative, iterative) adverbials, which introduce a multiplicity of events/situations (cf. De Mey 1981):

  (i) a. John often wears loud neckties. (Roberts 1990, attributed to B. Partee)
      b. John always introduces his girlfriends to his mother.

On the most salient reading, sentence (ia) states that there is a set of frequently occurring events which involve John wearing a loud necktie. Crucially, this sentence does not imply that John necessarily wears more than one necktie on each occasion. Similarly, in (ib) the quantificational adverb *always* serves as a licensor for the plural possessive DP *his girlfriends*. On the most natural reading, this sentences states that on each relevant occasion John introduces his one current girlfriend to his mother. It isn't necessary for John to be in a relationship with more than one woman on every (or any) relevant occasion for this sentence to be judged true. This indicates that we are again dealing with a dependent plural reading. For reasons of space, in this paper I will restrict attention to dependent plural constructions with nominal licensors (see however Minor (2017) for an analysis of dependent plurals with adverbial licensors).

2 Core properties of dependent plurals

2.1 Dependent plurals versus singular indefinites

As illustrated in (1a), bare plurals can be interpreted co-distributively with higher-scoping *all*-DPs. This means that the multiplicity requirement normally associated with plurals is not applied with respect to each member of the set quantified over by the licensor DP. However, the question remains whether the opposite, 'singularity', requirement is applied distributively. In other words, at this point we don't know whether the correct interpretation of (1a) should be as in (5a) or (5b):

(5) a. 'Each girl was wearing *one* hat.'
     b. 'Each girl was wearing *one or more* hats.'

If dependent plurals have the same underlying semantics as singular indefinites, the interpretation in (5a) should be correct. Kamp and Reyle (1993) demonstrate that this is in fact not the case. The following is a slightly modified version of the example they discuss:

(6) All the students bought books that would keep them fully occupied during the next two weeks.

This example can be contrasted with that in (7):

(7) All the students bought a book that would keep them fully occupied during the next two weeks.

If dependent plurals had the same interpretation as singular indefinites we would expect these sentences to have the same truth conditions. And indeed, there are contexts where both of these sentences will be judged true, e.g. if each student bought one book such that this single book would keep them fully occupied for two weeks. However, (6) on its dependent plural reading would be judged true in a wider range of contexts than (7). Consider the following scenario: There are three students, Alan, George, and Miriam, who each bought one or more book. Specifically, Alan bought one book, George bought three, and Miriam four. In each case the book or books that the student bought would keep the buyer fully occupied for two weeks. In this scenario, (7) would be false because it is not true that each student bought a single book that would keep them occupied for two weeks. On the other hand (6) would be judged true in this situation.

This example strongly suggests that dependent plurals are in fact *number-neutral* with respect to each member of the licensor-set, i.e. the semantics of (6) is closer to that of (8) than to (7):

(8) All the students bought one or more books that would keep them fully occupied during the next two weeks.

The contrast in truth conditions between (6) and (7) demonstrates that dependent plurals are in one respect less restrictive than singular indefinites. However, it turns

out that in another way they are more restrictive. Consider the following examples, due to Zweig (2008, 2009) (cf. also De Mey 1981; Spector 2003 a.o. for similar observations):

(9) a. Ten students live in New York boroughs.
    b. Ten students live in a New York borough.

As Zweig (2008, 2009) points out, sentence (9a) can have a reading on which each student lives in just one New York borough. A similar reading is readily available for sentence (9b), on the low-scope interpretation of the indefinite object DP. However, these examples differ in their truth conditions: (9b) would be true in a scenario where all the students live in the same New York borough (e.g., Manhattan), while sentence (9a) would be judged false under this scenario. For sentence (9a) to be true, at least two of the students must live in different boroughs, i.e. more than one New York borough must be involved overall. Zweig (2008, 2009) calls this requirement associated with dependent plurals the *Multiplicity Condition*:

The Multiplicity Condition
More than one of the things referred to by a dependent plural must be involved overall.

Any adequate analysis of dependent plurals must account both for their number-neutrality on the 'fine-grained' level, i.e. with respect to each individual element in the set quantified over by the licensor, and the multiplicity requirement that they introduce on the 'global' level.

2.2 The licensors

Dependent plural readings of bare plural noun phrases can be licensed by a range of quantificational DPs, as the following examples demonstrate:

(10) a. All of the girls were wearing hats.
     b. Most of the girls were wearing hats.
     c. Both girls were wearing hats.

In all these examples the subject and the object can be interpreted co-distributively. For instance, (10b) will be judged true relative to a situation in which there was a majority of girls each wearing one or more hats, as long as they were wearing more than one hat overall.

Similarly, dependent plurals are licensed in the scope of floating *all* and *both*:

(11) a. The girls were all wearing hats.
     b. The girls were both wearing hats.

On the other hand, dependent plural reading are not licensed in the scope of DPs with the quantificational determiners (QDs) *each* and *every* (examples 12a–12b), and in the scope of floating *each* (example 13):

(12)   a. Each girl was wearing hats.

       b. Every girl was wearing hats.

(13)   The girls were each wearing hats.

These sentences entail that each girl was wearing more than one hat. Discussing similar data in Dutch, De Mey (1981) relates the contrast between examples like (10) and (12) to the number feature associated with the subject DP — plural quantificational DPs license dependent plurals, while singular quantificational DPs do not.

Given these facts, I will adopt the following generalization:

(14)   *Ban on Singular Licensors*
     DPs that involve complement NPs in the singular do not license dependent plurals.[3]

2.3 The dependents

The class of DPs that allow a dependent plural interpretation in the scope of plural quantificational DPs is quite broad and includes, apart from bare plural indefinites, possessive DPs (cf. De Mey 1981), definites (cf. Roberts 1990), and specific indefinites:

(15)   a. All the boys brought their fathers along.

       b. Most of these men married the ex-wives of their neighbours.

       c. Most of these groups live permanently along certain coastlines or bays and can therefore be spotted regularly.[4]

All these examples can be interpreted co-distributively. This indicates that the availability of the dependent plural interpretation is independent of the definiteness and specificity of the plural involved.

On the other hand, plural indefinites involving numerals and cardinality modifiers such as *several*, *a few*, *numerous*, *multiple* etc. cannot be interpreted co-distributively with plural quantificational DPs (cf. the observations in Zweig 2008, 2009):

---

[3] An alternative would be to formulate the restriction in terms of agreement, e.g. 'DPs that trigger singular agreement on the verb do not license dependent plurals'. This would successfully rule out dependent plural readings in examples like (12). However, there are cases which show that singular agreement is not in itself an impediment to dependent plural interpretation:

(1)   a. Everybody has cell phones these days.

       b. "Everyone has guns down there, it's like the wild West," Byrnes said.

In both of these examples the quantificational licensor in subject position triggers singular agreement on the verb, but a dependent plural interpretation is nevertheless available. Sentence (1a) will be judged true if every individual in a contextually specified set owns one or more cell phones. Similarly, Byrnes' claim in (1b), taken from the Corpus of Contemporary American English (COCA, cf. Davies 2008), is most naturally interpreted as stating that each individual in the relevant location has one or more guns, rather than asserting that each individual has at least two guns. Under the generalization in (14), these examples can be accounted for, assuming that synchronically the nominal roots in *everyone* and *everybody* do not function as independent NPs, and thus do not themselves carry a number feature.

[4] http://www.explore-the-big-island.com/swim-with-dolphins-in-hawaii.html.

(16)   a.   Both students made two mistakes.

         b.   All the students made several mistakes.

Sentence (16a) will be judged true if each of the two relevant students made two mistakes. However, it will be judged false if each of the two students made only one mistake, i.e. the students made two mistakes *in total*. Similarly, (16b) will be judged true iff each of the students made two or more mistakes. It is not sufficient for the total number of mistakes made by the students to be greater than one.[5]

These data lead to the following generalization:[6]

(17)   *Ban on Numerals*
       Plural quantificational DPs scoping over plurals involving numerals and cardinality modifiers cannot be interpreted co-distributively with these plurals.

As pointed out in the Introduction, plural non-quantificational DPs (e.g. plural definites and indefinites) are not subject to this restriction:

(18)   Ten/the students made twelve mistakes.

This sentence has a distributive reading under which each (relevant) student made ten mistakes. But it also has a non-scopal co-distributive reading, often referred to as a *cumulative* reading (cf. Scha 1984; Does 1993; Landman 2000; Beck and Sauerland 2001, among many others). Under this reading (18) will be judged true if there is a set of students *X*, a set of twelve mistakes *Y*, and each student in *X* made one or more mistakes in *Y*, and each mistake in *Y* was made by a student in *X*. Crucially, on this reading sentence (18) does not entail that each student made more than one mistake.

2.4 Interim summary

Table 2 summarizes the core observations. As before, checkmarks in the cells represent the availability of co-distributive interpretation for particular combinations of licensors and dependents. Specifically, we must distinguish between three types of

---

[5] Similar facts obtain for plurals in the scope of pluractional adverbs. Here, too, the co-distributive reading disappears if the plural contains a numeral or cardinal modifier:

(i)   John often wears several loud neckties.

In contrast to the examples discussed above in footnote 2, this sentence entails that John wears more than one loud necktie on each relevant occasion.

[6] It seems that some speakers allow a cumulative interpretation in examples like (i):

(i)   All the students made 15 mistakes.

This would suggest that for these speakers, *all* allows for an alternative, non-distributive, interpretation, perhaps as a homogeneity remover (cf. Križ 2016). Such cumulative readings appear to be more marginal in (iia) and impossible in (iib):

(ii)   a.   The students all made 15 mistakes.

        b.   Most of the students made 15 mistakes.

More research is needed to better understand the nature of this variation.

Table 2 Availability of co-distributive readings

	Non-Q	Plural Q	Singular Q
Plurals without cardinal modifiers	✓	✓	*
Plurals with cardinal modifiers	✓	*	*

licensors, which correspond to the columns in the table: non-quantificational plural DPs, plural quantificational DPs (and adverbial licensors) and singular quantificational DPs. We must also distinguish between two relevant types of dependents: plurals without numerals and other types of cardinal modifiers and plurals with such modifiers.[7] The contrast between the first two columns reflects the distinction between dependent plurals and cumulative predication. The contrast between the last two columns reflects the Ban on Singular Licensors. Finally, the contrast between the two cells in the middle column reflects the Ban on Numerals.

In the next section I briefly discuss three families of approaches to the semantics of dependent plural constructions, arguing that they all face significant challenges.

3 Previous approaches to dependent plurals

Existing approaches to dependent plurals can be broadly divided into two categories, based on whether they take dependent plural interpretations (e.g. 19a) to be a special case of distributive interpretations (as in 19b) or a special case of cumulative (or collective) interpretations (as in 19c).

(19)   a. All the girls were wearing hats.
      b. Every girl was wearing a hat.
      c. Five girls were wearing five hats.

I will refer to the first class of approaches as 'distributivity-based' and the latter as 'cumulativity-based'. The third approach that I will consider, which I refer to as the 'mixed' approach, assumes that dependent plural readings in examples like (19a) arise in the context of interpretations which combine the semantics of cumulativity and distributivity.

3.1 Distributivity-based approaches

The first class of accounts attempts to assimilate dependent plural constructions to garden variety distributive predication by assuming that the plural number feature on the dependent is somehow 'defective' or 'fake' in that it is not semantically interpreted as a plural.

---

[7] I will occasionally refer to plurals without cardinal modified as *bare plurals*, although as we have seen this class also includes plurals which are not traditionally considered bare, e.g. plural definites, specific indefinites with *certain*, etc.

The first account of this type was put forward by Partee (1975), who suggests that the number marking on the object in (19a) may result from the application of a syntactic agreement rule, which also determines the number agreement on the verb. Such a rule would apply to a syntactic structure such as (20), and result in plural marking both on the verb *was* and on the direct object *a hat*.

(20)   [All the girls]$_{NP}$ + [was wearing a hat]$_{VP}$.

Subsequently, accounts along these lines have been developed by Kamp and Reyle (1993) and Spector (2003).[8]

Distributive approaches are able to account for the co-distributive interpretation characteristic of dependent plurals. However, they accomplish this by assuming that the plural feature on the dependent is not interpreted, and consequently they fail to account for the overarching Multiplicity Condition associated with dependent plurals. For instance, in the following example the bare plural *New York boroughs* would be analysed as semantically singular/number-neutral. This would incorrectly predict that this sentence should be true in a situation where all the students live in the same New York borough:[9]

(21)   All the students live in New York boroughs.

3.2 Cumulativity-based approaches

A more common analysis of dependent plurals assumes that they are mereologically plural, i.e. refer to non-atomic sums/non-singleton sets of individuals. The co-distributive relation between the dependent and the licensor is then analysed as an instance of *cumulative predication*. Zweig (2008, 2009) provides the most comprehensive exposition of this type of analysis (see also Bosveld-de Smet 1998; Swart 2006; Beck 2000 for similar proposals).[10] Zweig's formalization is based on Landman's (2000) theory of plurality, with the dependent plural reading of sentence (22a) represented as in (22b):

(22)   a.   Five boys flew kites.
       b. $\exists e \exists X \exists Y[|X| = 5 \;\wedge\; {}^{*}\text{BOY}(X) \;\wedge\; |Y| > 1 \;\wedge\; {}^{*}\text{KITE}(Y) \;\wedge\; {}^{*}\text{FLEW}(e) \;\wedge\;$
            ${}^{*}\text{AGENT}(e)(X) \;\wedge\; {}^{*}\text{THEME}(e)(Y)]$

Here, capital letters stand for variables which range over both atomic and non-atomic individuals. The star * represents Link's (1983) pluralization operator when it

---

[8] Kamp and Reyle (1993) propose a rule according to which bare plurals can introduce number-neutral discourse referents in the context of another plural noun phrase. Spector (2003) proposes an analysis of dependent readings of French *des*-indefinites according to which they carry a morphological, but not semantic, plural feature. The morphological plural is taken to be licensed in the scope of an element carrying a semantic plural feature.

[9] Spector (2003) seems to be aware of this problem, but does not provide a solution within his account.

[10] De Mey (1981) assumes that both the licensor and the dependent are interpreted collectively, and thus 'the dependent reading is a special subcase of the collective–collective reading' (cf. also Scha 1984). This proposal is conceptually similar to cumulativity-based analyses in that dependent plurals are taken to be semantically plural, but does not invoke the formal mechanism of cumulative predication.

combines with a one-place predicates (e.g. FLEW), and the cumulative operator when it combines with two-place predicates (e.g. THEME). These can be defined as follows (cf. Krifka 1989; Sauerland 1998; Sternefeld 1998, a.o.):

(23)   a. $*P(x) \leftrightarrow \exists x_1, \ldots, x_n. \, x_1 \oplus \ldots \oplus x_n = x \, \wedge \, P(x_1) \, \wedge \, \ldots \, \wedge \, P(x_n)$
       b. $*Q(x)(y) \quad \leftrightarrow \quad \exists x_1, \ldots, x_n. \, \exists y_1, \ldots, y_n. \, x_1 \oplus \ldots \oplus x_n = x \, \wedge \\ y_1 \oplus \ldots \oplus y_n = y \, \wedge \, Q(x_1)(y_1) \, \wedge \, \ldots \, \wedge \, Q(x_n)(y_n),$
            where P is a one-place predicate and Q is a two-place predicate.

According to these definitions, (22b) will be true in a wide range of scenarios where there is a flying event whose total sum of agents is a sum of five boys, and whose total sum of themes is a sum of more than one kite. In particular, this interpretation is compatible with a co-distributive scenario where each of the five boys flew one kite.

Zweig (2008, 2009) integrates this account of dependent plurals with the independently supported assumption that bare plurals in general are underlyingly number-neutral (cf. Krifka 2004; Sauerland et al. 2005; Spector 2007). The multiplicity requirement associated with the dependent (represented by the conjunct $|Y| > 1$ in 22b) is analysed as a scalar implicature which arises in competition with the corresponding singular indefinite. Zweig's approach is thus able to account for the fact that in downward entailing contexts the multiplicity implicature associated with dependent plurals does not arise. Consider the following example from Zweig (2009):

(24)   John denied that the carpenters built rafts.

Sentence (24) is interpreted as 'John claimed that the carpenters did not build any rafts', rather than 'John claimed that the carpenters did not build more than one raft'. Suppose John is testifying in court. If the carpenter did build exactly one raft, and (24) is true, then it must be the case that John gave a false testimony.

Zweig's (2008, 2009) analysis successfully accounts both for the co-distributive interpretation and the Multiplicity Condition associated with dependent plurals, as well as for the status of the Multiplicity Condition as an implicature. However, Zweig does not provide an account of the contrast between singular and plural quantificational DPs in their role as licensors (cf. the Ban on Singular Licensors in 14), or the contrast between bare plurals and plurals with numerals and cardinality modifiers in their role as dependents (cf. the Ban on Numerals in 17).[11]

The Ban on Numerals poses an especially pertinent challenge for cumulativity-based analyses of dependent plurals. Indeed, if plural quantificational DPs can license dependent plurals and the semantics of such constructions is cumulative, then we *a priori* expect cumulative interpretations to be possible between plural quantificational DPs and DPs involving numerals and other cardinal modifiers. After all, the most garden variety examples of cumulative interpretations involve DPs with numerals, as in (25a). In fact, as we have seen, examples like (25b) cannot be interpreted co-distributively:

---

[11] In fact, Zweig (2008) does put forward an account of the contrast between dependent plural and cumulative readings under *all*, however it does not deliver the correct results as shown in Champollion (2010b: 204–205).

(25) a. Ten safary participants saw thirty zebras.

   b. All the safari participants saw thirty zebras.

Clearly, something more needs to be said about the semantics of quantificational items like *all* and *most*. I will discuss two solutions to this problem, both of which can be viewed as extensions of Zweig's proposal. The first was proposed by Champollion (2010b), and later revised in Champollion (2017). The second is due to Ivlieva (2013). I will discuss them in turn.

Champollion (2010b, 2017) proposes that the semantic contribution of *all* amounts to a presupposition formulated in terms of *Stratified Reference*. The following is the interpretation of the prenominal (i.e. non-floating) *all* in the agent position as given in Champollion (2017), and the definition of the relevant type of Stratified Reference:[12]

(26) a. $[\![\text{all}_{agent}]\!] = \lambda y. \lambda V_{\langle vt \rangle}. \lambda e : \text{SR}_{*agent, \varepsilon(*agent(e))}(V). [V(e) \wedge *\text{agent}(e) = y]$

   b. $\text{SR}_{*agent, \varepsilon(*agent(e))}(V) \overset{\text{def}}{=} \forall e'. \ V(e') \rightarrow e' \in *\lambda e''. \ (V(e'') \wedge \varepsilon(*\text{agent}(e))(*\text{agent}(e'')))$

   (*V* has stratified reference along the *agent* dimension with granularity $\varepsilon(*\text{agent}(e))$ iff every event in *V* can be divided into one or more events which are also in *V* and whose agents are each small in number compared to the agent of *e*.)

The idea is that DPs involving *all* only combine with predicates that are distributive down to (relatively) small sub-groups of participants (cf. Dobrovie-Sorin 2014; Kuhn 2020). More precisely, a DP involving *all* in the agent position (or the floating *all* which applies to the agent) only combines with an event predicate *P* if any event in *P* can be represented as a sum of events which are also in *P* and whose agents are relatively small sums of individuals.[13] In the simplest case, a DP involving *all* combines with a lexically distributive predicate such as *smile*:

(27) All the boys smiled.

The predicate *smile* satisfies the condition in (26b) because any smiling event can be divided into smaller smiling sub-events whose agents are minimal, i.e. atomic.

Consider now example (25b), which does not have a cumulative reading. In this example, the subject combines with the following event predicate:

(28) $\lambda e[*\text{see}(e) \wedge *\text{zebra}(*\text{th}(e)) \wedge |*\text{th}(e)| = 30]$

This event predicate does not satisfy the Stratified Reference presupposition: not every event of seeing thirty zebras is the sum of one or more events each of which

---

[12] Note that following Landman (2000), Champollion assumes that theta-roles are (partial) functions of type $\langle ve \rangle$, which map an event to the individual that bears a certain role in that event.

[13] Champollion (2010b) assigns a slightly different presupposition to *all*, which requires the event predicate to have stratified reference with granularity Atom, i.e. every event in the denotation of the predicate must be dividable into one or more events which are also in the denotation of that predicate and whose agents (themes, etc.) are *atomic*. This difference is not important for the current discussion.

has a relatively small agent and is itself an event of seeing thirty zebras. For instance, suppose there are three safari participants in the model, who each saw 10 zebras (and all the zebras were different). Then there is a cumulative event e of them seeing 30 zebras, i.e. (28) is true of e. However, event e cannot be divided into sub-events which are also events of seeing 30 zebras, whose agents have a cardinality smaller than 3. In fact there are no sub-events in e which are also events of seeing 30 zebras and which are distinct from e.

Thus, the unavailability of a cumulative reading in (25b) is explained by the fact that the denotation of the VP does not satisfy the presupposition imposed by *all*.[14]

What about dependent plurals? Consider the following example:

(29)   All the safari participants saw zebras.

Champollion follows Zweig (2008, 2009) in analysing dependent plurals as semantically number-neutral, with the multiplicity requirement added as a scalar implicature at a stage preceding the existential closure of the event variable. The key to Champollion's account of examples like (29) is the assumption that the presupposition associated with *all* is checked against the denotation of the VP without the added multiplicity implicature, e.g. the presupposition is checked *before* the implicature is added. Sentence (29) is interpreted in the following way:

(30) $\exists e[*\text{see}(e) \wedge *\text{ag}(e) = \bigoplus \text{safari.participant} \wedge *\text{zebra}(*\text{th}(e)) \wedge |*\text{th}(e)| > 1]$
Presupposition: $\text{SR}_{*agent, \varepsilon(*agent(e))}(\lambda e[*\text{see}(e) \wedge *\text{zebra}(*\text{th}(e))])$
(True iff every event in which one or more zebras are seen can be divided into sub-events such that each sub-event is an event of seeing one or more zebras whose agent is small compared to the agent of e.)

According to these truth conditions, sentence (29) asserts the existence of a cumulative event whose cumulative agent is the maximal sum of safari participants and whose cumulative theme is a non-atomic sum of zebras. The condition $|*\text{th}(e)| > 1$ in (30) represents the multiplicity implicature, and is not included in the denotation of the VP when the presupposition is checked. In the absence of the multiplicity requirement, the event predicate denoted by the VP satisfies the Stratified Reference presupposition: any event of seeing one or more zebra can be divided into sub-events of seeing one or more zebra involving minimal atomic agents, which would count as small compared to the plural, non-atomic agent in (29). Thus sentence (29) is predicted to have a cumulative interpretation.

Thus, Champollion's (2010b, 2017) presuppositional approach to the semantics of *all* accounts for the lack of cumulative readings between *all*-DPs and DPs involving unmodified numerals without giving up the central tenet of the mereological approach—that dependent plurality is essentially a sub-type of cumulativity.[15]

---

[14] Sentence (25b) does have a distributive reading on which each of the safari participants saw 30 zebras. This reading is derived by attaching a silent distributivity operator to the predicate. This yields a new predicate that is distributive down to atomic agents and thus satisfies the presupposition imposed by *all*, see Champollion (2017: 263–264).

[15] Champollion's (2010b, 2017) account of the contrast between *all* and *every/each* is less clearly motivated. In Champollion's theory, *every* and *each* encode stratified reference down to sub-events involving atomic

However, the presupposition that Champollion ([2010b](#), [2017](#)) ascribes to *all* (cf. [26](#)) turns out to be too strong, ruling out dependent plural readings in certain contexts when they are in fact available. Consider the following example:

(31)   All the students handed in their papers.

If the pronoun is interpreted as a variable bound by the subject, this sentence has a dependent plural interpretation whereby each student handed in their paper (or papers), and more than one paper was handed in overall. This is unexpected under Champollion's account because the predicate that the subject DP combines with does not satisfy the Stratified Reference presupposition imposed by *all*:

(32) $\lambda e.\ [\text{*hand-in}(e) \wedge \text{*th}(e) = \bigoplus(\lambda y.[\text{*paper}(y) \wedge \text{*POSS}(y)(\bigoplus \text{student})])]$

This predicate identifies a set of handing-in events whose theme is the sum of the students' papers. Suppose that we have an event where each student handed in their own paper. Then there are *no* sub-events within (and distinct from) that event which are also events of handing in the sum of all the student' papers. This entails that there is no way to represent that event as a sum of sub-events of handing in the students' papers whose agents are small relative to the whole sum of students, i.e. the presupposition in ([26](#)) is not satisfied and the sentence in ([31](#)) is predicted to lack a dependent plural reading, contrary to fact.

Note however, that the binding relation between the agent of the event and the possessor of the theme is not represented in ([32](#)). We may wonder if solving this problem may also help resolve the issue with the lack of Stratified Reference. In turns out that it doesn't.

One way to make *all*'s presupposition sensitive to the binding relation between the agent and the possessor is to re-define the semantics of *all* in such a way that it takes a relation between individuals and events rather than an event predicate as it's second argument:

(33)   a. $[\![\text{all}]\!] = \lambda y.\lambda V_{\langle e,\, vt \rangle}.\lambda e : \text{SR}_\varepsilon(V).[V(y)(e)]$

b. $\text{SR}_\varepsilon(V) \stackrel{\text{def}}{=} \forall x'.\forall e'.\ V(x')(e') \rightarrow \langle x', e' \rangle \in \text{*}\lambda x''.\lambda e''.\ (V(e'') \wedge \varepsilon(y)(x''))$

When applied to a relation between individuals and events V, Stratified Reference as defined in ([33b](#)) requires that any pair of individuals and events $\langle x, e \rangle$ in V must be decomposable into a set of pairs $\langle x_1, e_1 \rangle$, $\langle x_2, e_2 \rangle$, $\ldots$, $\langle x_n, e_n \rangle$, such that:

(34)   a) $x_1 \oplus x_2 \oplus \ldots \oplus x_n = x$ and $e_1 \oplus e_2 \oplus \ldots \oplus e_n = e$,
b) $\langle x_1, e_1 \rangle \in V$, $\langle x_2, e_2 \rangle \in V$, $\ldots$, $\langle x_n, e_n \rangle \in V$, and
c) $x_1, x_2, \ldots, x_n$ are small relative to x.

Footnote 15 continued
participants. To account for the fact that these items block dependent plural readings, Champollion stipulates that the multiplicity implicature associated with bare plurals must be calculated ('is trapped') in their distributive scope, leaving open the question of how to motivate this assumption.

Assuming that the subject DP in (31) quantifier-raises from its base thematic position, it will combine with the following individual-event relation, which now includes the binding relation between the agent of the handing-in event and the possessor of the papers:

(35) $\lambda x.\lambda e.\ [\text{*hand-in}(e) \wedge \text{*ag}(e) = x \wedge \text{*th}(e) = \bigoplus(\lambda y.[\text{*paper}(y) \wedge \text{*POSS}(y)(x)])]$

At first glance, this relation has a better chance of satisfying the Stratified Reference presupposition than the predicate in (32). Indeed, if we take a sum of students and a sum-event where each of the students handed in their own paper, this pair will satisfy the conditions listed in (34), i.e. we can decompose this pair into a set of pairs of the form $\langle s, e \rangle$, where s is an individual student and e is the event of that student handing in their paper. However, there are other pairs of individuals and events in (35) that do not satisfy these requirements. Suppose we have a sum of students S, and each student in S handed in a paper by another student in S. Let the sum of these events be E. Then $\langle S, E \rangle$ is in (35) because E is a sum of handing-in events whose cumulative agent is S and whose cumulative theme is the sum of papers that stand in a (cumulative) possessive relation with S. The pair $\langle S, E \rangle$ would not, however, satisfy the conditions in (34), because the individual handing-in events in E (whose agents are the individual students) are not in fact events of handing in one's own paper(s). This means that the relation in (35) does not satisfy the Stratified Reference presupposition in (33b), which again incorrectly rules out the dependent plural reading in (31).[16]

I will now move on to an alternative account of the Ban on Numerals proposed by Ivlieva (2013, 2020). Ivlieva's approach is more successful in accounting for dependent plural readings in examples like (31). I will argue, however, that it too faces problems in delineating the precise range of available co-distributive interpretations.

3.3 Mixed approach

Building on the analysis of dependent plurals in Zweig (2009), Ivlieva (2013, 2020) proposes an interpretation for *all* that explicitly combines the semantics of cumulativity and distributivity:

---

[16] Ivlieva (2020) provides another argument that Champollion's (2010b, 2017) theory is too restrictive based on examples involving so-called *mixed predicates*, i.e. predicates that have collective readings which can be distinguished from their distributive readings, such as *eat a pizza* (cf. Link 1983; Scha 1984; Roberts 1990; Dowty 1987; Winter 2000, 2001, a.o.). Ivlieva notes that examples like (i) allow for a dependent plural interpretation:

(i) All the boys ate pizzas.

This sentence will be judged true if, for example, each of the boys ate a single pizza. However, it turns out that the event predicate that *all* combines with in (i) does not satisfy the Stratified Reference presupposition:

(ii) $\lambda e[\text{*eat}(e) \wedge \text{*pizza}(\text{*th}(e))]$

Take an event of several boys eating a single pizza together. This event is itself an event of eating one or more pizzas, but cannot be represented as a sum of smaller events of eating one or more pizzas. This means that the predicate in (ii) lacks Stratified Reference, and the sentence in (i) is incorrectly predicted to lack a dependent plural interpretation.

(36) $[\![\text{all the boys}]\!] = \lambda P_{\langle e,vt \rangle}.\lambda e.\ P(\sigma x.*boy(x))(e)\ \wedge\ \forall y\ [y\ \leq\ \sigma x.*boy(x)\ \wedge$
$atom(y)\ \rightarrow\ \exists e'\ [e' \leq e\ \wedge\ P(y)(e')]]]$

The interpretation of (37a) before event closure is (37b):

(37) a. All the boys flew kites.

b. $\lambda e.\exists z\ [*kite(z)\ \wedge\ *fly(\sigma x.*boy(x))(z)(e)\ \wedge\ \forall y\ [y\ \leq\ \sigma x.*boy(x)\ \wedge$
$atom(y)\ \rightarrow\ \exists e'\ [e' \leq e \wedge \exists w\ [*kite(w)\ \wedge\ *fly(y)(w)(e')]]]]$

This predicate will be true of sums of flying events such that its cumulative agent is the sum of all the boys and its cumulative theme is a sum of kites, and each boy is the agent of a flying sub-event whose theme is a sum of kites. At this point the event predicate in (37b) is compared to the alternative in (38), obtained by replacing the bare plural object with a singular indefinite:

(38) $\lambda e.\exists z\ [*kite(z) \wedge atom(z) \wedge *fly(\sigma x.*boy(x))(z)(e) \wedge \forall y\ [y \leq \sigma x.*boy(x) \wedge$
$atom(y)\ \rightarrow\ \exists e'\ [e' \leq e\ \wedge\ \exists w\ [*kite(w)\ \wedge\ atom(w)\ \wedge\ *fly(y)(w)(e')]]]]$

The alternative in (38) is stronger than (37): any event where each of the boys flew the same kite is an event of the boys flying one or more kites, with each boy flying one or more kites. Hence, the alternative in (38) is negated, giving rise to a strengthened interpretation, which after event closure gives (39) as the truth conditions of sentence (37a):

(39) $\exists e.\exists z\ [*kite(z) \wedge \neg atom(z) \wedge *fly(\sigma x.*boy(x))(z)(e) \wedge \forall y\ [y \leq \sigma x.*boy(x)\ \wedge$
$atom(y)\ \rightarrow\ \exists e'\ [e' \leq e\ \wedge\ \exists w\ [*kite(w)\ \wedge\ *fly(y)(w)(e')]]]]$

These truth conditions adequately capture the dependent plural reading of (37a): (37a) is predicted to be true iff there is a flying event whose cumulative subject is the maximal sum of boys and whose cumulative theme is a non-atomic sum of kites, and for each individual boy there is a sub-event of that boy flying one or more kites, i.e. each boy must fly one or more kites, and more than one kite must be flown overall.[17]

This analysis correctly predicts that in contrast to (37a), sentence (40a) will not have a co-distributive reading:

(40) a. All the boys flew 10 kites.

b. $\exists e.\exists z\ [*kite(z)\ \wedge\ |z|\ =\ 10\ \wedge\ *fly(\sigma x.*boy(x))(z)(e)\ \wedge\ \forall y\ [y \leq$
$\sigma x.*boy(x)\ \wedge\ atom(y)\ \rightarrow\ \exists e'\ [e' \leq e\ \wedge\ \exists w\ [*kite(w)\ \wedge\ |w|\ =$
$10\ \wedge\ *fly(y)(w)(e')]]]]$

According to the truth conditions in (40b), sentence (40a) will be true iff there is a flying event whose cumulative agent is the sum of all the boys and whose cumulative

---

[17] Like Zweig (2008, 2009), Ivlieva assumes that scalar implicatures can be calculated at different levels of the structure. However, she rejects the principle, adopted by Zweig, that only the strongest resulting interpretation is chosen as the meaning of the sentence. Consequently, her system generates two further readings for sentence (37a): a distributive reading, whereby each boy flew more than one kite (in case the implicature is calculated below the subject), and a reading whereby there is an event of the boys flying one or more kites, but there is no event of the boys flying the same kite (in case the implicature is calculated above the event closure).

theme is the sum of ten kites, and each boy is the agent of a flying sub-event whose theme is, again, a sum of ten kites. In other words, sentence (40a) is predicted to be true only if *each of the boys* flew ten kites. A co-distributive interpretation is blocked thanks to the presence of a distributive component in the semantics of *all*.

This analysis is also able to handle examples like (31), repeated here:

(41) All the students handed in their papers.

This example is predicted to be true iff there is an event of all the students cumulatively handing in all their papers, which consists of sub-events where each student *x* hands in *x*'s paper(s). This appears to be correct.[18] There is however a major empirical challenge to both Champollion's and Ivlieva's theories, to which I turn in the next section.

3.4 The challenge of modified numerals

As noted above, a major challenge for cumulativity-based approaches to dependent plurals is to account for the Ban on Numerals, i.e. for the unavailability of co-distributive readings between (higher scoping) DPs involving plural quantifiers like *all* and (lower scoping) DPs involving numerals and other cardinality modifiers. Champollion (2010b, 2017) tackles this problem by equipping *all* with a Stratified Reference presupposition, while Ivlieva (2013, 2020) adds a distributive component to *all*'s assertive semantics. These solutions succeed in ruling out co-distributive readings in examples like (25b) and (40a), which involve *all*-DPs scoping over noun phrases with unmodified numerals. However, both of these approaches make incorrect predictions when it comes to sentences like (42):

(42) All the students made fewer than 10 mistakes.

The event predicate in (43) satisfies the Stratified Reference presupposition: any large event of making fewer than ten mistakes can be represented as a sum of multiple sub-events of making fewer than ten mistakes with minimal (atomic) agents.

(43) $\lambda e[\text{*make}(e) \wedge \text{*mistake}(\text{*th}(e)) \wedge |\text{*th}(e)| < 10]$

Thus, Champollion's approach predicts that *fewer than n* DPs should pattern with bare plurals in allowing cumulative readings with *all*-DPs. Sentence (42) should then have a reading on which each student made one or more mistakes and fewer than 10 mistakes were made overall.

---

[18] Ivlieva's (2013) account is also able to capture the contrast between *all* and *every*. Whereas the semantics of *all* is taken to include both a cumulative and a distributive component, the semantics of *every* is purely distributive:

(i) a. $[\![\text{every boy}]\!] = \lambda P_{\langle e, vt \rangle}.\lambda e.\forall y \ [\text{*boy}(y) \wedge atom(y) \ \rightarrow \ \exists e' \ [e' \leq e \ \wedge \ P(y)(e')]]$
     b. Every boy flew kites.

Relying on a number of specific assumptions about the mechanics of implicature calculation, Ivlieva (2013) is able to derive the fact that sentences like (ib) only allow a distributive reading (i.e. each boy must have flown *more than one* kite).

Ivlieva makes a similar prediction. Her system generates the following interpretation for sentence (42):

(44) $\exists e.\exists z\ [*mistake(z)\ \wedge\ |z| < 10\ \wedge\ *make(\sigma x.*student(x))(z)(e)\ \wedge\ \forall y\ [y \leq \sigma x.*student(x)\ \wedge\ atom(y)\ \rightarrow\ \exists e'\ [e' \leq e\ \wedge\ \exists w\ [*mistake(w)\ \wedge\ |w| < 10\ \wedge\ *make(y)(w)(e')]]]]$

According to these truth conditions, sentence (42) will be judged true iff there is an event where the students cumulatively made fewer than 10 mistakes, and each individual student made fewer than 10 mistakes. Once again, sentence (42) is predicted to have a reading where *fewer than 10* specifies the total number of mistakes made by the students. These predications are not borne out.

To see this clearly, consider the following scenario: A student competition is being held. The students are divided into teams, and each student is asked to spell several words. For each team, the number of mistakes made by the students on that team is summed up, giving the total sum of mistakes for the whole team. To succeed a team must make fewer than 10 mistakes in total. Now, suppose someone points at a particular team and asks the question in (45):

(45) Did that team succeed?

Now consider the sentence in (46):

(46) Well, all the students on that team made fewer than 10 mistakes.

Intuitively, sentence (46) cannot function as an informative answer to the question in (45) — it cannot be understood as providing the information about the *total* number of mistakes the team made, it can only be read as stating that each of the students made fewer than 10 mistakes. This indicates that *fewer than n* DPs pattern with DPs involving unmodified numerals in that they do not allow cumulative readings in the scope of *all*, contra Champollion's and Ivlieva's predications.[19]

---

[19] The scale of the challenge posed by modified numerals becomes even more apparent if we consider expressions like *one or more*, *one or several* or *a certain number of*, e.g.:

(i) a. The students are shown one or more French movies.
     b. The students are all shown one or more French movies.

On the standard approach to numerals as predicates of sums, *French movies* is coextensive with *one or more French movies*. Then, any theory that adopts such an approach and, furthermore, assigns a cumulative interpretation to (ia), should allow one for (ib) as well, i.e. it will predict that in (ib) *one or more* can be understood as specifying the total number of French movies watched by the students. Intuitively this is incorrect — the numeral in (ib) must be understood distributively, as specifying the number of French movies watched by each student. From the point of view of truth conditions, the cumulative and distributive interpretations of (ib) are equivalent. Nevertheless, there is evidence that (ib) can in fact be only understood distributively. Suppose groups of students are participating in a psychological experiment which examines the effect of watching French movies on cognitive processes. In the experiment, each student is shown one (and only one) French movie, and then subjected to a serious of cognitive tests. In some groups, all the students are shown the same movie, in other groups the students are shown different movies. In this context, (ia) (on its cumulative reading) adequately describes a feature of the experiment. Sentence (ib), on the other hand, is misleading, because it suggests that there may be a student who is shown more than one French movie. This indicates that (ib) must indeed be interpreted distributively.

In a recent paper, Kuhn ([2020](#)) has proposed to solve this puzzle by augmenting Champollion's account with a novel approach to the semantics of measurement predicates, including numerals. The idea is that measurement predicates introduce a special type of event which relates the entity being measured ($stuff(e)$) to the corresponding measurement ($\mu(e)$). A VP containing a noun phrase with a numeral is taken to denote a predicate of events which are sums of an 'action event' (e.g. *make*-events) and a measurement (*number*) event, e.g.:

(47) $\llbracket$make fewer than 10 mistakes$\rrbracket =$
$\lambda e. \exists e', e'' [e = e' \oplus e'' \wedge {}^{*}make(e') \wedge {}^{*}mistake({}^{*}\text{th}(e')) \wedge {}^{*}number(e'') \wedge$
${}^{*}\text{th}(e') = {}^{*}stuff(e'') \wedge \exists n < 10 [\mu(e'') = n]]$

Kuhn argues that thanks to the inclusion of the measurement event, the predicate in (47) no longer has Stratified Reference. Suppose we take a sum consisting of (a) an event e' where three students cumulatively made 7 mistakes (student A made 1 mistake, student B made 2 mistakes, and student C made 4 mistakes), and (b) a *number*-event e'' that measures the total number of mistakes made in e' ($\mu(e'') = 7$). Then, the event predicate in (47) is true of $e = e' \oplus e''$, however there is no way to represent this event as a sum of sub-events which also satisfy (47) and have small agents. Take the sub-events in e' with minimal agents (students A, B, and C) summed up with the corresponding measurement events ($e_1 = e'_1 \oplus e''_1$, $e_2 = e'_2 \oplus e''_2$, $e_3 = e'_3 \oplus e''_3$), these sums will also be witnesses of the predicate in (47). However, it turns our that their sum ($e_1 \oplus e_2 \oplus e_3$) is not actually equal to e. This is because the sum of the small measurement events is not the same as the measurement corresponding to the big event ($e''_1 \oplus e''_2 \oplus e''_3 \neq e''$). Indeed, the cumulative measure of $e''_1 \oplus e''_2 \oplus e''_3$ is the mereological sum of three degrees ($\mu(e''_1 \oplus e''_2 \oplus e''_3) = 1 \oplus 2 \oplus 4$), while the measure of e'' is an atomic degree ($\mu(e'') = 7$). Since the predicate in (47) does not have Stratified Reference, it will not be directly compatible with a subject involving *all*, thus ruling out a co-distributive interpretation in (42).

While clearly a step forward, Kuhn's analysis does not fully solve the puzzle of modified numerals. The system still overgenerates co-distributive readings in contexts where they are not in fact possible. Consider first how the predicate in (47) is derived compositionally. Kuhn ([2020](#)) does not discuss the details, but it seems fair to assume that the function of constructing the sum event is part of the semantics of the noun phrase:[20]

(48) $\llbracket$fewer than 10 mistakes$_{theme}\rrbracket =$
$\lambda V_{(vt)}.\lambda e. \exists e', e'' [e = e' \oplus e'' \wedge V(e') \wedge {}^{*}mistake({}^{*}\text{th}(e')) \wedge {}^{*}number(e'') \wedge$
${}^{*}\text{th}(e') = {}^{*}stuff(e'') \wedge \exists n < 10 [\mu(e'') = n]]$

Consider now the following example:

---

[20] In Champollion's ([2010b](#), [2017](#)) system thematic roles are introduced by separate heads of type ve. (48) can be generalized as follows:

(i) $\llbracket$fewer than 10 mistakes$\rrbracket =$
$\lambda \theta_{(ve)}.\lambda V_{(vt)}.\lambda e. \exists e', e'' [e = e' \oplus e'' \wedge V(e') \wedge {}^{*}mistake(\theta(e')) \wedge {}^{*}number(e'') \wedge \theta(e') =$
${}^{*}stuff(e'') \wedge \exists n < 10 [\mu(e'') = n]]$

(49) All the kids received gift bags containing fewer than 10 candies.

In this sentence, the modified numeral (*fewer than 10*) can only be understood as specifying the number of candies received by each individual kid, not the total number of candies in all the gift bags received by the kids. Assuming a standard semantic analysis of restrictive relative clauses, this fact is problematic for Kuhn's account. The reason is that the measurement event introduced by the numeral will be summed up with the *containing*-event inside the relative clause, and will not influence the relevant properties of the higher event predicate:

(50) ⟦receive gift bags containing fewer than 10 candies⟧ =
λe [*receive(e) ∧ *bag(*th(e)) ∧ $\exists e_2$ [*loc(e_2) = *th(e) ∧ $\exists e', e''$ [e_2 = $e' \oplus e''$ ∧ *contain(e') ∧ *candy(*th(e')) ∧ *number(e'') ∧ *th(e') = *stuff(e'') ∧ $\exists n < 10$ [$\mu(e'') = n$]]]]

This predicate has Stratified Reference despite the presence of a measurement event, because this measurement event is not summed up with the *receiving*-event variable, but rather with the *containing*-event variable introduced by the verb in the relative clause. Any cumulative event that satisfies (50) (i.e. an event of receiving gift bags which cumulatively contain fewer than 10 candies) will be a sum of smaller sub-events which also satisfy (50) and have minimal agents.[21]

Thus, unless some additional mechanism is introduced that would 'propagate' the measurement event up from inside the relative clause in examples like (49), Kuhn's analysis does not fully solve the problem that modified numerals pose to cumulativity-based approaches to dependent plurals.[22]

---

[21] I am disregarding collective readings here, which pose an independent problem for the Stratified Reference account, cf. footnote 16.

[22] There is another, more technical, issue with Kuhn's (2020) proposal. Take the example in (ia) with the event predicate in (ib):

(i)   a. All the students watched fewer than 2 movies.

      b. $\lambda e. \exists e', e''$ [$e = e' \oplus e''$ ∧ *watch(e') ∧ *movie(*th(e)) ∧ *number(e'') ∧ *th(e') = *stuff(e'') ∧ $\exists n < 2$ [$\mu(e'') = n$]]

Since the only natural number below 2 is 1, the predicate in (ib) will only include events where all the agents watched the same movie. (If we decide to include a bottom element into the domain of entities, as argued by Bylinina and Nouwen (2018), we can replace *fewer than 2* with *fewer than 2 and more than 0*.) It then follows, that if we take a cumulative event satisfying (ib), it will always be divisible into a set of sub-events satisfying (ib) and involving minimal agents. All these sub-events will have the same theme (since all the agents watched the same movie), and consequently the same measurement event. In other words, the predicate in (ib) retains Stratified Reference despite the presence of a modified numeral, which means that sentence (ia) should have a reading where *fewer than 2* measures out the total number of movies watched by the students. Furthermore, sentence (ia) should contrast with (ii), where the modified numeral is predicted to lack this 'cumulative measurement' interpretation:

(ii) All the students watched fewer than 3 movies.

As far as I can see, these predictions are not borne out.

3.5 Summing up

We have seen that existing approaches to the semantics of dependent plurals suffer from significant drawbacks. Distributivity-based approaches successfully account for the contrasting properties of bare plurals and plurals with numerals under plural quantifiers, but fail to derive the overarching Multiplicity Condition associated with dependent plurals. Cumulativity-based approaches, as well as the mixed approach, have the opposite problem. They successfully account for the Multiplicity Condition, but fail to explain the Ban on Numerals.

From the point of view of the licensors, a successful solution should account for the 'distributive flavour' of plural quantifiers like *all* and *most*, and at the same time explain how this type of distributivity is different from the distributivity of singular quantifiers like *each* and *every* (cf. the Ban on Singular Licensors).

From the point of view of the dependents, what is needed is a unified semantic account of cardinality modifiers that would explain the broad contrast between 'bare' and (many different types of) 'measured' plurals. An account along these lines has recently been proposed by Kuhn (2020). In what follows, I present an alternative way of fleshing out this intuition, which is able to overcome the limitations of Kuhn's proposal.

4 Semantic framework: core features

The analysis I will propose is couched within an extended version of Plural Compositional DRT (PCDRT) of Brasoveanu (2007, 2008), which is itself an extension of Muskens' (1996) Compositional DRT. The main innovation of PCDRT in comparison with Muskens' (1996) system is the introduction of *plural information states* (or info states), as originally proposed by van den Berg (1994, 1996) (see also Nouwen 2003; Brasoveanu and Farkas 2011; Henderson 2014 for applications of related frameworks). A plural information state is a set of assignments which can be represented as a matrix where the rows correspond to individual assignments, and the columns correspond to variables, or *discourse referents* (drefs). The cells in this matrix contain values of discourse referents with respect to assignments, e.g. a cell in row i_m and column u_n will store the value of the dref u_n with respect to the assignment i_m:

(51)

Info state I	...	u_1		u_2		u_3		...
i_1	...	x_1	$(= u_1 i_1)$	y_1	$(= u_2 i_1)$	z_1	$(= u_3 i_1)$	...
i_2	...	x_2	$(= u_1 i_2)$	y_2	$(= u_2 i_2)$	z_2	$(= u_3 i_2)$	...
i_3	...	x_3	$(= u_1 i_3)$	y_3	$(= u_2 i_3)$	z_3	$(= u_3 i_3)$	...
...	...	...		...		...		...

Thus, a plural info state stores multiple values for each dref (the columns in 51), and the correspondence between the values of multiple drefs (the rows in 51).

My aim in what follows is to demonstrate that this semantic framework, initially proposed to account for quite different types of phenomena, is perfectly suited to represent the contrasts and generalizations observed in the realm of (co-)distributive readings. Informally, a framework like PCDRT allows us to talk about three distinct types (or levels) of multiplicity in the semantics of natural language: *mereological plurality* on the level of the individual values of the drefs (e.g. x_3 in (51) can stand for a non-atomic sum of individuals); multiplicity of values of a dref within a single info state, or *state-level plurality* (e.g. in (51) x_1, x_2 and x_3 are all values of dref u_1 in info state I, but they can stand for different individuals); and *distributive multiplicity*, i.e. multiplicity of values of a dref across multiple info states (analogous to having multiple values for a variable relative to different assignments in familiar semantic frameworks). Two of these levels, mereological plurality and distributive multiplicity, have correspondents in standard semantic frameworks without plural info states. But, as I will argue, it is the addition of a third, intermediate, level of multiplicity in systems like PCDRT that gives them the expressive power necessary to account for the full range of contrasts related to (co-)distributivity in natural language.

Specifically, I will argue that the contrast between bare plurals and plurals involving cardinality modifiers amounts to a contrast between state-level and mereological plurality. On the other hand, the contrast between plural and singular quantifiers is best understood in terms of two types of distributive operators: those that introduce state-level plurality by distributing the values of a dref across assignments in a single info state (I will refer to this as *weak distributivity*), and those that distribute the values of a dref across multiple info states (*strong distributivity*). All the generalizations discussed above are accounted for by the combination of these two assumptions.[23] I will start by presenting the core features of the semantic framework, and refer the reader to Brasoveanu (2007, 2008) for a more detailed exposition of the formal underpinnings.

4.1 Types

Brasoveanu's (2007, 2008) PCDRT has three basic types: t (truth-values); e (atomic and non-atomic individuals); and s (variable assignments). The domain of type t is the set of two values $\{0,1\}$. Variable assignments are modelled as basic entities of type s. The domain of type e, $\mathcal{D}_e$, is the powerset of a non-empty set of entities **IN** minus the empty set: $\mathcal{D}_e = \wp(\mathbf{IN}) \backslash \emptyset$. The sum operation is identified with set union: the sum

---

[23] A previous version of this analysis was presented in Minor (2017). A conceptually similar account of the contrast between bare and 'measured' plurals was later independently proposed by Champollion et al. (2017), focusing primarily on the properties of *for*-adverbials. I will not discuss *for*-adverbials in this paper.

$x_e \oplus y_e$ is the union of sets x and y. Similarly, the part-of relation $\leq$ over individuals is identified with the subset relation $\subseteq$ over $\mathcal{D}_e$. Thus, we follow Brasoveanu (2008) in allowing for *mereological plurality*.[24]

To the basic types t, e and s, we add type v for events. The domain of type v, $\mathcal{D}_v$, is defined in the same way as the domain of individuals, i.e. as the powerset of the set of atomic events **EV** minus the empty set: $\mathcal{D}_v = \wp\,(\mathbf{EV}) \backslash \emptyset$. The sum and part-of relations over events are analogous to the corresponding relations over individuals. I will refer to this version of PCDRT enriched with the type for events as PCDRT^e.

The set of basic types and the set of static types are defined in the following way:

(52)   a.   The set of basic static types **BasSTyp**: $\{t,\ e,\ v\}$ (truth-values, individuals and events)

      b.   The set of static types **STyp**: the smallest set including **BasSTyp** such that if $\sigma, \tau \in \mathbf{BasSTyp}$, then $(\sigma\tau) \in \mathbf{STyp}$

4.2 Drefs and DRSs

Discourse referents, or drefs, are modelled as functions of type $s\tau$, where τ is a static type. For instance, individual drefs are functions from assignments to individuals. Thus, a dref u_{se} applied to an assignment i_s, written as ui, returns an individual of type e. I will write uJ to mean the set of values that the dref u returns when applied to the assignments in the plural info state J, i.e. $uJ = \{x : \exists j \in J.\ uj = x\}$. I will also use $\oplus uJ$ to mean the sum of these values. Similarly, an event dref ε_{sv} applied to an assignment i_s, i.e. εi, returns an event of type v.[25]

The set of dref types is defined as follows:

(53)   The set of dref types **DRefTyp**: the smallest set such that if $\tau \in \mathbf{STyp}$, then $(s\tau) \in \mathbf{DRefTyp}$

Each sentence is interpreted as a Discourse Representation Structure (DRS), which is taken to be a function of type $(st)((st)t)$. In other words, a sentence denotes a relation between two sets of assignments, which correspond to the input plural info state and the output plural info state. A standard DRS can fulfil two functions—introduce new drefs and impose conditions on the output info state:

(54) $\lambda I_{st}.\lambda J_{st}.\ I[\textbf{new drefs}]J\ \wedge\ \textbf{conditions}\,J$

---

[24] Brasoveanu (2007), on the other hand, follows van den Berg (1996) in assuming that the domain of individuals $\mathcal{D}_e$ is restricted to atomic individuals. In this kind of system plurality is uniformly modelled as *state-level plurality*, i.e. the existence of multiple distinct values for a dref in a plural info state. In the analysis presented here the distinction between mereological and state-level plurality will play an important role.

[25] As a convention, I will use the the symbols u, u', u'', ... and u_1, u_2, u_3, ... both for individual dref constants of type se, and dref constants in general, and the symbols ε, ε', ε'', ... and ε_1, ε_2, ε_3, ... for event dref constants of type sv. I will specify the type of drefs using subscripts when necessary to avoid confusion. I will also use v, v', v'', ... and v_1, v_2, v_3, ... for variables of type se, and ζ, ζ', ζ'', ... and ζ_1, ζ_2, ζ_3, ... for variables of type sv.

This is abbreviated in the following way:[26]

(55) [**new drefs** | **conditions**]

The following is a simplified DRS corresponding to the sentence *A student chose a film*:[27]

(56) $[u, u', \varepsilon \mid student\{u\}, film\{u'\}, choose\{\varepsilon\}, \mathbf{Ag}\{u, \varepsilon\}, \mathbf{Th}\{u', \varepsilon\}]$
       $:= \lambda I_{st}.\lambda J_{st}.\ I[u, u', \varepsilon]J \ \wedge\ student\{u\}J \ \wedge\ film\{u'\}J \ \wedge\ choose\{\varepsilon\}J \ \wedge$
       $\mathbf{Ag}\{u, \varepsilon\}J \ \wedge\ \mathbf{Th}\{u', \varepsilon\}J$

This DRS introduces two new individual drefs, u and u', and one new event dref ε, and places a set of conditions on the output info state J: the dref u applied to the assignments in J must return a student-individual, dref u' must return a film-individual, and ε must return a choosing event, whose agent is the student returned by u and whose theme is the film returned by u'. I will write DIJ to mean $D(I)(J)$, where D is a DRS and I and J are info states.

DRSs that do not introduce any new drefs while imposing conditions on the output info state are called *tests*, and have the following form:

(57) [**conditions**] $:= \lambda I_{st}.\lambda J_{st}.\ I = J \ \wedge\ \mathbf{conditions}J$

I adopt the following definition of the truth of a DRS:

(58) *Truth of a DRS*
      A DRS D is true with respect to an input info state I iff there is an output info state J, such that $DIJ = 1$.

In what follows I will assume that the input info states for the DRSs are singleton. This greatly simplifies the exposition, and as I will suggest, reflects the default mode of discourse interpretation (see Sect. 6.5).

4.3 Introduction of new drefs

The introduction of a new dref is modelled as an arbitrary reassignment of the values of that dref. In other words, the introduction of a new dref u means that the output info state is allowed to differ from the input state with respect to the values of u. In familiar dynamic systems that do not involve plural info states, this is formalized with the help of the following two-place predicate over assignments:

(59) $[u] := \lambda g_s.\lambda h_s.\ \forall v_{se}\ (vg \neq vh \rightarrow v = u)$

---

[26] In the following, I will try to provide the abbreviated form of expressions when possible, since this format simplifies the presentation and should be more familiar to readers acquainted with regular DRT. However, the abbreviated language is less expressive than full PCDRT, and consequently not all expressions that we encounter will allow for an abbreviated form.

[27] Here and throughout the paper I disregard the semantics of tense and aspect. I am also for now disregarding the role of number.

Informally, $g_s[u]h_s$ means that assignments g and h differ at most with respect to the value for u.

Our system allows for non-singleton info states, which means that the predicate in (59) is not directly applicable. Instead, we will adopt the following definition:

(60) *New Dref Introduction*
    $[u] := \lambda I_{st}.\lambda J_{st}. \exists f_{ss}. \forall j \in J. \exists i \in I. (f(i) = j) \land \forall i \in I. (f(i) \in J \land i[u]f(i))$,
    where f is a function from $\mathcal{D}_s$ to $\mathcal{D}_s$.

Under this definition, a DRS introducing a new dref u maps every assignment in the input info state onto a single u-different assignment in the output info state.[28]

Multiple dref introduction (as in example (56) above) is defined on the basis of dynamic conjunction of multiple introduction predicates. The definition of dynamic conjunction and multiple dref introduction is given in (61):

(61) a. $D_{st((st)t)}; D'_{st((st)t)} := \lambda I_{st}.\lambda J_{st}. \exists H_{st}. (DIH \land D'HJ)$
    b. $[u_1, \ldots, u_n] := [u_1]; \ldots ; [u_n]$

Then, for example (56) the following holds:

(62) $I[u, u', \varepsilon]J := \exists H_{st}. \exists H'_{st}. I[u]H \land H[u']H' \land H'[\varepsilon]J$

4.4 Conditions and lexical relations

The second part of a DRS contains conditions on the output info state, i.e. terms of type $(st)t$ applied to the output info state. In example (56) above, these include lexical predicates, i.e. *student*$\{u\}J$, *film*$\{u'\}J$ and *choose*$\{\varepsilon\}J$, as well as thematic relations, **Ag**$\{u, \varepsilon\}J$ and **Th**$\{u', \varepsilon\}J$. All these conditions are evaluated distributively relative to each assignment in the output info state. More formally, for any non-logical constant R of type $e^n t$, the following convention holds:[29]

(63) $R\{u_1, \ldots, u_n\} := \lambda I_{st}. I \neq \varnothing \land \forall i_s \in I.(R(u_1i, \ldots, u_ni))$

---

[28] This way of implementing the introduction of new drefs is different from that adopted in Brasoveanu (2007, 2008):

(i) $[u] := \lambda I_{st}.\lambda J_{st}. \forall i_s \in I (\exists j_s \in J (i[u]j)) \land \forall j_s \in J (\exists i_s \in I (i[u]j))$

(i) requires for every individual assignment in the input info state to have a corresponding u-different assignment in the output state, and conversely, for every individual assignment in the output info state to have a corresponding u-different assignment in the input state. No additional restrictions are placed on the output info state. Our definition in (60) is more restrictive: it ensures that a DRS introducing a new dref cannot return an output info state of a cardinality greater than that of the input info state. Thus, in our system the function of increasing the cardinality of an info state is reserved for distributivity operators and quantificational expressions (to be discussed below). This greatly simplifies all the calculations within the system.

[29] $e^n t$ is defined as the smallest set of types s.t. $e^0 t := t$ and $e^{m+1}t := e(e^m t)$, cf. Muskens (1996), Brasoveanu (2007, 2008).

Thus, *student*$\{u\}J$ will be true iff *student*(uj) is true for every assignment j in J, i.e. the dref u maps every assignment j in J to an individual who is a student (or more accurately, a sum of students, cf. the discussion of lexical cumulativity in Sect. 4.5). Similarly, *walk*$\{\zeta\}J$ will be true iff *walk*(ζj) is true for every j in J.

The translations of most common nouns and verbs involve lexical relations of this type, e.g.:[30]

(64) a. student $\leadsto \lambda v_{se}.[student\{v\}] := \lambda v_{se}.\lambda I_{st}.\lambda J_{st}.\ I = J\ \wedge\ student\{v\}J$
         $:= \lambda v_{se}.\lambda I_{st}.\lambda J_{st}.\ I = J\ \wedge\ \forall j \in J.(student(vj))$

      b. walk $\leadsto \lambda v_{se}.\lambda \zeta_{sv}.[walk\{\zeta\},\ \mathbf{Ag}\{v,\ \zeta\}]$
         $:= \lambda v_{se}.\lambda \zeta_{sv}.\lambda I_{st}.\lambda J_{st}.\ I = J\ \wedge\ walk\{\zeta\}J\ \wedge\ \mathbf{Ag}\{v,\ \zeta\}J$
         $:= \lambda v_{se}.\lambda \zeta_{sv}.\lambda I_{st}.\lambda J_{st}.\ I = J\ \wedge\ \forall j \in J.(walk(\zeta j)\ \wedge\ \mathbf{Ag}(vj,\ \zeta j))$

In dealing with the types of expressions in PCDRT^e, it will be useful to adopt a simplifying convention proposed in Brasoveanu (2007, 2008): let us use **e** to stand for the type of individual drefs (se), **v** to stand for the type of event drefs (sv), and **t** to stand for the type of DRSs ($(st)((st)t)$). Then the type of common nouns $(se)((st)((st)t))$ (i.e. functions from individual drefs of type se to DRSs of type $(st)((st)t)$) can be abbreviated as **et**, the familiar type for common nouns and intransitive verbs in the standard Montgovian system. Similarly, determiners such as *the* and *every* translate into functions of type **(et)((et)t)**, intransitive verbs are translated into functions of type **e(vt)**, and transitive verbs are translated as functions of type **e(e(vt))**:

(65) choose $\leadsto \lambda v_{\mathbf{e}}.\lambda v'_{\mathbf{e}}.\lambda \zeta_{\mathbf{v}}.\ [choose\{\zeta\},\ \mathbf{Th}\{v,\ \zeta\},\ \mathbf{Ag}\{v',\ \zeta\}]$

I adopt the Neo-Davidsonian system of verb interpretation, where verbs are taken to introduce predicates over events, and arguments are related to events via thematic relations (cf. Parsons 1990). I will also assume Role Uniqueness, which states that for any thematic relation Θ and any event e, there is a unique individual x such that $\Theta(x,\ e)$ (cf. Carlson 1984; Parsons 1990; Landman 1996, 2000).

DPs are uniformly translated into functions of type **(et)t**. This means that in order to combine with their arguments verbs need to be type-shifted to a higher type. I will use an adapted version of the LIFT type-shifters from Landman (2000):

(66) a. **Intransitive Lift:** $\lambda v_{se}.\lambda \zeta_{sv}.[...] \Rightarrow \lambda Q_{(\mathbf{et})\mathbf{t}}.\lambda \zeta.\ Q(\lambda v.\ [...])$

      b. **Transitive Lift:** $\lambda v_{se}.\lambda v'_{se}.\lambda \zeta_{sv}.[...] \Rightarrow \lambda Q_{(\mathbf{et})\mathbf{t}}.\lambda v'.\lambda \zeta.Q(\lambda v[...])$

4.5 Lexical cumulativity and lexical distributivity

I will assume that most lexical predicates, e.g. *boy*, *girl*, *walk*, as well as thematic relation such as **Ag** and **Th** are closed under the sum operation, i.e. that they are *cumulative* at the domain level. Cumulativity for one-place and two-place lexical relations is defined as follows (cf. e.g. Krifka 1989; Landman 1996):

---

[30] Collective predicates such as *gather* may be an exception to this, imposing conditions on the set of values of a dref across assignments in a plural info state, cf. Minor (2017) for discussion.

(67) *Lexical Cumulativity*

$\forall x, y \; (R(x) \;\wedge\; R(y) \rightarrow R(x \oplus y))$

$\forall x_1, x_2, y_1, y_2. \; (R(x_1, y_1) \;\wedge\; R(x_2, y_2) \rightarrow R(x_1 \oplus x_2, \; y_1 \oplus y_2))$

Thus, the predicate *boy* applies to sums of boys, as well as to individual boys. Similarly, *walk* applies both to individual walking events, and to sums of such events.

We will also make reference to the *lexical distributivity* of predicates. This is the property that ensures e.g. that whenever the predicate *dog* is true of a sum of individuals it is also necessarily true of each atomic sub-individual in that sum, and when there is an event of a sum of dogs barking then for each atomic sub-individual d in that sum there must be an event of d barking. Generally, verbal predicates may or may not be lexically distributive with respect to a particular argument (i.e. theta-role). For instance, the transitive verb *carry* is lexically distributive with respect to its theme (i.e. if an individual carries a sum of boxes that individual necessarily carries each individual box), however it is not lexically distributive with respect to its agent (i.e. if a sum of individuals carries a box it does not follow that each of these individuals carries that box, because they may be carrying it together).

Formally, the lexical distributivity of particular predicates is encoded as a set of constraints (i.e. axioms or meaning postulates) on appropriate models, in the following way:

(68) a. $\forall x. \; (dog(x) \rightarrow \forall x'.(x' \leq x \;\wedge\; \mathbf{atom}(x') \rightarrow dog(x')))$

b. $\forall x.\forall e. \; (bark(e) \;\wedge\; \mathbf{Ag}(x, e) \rightarrow \forall x'. \; (x' \leq x \;\wedge\; \mathbf{atom}(x') \rightarrow \exists e'. \; (e' \leq e \;\wedge\; bark(e') \;\wedge\; \mathbf{Ag}(x', e'))))$

c. $\forall x.\forall y.\forall e. \; (carry(e) \;\wedge\; \mathbf{Ag}(x, e) \;\wedge\; \mathbf{Th}(y, e) \rightarrow \forall y'. \; (y' \leq y \;\wedge\; \mathbf{atom}(y') \rightarrow \exists x'.\exists e'. \; (x' \leq x \;\wedge\; e' \leq e \;\wedge\; carry(e') \;\wedge\; \mathbf{Ag}(x', e') \;\wedge\; \mathbf{Th}(y', e'))))$

4.6 Compositionality

Non-terminal syntactic constituents are translated with the help of a set of rules which define the translation of the mother node based on the translations of its daughter nodes:

Non-Branching Nodes (NN)
If A $\rightsquigarrow \alpha$ and A is the only daughter of B, then B $\rightsquigarrow \alpha$.

Functional Application (FA)
If A $\rightsquigarrow \alpha$ and B $\rightsquigarrow \beta$ and A and B are the only daughters of C, then C $\rightsquigarrow \alpha(\beta)$, provided that this is a well-formed term.

Generalized Sequencing (GSeq) (Sequencing + Predicate Modification)
If A $\rightsquigarrow \alpha$, B $\rightsquigarrow \beta$, A and B are the only daughters of C in that order, and α and β are of the same type τ of the form $\mathbf{t}$ or $(\sigma\mathbf{t})$ for some type σ, then C $\rightsquigarrow \alpha; \; \beta$ if $\tau = \mathbf{t}$ or C $\rightsquigarrow \lambda v_\sigma.\alpha(v); \; \beta(v)$ if $\tau = (\sigma\mathbf{t})$, provided that this is a well-formed term.

4.7 Indexing, traces, and *Quantifying-In*

I will follow Muskens (1996) in assuming that the syntactic component provides indexation for all determiners, pronouns, and traces. However, following Brasoveanu (2007, 2008) and unlike Muskens (1996), I will take drefs to serve as indices directly:

(69)   a.   The^u girl saw a$^{u'}$ rabbit. It$_{u'}$ was eating.
       b.   Every^u boy saw himself$_u$ in a$^{u'}$ mirror.

    Indices on determiners that introduce new discourse referents are written as superscripts, while indices on pronouns and traces link back to existing drefs, and are written as subscripts.

    Dislocated (moved) DPs are assigned an additional index by the movement rule, and leave behind co-indexed traces. Traces and the corresponding dislocated DPs differ from other index-bearing items in that their, and only their, indices are *variables* of the dref type (se), rather than constants:

(70)   The^u boy found a$^{u'}$ letter which^v the$^{u''}$ girl had lost t$_v$.

Traces are translated as the dref variables they are indexed with (i.e. t$_v \rightsquigarrow v$). The Quantifying-In translation rule can then be stated as follows:

(71)   *Quantifying-In (QIn) (Predicate Abstraction)*
      If DP$^v \rightsquigarrow \alpha$, B $\rightsquigarrow \beta$ and DP^v and B are daughters of C, then C $\rightsquigarrow \alpha(\lambda v.\beta)$, provided that this is a well-formed term.

4.8 Event closure

Binding of the event dref variable introduced by the verb is performed by a designated $\exists_{ev}$ operator:

(72) $\exists_{ev}^{\varepsilon} \rightsquigarrow \lambda V_{\mathbf{vt}}. \ [\varepsilon]; \ V(\varepsilon)$

    The event closure operator introduces a new event dref, ε in (72), and applies the verbal predicate to this event dref. I will assume that this operator is inserted at some level above the vP, after the verb has combined with all of its arguments, but I will remain agnostic about its exact position with respect to other functional heads.[31]

5 Analysis

5.1 The semantics of number and cardinal modifiers

I will assume that DPs have the following syntactic structure:

---

[31] As a notational convention, I will indicate the event dref introduced by the event closure operator as a superscript on the verb, e.g.:

   (i)   A^u student chose$^{\varepsilon}$ a$^{u'}$ film.

(73) *Structure of DP*
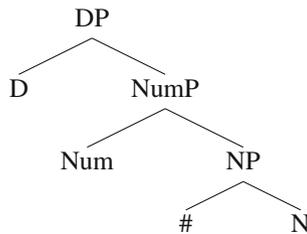
I will take grammatical number, #, to be interpreted in the position where it is spelled-out phonologically, i.e. adjacent to the noun.[32] The #-head has two variants in English: #:sg and #:pl, which are translated as predicates of type **et**, and combine with their sister nouns via Generalized Sequencing (cf. Sect. 4.6).

The singular imposes two conditions on its argument dref: it requires it to be atomic at the assignment-level (i.e. the value that the dref returns for each assignment in the output plural info state must be atomic), and unique. Uniqueness is satisfied if the dref returns the same value for all the assignments in the output info state. Together, these conditions are equivalent to stating that the sum of values of the dref in the output info state must be atomic (i.e. the singular encodes a *state-level* atomicity condition).

(74) #:sg $\leadsto \lambda v_\mathbf{e}.$ [**atom**$\{v\}$]; [**unique**$\{v\}$]

(75) a. **atom**$\{u\} := \lambda I_{st}. \forall i \in I.$ **atom**(ui)

  b. **unique**$\{u\} := \lambda I_{st}. \forall i, i' \in I. ui = ui'$

Plural number, on the other hand, is semantically vacuous, and thus plural noun phrases are taken to be underlyingly number-neutral (following Krifka 1989, 2004; Sauerland et al. 2005; Sauerland 2003; Spector 2007; Zweig 2008, 2009, a.o.). The multiplicity semantics associated with plurals in non-downward entailing context will be derived as an implicature (cf. Sect. 5.3).

(76) #:pl $\leadsto \lambda v_\mathbf{e}.\lambda I_{st}.\lambda J_{st}. I = J$

Numerals and cardinality modifiers attach above # and are translated as expressions of type **et**, imposing assignment-level cardinality conditions on the values of a dref:

(77) a. two $\leadsto \lambda v_\mathbf{e}.$ [**2_atoms**$\{v\}$]

  b. **2_atoms**$\{u\} := \lambda I_{st}. \forall i \in I.$ **2_atoms**(ui),
    where **2_atoms**$(x_e) := |\{y_e : y \leq x \ \wedge \ \mathbf{atom}(y)\}| = 2$.

---

[32] In (73) I represent number as a syntactic head directly above the noun. However, nothing in my analysis prevents number to be analysed as an interpretable feature on the noun itself.

(78) a. several $\rightsquigarrow \lambda v_{\mathbf{e}}.$ [**several_atoms**$\{v\}$]

   b. **several_atoms**$\{u\} := \lambda I_{st}. \forall i \in I.$ **several_atoms**(ui),
       where **several_atoms**$(x_e) := |\{y_e : y \leq x \ \wedge \ \mathbf{atom}(y)\}| > 1.$

5.2 Indefinite DPs

DPs are headed by determiners, including articles and quantificational determiners (QDs). The sole function of the indefinite article is to introduce a new dref. Each indefinite article is indexed with the name of the dref that it introduces, represented as a superscript. I will assume that the indefinite article in English has two morphological forms, depending on its syntactic environment: *a* combines only with phrases headed by #:sg, while a phonologically null indefinite article, Indef, combines with phrases headed by #:pl and NPs with numerals and cardinal modifiers.

(79) a^u, Indef$^u \rightsquigarrow \lambda P_{\mathbf{et}}.\lambda P'_{\mathbf{et}}.$ [u]; $P(u)$; $P'(u)$

The following examples illustrate the compositional translation of singular and plural indefinite DPs:[33]

(80)

a^u student
$\lambda P'.$ [u]; [**atom**$\{u\}$]; [**unique**$\{u\}$]; [*student*$\{u\}$]; $P'(u)$

a^u
$\lambda P.\lambda P'.$ [u]; $P(u)$; $P'(u)$

$\lambda v.$ [**atom**$\{v\}$]; [**unique**$\{v\}$]; [*student*$\{v\}$]

#:sg
$\lambda v.$[**atom**$\{v\}$]; [**unique**$\{v\}$]

student
$\lambda v.$ [*student*$\{v\}$]

(81)

students^u
$\lambda P'.$ [u]; [*student*$\{u\}$]; $P'(u')$

Indef^u
$\lambda P.\lambda P'.$ [u]; $P(u)$; $P'(u)$

$\lambda v.$ [*student*$\{v\}$]

#:pl
$\lambda v.\lambda I.\lambda J.\ I=J$

student
$\lambda v.$ [*student*$\{v\}$]

---
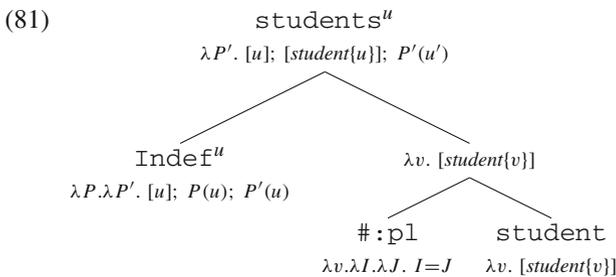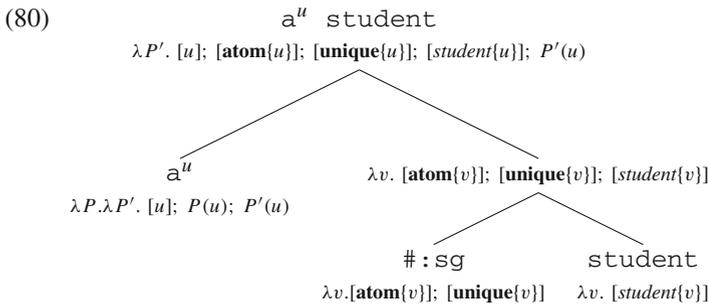
[33] As a notational convention, I will indicate the dref introduced by the null indefinite article as a superscript on the noun when providing noun phrase translations.

(82)

$$\text{two students}^u$$
$$\lambda P'. \; [u]; \; [\mathbf{2\_atoms}\{u\}]; \; [student\{u\}]; \; P'(u')$$

$$\text{Indef}^u$$
$$\lambda P.\lambda P'. \; [u]; \; P(u); \; P'(u)$$

$$\lambda v. \; [\mathbf{2\_atoms}\{v\}]; \; [student\{v\}]$$

$$\text{two}$$
$$\lambda v.[\mathbf{2\_atoms}\{v\}]$$

$$\lambda v. \; [student\{v\}]$$

$$\text{\#:pl} \qquad\qquad \text{student}$$
$$\lambda v.\lambda I.\lambda J. \; I{=}J \qquad \lambda v. \; [student\{v\}]$$
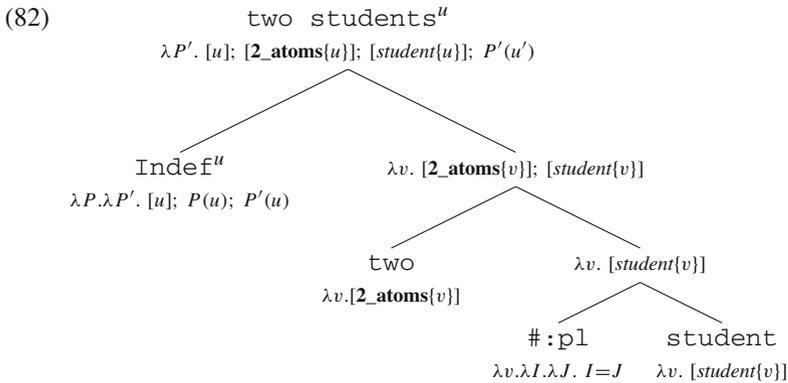
5.3 The multiplicity implicature

5.3.1 Constraints on *Exh*-insertion

I will follow Spector (2007), Zweig (2008, 2009) and Ivlieva (2013) and derive the multiplicity semantics associated with the plural number feature as a *scalar implicature* which arises in competition with a semantically more restrictive singular number feature. On this account, in order to derive the multiplicity implicature for a plural DP in structure α we must show that the corresponding structure β, where the plural DP has been replaced with its singular counterpart, has a *stronger* interpretation than α. Then, the interpretation of α can be strengthened (or *enriched*) via the negation of the stronger alternative.

Following Zweig (2008, 2009) and Ivlieva (2013), I will assume that the multiplicity implicature can be calculated at various points of the semantic derivation (cf. Chierchia 2004, 2006; Fox 2007; Chierchia et al. 2012; Fox and Spector 2018, a.o.). The implicature is incorporated into the semantic form by means of *exhaustification*— an operation that involves (a) comparing the logical form of an expression to a set of alternatives; (b) determining which alternatives are stronger than that logical form; and (c) adding the negation of the stronger alternatives to the logical form. For concreteness, I will assume that exhaustification is encoded as the semantics of a covert exhaustivity operator *Exh*, which can be inserted at various levels in the syntactic structure (cf. Chierchia et al. 2012, a.o.).[34]

Previous analyses disagree on the correct way of setting up the system of implicature calculation for bare plurals. Zweig (2008, 2009) (following Chierchia 2004, 2006) adheres to what Ivlieva (2020) calls the *Strongest Candidate Principle*. According to this principle, if there are multiple levels in the structure of a sentence at which the scalar implicature can be calculated, the one that produces the strongest overall interpretation is chosen as the meaning of the sentence. Chierchia et al. (2012)

---

[34] This assumption is not critical to the proposed account of dependent plurals. Alternative ways of deriving the multiplicity implicature can be formulated within a system of recursive implicature calculation such as that proposed in Chierchia (2004, 2006), which does not rely on syntactically represented exhaustification operators.

provide arguments against adopting this as a general principle of calculating scalar implicatures. Furthermore, Ivlieva (2020) argues that, contra Zweig (2008, 2009), this principle produces incorrect results when applied to sentences involving bare plurals in cumulative construction like (83a):

(83)   a.   My three friends attend good schools.

      b. $\exists e \exists y$ [*good-school$(y) \wedge$ *attend$(e)(\sigma x.$ my-3-friends$(x))(y)] \wedge$
         $\neg \exists e \exists y$ [*good-school$(y) \wedge$ atom$(y) \wedge$ *attend$(e)(\sigma x.$ my-3-friends$(x))(y)]$

      c. $\exists e \exists y$ [*good-school$(y) \wedge \neg$atom$(y) \wedge$ *attend$(e)(\sigma x.$ my-3-friends$(x))(y)]$

In Zweig's system, calculating the scalar implicature above event closure in (83a) results in the interpretation in (83b) according to which sentence (83a) will be true iff there is a sum of schools which the speaker's 3 friends cumulatively attend, but there is no single school that all these friends attend. As Ivlieva (2020) shows, this interpretation is stronger than the standard cumulative interpretation in (83c), which can be derived by calculating the multiplicity implicature below the event closure. Consequently, in Zweig's system the cumulative interpretation is blocked by the Strongest Candidate Principle, contrary to fact.

Ivlieva (2020) proposes to replace the Strongest Candidate Principle with a *Non-Weakening Condition* (cf. Fox and Spector 2018):

(84)   *Non-Weakening Condition*
      Do not introduce *Exh* in a structure *S* if that would lead to a sentence meaning which is entailed by the meaning of *S* without *Exh*.

Now, it turns out that in a dynamic framework like the one adopted here, sentences like (83a) do not pose a problem for the Strongest Candidate Principle (see Sect. 6.1). In fact, nothing in the data that I address in this paper is directly incompatible with the Strongest Candidate Principle. However, given the independent arguments against this principle in Chierchia et al. (2012), I will follow Ivlieva (2020) in adopting the condition in (84) instead. Nevertheless, it is worth noting that if the analysis presented here is on the right tract, cumulative and dependent plural constructions do not in themselves present an argument against the Strongest Candidate Principle (contra Ivlieva 2020).[35]

Finally, following Ivlieva (2013), I will take *Exh*-insertion to be obligatory in the context of bare plural DPs (see also Chierchia et al. 2012):

(85)   A bare plural DP must be c-commanded by an exhaustification operator, whose restrictor contains the alternative obtained by replacing the plural with the corresponding singular.

This accounts for the obligatory emergence of the multiplicity implicature in non-downward entailing contexts.[36]

---

[35] The Strongest Candidate Principle and the Non-Weakening Condition do make contrasting predictions with respect to the availability of certain marginal readings of sentences involving quantificational items. However the empirical status of these readings is unclear, see footnote 54.

[36] As Ivlieva (2020) notes, there is a tension between the Non-Weakening Condition in (84), which bars *Exh*-insertion in positions where it leads to weakening or vacuous exhaustification, and the condition in (85),

5.3.2 *Exh*-operators and strength

Formally, I will define a family of **Exh**-operators in PCDRT^e in the following way:

(86) *Generalized Exhaustification*
    For any conjoinable type α, such that $\alpha = (\tau_1, (\tau_2(\ldots \tau_n \mathbf{t}) \ldots))$,
    $\mathbf{Exh}_{\text{ALT} \langle \alpha, \alpha \rangle} := \lambda Q_\alpha . \lambda k^1_{\tau_1} . \lambda k^2_{\tau_2} \ldots \lambda k^n_{\tau_n} . \lambda I_{st} . \lambda J_{st} . \ Q(k^1) \ldots (k^n) I J \wedge \forall Q'_\alpha .$
    $(Q' \in \text{ALT} \wedge Q' \succ Q \rightarrow \neg Q'(k^1) \ldots (k^n) I J)$,
    where ALT is the set of alternatives to Q, and $Q' \succ Q$ means that Q' is stronger than Q.

For instance, if **Exh** combines with a DRS of type $\mathbf{t}$, the result is a DRS of the following form:

(87) $\mathbf{Exh}_{\text{ALT} \ \mathbf{tt}}(D) := \lambda I_{st} . \lambda J_{st} . \ D I J \wedge \forall D' . (D' \in \text{ALT} \wedge D' \succ D \rightarrow \neg D' I J)$

The syntactic *Exh* operator is translated as one of the **Exh** operators as defined in (86).

Two comments are in order. First, the definition in (86) makes reference to *conjoinable types*. The definition of conjoinable types is modelled on that in Partee and Rooth (1983):

(88) *Conjoinable Type*
    (i) $\mathbf{t}$ (i.e. $st((st)t)$) is a conjoinable type
    (ii) if σ is a conjoinable type, then for all types $\tau \in \mathbf{DRefTyp}$, $(\tau\sigma)$ is a conjoinable type.

Second, the definition of strength $\succ$ must be made explicit and sufficiently general to be applicable to terms of all conjoinable types. I will adopt the following definition:

(89) *Generalized Strength*
    For any conjoinable type α, such that $\alpha = (\tau_1, (\tau_2(\ldots \tau_n \mathbf{t}) \ldots))$,
    $Q'_\alpha \succ Q_\alpha$ iff:

Footnote 36 continued

which requires for each plural DP to be c-commanded by *Exh*. Specifically, taken together these conditions rule out all the alternative options for *Exh*-(non)insertion in sentences where plurals occur in downward entailing contexts, e.g.:

  (i) John did not carry boxes.

In this case, inserting *Exh* below negation would lead to the weakening of the sentence's truth conditions, inserting it above negation leads to vacuous exhaustification, and leaving *Exh* out of the structure altogether would result in the plural DP *boxes* lacking a c-commanding *Exh*-operator. Ivlieva (2013) proposes a way to solve this problem by adopting (a) a looser version of the Non-Weakening Condition (ruling out *Exh*-insertion which leads to asymmetric weakening of the truth conditions, but not vacuous exhaustification); and (b) an assumption that implicature calculation is blind to the lexical distributivity of predicates (cf. Ivlieva 2013 for detailed discussion, and Minor 2017 for some criticism of this approach). Alternatively, we may assume that one of the two conditions in (84) and (85) is violable, and must be applied as long as the other condition is satisfied. I will leave this question for future research.

(a) For any appropriate model M and assignment function g, and any a_1 of type τ_1, ..., a_n of type τ_n, input info state I_{st} and output info state J_{st} :
if $[\![Q'(a_1)\dots(a_n)IJ]\!]^{M,g} = 1$, then $[\![Q(a_1)\dots(a_n)IJ]\!]^{M,g} = 1$.
(b) There is an appropriate model M, assignment function g, a_1 of type τ_1, ..., a_n of type τ_n, info states I_{st} and J_{st}, such that $[\![Q(a_1)\dots(a_n)IJ]\!]^{M,g} = 1$ and $[\![Q'(a_1)\dots(a_n)IJ]\!]^{M,g} = 0$.

As a limiting case, for two DRSs D and D', $D' \succ D$ will be true iff for any appropriate model and info states I and J, if $D'IJ = 1$ then $DIJ = 1$, while the converse does not hold. I will assume that *appropriate models* are those that respect the meaning postulates associated with lexical items including those encoding the lexical cumulativity and lexical distributivity of predicates.

5.4 Weak and strong distributivity

In this section I introduce the distinction between weak and strong distributivity. This distinction will play a crucial role in our analysis of the properties of dependent plurals. Somewhat informally, weak distributivity involves distribution of individuals in a sum across the assignments in a single info state, while strong distributivity involves distribution across multiple info states. We will start by defining a weak and a strong distributivity operator in PCDRT^e, and then use these to provide the translations of syntactic distributivity operators. I will then show how the contrasting properties of singular and plural quantificational determiners can be captured in terms of weak and strong distributivity.

5.4.1 Weak and strong distributivity operators

Before I provide the definitions of the distributivity operators themselves, I will define an auxiliary relation between info states:

(90) $\langle u \rangle := \lambda I_{st}.\lambda J_{st}. \exists f. \forall j_s \in J. \exists i_s \in I. (j \in f(i)) \wedge$
$\forall i_s \in I. \forall H_{st}. (f(i) = H \rightarrow H \subseteq J \wedge \forall h_s \in H. (i[u]h \wedge \mathbf{atom}(uh)) \wedge$
$\oplus u H = ui)$,
where f is a function from the domain of assignments $\mathcal{D}_s$ to the set of info states $\wp(\mathcal{D}_s)$.

I will write $I\langle u \rangle J$ to mean that $\langle u \rangle$ applies to I and J. In a sense, what $\langle u \rangle$ does is 'split' each assignment in the input info state into multiple assignments, where the values for u are the atomic sub-parts of its value for the original assignment. For instance, suppose an input info state I contains one assignment i, such that ui returns the sum individual *john* $\oplus$ *mary*. Then an info state J such that $I\langle u \rangle J$ will contain two assignments, j_1, j_2, where j_1 and j_2 are identical to i, except that uj_1 is *john* and uj_2 is *mary*:[37]

---

[37] This generalizes to info states of higher cardinality. For instance, if I contains two assignments i_1 and i_2, such that ui_1 is *john* $\oplus$ *mary* and ui_2 is *jane* $\oplus$ *bob*, then J s.t. $I\langle u \rangle J$ will contain four assignments, j_1, j_2, j_3, j_4, where uj_1 is *john*, uj_2 is *mary*, uj_3 is *jane* and uj_4 is *bob*.

(91)

Info state I		...	u	...
i		...	*john* $\oplus$ *mary*	...

$$\Downarrow$$

Info state J, such that $I\langle u\rangle J$		...	u	...
j_1	(such that $j_1[u]i$)	...	*john*	...
j_2	(such that $j_2[u]i$)	...	*mary*	...

The following equivalence will be relevant for us:

(92) $\langle u\rangle; \langle u\rangle := \langle u\rangle$

It is easy to see why this should be the case. Applying the $\langle\rangle$-operator to the values of a dref will split these values into their atomic parts, and distribute them across the assignments in a plural info state. Applying the $\langle\rangle$-operator to the values of the same dref again will return exactly the same info state, because all the values are already atomic and thus cannot be split any further.

We can now use the $\langle\rangle$-operator to define two distinct distributivity operators:

(93) a. $\mathbf{dist}_w(D)(u) := \lambda I_{st}.\ \lambda J_{st}.\ (\langle u\rangle;\ D)IJ$
$\qquad := \lambda I_{st}.\ \lambda J_{st}.\exists H_{st}.\ I\langle u\rangle H\ \wedge\ DHJ$
b. $\mathbf{dist}_s(D)(u) := \lambda I_{st}.\ \lambda J_{st}.\ (\langle u\rangle;\ \mathbf{dist}(D))IJ$
$\qquad := \lambda I_{st}.\ \lambda J_{st}.\exists H_{st}.\ I\langle u\rangle H\ \wedge\ \mathbf{dist}(D)HJ$

The definition of the **dist** operator used in (93b) is the same as in Brasoveanu (2008):

(94) $\mathbf{dist}(D) := \lambda I_{st}.\lambda J_{st}.\ \exists R_{s((st)t)} \neq \varnothing\ (I = \mathbf{Dom}(R)\ \wedge\ J = \bigcup \mathbf{Ran}(R)\ \wedge$
$\forall k_s \forall L_{st}\ (RkL \rightarrow D\{k\}L))$,
where D is a DRS; $\mathbf{Dom}(R) := \{k_s\ :\ \exists L_{st}(RkL)\}$; $\mathbf{Ran}(R) := \{L_{st}\ :\ \exists k_s(RkL)\}$; and $\{k\}$ is the singleton info state containing k.

The *weak distributivity operator* $\mathbf{dist}_w$ combines with a dref u and a DRS D, and returns a DRS which applies to an input info state I and an output info state J iff J is the result of 'splitting' the value(s) for u in I with the help of the $\langle\rangle$-operator, and then updating the resulting info state with D. For instance, consider again info state I in (91), repeated in (95). If we update it with the DRS $\mathbf{dist}_w(D)(u)$, the values for u in I will first be split across multiple assignments, generating an intermediate info state H, as shown in (95). This is the result of updating I with $\langle u\rangle$. Then this info state H will be updated with D to yield the output info state.

(95)

Info state I		...	u	...
i		...	*john* $\oplus$ *mary*	...

$$\Downarrow$$

Info state H, such that $I\langle u\rangle H$	...	u	...
h_1 (such that $h_1[u]i$)	...	*john*	...
h_2 (such that $h_2[u]i$)	...	*mary*	...

The *strong distributivity operator* $\textbf{dist}_s$ also splits the values for u in I with the help of the $\langle\rangle$-operator, yielding an intermediate info state H. But instead of updating H directly with D, it further splits H into a set of singleton info state, updates each of these info states independently with D, and then 'glues' the resulting info states back together into J.

Take once again info state I in (96). Updating I with $\textbf{dist}_s(D)(u)$ will first generate the same intermediate info state H as in (95), which is the result of applying the $\langle\rangle$-operator. But then, following the semantics of $\textbf{dist}$ in (94), H will be split into two singleton info states, H_1 and H_2, as shown in (96), and each of these info states will be independently updated with D, yielding two new info states, call them H'_1 and H'_2. The output info state will then be the union of H'_1 and H'_2, i.e. $J = H'_1 \cup H'_2$.

(96)

Info state I	...	u	...
i	...	*john* $\oplus$ *mary*	...

$\Downarrow$

Info state H, such that $I\langle u\rangle H$	...	u	...
h_1 (such that $h_1[u]i$)	...	*john*	...
h_2 (such that $h_2[u]i$)	...	*mary*	...

$\Downarrow$

Info state H_1	...	u	...
h_1 (such that $h_1[u]i_1$)	...	*john*	...

Info state H_2	...	u	...
h_2 (such that $h_2[u]i_1$)	...	*mary*	...
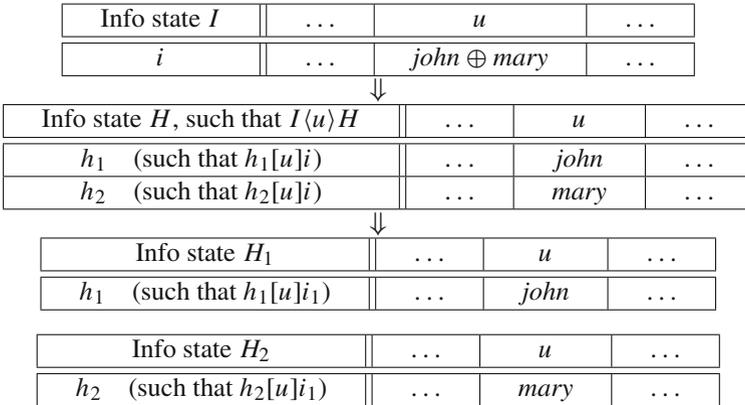
I will also define an 'externally static' version of the distributivity operators as follows:

(97) a. $\textbf{dist'}_w(D)(u) := \lambda I_{st}.\, \lambda J_{st}.\, I = J \ \wedge\ \exists J'.\, (\langle u\rangle;\ D)IJ'$

b. $\textbf{dist'}_s(D)(u) := \lambda I_{st}.\, \lambda J_{st}.\, I = J \ \wedge\ \exists J'.\, (\langle u\rangle;\ \textbf{dist}(D))IJ'$

In contrast to the 'externally dynamic' operators in (93), the operators in (97) return an output info state that is identical to the input info state, i.e. they do not update the running info state directly. Instead, they function as tests that ensure that the input info state *can* be updated in the relevant way (which entails that the relevant DRS is in fact true). These modified versions of the distributivity operators will allow us to define a more elegant semantics for quantificational determiners.

An important property of distributivity operators is that if they are 'stacked', the following equivalences hold:

(98)   a. $\mathbf{dist}_w(\mathbf{dist}_w(D)(u))(u) := \mathbf{dist}_w(D)(u)$

     b. $\mathbf{dist}_s(\mathbf{dist}_s(D)(u))(u) := \mathbf{dist}_s(D)(u)$

     c. $\mathbf{dist}_w(\mathbf{dist}_s(D)(u))(u) := \mathbf{dist}_s(D)(u)$

     d. $\mathbf{dist}_s(\mathbf{dist}_w(D)(u))(u) := \mathbf{dist}_s(D)(u)$

It is not hard to see how these equivalences follow from the definitions in (93) and the equivalence in (92).

We will use the weak and strong distributivity operators defined in this section to provide the translation of two classes of lexical items: syntactic distributivity operators/floating quantifiers and quantificational determiners.

5.4.2 Syntactic distributivity operators

Following Link (1987) and Roberts (1990), I will assume that predicates can combine with distributivity operators, which modify the way the predicate is applied to its argument (cf. also Landman 1989, 2000; Schwarzschild 1996; Lasersohn 1998; Kratzer 2007, a.o.). Furthermore, I will assume that distributivity operators exist as syntactic objects which can be attached to any constituent along the verbal spine. I will posit two types of such operators: weak and strong. I will take the floating quantifier *all* to be an instantiation of the weak syntactic distributitivy operator, and the floating quantifier *each* to be an instantiation of the strong syntactic distributitivy operator. Furthermore, I will assume that both of these operators have phonologically null counterparts, represented as δ_w and δ_s respectively.

The two types of syntactic distributivity operators receive the following translations:

(99)   *Distributivity operators*

     a. $\delta_w,\ all \rightsquigarrow \lambda P_{\mathbf{et}}.\lambda v_{\mathbf{e}}.\ \mathbf{dist}_w(P(v))(v)$
   $:= \lambda P_{\mathbf{et}}.\lambda v_{\mathbf{e}}.\lambda I_{st}.\lambda J_{st}.\ \exists H_{st}.[I\langle v\rangle H\ \wedge\ (P(v))HJ]$

     b. $\delta_s,\ each \rightsquigarrow \lambda P_{\mathbf{et}}.\lambda v_{\mathbf{e}}.\ \mathbf{dist}_s(P(v))(v)$
   $:= \lambda P_{\mathbf{et}}.\lambda v_{\mathbf{e}}.\lambda I_{st}.\lambda J_{st}.\ \exists H_{st}.[I\langle v\rangle H\ \wedge\ (\mathbf{dist}(P(v)))HJ]$

Both syntactic distributivity operators are translated as functions that take a one-place predicate as argument, and return another one-place predicate. These functions combine with a predicate P and dref u, and update the input info state either with $\mathbf{dist}_w(P(u))(u)$ or with $\mathbf{dist}_s(P(u))(u)$, depending on the strength of the operator.[38]

---

[38] Given these definitions, it's predicted that both *each* and *all* will distribute down to the atoms. In fact, however, in certain contexts *all* allows for 'intermediate' interpretations as in (i), modelled on an example from Lasersohn (1998):

  (i)   The shoes all cost $50.

This examples can be understood as stating that each *pair of shoes* costs $50, rather than each individual shoe. Folllowing Gillon (1987) and Schwarzschild (1996), such examples can be accommodated in the current system by relativizing the $\langle\rangle$-relation to a contextually specified cover C, along the following lines:

  (ii) $\langle u\rangle_C := \lambda I_{st}.\lambda J_{st}.\ \exists f.\ \forall j_s \in J.\ \exists i_s \in I.\ (j \in f(i))\ \wedge$
   $\forall i_s \in I.\ \forall H_{st}.\ (f(i) = H \rightarrow H \subseteq J\ \wedge\ \forall h_s \in H.\ (i[u]h\ \wedge\ uh \in C(ui))\ \wedge\ \oplus uH = ui),$

Now, given the equivalences in (98), it follows that if a strong distributivity operator δ_s is stacked on top of a weak distributivity operator δ_w / *all*, or the other way round, the translation of the structure will be equivalent to that involving a single strong distributivity operator. For instance, the translation of (100) will be equivalent to that of (101):

(100)

      DP

        all δ_s XP

(101)

     DP δ_s XP
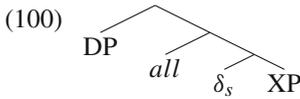
I will assume that stacking as in (100) accounts for strong distributive readings of structures involving overt weak distributivity operators.[39,40]

Footnote 38 continued

    where f is a function from the domain of assignments $\mathcal{D}_s$ to the set of info states $\wp(\mathcal{D}_s)$, and C is a contextually specified function that combines with a sum x and returns a set of sums X such that $\oplus X = x$.

The definition of **dist**$_w$ can then be re-stated in terms of $\langle\rangle_C$. In what follows, I will put aside intermediate readings.

[39] Syntactic constraints should restrict the availability of stacking, ruling out sentences like (i):

    (i)   *Three students all each carried a box.

[40] The definition of the syntactic distributivity operators may need to be modified to account for the ambiguity of pronouns in their scope. For instance, the following example can have a distributive interpretation whereby each lawyer hired a possibly different secretary:

    (i)   The^u lawyers δ_s hired$^\varepsilon$ a$^{u'}$ secretary they liked$^{\varepsilon'}$.

As Kamp and Reyle (1993) point out, the reference of the plural pronoun *they* in this example is ambiguous: it can either refer to each individual lawyer or to the whole group of lawyers referred to by the subject. In the current framework, the distributive interpretation of examples like (i) is derived by positing a null distributivity operator below the subject, which splits the value of the subject dref (u) into atomic part. However, this would mean that the original reference of u to the sum of all the lawyers would no longer be accessible in the scope of the distributivity operator.

We can solve this problem by assuming, in the spirit of Kamp and Reyle (1993), that the distributivity operator does not directly 'split' the referent of the DP it combines with. Instead, it introduces a new dref which returns the same value for the original info state as the dref it takes as argument, and then splits this newly introduced dref:

    (ii) $\delta_s^{u'}$, *each*$^{u'}$ $\leadsto \lambda P_{\mathbf{et}}.\lambda v_{\mathbf{e}}.\ [u' \mid u' = v];\ \mathbf{dist}_s(P(u'))(u')$

Equality between drefs ($u = v$) is defined as follows:

    (iii) $u = u' := \lambda I_{st}.\ I \neq \varnothing\ \wedge\ \forall i \in I.\ (ui = u'i)$

Then, the distributive and group interpretations of plural pronouns in examples like (i) can be distinguished by co-indexing the pronoun either directly with the antecedent DP (for a group interpretation) or with the distributivity operator (for a distributive interpretation). In the following, I will continue to use the simpler translations of the distributivity operators in (99).

5.4.3 *Distributive Quantifying-In*

As things stand at the moment, the system predicts that distributivity operators can only combine with predicates of individual drefs, i.e. with expressions of type **et**. However, given that verbs are predicated over event drefs in the current system, there is no way of directly combining distributivity operators with verbal projections below event closure. Inserting distributivity operators above event closure also results in a type mismatch:

(102)
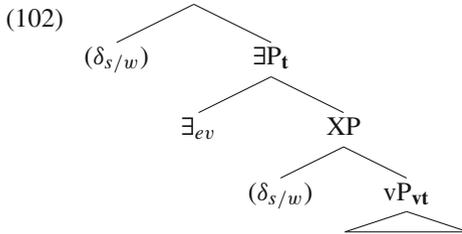
To solve this problem, I will introduce a new rule of translation, *Distributive Quantifying-In (DistrQIn)*, that targets structures like (103), where a distributivity operator is inserted below a DP that has undergone syntactic movement:

(103)
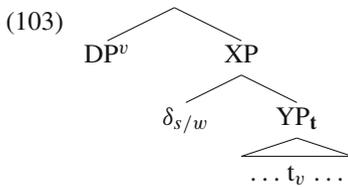
(104)   *Distributive Quantifying-In (DistrQIn)*
If $DP^v \rightsquigarrow \alpha$, B is a distributivity operator (i.e. B $\in \{all, each, \delta_w, \delta_s\}$), B $\rightsquigarrow \delta$, C $\rightsquigarrow \beta$, B and C are daughters of D, and DP^v and D are daughters of E, then E $\rightsquigarrow \alpha(\delta(\lambda v.\beta))$, provided that this is a well-formed term.

5.4.4 Two classes of quantificational determiners

Following Brasoveanu (2008), I will define the translation of quantificational determiners via a general schema which links the semantics of dynamic quantifiers to their standard static counterparts. My treatment of quantificational determiners will be close to that of Krifka (1996), the major innovation being a contrasting analysis of singular and plural QDs.

Both singular and plural quantificational determiners will take two predicates of type **et** as arguments, and return a DRS of the following simplified form, where Ψ is derived from the restrictor predicate of the QD, and Ψ' is derived from its nuclear scope predicate:

(105) $[u', u]$; $[\mathbf{DET}\{u', u\}]$; $\mathbf{max}_{u'}(\Psi(u'))$; $\mathbf{max}_u(\Psi(u); \Psi'(u))$

This schema makes use of the maximization operator **max**, defined in the following way:[41]

(106) $\mathbf{max}_u(D) := \lambda I_{st}.\lambda J_{st}.\, D I J \wedge \forall I'_{st}\, (I[u]I' \wedge \exists J'_{st}.\, (D I' J') \rightarrow \forall i_s \in I.\, \forall i'_s \in I'.\, (i[u]i' \rightarrow ui' \leq ui))$

The **max**-operator is indexed with a dref u, and its function is to update the input state I with a DRS D while ensuring that uI returns the maximal values which render D true with respect to I. In other words, the **max** operator ensures that there is no way of replacing the values that u returns for the info state I for mereologically larger ones, such that D would still be true with respect to this modified input info state.[42]

Consider now the DRS in (105). It does several things. First, it introduces two new drefs. Then, it imposes a condition on the relation between these drefs which invokes the static quantifier (**DET**) corresponding to the dynamic quantifier that is being translated. For instance, the dynamic translation for *most* will involve the static quantifier **MOST**, the translation for *all* will involve the static quantifier **ALL**, etc. Importantly, like most other lexical relations, the static quantifiers themselves are interpreted distributively with respect to the info state they apply to, e.g.:[43]

(107) a. $\mathbf{MOST}(u', u) := \lambda I_{st}.\, I \neq \varnothing \wedge \forall i \in I.\, (\,|ui| > |u'i - ui|\,)$
       b. $\mathbf{ALL}(u', u),\ \mathbf{EVERY}(u', u) := \lambda I_{st}.\, I \neq \varnothing \wedge \forall i \in I.\, (\,|u'i| = |ui|\,)$

Finally, the DRS in (105) ensures that the first of the newly introduced drefs (u') stores the maximal possible values that satisfy the predicate Ψ, derived from the restrictor predicate, while the second (u) stores the maximal possible values that satisfy the (dynamic) conjunction of predicates Ψ and Ψ', the latter derived from the nuclear scope predicate.[44]

The difference between plural and singular quantificational determiners is captured by the way Ψ and Ψ' are defined for the schema in (105). Specifically, in the case of

---

[41] I am grateful to Jakub Dotlačil for helpful suggestions regarding this definition.

[42] The maximization operator defined in (106) requires that each value of a dref in the input info state be maximal with respect to a certain DRS, i.e. it ensures maximality at the assignment level. This can be contrasted with state-level maximality, which would require that the total set of values of a dref in an info state (or the total sum of these values) be maximal (cf. Brasoveanu 2008). Treating maximality as an assignment-level condition ensures that our system correctly handles cases of multiple quantification, as in the following example:

(i) All the students read most of the papers they were assigned.

To get the correct truth conditions for this example, we need the maximal sums of assigned papers and the maximal sums of papers that were read to be calculated separately for each student. In our system, this will be accomplished by calculating maximal sums of papers for each assignment in a plural info state, where each of the assignments will store a separate atomic student as the value of the subject dref (see below on the semantics of plural QDs).

[43] Note that I take the static quantifier to relate sets (or, rather, sums) of individuals, rather than sets of assignments. This allows us to avoid the proportion problem discussed at length in Brasoveanu (2007).

[44] The introduction of two distinct new drefs in the translation of QDs is motivated by the existence of two types of reference to presuppositional quantificational DPs, discussed in detail by Nouwen (2003): reference to the *refset* and to the *maxset* (cf. also related observations in Krifka 1996). These can be illustrated with the help of the following examples from Nouwen (2003):

plural quantifiers such as *all* and *most*, Ψ and Ψ' are obtained via the application of the weak distributivity operator $\mathbf{dist}_w$, in the following way:

(108) $\texttt{det\_pl}^{u,\,u'} \rightsquigarrow \lambda P_{\mathbf{et}}.\lambda P'_{\mathbf{et}}.\,[u',u];\,[\mathbf{DET}\{u',u\}];\,\mathbf{max}_{u'}(\mathbf{dist'}_w(P(u'))(u'));$
$\mathbf{max}_u(\mathbf{dist}_w(P(u);\ P'(u))(u))$

For singular quantifiers such as *each* and *every*, Ψ and Ψ' correspond to the restrictor and nuclear scope predicates taken under the strong distributivity operator $\mathbf{dist}_s$:

(109) $\texttt{det\_sg}^{u,\,u'} \rightsquigarrow \lambda P_{\mathbf{et}}.\lambda P'_{\mathbf{et}}.\,[u',u];\,[\mathbf{DET}\{u',u\}];\,\mathbf{max}_{u'}(\mathbf{dist'}_s(P(u'))(u'));$
$\mathbf{max}_u(\mathbf{dist}_s(P(u);\ P'(u))(u))$

Thus, quantificational determiners introduce two new drefs. The first stores the maximal values that distributively satisfy the restrictor predicate alone, while the second stores the maximal values that distributively satisfy the conjunction of the restrictor and nuclear scope predicates.[45] The difference between plural and singular QDs is the type of distributivity that is involved (weak vs strong). Note that in both (108) and (109), the distributivity operator used to restrict the values of the 'maxset' dref u' is externally static (cf. the definitions in 97), which means that the output info state will not store the 'split' values of u'. On the other hand, the distributivity operator used to restrict the 'refset' dref u is externally dynamic (cf. 93). This entails that the output info states of (108) and (109) will store the individual 'split' values of u, as well as (potentially) the values of other drefs introduced in the restrictor and nuclear scope of the determiner. Consequently, the proposed system is compatible

---

Footnote 44 continued

    (i) a. Few senators admire Kennedy; and they are very junior.

         b. Few senators admire Kennedy. Most of them prefer Carter.

In (ia) the plural pronoun in the second clause refers to the set of senators who admire Kennedy, i.e. the maximal individual that satisfies both the the restrictor and the nuclear scope predicate of the quantifier *few* in the first clause (i.e. the refset). Example (ib), on the other hand, is an instance of reference to the maxset: the plural pronoun in the second sentence picks up the whole set of senators as its referent, i.e. the maximal individual that satisfies the restrictor predicate of the quantifier *few*.

[45] Syntactically, many QDs can combine either directly with NP restrictors, or with PPs headed by the preposition *of*, e.g *all/most/each of the students*. Definite DPs are translated with the same $\mathbf{max}$ operator in (106):

    (i) $\texttt{the}^u$ $\texttt{students} \rightsquigarrow \lambda Q_{\mathbf{et}}.\,[u];\,\mathbf{max}_u([student\{u\}]);\,Q(u).$

Then, the preposition *of* can be translated as a function that takes DPs of type $\mathbf{(et)t}$ and turns them into predicates of type $\mathbf{et}$ in the following way (similar to the *BE* type shifter of Partee 1987):

    (ii) $\texttt{of} \rightsquigarrow \lambda Q_{\mathbf{(et)t}}.\lambda v_{\mathbf{e}}.\,Q(\lambda v'.\,[v \leq v']),$
         where $u \leq u' := \lambda I.\,I \neq \varnothing\ \wedge\ \forall i \in I\ (ui \leq u'i).$

This translation is essentially a dynamic way of saying that *of* extracts the sub-parts of the sum referred to by its complement definite DP. The predicate returned by *of* is then of the right type to combine with a QD. In what follows, I will for simplicity translate all restrictors of QDs as simple NPs, e.g.:

    (iii) $\texttt{of the students} \rightsquigarrow \lambda v.\,[student\{v\}].$

with previous proposals that make use of related semantic frameworks to account for complex cases of cross-sentential anaphora, such as quantificational subordination (cf. e.g. Krifka 1996; Nouwen 2003; Brasoveanu 2007, and discussion in Sect. 6.5). At the same time, the use of externally static distributivity operators to impose conditions on the maxset-drefs allows us to avoid unnecessarily inflating the output info state.[46]

6 Putting it all together

6.1 Cumulative readings

Consider first example (110):

(110) Three^u students carried$^\varepsilon$ boxes$^{u'}$.

This sentence allows for a cumulative interpretation, whereby there are three students, several boxes and a set of carrying events, such that the students are the cumulative agent of these events and the boxes are the cumulative theme of these events. Let us consider how these truth conditions are derived in PCDRT^e.

Assuming that no covert distributivity operators are inserted, sentence (110) is assigned the following structure:

(111)
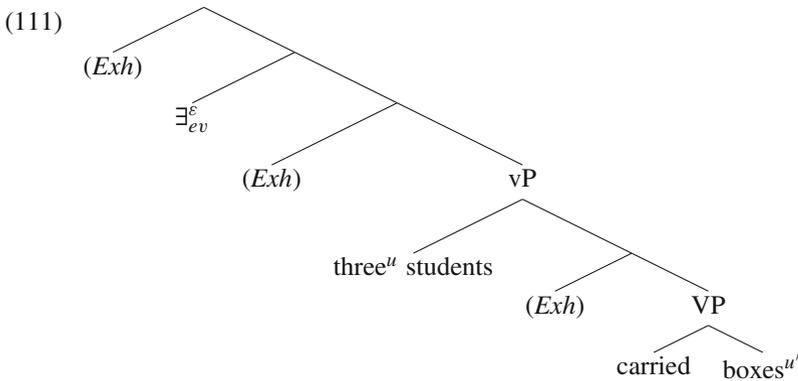
Following our assumptions (cf. 85), an exhaustification operator Exh must be inserted in (111). There are three potential sites: immediately above the VP, immediately above the vP, or above the event closure operator. Suppose Exh is inserted immediately above the VP.

The translation of the VP is the following:

(112) $[_{VP}$ carried boxes$^{u'}] \rightsquigarrow \lambda v'.\lambda \zeta.\lambda I.\lambda J.\ I[u']J\ \wedge\ box\{u'\}J\ \wedge$
$carry\{\zeta\}J\ \wedge\ \mathbf{Th}\{u',\ \zeta\}J\ \wedge\ \mathbf{Ag}\{v',\ \zeta\}J$

---

[46] If we replace the externally static distributivity operators with their externally dynamic counterparts in (108) and (109), the output info states would contain assignments corresponding to each possible combination of the 'split' values of u' and the 'split' values of u. Although this wouldn't in principle be problematic, it seems to be an unnecessary complication of the structure of the info state.

Exh is translated as an **Exh** operator. A generalized definition of **Exh** operators was given above in (86). In combination with a term of type **e(vt)**, **Exh** is defined as follows:

(113) $\mathbf{Exh}_{\text{ALT (e(vt))(e(vt))}}(V_{\mathbf{e(vt)}}) := \lambda v_{\mathbf{e}}.\lambda \zeta_{\mathbf{v}}.\lambda I_{st}.\lambda J_{st}.\ V(v)(\zeta)IJ \wedge$
$\forall V'_{\mathbf{e(vt)}}.\ (V' \in \text{ALT} \wedge V' \succ V \rightarrow \neg V'(v)(\zeta)IJ),$
where ALT is the set of alternatives for V.

The definition in (113) requires for all stronger alternatives to be negated. The set of alternatives for (112) includes the following term, which is the translation of the corresponding VP with the plural indefinite direct object replaced by a singular indefinite :

(114) $\lambda v'.\lambda \zeta.\lambda I.\lambda J.\ I[u']J \wedge \mathbf{atom}\{u'\}J \wedge \mathbf{unique}\{u'\}J \wedge box\{u'\}J \wedge carry\{\zeta\}J \wedge$
$\mathbf{Th}\{u',\ \zeta\}J \wedge \mathbf{Ag}\{v',\ \zeta\}J$

To perform exhaustification, we must determine whether the alternative in (114) is stronger than (112). Since (114) only differs from (112) in that it imposes some additional restrictions on the values of u' in the output info state (i.e. it requires for u' to return the same atomic individual for all the assignment), it follows that for any info states I and J, event dref ε and individual dref u, if u, ε, I and J satisfy (114), then they also satisfy (112). The converse, however, does not hold. Hence, following the definition of strength in (89), we can conclude that (114) is indeed stronger than (112).

The expression in (112) must then be strengthened via the negation of the stronger alternative:

(115) $\lambda v'.\lambda \zeta.\lambda I.\lambda J.\ I[u']J \wedge box\{u'\}J \wedge carry\{\zeta\}J \wedge \mathbf{Th}\{u',\ \zeta\}J \wedge \mathbf{Ag}\{v',\ \zeta\}J \wedge$
$\neg(I[u']J \wedge \mathbf{atom}\{u'\}J \wedge \mathbf{unique}\{u'\}J \wedge box\{u'\}J \wedge carry\{\zeta\}J \wedge$
$\mathbf{Th}\{u',\ \zeta\}J \wedge \mathbf{Ag}\{v',\ \zeta\}J)$
$:= \lambda v'.\lambda \zeta.\lambda I.\lambda J.\ I[u']J \wedge box\{u'\}J \wedge carry\{\zeta\}J \wedge \mathbf{Th}\{u',\ \zeta\}J \wedge$
$\mathbf{Ag}\{v',\ \zeta\}J \wedge (\neg\mathbf{atom}\{u'\}J \vee \neg\mathbf{unique}\{u'\}J)$

We then arrive at the following DRS as the translation of (112):

(116) $\lambda I_{st}.\lambda J_{st}.\ I[\varepsilon,\ u,\ u']J \wedge \mathbf{3\_atoms}\{u\}J \wedge student\{u\}J \wedge box\{u'\}J \wedge$
$carry\{\varepsilon\}J \wedge \mathbf{Th}\{u',\ \varepsilon\}J \wedge \mathbf{Ag}\{u,\ \varepsilon\}J \wedge (\neg\mathbf{atom}\{u'\}J \vee \neg\mathbf{unique}\{u'\}J)$

This DRS introduces three new drefs: an event-dref ε, and two individual-drefs u and u', and imposes conditions on the values of these drefs in the output info state J. Assuming a singleton input info state I, the output state J will have the following form:

(117)

Info state J		...	ε	u	u'	...
j		...	e	$s_1 \oplus s_2 \oplus s_3$	$b_1 \oplus \ldots \oplus b_n$	...

Given our definition of new dref introduction ([], cf. 60), if I is singleton, J must also be singleton, i.e. $J = \{j\}$. Then, according to (116), and given that we assume lexical predicates to be cumulative, εj must be a sum of carrying events, uj must be a sum students, and $u'j$ must be a sum of books. The numeral in (110) is translated as a cardinality predicate, which requires for uj to be a sum of 3 individuals. Furthermore, given that the thematic relations (**Ag** and **Th**) are also cumulative, it follows that the three students (uj) are the cumulative agent in e, and the books ($u'j$) are the cumulative theme in e. Finally, (116) states that u' must either return different values for some two assignments in J or return a non-atomic individual for some assignment in J. Given that J is singleton, it follows that $u'j$ must be a non-atomic sum of boxes. The DRS in (116) will be true (relative to input state I) iff such a J exists, which is equivalent to saying that there exists a sum of carrying events, a sum of three students and a non-atomic sum of boxes that stand in the defined cumulative relations to each other. We have thus derived the cumulative interpretation of sentence (110).

Interestingly, the same truth conditions are derived in our system if Exh is inserted above the vP or above event closure in (111). Consider the latter option. In this case, the Exh operator combines with the following DRS:

(118) $\lambda I_{st}.\lambda J_{st}.\ I[\varepsilon,\ u,\ u']J\ \wedge\ \mathbf{3\_atoms}\{u\}J\ \wedge\ student\{u\}J\ \wedge\ box\{u'\}J\ \wedge\ carry\{\varepsilon\}J\ \wedge\ \mathbf{Th}\{u',\ \varepsilon\}J\ \wedge\ \mathbf{Ag}\{u,\ \varepsilon\}J$

This is compared to the alternative in (119):

(119) $\lambda I_{st}.\lambda J_{st}.\ I[\varepsilon,\ u,\ u']J\ \wedge\ \mathbf{3\_atoms}\{u\}J\ \wedge\ student\{u\}J\ \wedge\ \mathbf{atom}\{u'\}J\ \wedge\ \mathbf{unique}\{u'\}J\ \wedge\ box\{u'\}J\ \wedge\ carry\{\varepsilon\}J\ \wedge\ \mathbf{Th}\{u',\ \varepsilon\}J\ \wedge\ \mathbf{Ag}\{u,\ \varepsilon\}J$

The alternative in (119) is stronger than (118): any pair of info states that satisfy (119) would also satisfy (118). Hence, following the translation of the Exh-operator in (120), the stronger alternative is negated, once again yielding (116) as the translation of the structure in (111).

(120) $\mathbf{Exh}_{\text{ALT}\,\mathbf{tt}}(D_{\mathbf{t}}) := \lambda I_{st}.\lambda J_{st}.\ DIJ\ \wedge\ \forall D'.\ (D' \in \text{ALT} \wedge D' \succ D \rightarrow \neg D'IJ)$, where ALT is the set of alternatives for D.

6.2 Dependent plurals

6.2.1 Deriving co-distributivity

Consider now sentence (121) that minimally differs from (110) it that a weak distributivity operator has been inserted below the subject:

(121) Three^u students [all/δ_w carried$^\varepsilon$ boxes$^{u'}$]

Ignoring the exhaustification operator for the time being, this sentence is assigned the structure in (122), with the subject raised from its base position to a position above the distributivity operator:

(122)

$$
\begin{array}{l}
\text{[Three}^u \text{ students]}^v \\
\qquad \text{all}/\delta_w \\
\qquad\qquad \exists^\varepsilon_{ev} \qquad \text{vP} \\
\qquad\qquad\qquad\qquad t_v \qquad \text{VP} \\
\qquad\qquad\qquad\qquad\qquad\qquad \text{carried} \quad \text{boxes}^{u'}
\end{array}
$$
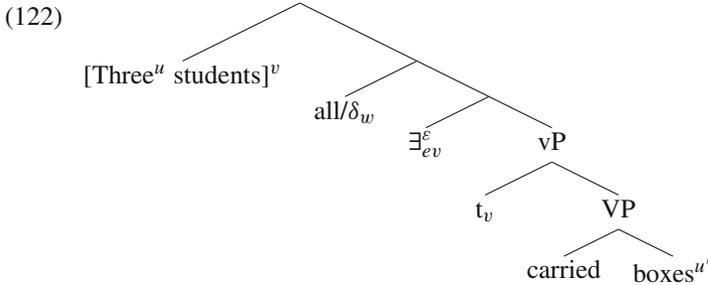
The compositional translation of the vP combined with the event closure operator is the following:

(123)

$[\varepsilon]$; $[u']$; $[box\{u'\}]$;
$[carry\{\varepsilon\}$, $\mathbf{Th}\{u', \varepsilon\}$, $\mathbf{Ag}\{v, \varepsilon\}]$

$\exists^\varepsilon_{ev}$ vP
$\lambda V. [\varepsilon]; V(\varepsilon)$ $\lambda\zeta. [u']$; $[box\{u'\}]$;
$[carry\{\zeta\}$, $\mathbf{Th}\{u', \zeta\}$, $\mathbf{Ag}\{v, \zeta\}]$

t_v VP
v $\lambda v'.\lambda\zeta. [u']$; $[box\{u'\}]$;
$[carry\{\zeta\}$, $\mathbf{Th}\{u', \zeta\}$, $\mathbf{Ag}\{v', \zeta\}]$
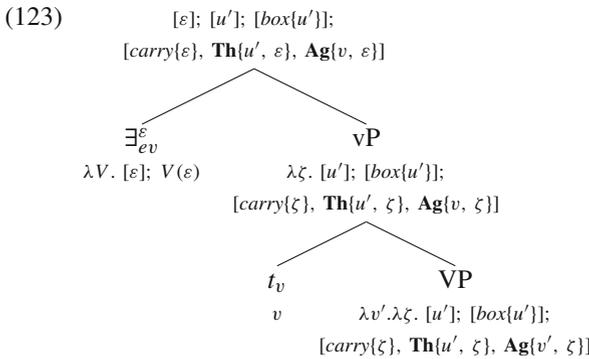
This is then combined with the distributivity operator and the raised subject. The resulting structure is translated via the Distributive Quantifying-In rule in (104) to yield the following DRS:

(124)

$\texttt{three}^u \texttt{ students } \delta_w \texttt{ carried}^\varepsilon \texttt{ boxes}^{u'}$

$[u]$; $[\mathbf{3\_atoms}\{u\}]$; $[student\{u\}]$;
$\mathbf{dist}_w([\varepsilon]$; $[u']$; $[box\{u'\}]$; $[carry\{\varepsilon\}$, $\mathbf{Th}\{u', \varepsilon\}$, $\mathbf{Ag}\{u, \varepsilon\}])(u)$

$[\texttt{three}^u \texttt{ students}]^v$
$\lambda P'. [u]$; $[\mathbf{3\_atoms}\{u\}]$;
$[student\{u\}]$; $P'(u)$ δ_w $[\varepsilon]$; $[u']$; $[box\{u'\}]$;
$\lambda P.\lambda v'. \mathbf{dist}_w(P(v'))(v')$ $[carry\{\varepsilon\}$, $\mathbf{Th}\{u', \varepsilon\}$, $\mathbf{Ag}\{v, \varepsilon\}]$
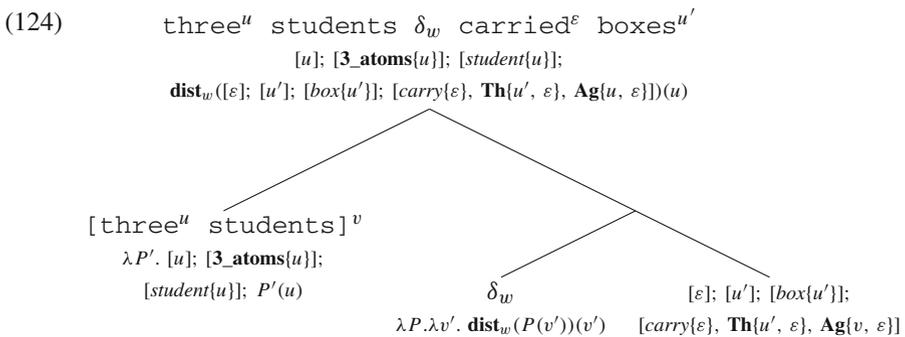
Unpacking the $\mathbf{dist}_w$ operator, we obtain the following DRS:

(125) $\lambda I.\lambda J.\exists H$ ([u]; [**3_atoms**{u}]; [$student${u}])IH $\wedge$
      $\exists H'.(H\langle u\rangle H' \wedge$ ([ε]; [u']; [box{u'}]; [$carry${ε}, **Th**{u', ε}, **Ag**{u, ε}])$H'J$)

Let us consider the truth conditions of this DRS in detail. It will be true with respect to a singleton input info state I iff there exists an info state J such that the following conditions are met:

(a) there is an info state H such that the assignments in H differ from the assignments in I at most with respect to the values for u, and u returns a sum of three students for each assignment in H. Given that I is singleton, H is also singleton:

(126)

Info state H, s.t. $I[u]H$	$\ldots$	u	$\ldots$
h	$\ldots$	$s_1 \oplus s_2 \oplus s_3$	$\ldots$

(b) There exists an info state H', such that $H\langle u\rangle H'$, i.e. the value of u in H is 'split' into its atomic parts, which are returned as values of u in H':

(127)

Info state H', s.t. $H\langle u\rangle H'$	$\ldots$	u	$\ldots$
h'_1 (s.t. $h_1[u]j$)	$\ldots$	s_1	$\ldots$
h'_2 (s.t. $h_2[u]j$)	$\ldots$	s_2	$\ldots$
h'_3 (s.t. $h_3[u]j$)	$\ldots$	s_3	$\ldots$

(c) The output info state J differs from H' at most with respect to the values for ε and u', such that for every assignment j in J, $u'j$ is a (possibly atomic) sum of boxes, εj is a (possibly atomic) sum of carrying events, uj is the agent of εj, and $u'j$ is the theme of εj. This is illustrated in (128):

(128)

Info state J, s.t. $H'[u', \varepsilon]J$	$\ldots$	u	ε	u'	$\ldots$
j_1 (s.t. $j_1[u', \varepsilon]h'_1$)	$\ldots$	s_1	e_1	b_1	$\ldots$
j_2 (s.t. $j_2[u', \varepsilon]h'_2$)	$\ldots$	s_2	e_2	b_2	$\ldots$
j_3 (s.t. $j_3[u', \varepsilon]h'_3$)	$\ldots$	s_3	e_3	b_3	$\ldots$

Note, that the values of u' with respect to the assignments in J (b_1, b_2 and b_3 in 128) can be either atomic or non-atomic on the assignment level, and can be either distinct or identical on the state level. For instance, the truth conditions of the DRS in (125) are compatible with a scenario where three students each carried a different box. We have thus derived the co-distributive reading of sentence (121).

6.2.2 Deriving the multiplicity condition

To fully capture the dependent plural interpretation we must also derive the overarching Multiplicity Condition associated with the dependent plural: (121) will not be judged

true if all the students carried the same box. Thus, we must ensure that the sum of the boxes carried by the students is greater than one. I will now demonstrate that the Multiplicity Condition is derived as an implicature in the current system, in exactly the same way as the non-atomicity requirement associated with non-dependent bare plurals (cf. the discussion of example (110) above).

To satisfy the requirement in (85), an exhaustification operator must be inserted into the structure in (122). There are four potential insertion sites for *Exh*: above the VP, above the vP, above event closure, and at the root above the raised subject.[47] Inserting the *Exh* in the lowest position, we again derive (115) as the strengthened translation of the VP, and the DRS in (129) as the translation of example (121):

(129) $\lambda I.\lambda J.\exists H.\ I[u]H\ \wedge\ \textbf{3\_atoms}\{u\}H\ \wedge\ student\{u\}H\ \wedge$
$\exists H'.\ (H\langle u\rangle H'\ \wedge\ H'[\varepsilon,\ u']J\ \wedge\ box\{u'\}J\ \wedge\ carry\{\varepsilon\}J\ \wedge\ \textbf{Th}\{u',\ \varepsilon\}J\ \wedge$
$\textbf{Ag}\{u,\ \varepsilon\}J\ \wedge\ (\neg\textbf{atom}\{u'\}J\ \vee\ \neg\textbf{unique}\{u'\}J))$

This DRS is very similar to the one in (125), except that (129) requires for the values of u' to be either non-atomic or non-unique with respect to the output info state J, i.e. u' must either return a non-atomic individual for some assignment in J or return different individuals for some assignments in J. For convenience, I reproduce the table representing J here:

(130)

Info state J	$\ldots$	u	ε	u'	$\ldots$
j_1	$\ldots$	s_1	e_1	b_1	$\ldots$
j_2	$\ldots$	s_2	e_2	b_2	$\ldots$
j_3	$\ldots$	s_3	e_3	b_3	$\ldots$

Thus, the DRS in (129) will be true iff there exist three atomic students s_1, s_2 and s_3, three (sums of) carrying events e_1, e_2 and e_3, and three sums of boxes b_1, b_2 and b_3, such that s_1 carried b_1 in e_1, s_2 carried b_2 in e_2, and s_3 carried b_3 in e_3. Moreover, following the conditions in (129), it must be the case that either b_1, b_2 or b_3 is a non-atomic sum of boxes, or $\neg(b_1 = b_2 = b_3)$. In other words, (129) will be true iff there are three students such that they each carried one or more boxes, and more than one box was carried overall. This amounts to a co-distributive reading combined with a global Multiplicity Condition, i.e. a dependent plural reading.

If *Exh* is inserted directly above the vP, directly above the event closure operator or at the root in (122) the result is again the DRS in (129).[48] Thus, for sentences like (121) our system derives the overarching Multiplicity Condition irrespective of the choice of exhaustification site.

---

[47] Formally, there is a fifth option of inserting *Exh* immediately above the distributivity operator, below the subject. However, this structure would be uninterpretable in the current system since the exhaustivity operator would block the application of the Distributive Quantifying-In rule.

[48] For reasons of space, I leave the calculation of the relevant translations to the reader.

6.2.3 Dependent plurals under plural quantifiers

Consider now the interpretation of bare plural indefinites in the scope of plural quantifiers. Take example (131), with the structure in (132) (leaving out the Exh-operator for now):

(131) Most$^{u,u'}$ students carried$^{\varepsilon}$ boxes$^{u''}$.

(132)
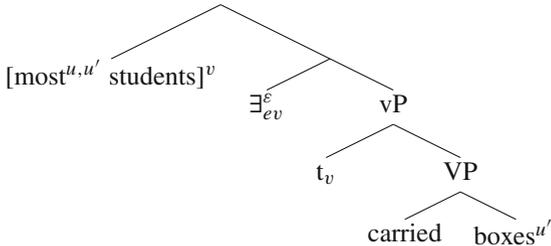
The translation of the subject DP, given the schema in (108), is the following:

(133) most$^{u,u'}$ students $\rightsquigarrow \lambda P'_{et}.\ [u', u];\ [\mathbf{MOST}\{u', u\}];\ \mathbf{max}_{u'}$
$(\mathbf{dist'}_w([student(u')])(u'));\ \mathbf{max}_u(\mathbf{dist}_w([student(u)];\ P'(u))(u));$

The translation of the VP was already calculated above. I repeat it here:

(134) [$_{VP}$ carried boxes$^{u''}$] $\rightsquigarrow \lambda v'.\lambda \zeta.\lambda I.\lambda J.\ I[u'']J\ \wedge\ box\{u''\}J\ \wedge$
$carry\{\zeta\}J\ \wedge\ \mathbf{Th}\{u'',\ \zeta\}J\ \wedge\ \mathbf{Ag}\{v',\ \zeta\}J$

In order to satisfy the condition in (85), an exhaustivity operator must be inserted in a position c-commanding the bare plural in (132). Suppose Exh is inserted directly above the VP, yielding the following strengthened term:

(135) $\lambda v'.\lambda \zeta.\lambda I.\lambda J.\ I[u'']J\ \wedge\ box\{u''\}J\ \wedge\ carry\{\zeta\}J\ \wedge\ \mathbf{Th}\{u'',\ \zeta\}J\ \wedge$
$\mathbf{Ag}\{v',\ \zeta\}J\ \wedge\ (\neg\mathbf{atom}\{u''\}J\ \vee\ \neg\mathbf{unique}\{u''\}J)$

This is then combined with the subject trace and the event closure operator to yield the following DRS:

(136) $\exists^{\varepsilon}_{ev}$ t$_v$ carry boxes$^{u''}$ $\rightsquigarrow \lambda I.\lambda J.\ I[\varepsilon, u'']J \wedge box\{u''\}J \wedge carry\{\varepsilon\}J \wedge$
$\mathbf{Th}\{u'',\ \varepsilon\}J\ \wedge\ \mathbf{Ag}\{v,\ \varepsilon\}J\ \wedge\ (\neg\mathbf{atom}\{u''\}J\ \vee\ \neg\mathbf{unique}\{u''\}J)$

Finally, after combining the subject in (133) with the DRS in (136) via the Quantifying-In rule, we arrive at the following translation for (131):

(137) $[u', u];\ [\mathbf{MOST}\{u', u\}];\ \mathbf{max}_{u'}(\mathbf{dist'}_w([student\{u'\}])(u'));$
$\mathbf{max}_u(\mathbf{dist}_w(\lambda I.\lambda J.\ student\{u\}I\ \wedge\ I[\varepsilon, u'']J\ \wedge\ box\{u''\}J\ \wedge\ carry\{\varepsilon\}J\ \wedge$
$\mathbf{Th}\{u'',\ \varepsilon\}J\ \wedge\ \mathbf{Ag}\{u,\ \varepsilon\}J\ \wedge\ (\neg\mathbf{atom}\{u''\}J\ \vee\ \neg\mathbf{unique}\{u''\}J))(u))$

Lets us consider the four parts of this DRS in turn, assuming a singleton input info state I. First, the DRS in (137) introduces two new drefs, u' and u, giving rise to a new info state, call it K:

(138)

Info state K, s.t. $I[u', u]K$	...	u'	u	...
k	...	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	$s_1 \oplus s_2 \oplus \ldots \oplus s_n$	...

The values that these drefs return for the assignment in K are then compared by the quantifier **MOST**. I repeat its definition here for convenience:

(139) $\mathbf{MOST}(u', u) := \lambda I_{st}.\ I \neq \varnothing \ \wedge\ \forall i \in I.\ (\ |ui| > |u'i - ui|\ )$

Following this definition, the sum of individuals that u returns for the assignment in K (the sum $s_1 \oplus s_2 \oplus \ldots \oplus s_n$ in 138) must have a cardinality greater than the cardinality of the complement of that sum in the sum of individuals returned by u' (the sum $s'_1 \oplus s'_2 \oplus \ldots \oplus s'_n$ in 138).

The static quantifier in (137) is followed by a maximization operator. For convenience, I repeat the definition of **max** from (106):

(140) $\mathbf{max}_u(D) := \lambda I_{st}.\lambda J_{st}.\ DIJ \ \wedge\ \forall I'_{st}\ (I[u]I' \wedge \exists J'_{st}.\ (DI'J') \rightarrow \forall i_s \in I.\ \forall i'_s \in I'.\ (i[u]i' \rightarrow ui' \leq ui))$

Consider the role of the first **max** operator in (137). Its argument is the DRS $\mathbf{dist'}_w([student\{u'\}])(u')$. Given the definitions in (97), this DRS can be unwrapped in the following way:

(141) $\mathbf{dist'}_w([student\{u'\}])(u')$
$:= \lambda I_{st}.\ \lambda J_{st}.\ I = J \ \wedge\ \exists J'.\ (\langle u'\rangle;\ [student\{u'\}])IJ'$

Combining it with the **max** operator produces the following DRS:

(142) $\mathbf{max}_{u'}(\mathbf{dist'}_w([student\{u'\}])(u')) :=$
$\lambda I_{st}.\lambda J_{st}.\ I = J \ \wedge\ \exists J'.\ (\langle u'\rangle;\ [student\{u'\}])IJ' \ \wedge$
$\forall I'_{st}\ (I[u']I' \ \wedge\ \exists J'_{st}.\ (I' = J' \ \wedge\ \exists J''.\ (\langle u'\rangle;\ [student\{u'\}])I'J'') \rightarrow$
$\forall i_s \in I.\ \forall i'_s \in I'.\ (i[u']i' \rightarrow u'i' \leq u'i))$

The DRS in (142) is a test, i.e. it does not re-assign the values of any drefs in the output info state. Instead, it requires for the values of u' in the input info state I to be such that if they are split into their atomic parts, each of those parts would be in the denotation of $student$, i.e. $u'i$ must return a sum of students for every $i \in I$. Furthermore, its requires that there be no way to re-assign the values of u', e.g. in I', such that u' returns larger sums of students in I' than for the corresponding assignments in I. When applied to the info state K in (138), this entails that $u'k$ must return the maximal sum of students.

The final part of (137) involves another maximization operator, but in this case it is applied to a conjunction of the nuclear scope and restrictor predicates taken under the weak distributivity operator. Replacing the **max** and $\mathbf{dist}_w$ operators with their definitions in (140) and (93), we arrive at the following DRS:

(143) $\lambda I.\lambda J.\exists H_{st}.\ I\langle u\rangle H \wedge student\{u\}H \wedge H[\varepsilon,\ u'']J \wedge box\{u''\}J \wedge carry\{\varepsilon\}J \wedge$
$\mathbf{Th}\{u'',\ \varepsilon\}J \ \wedge\ \mathbf{Ag}\{u,\ \varepsilon\}J \ \wedge\ (\neg\mathbf{atom}\{u''\}J \ \vee\ \neg\mathbf{unique}\{u''\}J) \wedge$

$$\forall I'_{st}. \ (I[u]I' \ \wedge \ \exists J'_{st}. \ \exists H'_{st}. \ I'\langle u\rangle H' \ \wedge \ student\{u\}H' \ \wedge \ H'[\varepsilon, \ u'']J' \ \wedge$$
$$box\{u''\}J' \ \wedge \ carry\{\varepsilon\}J' \ \wedge \ \mathbf{Th}\{u'', \ \varepsilon\}J' \ \wedge \ \mathbf{Ag}\{u, \ \varepsilon\}J' \ \wedge \ (\neg\mathbf{atom}\{u''\}J' \ \vee$$
$$\neg\mathbf{unique}\{u''\}J')) \rightarrow \forall i_s \in I. \ \forall i'_s \in I'. \ (i[u]i' \rightarrow ui' \le ui)$$

Because $\mathbf{dist}_w$ is externally dynamic, the DRS in (143) is not a test. Applied to the info state K in (138), it first splits the value for u in K into its component atoms, producing an new info state H:

(144)

Info state H, s.t. $K\langle u\rangle H$		$\ldots$	u'	u	$\ldots$
h_1		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_1	$\ldots$
h_2		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_2	$\ldots$
$\ldots$		$\ldots$	$\ldots$	$\ldots$	$\ldots$
h_n		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_n	$\ldots$

This info state is then updated with the event dref ε and individual dref u'' resulting in the output info state J:

(145)

Info state J		$\ldots$	u'	u	ε	u''	$\ldots$
j_1		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_1	e_1	b_1	$\ldots$
j_2		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_2	e_2	b_2	$\ldots$
$\ldots$		$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
j_n		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_n	e_n	b_n	$\ldots$

The DRS in (143) places a number of restrictions on the values of the drefs in J. Specifically, $s_1, s_2, \ldots, s_n$ are (atomic) students, $e_1, e_2, \ldots, e_n$ are carrying events, and $b_1, b_2, \ldots, b_n$ are possibly atomic sums of boxes. Next, s_1 is the agent of e_1, s_2 is the agent of e_2, ..., s_n is the agent of e_n. Similarly, b_1 is the theme of e_1, b_2 is the theme of e_2, ..., b_n is the theme of e_n. It must also be the case that the cardinality of the sum $b_1 \oplus b_2 \oplus \ldots \oplus b_n$ is greater than one. This output info state will exist if there is a set of students, where each student carried a (possibly atomic) sum of boxes, and more than one box was carried overall.

Finally, the DRS in (143) when applied to K requires that there be no K', which differs from K at most in the value for u such that: a) there exist info states H' and J' satisfying the same conditions as H and J above; and b) the sum of individuals that u returns for the assignment in K is a proper sub-sum of the sum of individuals that u returns for the assignment in K'. In other words, uk in (138), i.e. $s_1 \oplus s_2 \oplus \ldots \oplus s_m$, is the maximal sum of students such that each student carried a (possibly atomic) sum of boxes, and more than one box was carried overall.

Now, recall that the quantifier $\mathbf{MOST}$ in (137) requires for the sum of individuals that u returns for the assignment in K ($s_1 \oplus s_2 \oplus \ldots \oplus s_n$) to have a cardinality greater than the cardinality of the complement of that sum in the sum of individuals returned by u' in K ($s'_1 \oplus s'_2 \oplus \ldots \oplus s'_n$). This means that the cardinality of the maximal sum of students S such that each student in S carried a (possibly atomic) sum of boxes and together the students in S carried more than one box, must be greater than half the cardinality of the maximal sum of students.

Summing up, the DRS in (137) will be true iff there is a sum of students S, such that each student in S carried a possibly atomic sum of boxes and more than one box

was carried overall, S is the maximal sum of students satisfying this condition, and the cardinality of S is greater than half the cardinality of the maximal sum of students. This adequately captures the dependent plural interpretation of sentence (131).

Attaching the *Exh*-operator at other sites in (132) (e.g. above the vP, above the subject trace, above the event closure operator or at the root) yields the same DRS in (137) (cf. Appendix 1).

6.3 Accounting for the ban on numerals

6.3.1 Numerals and weak distributivity

Consider now the interpretation of plural DPs containing numerals and cardinal modifiers in the scope of weak distributivity operators and plural quantificational items.

Take the following example with the structure in (147):

(146) Three^u students *all* / δ_w carried$^\varepsilon$ four$^{u'}$ boxes.

(147)

[Three^u students]^v

δ_w

$\exists_{ev}^\varepsilon$ vP

t$_v$ VP

carried   four$^{u'}$ boxes
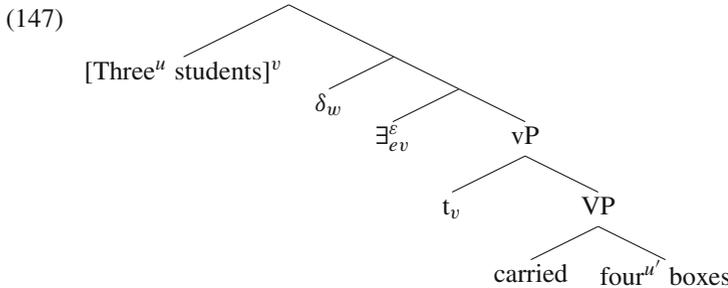
The translation of this sentence in PCDRT^e is parallel to that of (121) discussed above, except for the cardinality condition contributed by the numeral:

(148) $[u]$; $[\mathbf{3\_atoms}\{u\}]$; $[student\{u\}]$;
$\mathbf{dist}_w([\varepsilon]$; $[u']$; $[\mathbf{4\_atoms}\{u'\}]$; $[box\{u'\}]$; $[carry\{\varepsilon\}$, $\mathbf{Th}\{u', \varepsilon\}$,
$\mathbf{Ag}\{u, \varepsilon\}])(u)$

The DRS in (148) will be true (with respect to a singleton input info state I) iff there exists an output info state like the following, where the value for the subject dref u is distributed (by virtue of the weak distributivity operator) across multiple assignments:

(149)

Info state J	...	u	ε	u'	...
j_1	...	s_1	e_1	b_1	...
j_2	...	s_2	e_2	b_2	...
j_3	...	s_3	e_3	b_3	...

Here, s_1, s_2 and s_3 are distinct atomic sums of students, b_1, b_2 and b_3 are sums of boxes, and e_1, e_2 and e_3 are carrying events whose agents are s_1, s_2 and s_3, respectively, and whose themes are b_1, b_2 and b_3, respectively. The cardinality condition

4_atoms$\{u'\}$ in (148) is interpreted distributively, i.e. with respect to each assignment in J. This means that b_1, b_2 and b_3 are all sums of four boxes. An info state like this will exists iff there are three students who each carried four boxes, i.e. we end up with truth conditions for (147) where the subject has distributive scope over the direct object. Thus, the proposed analysis correctly captures the fact that floating quantifiers as in (146) block co-distributive interpretations between the subject and the direct object involving a numeral. In the proposed system this follows from the fact that distributivity operators split the values of the subject dref across multiple assignments or info states, while the cardinality conditions imposed by numerals apply to the values of the object dref for each assignment.[49]

Consider now example (150), and its translation in (151):

(150) Most$^{u,u'}$ students carried$^\varepsilon$ four$^{u''}$ boxes.

(151) $[u', u]$; $[\textbf{MOST}\{u', u\}]$; $\textbf{max}_{u'}(\textbf{dist'}_w([student\{u'\}])(u'))$;
$\textbf{max}_u(\textbf{dist}_w([student\{u\}]; [\varepsilon, u'']; [\textbf{4\_atoms}\{u''\}]; [box\{u''\}];$
$[carry\{\varepsilon\}]; [\textbf{Th}\{u'', \varepsilon\}]; [\textbf{Ag}\{u, \varepsilon\}])(u))$

Sentence (150) does not allow for a co-distributive interpretation, i.e. the numeral cannot be understood as specifying the total number of boxes carried by the students. Instead, this sentence must be interpreted distributively with the cardinality condition evaluated relative to the individual students. This result, again, follows directly from our analysis of QDs and numerals.

Consider the DRS in (151). It's truth conditions mirror those of (137) discussed above, up to the translation of the object DP. It introduces two new drefs, u' and u, as in (138), such that the cardinality of the sum of individuals returned by u (i.e. uk in 138) is greater than half the cardinality of the sum returned by u' ($u'k$). Moreover, u' must return the maximal sum of students. The lack of a co-distributive interpretation is explained by the conditions that the DRS in (151) imposes on the values of u and u''. It requires u to return the maximal sum of individuals uk such that there exists an output info state J of the following form:

(152)

Info state J	$\ldots$	u	ε	u''	$\ldots$
j_1	$\ldots$	s_1	e_1	b_1	$\ldots$
j_2	$\ldots$	s_2	e_2	b_2	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
j_n	$\ldots$	s_n	e_n	b_n	$\ldots$

Here, the value for u has been split by the weak distributivity operator into its atomic sub-parts $s_1, s_2, \ldots, s_n$. Each of these individuals must be a student. Next, ε and u'' have been introduced, such that the values of ε are carrying events whose agents are

---

[49] It is easy to see that the same truth conditions will obtain if we replace *all* with *each* in (146). The floating *each* is analysed as a strong distributivity operator, which will split the value of the subject dref into atomic parts distributed across multiple singleton info states. The cardinality condition on the direct object dref will then apply distributively to each of these singleton info states.

the values of u, and whose themes are the values of u''. Finally, u'' must return a sum of four boxes for each assignment in J. Crucially, since the cardinality condition imposed on the values of u'' occurs in the scope of the weak distributivity operator that splits the value of the agent dref, it will apply distributively relative to each atomic value of the agent dref, i.e. each student must have carried four boxes. Thus, the presence of the weak distributivity operator in the semantics of plural QDs, in combination with the fact that numerals encode assignment-level cardinality conditions, explains the lack of co-distributive interpretations in contexts involving plural quantificational items scoping over DPs with numerals.[50]

6.3.2 Modified numerals

The analysis extends naturally to all types of cardinality modifiers (e.g., modified numerals of various types, modifiers like *several* and *multiple*, etc.). As long as the cardinality conditions associated with these items are taken to apply at the level of the individual assignments, we predict that they will not be interpreted co-distributively in the scope of weak distributivity operators and plural quantificational DPs.

For concreteness, let us consider the semantics of comparative numerals like *fewer than 10*, which as we have seen pose problems for cumulativity-based approaches to the semantics of plural quantifiers. Take sentence (42), repeated in (153):

(153)   All the students made fewer than 10 mistakes.

In examples like this, modified numerals receive a *maximal interpretation*, i.e. sentence (153) can be paraphrased as stating that for each student the maximal number of mistakes that they made was less than 10.[51]

To account for this reading, I will incorporate an analysis of comparative numerals along the lines of Hackl (2000) and Heim (2000). First, we need to add a new basic type d for degrees. The domain of d is the set of non-negative integers, with the $<$ relation defined in the standard way. Next, we need to introduce the notion of drefs over degrees (functions from assignments to degrees, type sd). I will use $d, d_1, d_2 \ldots$ for constants of type sd, and $n, n_1, n_2 \ldots$ for variables of type sd.

Following Hackl (2000), I will assume that the structure of comparative numerals (154) involves a phonologically null parametrized determiner *many*, with the translation in (155):

(154)

---
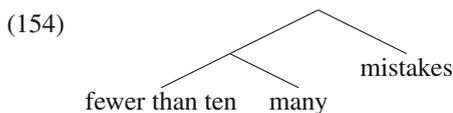
[50] The proposed analysis also derives only distributive readings for DPs with numerals in the scope of singular quantificational DPs like *each/every student*.

[51] In other contexts, comparative numerals can receive a non-maximal, existential interpretation. I will not address the complex issue of the source and nature of this variation (cf. the discussion and references in Buccola and Spector 2016). The aim of this section is simply to demonstrate how maximal readings of numerals can be treated in the proposed system.

(155) $\text{many}^u \rightsquigarrow \lambda n_{sd}.\lambda P_{\mathbf{et}}.\lambda P'_{\mathbf{et}}.\ [u];\ [\mathbf{card}\{u, n\}];\ P(u);\ P'(u),$
where $\mathbf{card}\{u, d\} := \lambda I_{st}.\ \forall i \in I.(|\{x_e : x \leq ui\ \wedge\ \mathbf{atom}(x)\}| = di)$

Defined in this way, *many* combines with a dref over degrees n and two dynamic predicates, P and P', and introduces a new individual dref u satisfying P and P', such that for each assignment i in the updated info state the cardinality of ui is equal to ni.

The comparative numeral is treated as a generalized quantifier over degrees: it combines with a predicate of degree drefs and places a condition on the value of the *maximal* degree that satisfies that predicate:

(156) $\text{fewer}^d\ \text{than}\ 10 \rightsquigarrow \lambda M_{(sd)\mathbf{t}}.\ [d];\ \mathbf{max}_d(M(d));\ [d < 10],$
where $d < 10 := \lambda I_{st}.\ I \neq \varnothing\ \wedge\ \forall i \in I.(di < 10)$

Maximality is encoded with the help of the same **max** operator that we used in the translations of definites and QDs (cf. the definition in 106). Recall, that this operator encodes *distributive maximality* for the values of a dref, i.e. it ensures that for each assignment in the current info state, the value of a dref is maximal with respect to a particular DRS. Furthermore, (156) includes a cardinality condition ($d < 10$) that applies distributively to each value of the degree dref in the plural info state.

The final piece of the analysis is the idea that comparative numerals can quantifier raise out of their base position leaving behind a degree-type trace. In our system, this means that the comparative numeral leaves a trace whose index is a variable of type sd, i.e. the type of degree drefs.

Sentence (153) is assigned the following structure:

(157)
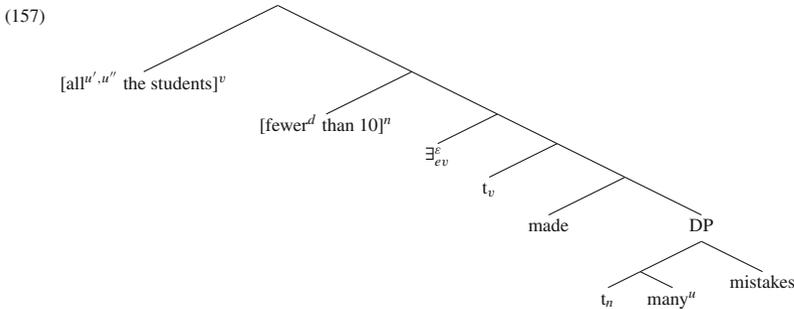
The comparative numeral (*fewer than 10*) must raise out of its base position in order to ensure that *many* combines with an expression of an appropriate type (sd), and land above the event closure operator.[52] The compositional translation of the structure in (157) results in the following DRS:

---

[52] Raising of the comparative numeral above the quantificational subject appears to be blocked in such constructions, a phenomenon sometimes referred to as *Kennedy's Generalization* (cf. Heim 2000, see also Lassiter 2010).

(158) $[u'', u']$; $[\mathbf{ALL}\{u'', u'\}]$; $\mathbf{max}_{u''}(\mathbf{dist'}_w([student\{u'\}])(u'))$;
$\mathbf{max}_{u'}(\mathbf{dist}_w([student\{u'\}]; [d]; \mathbf{max}_d([u]; [\mathbf{card}\{u, d\}]; [\varepsilon]; [mistake\{u\}];$
$[make\{\varepsilon\}]; [\mathbf{Th}\{u, \varepsilon\}]; [\mathbf{Ag}\{u', \varepsilon\}]); [d < 10])(u'))$

Consider the truth conditions of this DRS relative to a singleton input info state. First, it introduces two new drefs u'' and u':

(159)

Info state K, s.t. $I[u'', u']K$	...	u''	u'	...
k	...	$s_1 \oplus s_2 \oplus \ldots \oplus s_n$	$s_1 \oplus s_2 \oplus \ldots \oplus s_n$	...

Given the definition of **ALL** (cf. 107), the values for u'' and u' must have equal cardinality. The value for u'' is the maximal sum of students. The value for u' is the maximal sum of individuals such that there is an info state of the following form:

(160)

Info state J	...	u'	d	ε	u	...
j_1	...	s_1	N_1	e_1	m_1	...
j_2	...	s_2	N_2	e_2	m_2	...
...	...	...	...	...	...	...
j_n	...	s_n	N_n	e_n	m_n	...

Here, the value for u' in (159) has been split by the weak distributivity operator into its atomic parts, and each of these parts is a student. The values of d are degrees (i.e. natural numbers), the values for u are sums of mistakes, and the values for ε are making events, such that for each assignment j, the student $u'j$ is the agent of εj, the sum of mistakes uj is the theme of εj, and the degree dj is the cardinality of the sum of mistakes uj. Moreover, for each assignment j, d must return the maximal number such that such sums of mistakes and making events exist. Finally, d must return a number lower than 10 for each assignment in (160).

In other words, the DRS in (158) will be true iff the cardinality of the maximal sum of students who each made at most 9 mistakes is equal to the cardinality of the maximal sum of students, i.e. iff each student student made fewer than 10 mistakes. This captures the intended reading of sentence (153). A co-distributive interpretation of this sentence is ruled out by the combination of two factors. First, the subject's plural QD is weakly distributive, splitting the values of the subject dref into atomic parts and distributing them across multiple assignments in a plural info state. And second, the maximality and cardinality conditions encoded in the semantics of the comparative numeral apply distributively to the values of the object dref for each assignment.

6.3.3 Numerals at a distance

This approach also immediately accounts for the lack of co-distributive readings in examples like (161), which I have argued pose a problem for Kuhn's (2020) proposal:

(161)   All$^{u_1, u_2}$ the kids received$^{\varepsilon_1}$ gift bags^{u_3} containing$^{\varepsilon_2}$ fewer^d than 10^{u_4} candies.

This sentence translates into the following DRS (see Appendix 1 for a compositional translation of the complex DP):[53]

(162) $[u_1, u_2]$; $[\mathbf{ALL}\{u_1, u_2\}]$; $\mathbf{max}_{u_1}(\mathbf{dist'}_w([kid\{u_1\}])(u_1))$;
$\mathbf{max}_{u_2}(\mathbf{dist}_w([kid\{u_2\}]); [\varepsilon_1, u_3]; [gift\text{-}bag\{u_3\}]; [\neg\mathbf{atom}\{u_3\} \vee \neg\mathbf{unique}\{u_3\}]; [d];$
$\mathbf{max}_d([\varepsilon_2, u_4]; [\mathbf{card}\{u_4, d\}]; [candy\{u_4\}]; [contain\{\varepsilon_2\}]; [\mathbf{Th}\{u_4, \varepsilon_2\}]; [\mathbf{Loc}\{u_3, \varepsilon_2\}]);$
$[d < 10]; [receive\{\varepsilon_1\}]; [\mathbf{Th}\{u_3, \varepsilon\}]; [\mathbf{Ag}\{u_2, \varepsilon_1\}])(u_2))$

Let us evaluate the truth conditions of this DRS relative to an arbitrary singleton input info state I. We start by introducing two individual drefs u_1 and u_2:

(163)

Info state K, s.t. $I[u_1, u_2]K$		$\ldots$	u_1	u_2	$\ldots$
k		$\ldots$	$k_1 \oplus k_2 \oplus \ldots \oplus k_n$	$k_1 \oplus k_2 \oplus \ldots \oplus k_n$	$\ldots$

Here, dref u_1 must return the maximal sum of kids, and u_1 and u_2 must return sums of individuals of equal cardinality. Furthermore, the value for u_2 must be the maximal sum of individuals such that there is an info state like the following:

(164)

Info state J		$\ldots$	u_2	ε_1	u_3	d	u_4	ε_2	$\ldots$
j_1		$\ldots$	k_1	e_1	b_1	N_1	c_1	e'_1	$\ldots$
j_2		$\ldots$	k_2	e_2	b_2	N_2	c_2	e'_2	$\ldots$
$\ldots$		$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
j_n		$\ldots$	k_n	e_n	b_n	N_n	c_3	e'_n	$\ldots$

Here, the values for u_2 are atomic sums of kids, the values for ε_1 are receiving events, and the values for u_3 are sums of gift bags. The sum of values of u_3 across all the assignments is non-atomic (i.e. more than one gift bag must be involved overall). For each assignment j, kid $u_2 j$ is the agent of the receiving event $\varepsilon_1 j$, and the sum of gift bags $u_3 j$ is the theme of $\varepsilon_1 j$. Next, the values of d are degrees, the values of u_4 are sums of candies, and the values of ε_2 are containing events. For each assignment j, the sum of gift bags $u_3 j$ is the location of the containing event $\varepsilon_2 j$, the sum of candies $u_4 j$ is the theme of $\varepsilon_2 j$, and dj is the cardinality of $u_4 j$ (i.e. the number of candies). Moreover, for each assignment j, dj is the maximal number of candies contained in $u_3 j$. Finally, each value of d must be a number lower than 10.

In other words, to evaluate the truth of the DRS in (162) we must find the maximal sum of kids K_1 such that each kid k_i in K_1 received a sum of gift bags b_i containing at most 9 candies, and more that one gift bag was received overall. Then we need to compare K_1 to the maximal sum of all the kids K_2. If $|K_1| = |K_2|$, the DRS in (162) is true, otherwise false. This amounts to saying that every kid received a gift bag containing fewer than 10 candies, and more than one such gift bag was received overall.

---

[53] The abbreviated term $[\neg\mathbf{atom}\{u_3\} \vee \neg\mathbf{unique}\{u_3\}]$ stands for $\lambda I.\lambda J.\ I = J \wedge (\neg\mathbf{atom}\{u_3\}J \vee \neg\mathbf{unique}\{u_3\}J)$.

Crucial for our purposes is the fact that the modified numeral in the relative clause in (161) is interpreted as specifying the cardinality of candies relative to each individual kid, rather than the total number of candies that the kids received altogether. This is a direct consequence of the fact that the structure of the plural info state generated by the quantificational subject is passed down in the course of semantic composition to all expressions in the scope of that subject.

6.4 Accounting for the ban on singular licensors

6.4.1 Bare plurals in the scope of strong distributivity operators

Now consider the interpretation of bare plurals in the scope of strong distributivity operators:

(165)   Three^u students δ_s/*each* carried boxes$^{u'}$.

The example in (165) has the following structure (disregarding exhaustification for now), with the compositional translation in (167):

(166)

(167)
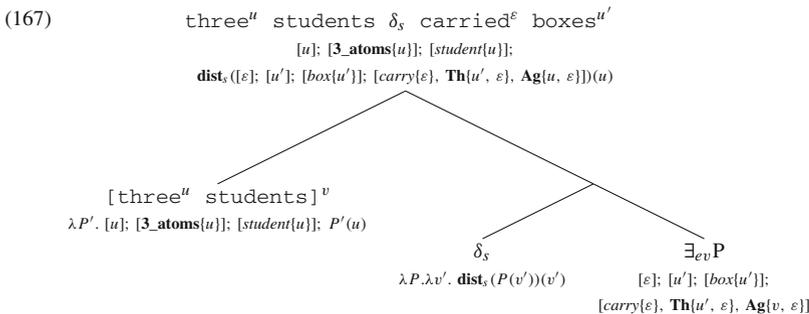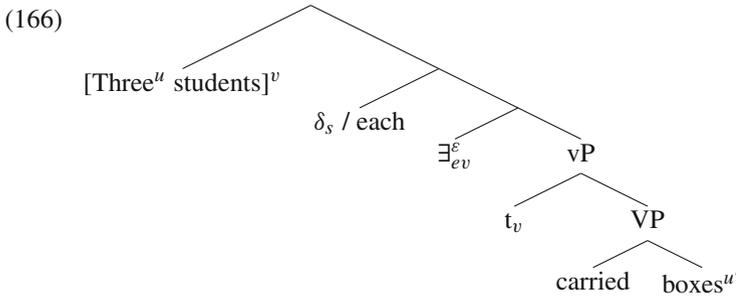
Replacing the **dist**$_s$ operator with its definition, we obtain the following DRS:

(168) $\lambda I . \lambda J . \exists H .$ ([u]; [**3_atoms**{u}]; [*student*{u}])$IH \; \wedge \; \exists H' . (H\langle u\rangle H' \wedge$ **dist**([ε]; [u']; [*box*{u'}]; [*carry*{ε}, **Th**{u', ε}, **Ag**{u, ε}])$H'J$)

The difference between the DRS in (168) and that in (125), discussed above, is that here the introduction of ε and u', as well as the application of the predicates *box* and

carry and the thematic relations occurs in the scope of a **dist**-operator. This difference in itself does not make a truth-conditional impact on the basic (i.e. non-enriched) meaning, i.e. (168) and (125) apply to the same set of input-output state pairs in any (appropriate) model. However, the introduction of a **dist**-operator makes a crucial difference when it comes to calculating the scalar implicature associated with the bare plural *boxes* in (166).

The four relevant sites for the insertion of *Exh* in (166) are above the VP, above the vP, above event closure, and above the raised subject. Suppose we apply *Exh* directly to the VP. We have already calculated this strengthened translation in (115), and I repeat it here:

(169) $\lambda v'.\lambda \zeta.\lambda I.\lambda J. I[u']J \wedge box\{u'\}J \wedge carry\{\zeta\}J \wedge \textbf{Th}\{u', \zeta\}J \wedge \textbf{Ag}\{v', \zeta\}J \wedge$
       $(\neg\textbf{atom}\{u'\}J \vee \neg\textbf{unique}\{u'\}J)$

This term is then combined with the subject trace, the event closure operator, the strong distributivity operator and the raised subject DP, resulting the following enriched sentential translation:

(170) $\lambda I.\lambda J.\exists H. I[u]H \wedge \textbf{3\_atoms}\{u\}H \wedge student\{u\}H \wedge \exists H'. (H\langle u\rangle H' \wedge$
       $\textbf{dist}(\lambda I'.\lambda J'. I'[\varepsilon, u']J' \wedge box\{u'\}J' \wedge carry\{\varepsilon\}J' \wedge \textbf{Th}\{u', \varepsilon\}J' \wedge$
       $\textbf{Ag}\{u, \varepsilon\}J' \wedge (\neg\textbf{atom}\{u'\}J' \vee \neg\textbf{unique}\{u'\}J'))H'J)$

Consider the truth conditions of this DRS. It will be true with respect to a singleton input info state I iff:

(a) there is an info state H such that the assignments in H differ from the assignments in I at most with respect to the values for u, and u returns a sum of three students for each assignment in H. Given that I is singleton, H can be represented as follows:

(171)

Info state H, s.t. $I[u]H$	$\ldots$	u	$\ldots$
h	$\ldots$	$s_1 \oplus s_2 \oplus s_3$	$\ldots$

(b) There exists an info state H', such that $H\langle u\rangle H'$, i.e. the value of u in H is split into its atomic parts in H':

(172)

Info state H', s.t. $H\langle u\rangle H'$	$\ldots$	u	$\ldots$
h'_1 (s.t. $h_1[u]j$)	$\ldots$	s_1	$\ldots$
h'_2 (s.t. $h_2[u]j$)	$\ldots$	s_2	$\ldots$
h'_3 (s.t. $h_3[u]j$)	$\ldots$	s_3	$\ldots$

(c) There is an output info state J that is the conjunction of singleton info states obtained in the following way: the info state H' in (172) is split into a set of singleton info states and each of these info states is updated with the event dref ε and individual dref u'. Moreover, for each of these singleton info states: the value for u is an atomic student s_n; the value for ε is a carrying event e_n, such that s_n is the agent of e_n; and the value for u' is a *non-atomic* sums of boxes b_n, such that b_n is the theme of e_n.

(173)

Info state H_1''	$\ldots$	u	ε	u'	$\ldots$
h_1''	$\ldots$	s_1	e_1	b_1	$\ldots$

Info state H_2''	$\ldots$	u	ε	u'	$\ldots$
h_2''	$\ldots$	s_2	e_2	b_2	$\ldots$

Info state H_3''	$\ldots$	u	ε	u'	$\ldots$
h_3''	$\ldots$	s_3	e_3	b_3	$\ldots$

$$\Downarrow$$

Info state J	$\ldots$	u	ε	u'	$\ldots$
j_1	$\ldots$	s_1	e_1	b_1	$\ldots$
j_2	$\ldots$	s_2	e_2	b_2	$\ldots$
j_3	$\ldots$	s_3	e_3	b_3	$\ldots$

The way the output info state J is constructed follows from the definition of the **dist** operator in (94). This operator first splits the info state H' in (172) into singleton info states, then updates each of these info states *separately*, and finally 'glues' them back together to produce the output info state J. Crucially, the state-level non-atomicity condition imposed on the values of u' is also applied separately to each of the singleton info states, and not to J directly. This means that *each* value for u' in J must be non-atomic. Thus, the DRS in (170) will be true iff there are three students who each carried more than one box. This correctly captures the truth conditions of sentence (165).

If we attach *Exh* directly above the vP or directly above the event closure operator in (166), we will again obtain the DRS in (170). On the other hand, inserting *Exh* at the root (above the raised subject) produces a DRS which will be true iff there are three students who each carried a sum of boxes, and at least one of these students carried more than one box. These truth conditions are weaker than those for (170), which require for each student to have carried more than one box. Whether this weaker reading exists for sentences like (165) is not completely clear, although it does seem to be available at least marginally for some speakers (cf. some relevant discussion in Sauerland 2003, Sauerland et al. 2005, Ivlieva 2013, and the next section). I will not have anything new to say on this matter, and will leave it for future research.[54]

6.4.2 Bare plurals in the scope of singular quantificational DPs

Consider now the interpretation of bare plural DPs in the scope of singular quantifiers, as in the following example:

(174)   Every$^{u,u'}$ student carried$^{\varepsilon}$ boxes$^{u''}$.

This sentence does not have a dependent plural interpretation, e.g. it will not be judged true if each student carried only a single box. Instead, it can be paraphrased

---

[54] Adopting the Strongest Candidate Principle would rule out this weaker reading. Thus, this is one of the contexts where the Non-Weakening Condition and the Strongest Candidate Principle make contrasting predictions.

as stating that each student carried *more than one* box. In the current system this is explained by the presence of a strong distributivity operator scoping over the nuclear scope predicate in the translation of singular QDs like *every*:

(175) $\text{every}^{u,\,u'} \rightsquigarrow \lambda P_{\mathbf{et}}.\lambda P'_{\mathbf{et}}.\ [u',u];\ [\mathbf{EVERY}\{u',u\}];\ \mathbf{max}_{u'}(\mathbf{dist'}_s(P(u'))(u'));$
$\mathbf{max}_u(\mathbf{dist}_s(P(u);\ P'(u))(u))$

Like before, let us first consider a structure where the Exh-operator applies to the VP:

(176)
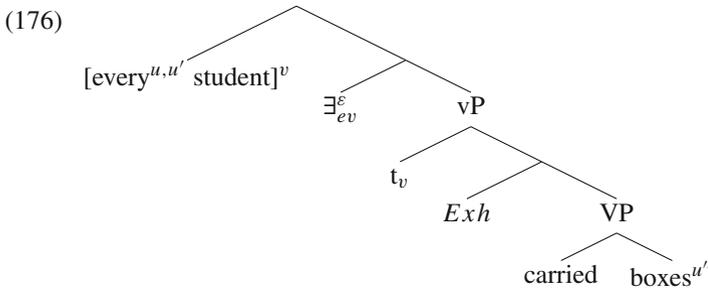
This structure is translated into the following DRS:

(177) $[u';u];\ [\mathbf{EVERY}\{u',u\}];\ \mathbf{max}_{u'}(\mathbf{dist'}_s([\mathbf{atom}\{u'\}];\ [\mathbf{unique}\{u'\}];\ [student\{u'\}])(u'));$
$\mathbf{max}_u(\mathbf{dist}_s(\lambda I.\lambda J.\ student\{u\}I\ \wedge\ \mathbf{atom}\{u\}I\ \wedge\ \mathbf{unique}\{u\}I\ \wedge\ I[\varepsilon,\ u'']J\ \wedge\ box\{u''\}J\ \wedge\ carry\{\varepsilon\}J\ \wedge\ \mathbf{Th}\{u'',\ \varepsilon\}J\ \wedge\ \mathbf{Ag}\{u,\ \varepsilon\}J\ \wedge\ (\neg\mathbf{atom}\{u''\}J\ \vee\ \neg\mathbf{unique}\{u''\}J))(u))$

Let us examine the truth conditions of (177) closer. Assuming a singleton input info state I, the DRS in (177) first introduces two new individual dref u' and u, giving rise to a new info state K of the following form:

(178)

Info state K,		$\ldots$	u'	u	$\ldots$
k		$\ldots$	$s_1' \oplus s_2' \oplus \ldots \oplus s_m'$	$s_1 \oplus s_2 \oplus \ldots \oplus s_n$	$\ldots$

The values of u' and u in K are then related by the quantifier **EVERY** with the following definition:

(179) $\mathbf{EVERY}(u',u) := \lambda I_{st}.\ I \neq \varnothing\ \wedge\ \forall i \in I.\ (\ |u'i| = |ui|\ )$

Applied to K, this operator ensures that the cardinality of the sum of individuals returned by u' ($s_1' \oplus s_2' \oplus \ldots \oplus s_m'$) is equal to the cardinality of the sum of individuals returned by u ($s_1 \oplus s_2 \oplus \ldots \oplus s_n$).

The first maximality operator in (177) applies to the restrictor predicate taken under the externally static version of the strong distributivity operator. The resulting DRS is a test. Applied to the info state in (178), it ensures that if the value of u is split into atomic sub-parts, each of these sub-parts is a student. Moreover, it states that the value of u in K (178) is the maximal sum of individuals that satisfies this condition, i.e. it is the maximal sum of students. Note, that since the uniqueness and atomicity conditions occur in the scope of the $\mathbf{dist}_s$ operator in (177), they will be applied separately to multiple singleton info states, and hence will be effectively neutralized.

Consider now the second maximality operator in (177):

(180) $\mathbf{max}_u(\mathbf{dist}_s(\lambda I.\lambda J.\, student\{u\}I \wedge \mathbf{atom}\{u\}I \wedge \mathbf{unique}\{u\}I \wedge I[\varepsilon,\, u'']J \wedge$
$box\{u''\}J \wedge carry\{\varepsilon\}J \wedge \mathbf{Th}\{u'',\, \varepsilon\}J \wedge \mathbf{Ag}\{u,\, \varepsilon\}J \wedge (\neg\mathbf{atom}\{u''\}J \vee$
$\neg\mathbf{unique}\{u''\}J))(u))$

It does two things. First, it updates the input info state K in (178) with the dynamic conjunction of the restrictor and nuclear scope predicates taken under the externally dynamic version of the strong distributivity operator:

(181) $\mathbf{dist}_s(\lambda I.\lambda J.\, student\{u\}I \wedge \mathbf{atom}\{u\}I \wedge \mathbf{unique}\{u\}I \wedge I[\varepsilon,\, u'']J \wedge$
$box\{u''\}J \wedge carry\{\varepsilon\}J \wedge \mathbf{Th}\{u'',\, \varepsilon\}J \wedge \mathbf{Ag}\{u,\, \varepsilon\}J \wedge (\neg\mathbf{atom}\{u''\}J \vee$
$\neg\mathbf{unique}\{u''\}J))(u) :=$
$\lambda I_{st}.\lambda J_{st}.\exists H.\ I\langle u\rangle H \wedge \mathbf{dist}(\lambda I'.\lambda J'.\, student\{u\}I' \wedge \mathbf{atom}\{u\}I' \wedge$
$\mathbf{unique}\{u\}I' \wedge I'[\varepsilon,\, u'']J' \wedge box\{u''\}J' \wedge carry\{\varepsilon\}J' \wedge \mathbf{Th}\{u'',\, \varepsilon\}J' \wedge$
$\mathbf{Ag}\{u,\, \varepsilon\}J' \wedge (\neg\mathbf{atom}\{u''\}J' \vee \neg\mathbf{unique}\{u''\}J'))H J$

Applied to K in (178), this DRS first splits the value for u into its atomic parts giving rise to info state H:

(182)

Info state H, s.t. $K\langle u\rangle H$		$\ldots$	u'	u	$\ldots$
h_1		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_1	$\ldots$
h_2		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_2	$\ldots$
$\ldots$		$\ldots$	$\ldots$	$\ldots$	$\ldots$
h_n		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_n	$\ldots$

This info state is then split into multiple singleton info states by the **dist** operator, and each of these info state is updated with an event dref ε and individual dref u'':

(183)

Info state H'_1		$\ldots$	u'	u	ε	u''	$\ldots$
h'_1		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_1	e_1	b_1	$\ldots$

Info state H'_2		$\ldots$	u'	u	ε	u''	$\ldots$
h''_2		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_2	e_2	b_2	$\ldots$

$\ldots$

Info state H'_n		u'	u	ε	u''	$\ldots$	
h'_n		$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_n	e_n	b_n	$\ldots$

Several conditions are applied separately to each of these singleton info states. For each H'_k in (183), the value for u is an atomic student s_k; the value for ε is a carrying event e_k, such that s_k is the agent of e_k; and the value for u'' is a *non-atomic* sum of boxes b_k, such that b_k is the theme of e_k.[55]

Following the definition of the **dist** operator, the singleton info states in (183) are then 'glued' back together to produce the following output info state J:

---

[55] Note that the non-uniqueness condition on the values of u'' in (181) when applied separately to each singleton info state is trivially false, which means that the value of u'' in each info state in (183) must be non-atomic.

(184)

Info state J	$\ldots$	u'	u	ε	u''	$\ldots$
j_1	$\ldots$	$s_1' \oplus s_2' \oplus \ldots \oplus s_m'$	s_1	e_1	b_1	$\ldots$
j_2	$\ldots$	$s_1' \oplus s_2' \oplus \ldots \oplus s_m'$	s_2	e_2	b_2	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
j_n	$\ldots$	$s_1' \oplus s_2' \oplus \ldots \oplus s_m'$	s_n	e_n	b_n	$\ldots$

This output state will exist iff there is a set of individuals $s_1, s_2, \ldots, s_n$ who are (atomic) students, who each carried a non-atomic sum of boxes.

The second function of the maximality operator in (180) is to ensure that the value of u in K is the maximal sum of students where each student carried more than one box.

Now, given that the cardinality of the value of u in K must be equal to the cardinality of u' in K (cf. 179), it follows that the DRS in (177) will be true iff the cardinality of the maximal sum of students who each carried more than one box is equal to the cardinality of the maximal sum of students, i.e. each student carried more than one box. These are the desired truth conditions for sentence (174).

Inserting the *Exh*-operator anywhere else below the raised subject in (176) (i.e. immediately above the vP or the event closure operator) does not change the resulting truth conditions. On the other hand, applying exhaustification at the root (above the subject) gives rise to weaker truth conditions whereby each student carried a sum of boxes, and at least one student carried more than one box (cf. Appendix 1). As discussed in the previous section, the status of such weakened readings is uncertain. For instance, Ivlieva (2013) cites the following example:

(185)   Every professor in our department has students.

Ivlieva reports that at least some speakers judge sentence (185) to be true in a situation where some of the professors have just one student, and one or more professors have more than one student (cf. also the discussion of similar 'mixed' readings in Sauerland 2003; Sauerland et al. 2005). She goes on to explore a more complex semantics for the event closure operator in order to derive such readings. In our system these readings do not require any additional assumptions, and are derived automatically by applying exhaustification at the highest level.[56]

To conclude, we have seen that bare plural DPs in the scope of weak distributivity operators are interpreted as dependent plurals, with an overarching Multiplicity Condition derived as a scalar implicature. On the other hand, when a bare plural DP occurs in the scope of a strong distributivity operator, the multiplicity requirement is calculated relative to each atomic individual in the distributed sum. This accounts for the contrast between weak and strong syntactic distributivity operators (e.g. floating *all* and *each*), as well as that between plural and singular QDs.

6.4.3 Deriving the ban on singular licensors

I have shown how the contrast between singular and plural QDs is captured in the current system by analysing the semantics of singular QDs in terms of strong distribu-

---

[56] The Strongest Candidate Principle would rule out the weaker 'mixed' readings for sentences like (174) and (185). I leave it to future research to firmly establish the empirical status of such readings.

tivity, and that of plurals QDs in terms of weak distributivity. Now I would like to address a broader question, namely whether the link between the number marking on the restrictor NP and the type of distributivity operator involved in a QD's translation is accidental, or conversely, can be derived in a principled way.

Specifically, we may ask whether we expect to find a language that possesses a singular QD, i.e. a QD combining with a restrictor NP carrying a singular number feature, and at the same time is weakly distributive with respect to its nuclear scope predicate, thus licensing dependent plurals in its scope. Logically, such a QD may exists. Consider for instance a hypothetical determiner *every**, which by assumption combines with a singular restrictor NP and has the following translation:

(186) $\texttt{every*}^{u',u} \rightsquigarrow \lambda P_{\mathbf{et}}.\lambda P'_{\mathbf{et}}. [u', u]; [\mathbf{EVERY*}\{u', u\}]; \mathbf{max}_{u'}(\mathbf{dist}'_s(P(u'))(u'));$
$\mathbf{max}_u(\mathbf{dist}_w(P'(u))(u))$
where $\mathbf{EVERY*}(u', u) := \lambda I_{st}. \forall i \in I. (u'i \leq ui)$

This QD introduces a maximal dref that satisfies the restrictor predicate taken under a strong distributivity operator, and compares it to the maximal dref that satisfies the nuclear scope predicate, taken under a weak distributivity operator. Such a quantifier would violate the Ban on Singular Licensors as formulated above, allowing for a singular restrictor NP (due to the fact that the restrictor predicate is placed under a strong distributivity operator) and at the same time licensing dependent plurals as part of its nuclear scope (since the nuclear scope predicate occurs under a weak distributivity operator).

As far as I know, quantificational determiners of this type have never been identified, and I would like to suggest that there are theoretical reasons to expect that such quantifiers should not exist in natural language. These reasons have to do with the *Conservativity Universal*, first proposed (albeit, in slightly different terms) by Barwise and Cooper (1981) (see also Keenan and Stavi 1986, and much subsequent work):

(187) *Conservativity Universal*
    For all natural language determiners the following holds:
    $D(P)(Q) \leftrightarrow D(P)(P \cap Q),$
    where D is the interpretation of the determiner, and P and Q are sets.

The following examples illustrate that this generalization holds of the English determiners *every* and *most*:

(188) a. Every box is red. $\leftrightarrow$ Every box is such that it is a box and it is red.
    b. Most boxes are red. $\leftrightarrow$ Most boxes are such that they are a box and they are red.

In a compositional dynamic framework, the *Conservativity Universal* can be reformulated in the following way, adapted from Chierchia (1995):

(189) *Dynamic Conservativity Universal*
    For all natural language determiners and all models M the following holds:
    $[\![D(P)(Q)]\!]^M = [\![D(P)(\lambda v. P(v); Q(v))]\!]^M,$

where D is the translation of the determiner, P and Q are translations of the restrictor and nuclear scope constituents, respectively, and ; is dynamic conjunction.

Existing English QDs conform to this universal. On the other hand, the hypothetical QD *every** defined in (186) violates it. To see why consider the translation of the following example involving this hypothetical determiner:

(190) a. Every* $^{u,u'}$ student is thinking$^\varepsilon$.

   b. $[u', u]$; $[\textbf{EVERY*}\{u', u\}]$; $\textbf{max}_{u'}(\textbf{dist}'_s([\textbf{atom}\{u'\}]; [\textbf{unique}\{u'\}];$
       $[student\{u'\}])(u'))$; $\textbf{max}_u(\textbf{dist}_w([\varepsilon]; [think\{\varepsilon\}, \textbf{Exp}\{u, \varepsilon\}])(u))$

Let us examine what it means for the DRS in (190b) to apply to a pair of input and output info state. For simplicity, let's take a singleton input info states $I = \{i\}$. The DRS in (190b) first introduces two new drefs u' and u such that the value for u' is a (non-proper) sub-part of the value for u:

(191)

Info state K, s.t. $I[u', u]K$	$\ldots$	u'	u	$\ldots$
k	$\ldots$	$s_1 \oplus s_2 \oplus \ldots \oplus s_m$	$x_1 \oplus x_2 \oplus \ldots \oplus x_n$	$\ldots$

Next, it states that the value for u' ($u'k$) must be the maximal sum such that there is a set of singleton info states where the values for u' are the atomic parts of $u'k$, and each of these values is a student:

(192)

Info state H_1	$\ldots$	u'	$\ldots$
h_1	$\ldots$	s_1	$\ldots$

Info state H_2	$\ldots$	u'	$\ldots$
h_2	$\ldots$	s_2	$\ldots$

$\ldots$

Info state H_1	$\ldots$	u'	$\ldots$
h_m	$\ldots$	s_m	$\ldots$

Note, that by virtue of the strong distributivity operator, the uniqueness condition in (190b), derived from the singular number feature on the restrictor NP, is applied separately to each of these singleton info states and is thus trivially satisfied. It follows, that $s_1 \oplus s_2 \oplus \ldots \oplus s_m$ is the maximal sum of students.

Finally, the DRS in (190b) states that uk in (191) is the maximal sum of individuals which can be split into atomic parts $x_1, x_2, \ldots, x_n$, and assigned to the values of u in a plural info state of the following form, where $e_1, \ldots, e_n$ are thinking-events whose experiencers are $x_1, \ldots, x_n$, respectively:

(193)

Info state J		$\ldots$	u	ε	$\ldots$
j_1		$\ldots$	x_1	e_1	$\ldots$
j_2		$\ldots$	x_2	e_2	$\ldots$
$\ldots$		$\ldots$	$\ldots$	$\ldots$	$\ldots$
j_n		$\ldots$	x_n	e_n	$\ldots$

Thus, $x_1 \oplus x_2 \oplus \ldots \oplus x_n$ is the maximal sum of individuals who are thinking. Since (190b) states that $s_1 \oplus s_2 \oplus \ldots \oplus s_m$ (the maximal sum of students) must be a sub-part of $x_1 \oplus x_2 \oplus \ldots \oplus x_n$ (the maximal sum of thinking individuals), is follows that the hypothetical sentence in (190a) will be true iff every student is thinking.

Now if *every** conforms to the Dynamic Conservativity Universal in (189), the DRS in (190b) should be equivalent to the DRS that results from applying the same quantifier to the restrictor predicate and the dynamic conjunction of the restrictor and nuclear scope predicates. Namely, to the predicates in (194a) and (194b):

(194)   a. $\lambda v.$ [**atom**$\{v\}$]; [**unique**$\{v\}$]; [*student*$\{v\}$]
       b. $\lambda v.$ [**atom**$\{v\}$]; [**unique**$\{v\}$]; [*student*$\{v\}$]; [ε]; [*think*$\{\varepsilon\}$, **Exp**$\{v, \varepsilon\}$]

Applying *every** to these predicates yields the following DRS:

(195) $[u', u]$; [**EVERY***$\{u', u\}$]; $\mathbf{max}'_u(\mathbf{dist}'_s([\mathbf{atom}\{u'\}]; [\mathbf{unique}\{u'\}]; [student\{u'\}])(u'))$;
       $\mathbf{max}_u(\mathbf{dist}_w([\mathbf{atom}\{u\}]; [\mathbf{unique}\{u\}]; [student\{u\}]; [\varepsilon]; [think\{\varepsilon\}, \mathbf{Exp}\{u, \varepsilon\}])(u))$

This DRS again introduces two new drefs u' and u, giving rise to the info state K in (191). The conditions that it places on the value of u' and the relation between the values of u and u' are the same as in (190b), i.e. $u'k$ must return the maximal sum of students and this sum must be a non-proper sub-part of uk. However, the conditions imposed on the value of u are different. The second maximality operator in (195) applies to the dynamic conjunction of the restrictor and nuclear scope DRSs under a weak distributivity operator. This means that uk in (191) must be the maximal individual such that there exists an output info state J like the following:

(196)

Info state J		$\ldots$	u	ε	$\ldots$
h_1		$\ldots$	x_1	e_1	$\ldots$
h_2		$\ldots$	x_2	e_2	$\ldots$
$\ldots$		$\ldots$	$\ldots$	$\ldots$	$\ldots$
h_n		$\ldots$	x_n	e_n	$\ldots$

Here $x_1, \ldots, x_n$ are atomic individuals, the sum $x_1 \oplus \ldots \oplus x_n$ is equal to uk, and $e_1, \ldots, e_n$ are thinking-events whose experiencers are $x_1, \ldots, x_n$, respectively. This is similar to what we had in (193) above. However, the conjoined DRS under the $\mathbf{dist}_w$ operator in (195) places additional conditions on the values of u in J. Specifically, it requires *student* to be true of every x in $\{x_1, \ldots, x_n\}$, and, crucially, for all the individuals in $\{x_1, \ldots, x_n\}$ to be atomic and identical. Since, by definition of $\mathbf{dist}_w$,

$uk = x_1 \oplus \ldots \oplus x_n$, it follows that uk in (191) must be atomic. Moreover, since uk is required to be the maximal sum such that an info state as in (196) exists (i.e. every individual that is a student and is thinking must be part of or equal to uk), such a sum will only exist in a model that contains only a single thinking student. And since the maximal sum of students must be a non-proper sub-part of uk, it follows that there must be only one student overall.

In sum, (195) will only be true with respect to any pair of input and output info states in a model where there is only one student, and that student is thinking. Clearly, then, the DRS in (195) is not equivalent to that in (190b), and we may conclude that *every** violates the Dynamic Conservativity Universal.

The reason that the Dynamic Conservativity Universal is violated in this case is that placing the translation of the singular restrictor NP under a weak distributivity operator results in an undesired global atomicity condition on the value of the maximal dref satisfying the nuclear scope predicate. This problem does not arise, however, if the translation of the QD involves a strong distributivity operator scoping over the nuclear scope predicate, because in that case the global atomicity condition applies distributively to multiple singleton info states, and is thus in effect neutralized.

In fact, quite generally, if a QD combines with a restrictor carrying a singular number feature, it must be interpreted as involving a strong distributivity operator scoping over both its restrictor and nuclear scope, otherwise the unwelcome global atomicity effect obtains and the Dynamic Conservativity Universal is violated. This in turn means that such QDs will not be able to license dependent plurals as part of their nuclear scope. I conclude that in the current system the Ban on Singular Licensors follows directly form the Dynamic Conservativity Universal, as stated in (189), which restricts the class of quantificational determiners possible in natural language.

6.5 A note on discourse interpretation

This paper has focused exclusively on intra-sentential phenomena related to the semantics of quantificational items, grammatical number and cardinality modifiers. Before I conclude, I would like to briefly address the issue of discourse interpretation, and specifically the way information is passed on between sentences. As noted above, the notion of context as a set of assignments, i.e. the idea of dynamic semantics with plural info states, was originally developed to account for certain complex cases of cross-sentential anaphora which posed problems for existing compositional dynamic semantic systems (cf. van den Berg 1990, 1993, 1994, 1996). One such phenomenon is *quantificational subordination*, illustrated in (197) (cf. detailed discussion in Karttunen 1976; Krifka 1996; Nouwen 2003; Brasoveanu 2007):

(197) Every student wrote an article. They each sent it to L&P.

In this example, the second sentence can be understood as stating that each of the students mentioned in the first sentence sent the article that they wrote to L&P. Thus, for each student the singular pronoun *it* is able to pick up the value of the referent introduced by the indefinite *an article* corresponding to that student. In systems that model context as a set of assignments, the dependency between the two variables

(the students and the articles) is established by the first sentence, and passed on as the input context to the second sentence.[57] Formally, it is standardly assumed that the semantic representation of a discourse is obtained by dynamically conjoining the DRSs corresponding to the sentences in that discourse:

(198) If Σ is a sequence of sentences such that $\Sigma \rightsquigarrow D$, and S is a sentence such that $S \rightsquigarrow D'$, then:
$\langle \Sigma, S \rangle \rightsquigarrow \lambda I. \lambda J. \exists H. (DIH \wedge D'HJ)$

I have chosen to formulate my account of the semantics of distributivity in such a way as to make it compatible with this type of approach to cross-sentential anaphora. Specifically, this entails that distributive items are taken to be *externally dynamic*, i.e. the dependencies they construct are encoded directly in their output info state, which (assuming 198) can be passed on to the subsequent discourse.

At the same time, throughout the paper I have evaluated the truth conditions of the DRSs we've encountered assuming *singleton* input info state. I would now like to suggest that this assumption is not just a simplification adopted for presentational purposes, but rather a reflection of existing properties of discourse interpretation.

One thing to note is that the accessibility of quantificational dependencies like that in (197) seems to be rather short-lived, rarely stretching beyond the sentence immediately following the one where the dependency is first established.[58] For instance, the following discourse is harder to interpret than (197):

(199) Every student wrote an article. They spent a few days considering different journals. In the end, they each sent it to L&P.

Given this tentative observation, I would like to suggest that if a sentence produces a complex, non-singleton info state as its output, this info state is by default simplified before it is fed as input to the subsequent sentence. Specifically, I would like to suggest the following as the default rule of discourse interpretation:

(200) If Σ is a sequence of sentences such that $\Sigma \rightsquigarrow D$, and S is a sentence such that $S \rightsquigarrow D'$, then:
$\langle \Sigma, S \rangle \rightsquigarrow \lambda I. \lambda J. \exists H. (DIH \wedge \exists K. (K = \oplus H \wedge D'KJ))$,
where $K = \oplus H := \exists k_s. (K = \{k\} \wedge \forall v. (vk = \oplus vH))$

According to this, the output info state of a sentential DRS is by default collapsed into a singleton info state, with all the values of the drefs summed up, and only then passed on to the next DRS in the discourse. Then the more direct mode of discourse interpretation in (198) (where the output of one sentential DRS is directly fed as input to the next one) is treated as a marked option, which is only employed when a

---

subsequent DRS makes direct reference to the dependencies established in the output info state of the previous DRS (e.g. in cases involving quantificational subordination). From a processing point of view, this can be regarded as a functional strategy on the part of the interpreter aimed at reducing memory load.

A consequence of adopting (200) as the default rule for discourse interpretation in the current system is that (putting aside contexts involving quantificational subordination) non-singleton info states are predicted to arise only in the scope of distributivity operators and quantificational determiners.

7 Conclusion

I have presented an account of the semantics of grammatical number, quantificational items, and cardinality predicates couched within PCDRT^e, a dynamic semantic framework with plural info states. I have shown how this analysis is able to treat dependent plural readings as distinct from both cumulative and distributive readings, in the classical sense, thus overcoming the challenges facing previous approaches. The proposed solution has two components. First, I introduced a distinction between two types of distributivity: weak distributivity across the assignments within a single plural info state and strong distributivity across multiple info states. I have argued that both of these types of distributivity play a role in the semantics of natural language, accounting for the contrasting properties of 'singular quantifiers', such as *each* and *every*, and 'plural quantifiers', such as *all* and *most*. Second, I analysed numerals (and other cardinal expressions) as imposing *assignment-level* cardinality conditions, i.e. restricting the cardinality of each value returned by a dref in a plural info state. The singular number feature, on the other hand, was analysed as introducing a *state-level* atomicity condition. The multiplicity condition associated with (underlyingly number-neutral) bare plurals was then derived as a state-level non-atomicity condition, in competition with the singular alternative. Together, these two distinctions (weak vs strong distributivity and assignment-level vs state-level plurality) give rise to three distinct ways of representing the notion of multiplicity in the semantics of natural language, and are sufficient to explain the observed patterns of (co-)distributive interpretation.

Appendix 1. Some compositional translations

Complex DP with Embedded Numeral

(201) All^{u_1,u_2} the kids received$^{\varepsilon_1}$ gift bags^{u_3} containing$^{\varepsilon_2}$ fewer^d than 10^{u_4} candies.

Structure of DP:

(202) gift bags containing fewer than 10 candies
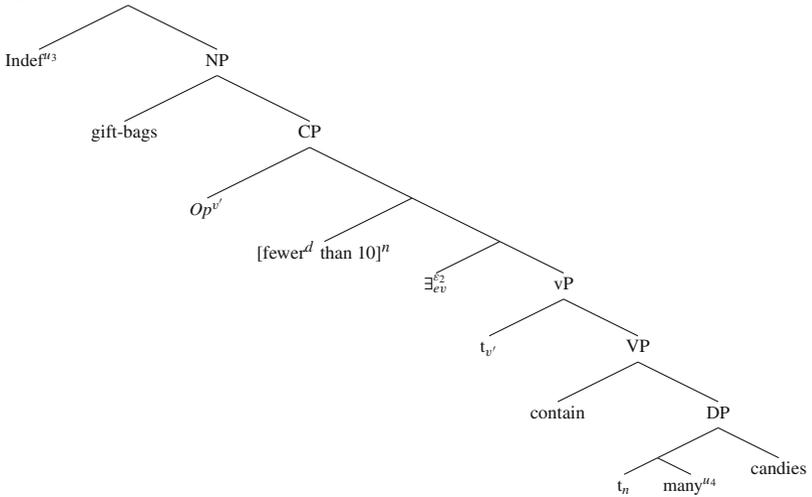
Stepwise compositional translation:[59]

(203)

---
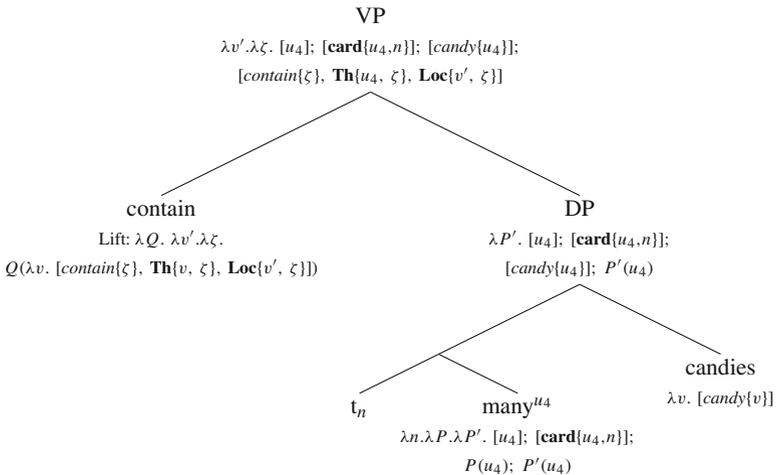
[59] In this tree and the following, I indicate the lifted translations of Vs, VPs and vPs when necessary.

(204)
$$\exists_{ev}P$$

$[\varepsilon_2]$; $[u_4]$; $[\mathbf{card}\{u_4,n\}]$; $[candy\{u_4\}]$;
$[contain\{\varepsilon_2\},\ \mathbf{Th}\{u_4,\ \varepsilon_2\},\ \mathbf{Loc}\{v',\ \varepsilon_2\}]$

$\exists_{ev}^{\varepsilon_2}$
$\lambda V.\ [\varepsilon_2];\ V(\varepsilon_2)$

$\lambda\zeta.\ [u_4];\ [\mathbf{card}\{u_4,n\}];\ [candy\{u_4\}];$
$[contain\{\zeta\},\ \mathbf{Th}\{u_4,\ \zeta\},\ \mathbf{Loc}\{v',\ \zeta\}]$

$t_{v'}$

VP
$\lambda v'.\lambda\zeta.\ [u_4];\ [\mathbf{card}\{u_4,n\}];\ [candy\{u_4\}];$
$[contain\{\zeta\},\ \mathbf{Th}\{u_4,\ \zeta\},\ \mathbf{Loc}\{v',\ \zeta\}]$

(205)
CP
$\lambda v'.\ [d];$
$\mathbf{max}_d([\varepsilon_2];\ [u_4];\ [\mathbf{card}\{u_4,d\}];\ [candy\{u_4\}];\ [contain\{\varepsilon_2\},\ \mathbf{Th}\{u_4,\ \varepsilon_2\},\ \mathbf{Loc}\{v',\ \varepsilon_2\}]);$
$[d<10]$

$Op^{v'}$
$\lambda P.P$

$[d];\ \mathbf{max}_d([\varepsilon_2];\ [u_4];\ [\mathbf{card}\{u_4,d\}];\ [candy\{u_4\}];$
$[contain\{\varepsilon_2\},\ \mathbf{Th}\{u_4,\ \varepsilon_2\},\ \mathbf{Loc}\{v',\ \varepsilon_2\}]);$
$[d<10]$

$[fewer^d\ than\ 10]^n$
$\lambda M.\ [d];\ \mathbf{max}_d(M(d));\ [d<10]$

$\exists_{ev}P$
$[\varepsilon_2];\ [u_4];\ [\mathbf{card}\{u_4,n\}];\ [candy\{u_4\}];$
$[contain\{\varepsilon_2\},\ \mathbf{Th}\{u_4,\ \varepsilon_2\},\ \mathbf{Loc}\{v',\ \varepsilon_2\}]$

(206)
DP
$\lambda P'.\ [u_3];\ [gift\text{-}bag\{u_3\}];\ [d];$
$\mathbf{max}_d([\varepsilon_2];\ [u_4];\ [\mathbf{card}\{u_4,d\}];\ [candy\{u_4\}];\ [contain\{\varepsilon_2\},\ \mathbf{Th}\{u_4,\ \varepsilon_2\},\ \mathbf{Loc}\{u_3,\ \varepsilon_2\}]);$
$[d<10];\ P'(u_3)$

Indef^{u_3}
$\lambda P.\lambda P'.\ [u_3];\ P(u_3);\ P'(u_3)$

NP
$\lambda v'.\ [gift\text{-}bag\{v'\}];\ [d];$
$\mathbf{max}_d([\varepsilon_2];\ [u_4];\ [\mathbf{card}\{u_4,d\}];\ [candy\{u_4\}];$
$[contain\{\varepsilon_2\},\ \mathbf{Th}\{u_4,\ \varepsilon_2\},\ \mathbf{Loc}\{v',\ \varepsilon_2\}]);$
$[d<10]$

gift-bags
$\lambda v'.\ [gift\text{-}bag\{v'\}]$

CP
$\lambda v'.\ [d];$
$\mathbf{max}_d([\varepsilon_2];\ [u_4];\ [\mathbf{card}\{u_4,d\}];\ [candy\{u_4\}];$
$[contain\{\varepsilon_2\},\ \mathbf{Th}\{u_4,\ \varepsilon_2\},\ \mathbf{Loc}\{v',\ \varepsilon_2\}]);$
$[d<10]$

Exhaustifying at the Root

(207) Most$^{u,u'}$ students carried$^{\varepsilon}$ boxes$^{u''}$.

Structure:

(208)

Stepwise compositional translation:

(209)

$\exists_{ev}P$

[ε]; [u'']; [$box\{u''\}$];
[$carry\{\varepsilon\}$, **Th**$\{u'', \varepsilon\}$, **Ag**$\{v, \varepsilon\}$]

$\exists_{ev}^{\varepsilon}$
$\lambda V.$ [ε]; $V(\varepsilon)$

vP
$\lambda\zeta.$ [u'']; [$box\{u''\}$];
[$carry\{\zeta\}$, **Th**$\{u'', \zeta\}$, **Ag**$\{v, \zeta\}$]

t_v

VP
$\lambda v'.\lambda\zeta.$ [u'']; [$box\{u''\}$];
[$carry\{\zeta\}$, **Th**$\{u'', \zeta\}$, **Ag**$\{v', \zeta\}$]

carried
Lift: $\lambda Q.\ \lambda v'.\lambda\zeta.$
$Q(\lambda v.\ [carry\{\zeta\}$, **Th**$\{v, \zeta\}$, **Ag**$\{v', \zeta\}])$

boxes$^{u''}$
$\lambda P'.$ [u'']; [$box\{u''\}$]; $P'(u'')$

(210)

Exh_{ALT}
$\lambda D.\lambda I.\lambda J.\ DIJ\ \wedge$
$\forall D'.\ (D'\in\text{ALT}\ \wedge\ D'\succ D\rightarrow\neg D'IJ)$

$[u',u];\ [\textbf{MOST}\{u',u\}];$
$\textbf{max}_{u'}(\textbf{dist'}_w([student(u')])(u'));$
$\textbf{max}_u(\textbf{dist}_w([student(u)];\ [\varepsilon];\ [u''];\ [box\{u''\}];$
$[carry\{\varepsilon\},\ \textbf{Th}\{u'',\ \varepsilon\},\ \textbf{Ag}\{u,\ \varepsilon\}])(u))$

$[\text{most}^{u,u'}\ \text{students}]^v$
$\lambda P'.\ [u',u];\ [\textbf{MOST}\{u',u\}];$
$\textbf{max}_{u'}(\textbf{dist'}_w([student(u')])(u'));$
$\textbf{max}_u(\textbf{dist}_w([student(u)];\ P'(u))(u))$

$\exists_{ev}P$
$[\varepsilon];\ [u''];\ [box\{u''\}];$
$[carry\{\varepsilon\},\ \textbf{Th}\{u'',\ \varepsilon\},\ \textbf{Ag}\{v,\ \varepsilon\}]$
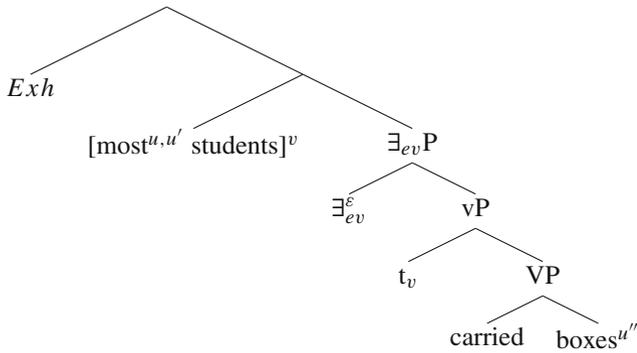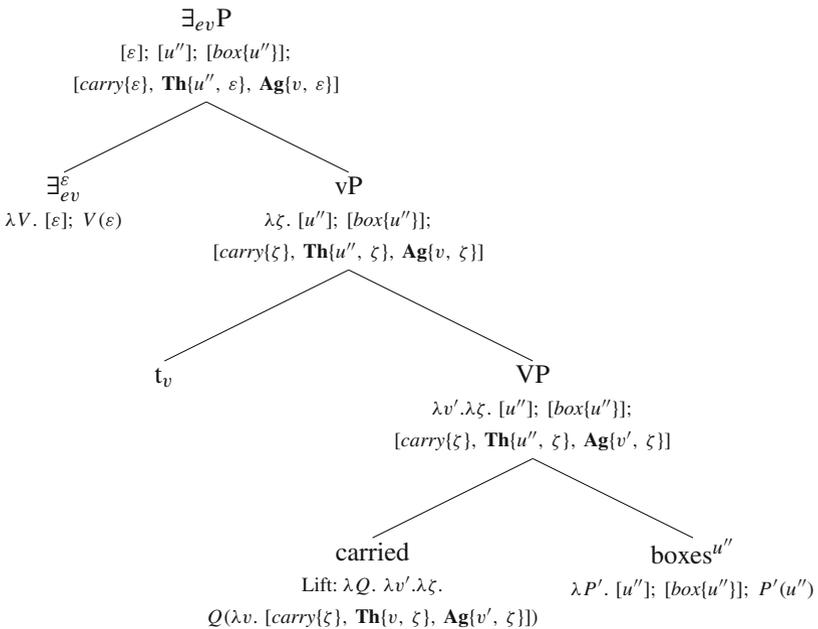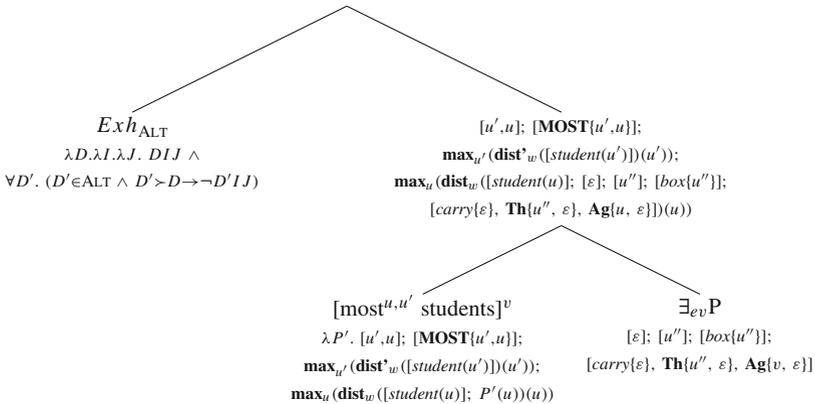
The alternative to (210) is the following DRS, obtained by replacing the plural object with its singular indefinite counterpart:

(211) $[u',u];\ [\textbf{MOST}\{u',u\}];\ \textbf{max}_{u'}(\textbf{dist'}_w([student(u')])(u'));$
$\textbf{max}_u(\textbf{dist}_w([student(u)];\ [\varepsilon];\ [u''];\ [atom\{u''\}];\ [\textbf{unique}\{u''\}];\ [box\{u''\}];$
$[carry\{\varepsilon\},\ \textbf{Th}\{u'',\ \varepsilon\},\ \textbf{Ag}\{u,\ \varepsilon\}])(u))$

The alternative in (211) is stronger than the DRS derived in (210). For any model M, (211) will be true iff the maximal sum of students S_1 in M who all carried the same box is greater than half of the total sum of students S_2 in M. Given the definition of maximality in (106) and the fact that *carry* is lexically distributive with respect to its theme, S_1 will only exist in a model M where all the students who carried any boxes carried the same box. Consequently, in any model M, (210) will return 1 for every input-output info states for which (211) returns 1. The converse is obviously not the case. Hence, (211) is stronger than (210), and the translation is strengthened by the negation of (211):

(212) $\lambda I.\lambda J.\ ([u',u];\ [\textbf{MOST}\{u',u\}];\ \textbf{max}_{u'}(\textbf{dist'}_w([student(u')])(u'));$
$\textbf{max}_u(\textbf{dist}_w([student(u)];\ [\varepsilon];\ [u''];\ [box\{u''\}];$
$[carry\{\varepsilon\},\ \textbf{Th}\{u'',\ \varepsilon\},\ \textbf{Ag}\{u,\ \varepsilon\}])(u)))IJ\ \wedge$
$\neg([u',u];\ [\textbf{MOST}\{u',u\}];\ \textbf{max}_{u'}(\textbf{dist'}_w([student(u')])(u'));$
$\textbf{max}_u(\textbf{dist}_w([student(u)];\ [\varepsilon];\ [u''];\ [atom\{u''\}];\ [\textbf{unique}\{u''\}];\ [box\{u''\}];$
$[carry\{\varepsilon\},\ \textbf{Th}\{u'',\ \varepsilon\},\ \textbf{Ag}\{u,\ \varepsilon\}])(u)))IJ :=$
$\lambda I.\lambda J.\exists K.\ ([u',u];\ [\textbf{MOST}\{u',u\}];\ \textbf{max}_{u'}(\textbf{dist'}_w([student(u')])(u')))IK\ \wedge$
$\textbf{max}_u(\textbf{dist}_w([student(u)];\ [\varepsilon];\ [u''];\ [box\{u''\}];\ [carry\{\varepsilon\},\ \textbf{Th}\{u'',\ \varepsilon\},$
$\textbf{Ag}\{u,\ \varepsilon\}])(u))KJ\ \wedge$
$\neg\textbf{max}_u(\textbf{dist}_w([student(u)];\ [\varepsilon];\ [u''];\ [atom\{u''\}];\ [\textbf{unique}\{u''\}];\ [box\{u''\}];$
$[carry\{\varepsilon\},\ \textbf{Th}\{u'',\ \varepsilon\},\ \textbf{Ag}\{u,\ \varepsilon\}])(u))KJ$

This DRS will be true iff the sum of students who carried any boxes is greater than half of the total sum of students, and it is not the case that all the students who carried any boxes carried (only) one and the same box. In other words, most of the students carried one or more boxes, and they carried more than one box overall.

(213) Every$^{u,u'}$ student carried$^{\varepsilon}$ boxes$^{u''}$.

(214)
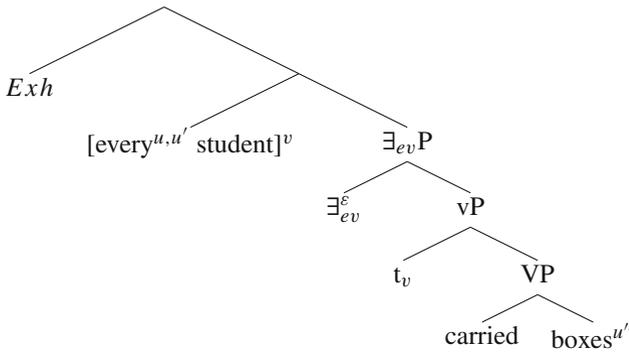
(215)
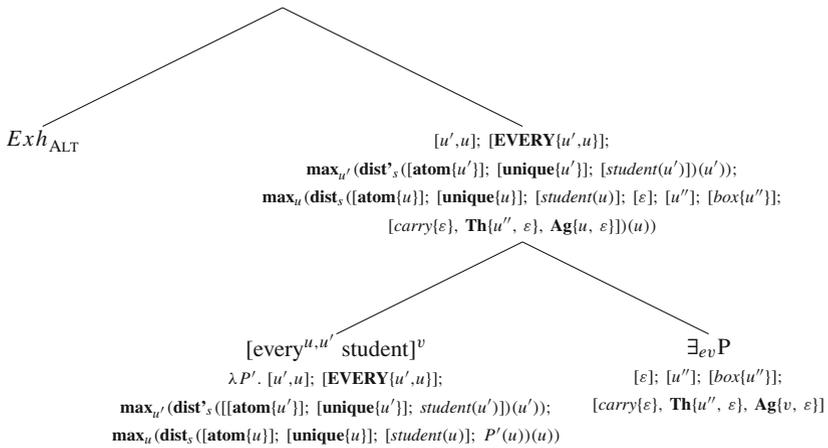
The alternative to (215) is the following DRS:

(216) $[u', u]$; [**EVERY**$\{u', u\}$]; **max**$_{u'}$(**dist'**$_s$([**atom**$\{u'\}$]; [**unique**$\{u'\}$]; [*student*(u')])(u'))); **max**$_u$(**dist**$_s$([**atom**$\{u\}$]; [**unique**$\{u\}$]; [*student*(u)]; $[\varepsilon]$; $[u'']$; [**atom**$\{u''\}$]; [**unique**$\{u''\}$]; [*box*$\{u''\}$]; [*carry*$\{\varepsilon\}$, **Th**$\{u'', \varepsilon\}$, **Ag**$\{u, \varepsilon\}$])(u)))

Consider the following output info state:

(217)

Info state J	$\ldots$	u'	u	ε	u''	$\ldots$
j_1	$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_1	e_1	b_1	$\ldots$
j_2	$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_2	e_2	b_2	$\ldots$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
j_n	$\ldots$	$s'_1 \oplus s'_2 \oplus \ldots \oplus s'_m$	s_n	e_n	b_n	$\ldots$

Both (215) and (216) will be true with respect to an input info state I and output info state J iff for each assignment in J, u' returns the sum of all the students, u

returns an atomic sum of students, ε returns a carrying event whose agent is the value for u and whose theme is the value for u'', and u'' returns a sum of boxes. Now, (215) requires for the sum of values of u in J to be the maximal sum of students such that a state like J exists, i.e. it must be the maximal sum of students who each carried one or more box. Then, the static quantifier **EVERY** ensures than the cardinality of this sum is equal to the cardinality of the total sum of students, i.e. each student must have carried one or more box.

The difference in (216) is that this DRS requires for each value of u'' in J to be an atomic sum of boxes. This means, that the sum of values of u in J must be the maximal sum of students where each student carried one box, and the cardinality of this sum must be equal to the cardinality of the total sum of students. Given that the verb *carry* is lexically distributive with respect to its theme, for any model M, any I and J that satisfy (216) will also satisfy (215), while the converse does not hold. Then, by definition, (216) is stronger than (215), and the translation of (214) is strengthened. The resulting DRS (cf. 218) will apply to an input info state I, and an output info state J as in (217) iff all the conditions in (215) hold, and moreover it is not the case that all the values of u'' in J are atomic. In other words, it will be true iff each student carried one or more boxes, and at least one student carried more than one box.

(218) $\lambda I.\lambda J. \;([u', u];\; [\mathbf{EVERY}\{u', u\}];$
$\mathbf{max}_{u'}(\mathbf{dist'}_s([\mathbf{atom}\{u'\}];\; [\mathbf{unique}\{u'\}];\; [student(u')])(u'));$
$\mathbf{max}_u(\mathbf{dist}_s([\mathbf{atom}\{u\}];\; [\mathbf{unique}\{u\}];\; [student(u)];\; [\varepsilon];\; [u''];\; [candy\{u''\}];$
$[carry\{\varepsilon\},\; \mathbf{Th}\{u'',\; \varepsilon\},\; \mathbf{Ag}\{u,\; \varepsilon\}])(u)))IJ \;\wedge$
$\neg([u', u];\; [\mathbf{EVERY}\{u', u\}];\; \mathbf{max}_{u'}(\mathbf{dist'}_s([\mathbf{atom}\{u'\}];\; [\mathbf{unique}\{u'\}];\; [student(u')])(u'));$
$\mathbf{max}_u(\mathbf{dist}_s([\mathbf{atom}\{u\}];\; [\mathbf{unique}\{u\}];\; [student(u)];\; [\varepsilon];\; [u''];\; [\mathbf{atom}\{u''\}];\; [\mathbf{unique}\{u''\}];$
$[candy\{u''\}];\; [carry\{\varepsilon\},\; \mathbf{Th}\{u'',\; \varepsilon\},\; \mathbf{Ag}\{u,\; \varepsilon\}])(u)))IJ$

Appendix 2. Schein–Kratzer sentences

Quantificational DPs, both plural and singular, allow for cumulative readings with other plurals as long as the base position of the plural c-commands the base position of the quantificational DP (cf. Schein 1993; Kratzer 2000; see also Zweig 2008, 2009; Champollion 2010a for relevant observations):

(219) a. Three copy editors found every mistake in the manuscript.
       b. Three copy editors found all the mistakes in the manuscript.

Both of these sentences can be understood as stating that there were three copy editors who *between them* identified the maximal set of mistakes, i.e. each copy editor may have identified only as subset of this maximal set, as long as the union of these subsets is equal to the maximal set.

Kratzer's (2000) analysis of such examples relies on two assumptions: first, the external argument (i.e. the agent in 219) is introduced by a separate thematic head, rather than by the lexical verb itself. Second, quantificational DPs must be allowed to scope below the closure of the event variable. In this appendix I will sketch out how an analysis along these lines can be incorporated into the account argued for here, while preserving its core results.

Fist, I will re-define the translation of transitive verbs as relations between the internal argument and the event:

(220) find $\leadsto \lambda v_\mathbf{e}.\lambda \zeta_\mathbf{v}.\ [find\{\zeta\},\ \mathbf{Th}\{v,\ \zeta\}]$

Agents will be introduced by a separate syntactic head, which I identify with v:

(221) $v \leadsto \lambda P_\mathbf{vt}.\lambda v_\mathbf{e}.\lambda \zeta_\mathbf{v}.\ [\mathbf{Ag}\{v,\ \zeta\}];\ P(\zeta)$

Finally, I will assume that quantificational DPs are able to combine directly with relations of type (**e,vt**). To make this work, we will need to generalize the $\langle\rangle$-relation, previously defined for a single dref (cf. 90):

(222) $\langle u_1, \ldots, u_n \rangle := \lambda I_{st}.\lambda J_{st}.\ \exists f.\ \forall j_s \in J.\ \exists i_s \in I.\ (j \in f(i)) \wedge$
$\forall i_s \in I.\ \forall H_{st}.\ (f(i) = H \rightarrow H \subseteq J \wedge \forall h_s \in H.\ (i[u_1,\ldots,u_n]h \wedge$
$\mathbf{atom}(u_1 h) \wedge \ldots \wedge \mathbf{atom}(u_n h)) \wedge \oplus u_1 H = ui \wedge \ldots \wedge \oplus u_n H = u_n i)),$
where f is a function from the domain of assignments $\mathcal{D}_s$ to the set of info states $\wp(\mathcal{D}_s)$.

The $\langle\rangle$-relation is now able to split multiple drefs into their atomic parts, and 'co-distribute' these parts across the assignments in a plural info state. Based on (222) we can define three-place versions of the distributivity operators:

(223) a. $\mathbf{dist}_w(D)(u)(u') := \lambda I_{st}.\ \lambda J_{st}.\ (\langle u, u' \rangle;\ D) I J$
b. $\mathbf{dist}_s(D)(u)(u') := \lambda I_{st}.\ \lambda J_{st}.\ (\langle u, u' \rangle;\ \mathbf{dist}(D)) I J$

We can then adopt the following modified translations of *every* and *all*:

(224) a. all$^{u,u'} \leadsto \lambda P_\mathbf{et}.\lambda P'_\mathbf{e,vt}.\lambda \zeta_\mathbf{v}.\ [u',u];\ [\mathbf{ALL}\{u',u\}];$
$\mathbf{max}_{u'}(\mathbf{dist'}_w(P(u'))(u'));$
$\mathbf{max}_u(\mathbf{dist}_w(P(u);\ P'(u)(\zeta))(u)(\zeta))$
b. every$^{u,u'} \leadsto \lambda P_\mathbf{et}.\lambda P'_\mathbf{e,vt}.\lambda \zeta_\mathbf{v}.\ [u',u];\ [\mathbf{EVERY}\{u',u\}];$
$\mathbf{max}_{u'}(\mathbf{dist'}_s(P(u'))(u'));$
$\mathbf{max}_u(\mathbf{dist}_s(P(u);\ P'(u)(\zeta))(u)(\zeta))$

Consider now the compositional translation of (219a):

(225)

VP
$\lambda \zeta.\ [u',u];\ [\mathbf{EVERY}\{u',u\}];$
$\mathbf{max}_{u'}(\mathbf{dist'}_s([\mathbf{atom}\{u'\}];\ [\mathbf{unique}\{u'\}];\ [mistake\{u'\}])(u'));$
$\mathbf{max}_u(\mathbf{dist}_s([\mathbf{atom}\{u\}];\ [\mathbf{unique}\{u\}];\ [mistake\{u\}];\ [find\{\zeta\},\ \mathbf{Th}\{u,\ \zeta\}])(u)(\zeta))$

find
$\lambda v.\lambda \zeta.\ [find\{\zeta\},\ \mathbf{Th}\{v,\ \zeta\}]$

every$^{u,u'}$ mistake
$\lambda P.\lambda \zeta.\ [u',u];\ [\mathbf{EVERY}\{u',u\}];$
$\mathbf{max}_{u'}(\mathbf{dist'}_s([\mathbf{atom}\{u'\}];\ [\mathbf{unique}\{u'\}];\ [mistake\{u'\}])(u'));$
$\mathbf{max}_u(\mathbf{dist}_s([\mathbf{atom}\{u\}];\ [\mathbf{unique}\{u\}];\ [mistake\{u\}];\ P(u)(\zeta))(u)(\zeta))$

(226)

$$\text{vP}$$

$$\lambda\zeta.\ [u'']; [\textbf{3\_atoms}\{u''\}]; [\textit{copy-editor}\{u''\}]; [\textbf{Ag}\{u'',\ \zeta\}];$$
$$[u',u]; [\textbf{EVERY}\{u',u\}];$$
$$\textbf{max}_{u'}(\textbf{dist'}_s([\textbf{atom}\{u'\}]; [\textbf{unique}\{u'\}]; [\textit{mistake}\{u'\}])(u'));$$
$$\textbf{max}_u(\textbf{dist}_s([\textbf{atom}\{u\}]; [\textbf{unique}\{u\}]; [\textit{mistake}\{u\}]; [\textit{find}\{\zeta\},\ \textbf{Th}\{u,\ \zeta\}])(u)(\zeta))$$

```
three^{u''} copy editors
```
$$\lambda P'.\ [u'']; [\textbf{3\_atoms}\{u''\}];$$
$$[\textit{copy-editor}\{u''\}];\ P'(u'')$$

$$\text{v'}$$

$$\text{Lifted: } \lambda Q.\lambda\zeta.\ Q(\lambda v.\ [\textbf{Ag}\{v,\ \zeta\}]; [u',u]; [\textbf{EVERY}\{u',u\}];$$
$$\textbf{max}_{u'}(\textbf{dist'}_s([\textbf{atom}\{u'\}]; [\textbf{unique}\{u'\}]; [\textit{mistake}\{u'\}])(u'));$$
$$\textbf{max}_u(\textbf{dist}_s([\textbf{atom}\{u\}]; [\textbf{unique}\{u\}]; [\textit{mistake}\{u\}]; [\textit{find}\{\zeta\},\ \textbf{Th}\{u,\ \zeta\}])(u)(\zeta));$$

$$v$$

$$\lambda P.\lambda v.\lambda\zeta.\ [\textbf{Ag}\{v,\ \zeta\}];\ P(\zeta)$$

$$\text{VP}$$

```
find every^{u,u'} mistake
```

After event closure applies, we obtain the following DRS:

(227) $[\varepsilon]; [u'']; [\textbf{3\_atoms}\{u''\}]; [\textit{copy-editor}\{u''\}]; [\textbf{Ag}\{u'',\ \varepsilon\}]; [u',u];$
$[\textbf{EVERY}\{u',u\}];$
$\textbf{max}_{u'}(\textbf{dist'}_s([\textbf{atom}\{u'\}]; [\textbf{unique}\{u'\}]; [\textit{mistake}\{u'\}])(u'));$
$\textbf{max}_u(\textbf{dist}_s([\textbf{atom}\{u\}]; [\textbf{unique}\{u\}]; [\textit{mistake}\{u\}]; [\textit{find}\{\varepsilon\},$
$\textbf{Th}\{u,\ \varepsilon\}])(u)(\varepsilon))$

This DRS will be true with respect to a singleton input info state I iff there is an output info state J constructed in the following way. First, we introduce three individual drefs u, u' and u'', and an event dref ε:

(228)

Info state K, s.t. $I[u, u', u'', \varepsilon]K$	...	u	u'	u''	ε	...
k	...	m	m'	$c_1 \oplus c_2 \oplus c_3$	e	...

Here, $u''k$ must be a sum of three copy editors, εk is a sum of events, $u''k$ is the (cumulative) agent of εk, and uk and $u'k$ must return sums of equal cardinality. Moreover, u' must return the maximal sum of mistakes.

Next, following the definition of $\textbf{dist}_s$ in (223b) and the $\langle\rangle$-relation in (222), we split both uk and εk into atomic parts, and distribute these across multiple singleton info states:

(229)

Info state H_1'		. . .	u	u'	u''	ε	. . .
h_1'		. . .	m_1	m'	$c_1 \oplus c_2 \oplus c_3$	e_1	. . .

Info state H_2'		. . .	u	u'	u''	ε	. . .
h_2''		. . .	m_2	m'	$c_1 \oplus c_2 \oplus c_3$	e_2	. . .

. . .

Info state H_n'		. . .	u	u'	u''	ε	. . .
h_n'		. . .	m_n	m'	$c_1 \oplus c_2 \oplus c_3$	e_n	. . .

Here, for each H_k' the value for u is an atomic mistake m_k, and the value for ε is an atomic finding event whose theme is m_k. All these info states are then 'glued' together to produce the output info state J:

(230)

Info state J		. . .	u	u'	u''	ε	. . .
j_1		. . .	m_1	m'	$c_1 \oplus c_2 \oplus c_3$	e_1	. . .
j_2		. . .	m_2	m'	$c_1 \oplus c_2 \oplus c_3$	e_2	. . .
. . .		. . .	. . .	. . .	. . .	. . .	. . .
j_n		. . .	m_n	m'	$c_1 \oplus c_2 \oplus c_3$	e_n	. . .

The second maximality operator in (227) requires for uk in (228) to be the maximal sum such that an info state like J exists.

Summing up, the DRS in (227) will be true iff there is a sum of three copy-editors who were the cumulative agent of a sum of finding events E, the theme of each finding event in E was a mistake, and the cardinality of the total sum of mistakes found in E was equal to the total cardinality of mistakes. In other words, three copy-editors between them found all the mistakes. This adequately captures the cumulative reading of sentence (219a). The cumulative reading of (219b) will be derived in an analogous way, expect that the atomic sub-parts of uk and εk will be distributed across the assignments in a single info state.

Note, that the modified translations of *all* and *every* in (224) still include contrasting types of distributive operators scoping over the nuclear scope predicate. This means that we retain the prediction that only DPs with *all*, but not those with *every*, will license dependent plurals in their scope. Moreover, neither type of quantificational DPs is predicted to allow co-distributive readings with lower-scoping plurals involving numerals.

References

Asher, N. & Wang, L. (2003). Ambiguity and anaphora with plurals in discourse. In R. B. Young, & Y. Zhou (Eds.), *Procedings of SALT 13* (pp. 19–36). https://doi.org/10.3765/salt.v13i0.2879.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2), 159–219.

Beck, S. (2000). Star operators. Episode 1: Defense of the double star. In K. Kusumoto, & E. Villalta (Eds.), *UMass occasional papers in linguistics 23: Issues in semantics* (pp.1–23). Amherst, MA: GLSA.

Beck, S., & Sauerland, U. (2001). Cumulation is needed: A reply to Winter (2000). *Natural Language Semantics*, *8*, 349–371.

Bosveld-de Smet, L. (1998). *On mass and plural quantification: The case of French des/du NPs*. PhD thesis, University of Groningen.

Brasoveanu, A. (2007). *Structured nominal and modal reference*. PhD thesis, Rutgers, The State University of New Jersey.

Brasoveanu, A. (2008). Donkey pluralities: Plural information states versus non-atomic individuals. *Linguistics and Philosophy*, *31*(2), 129–209. https://doi.org/10.1007/s10988-008-9035-0.

Brasoveanu, A., & Farkas, D. F. (2011). How indefinites choose their scope. *Linguistics and Philosophy*, *34*(1), 1–55. https://doi.org/10.1007/s10988-011-9092-7.

Buccola, B., & Spector, B. (2016). Modified numerals and maximality. *Linguistics and Philosophy*, *39*(3), 151–199. https://doi.org/10.1007/s10988-016-9187-2.

Bylinina, L., & Nouwen, R. (2018). On "zero" and semantic plurality. *Glossa: A Journal of General Linguistics, 3*(1), 1–23. https://doi.org/10.5334/gjgl.441.

Carlson, G. N. (1984). Thematic roles and their role in semantic interpretation. *Linguistics*, *22*(3), 259–280.

Champollion, L. (2010a). cumulative readings of every do not provide evidence for events and thematic roles. In M. Aloni, H. Bastiaanse, T. de Jager, & K. Schulz (Eds.), *Logic, language and meaning* (pp. 213–222). Berlin: Springer.

Champollion, L. (2010b). *Parts of a whole: Distributivity as bridge between aspect and measurement*. PhD thesis, University of Pennsylvania.

Champollion, L. (2017). *Parts of a whole*. Oxford: Oxford University Press. https://doi.org/10.1093/oso/9780198755128.001.0001.

Champollion, L., Bledin, J. & Li, H. (2017). Rigid and flexible quantification in plural predicate logic. In D. C. Burgdorf, J. Collard, S. Maspong, & B. Stefánsdóttir (Eds.), *Proceedings of SALT 27* (pp. 418–437). https://doi.org/10.3765/salt.v27i0.4190.

Chierchia, G. (1995). *Dynamics of meaning: Anaphora, presupposition, and the theory of grammar*. Chicago, IL: University of Chicago Press.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond* (pp. 39–103). Oxford: Oxford University Press.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "Logicality" of language. *Linguistic Inquiry*, *37*(4), 535–590. https://doi.org/10.1162/ling.2006.37.4.535.

Chierchia, G., Fox, D., & Spector, B. (2012). Scalar implicature as grammatical phenomenon. In P. Portner, C. Maienborn, & K. von Heusinger (Eds.), *Semantics: An international handbook of natural language meaning* (Vol. 3, pp. 2297–2331). Berlin: Mouton de Gruyter.

Davies, M. (2008). *The corpus of contemporary American English: 450 million words, 1990-present*. http://corpus.byu.edu/coca/.

De Mey, S. (1981). The dependant plural and the analysis of tense. In V. A. Burke, & J. Pustejovsky (Eds.), *Proceedings of NELS 11* (pp. 58–78). Amherst, MA: GLSA.

Dobrovie-Sorin, C. (2014). Collective quantification and the homogeneity constraint. In T. Snider, S. D'Antonio, & M. Weigand (Eds.), *Proceedings of SALT 24* (pp. 453–472). https://doi.org/10.3765/salt.v24i0.2428.

Does, J. (1993). Sums and quantifiers. *Linguistics and Philosophy, 16*, 509–550.

Dowty, D. (1987). Collective predicates, distributive predicates, and *all*. In F. Marshall (Ed.), *Proceedings of the Third Eastern States Conference on Linguistics (ESCOL 3)* (pp. 97–115). Columbus, OH: The Ohio State University.

Elbourne, P. (2001). E-type anaphora as NP-deletion. *Natural Language Semantics*, *9*(3), 241–288. https://doi.org/10.1023/A:1014290323028.

Elbourne, P. (2005). *Situations and individuals*. Cambridge, MA: MIT Press. https://doi.org/10.1215/00318108-2007-046.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland, & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 71–120). London: Palgrave-Macmillan.

Fox, D., & Spector, B. (2018). Economy and embedded exhaustification. *Natural Language Semantics*, *26*(1), 1–50. https://doi.org/10.1007/s11050-017-9139-6.

Gillon, B. S. (1987). The readings of plural noun phrases in English author. *Linguistics and Philosophy*, *10*(2), 199–219.

Hackl, M. (2000). *Comparative quantifiers*. PhD thesis, Massachusetts Institute of Technology.

Heim, I. (1990). E-type pronouns, i-sums, and donkey anaphora. *Linguistics and Philosophy*, *13*(2), 137–177. https://doi.org/10.1007/BF00630732.

Heim, I. (2000). Degree operator and scope. In B. Jackson, & T. Matthews (Eds.), Proceedings of SALT 10 (pp. 40–64). https://doi.org/10.3765/salt.v10i0.3102.

Henderson, R. (2014). Dependent indefinites and their post-suppositions. *Semantics and Pragmatics*, *7*(6), 1–58. https://doi.org/10.3765/sp.7.6.

Ivlieva, N. (2013). *Scalar implicatures and the grammar of plurality and disjunction*. PhD thesis, MIT.

Ivlieva, N. (2020). Dependent plurality and the theory of scalar implicatures: Remarks on Zweig 2009. *Journal of Semantics*, *7*(3), 425–454. https://doi.org/10.1093/jos/ffaa004.

Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to model theoretic semantics of natural language, formal logic and discourse representation theory*. Dordrecht: Kluwer.

Karttunen, L. (1976). Discourse referents. In J. D. McCawley (Ed.), *Syntax and semantics7: Notes from the linguistic underground*. (pp. 363–385). New York: Academic Press.

Keenan, E. L., & Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, *9*, 253–326.

Keshet, E. (2018). Dynamic update anaphora logic: A simple analysis of complex Anaphora. *Journal of Semantics*, *35*(2), 263–303. https://doi.org/10.1093/jos/ffx020.

Kratzer, A. (2000). The event argument and the semantics of verbs. Ms., University of Massachusetts at Amherst.

Kratzer, A. (2007). On the plurality of verbs. In J. Dölling, T. Heyde-Zybatow, & M. Schäfer (Eds.), *Event structures in linguistic form and interpretation* (pp. 269–300). Berlin, Boston: de Gruyter.

Krifka, M. (1989). *Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen*. München: Fink.

Krifka, M. (1996). Parametrized sum individuals for plural anaphora. *Linguistics and Philosophy*, *19*, 555–598.

Krifka, M. (2004). Bare NPs: Kind-referring, indefinites, both, or neither? In R. B. Young, & Y. Zhou (Eds.), *Proceedings of SALT 13*, (pp. 180–203). https://doi.org/10.3765/salt.v13i0.2880.

Križ, M. (2016). Homogeneity, non-maximality, and all. *Journal of Semantics*, *33*(3), 493–539. https://doi.org/10.1093/jos/ffv006.

Kuhn, J. (2020). *Gather/numerous* as a mass/count opposition. *Natural Language Semantics*, 28, 225–253. https://doi.org/10.1007/s11050-020-09163-x.

Landman, F. (1989). Groups I. *Linguistics and Philosophy*, *12*(5), 559–605. https://doi.org/10.1007/BF00627774.

Landman, F. (1996). Plurality. In S. Lappin (Ed.), *The handbook of contemporary semantic theory* (pp. 425–457). Oxford: Blackwell.

Landman, F. (2000). *Events and plurality: The Jerusalem lectures*. Dordrecht: Springer.

Lasersohn, P. (1998). Generalized distributivity operators. *Linguistics and Philosophy*, *21*, 83–93.

Lassiter, D. (2010). The algebraic structure of amounts: Evidence from comparatives. In T. Icard, & R. Muskens (Eds.), *Interfaces: Explorations in logic, language and computation* (pp. 38–56). Berlin: Springer. https://doi.org/10.1007/978-3-642-14729-6_4.

Link, G. (1983). The logical anlysis of plurals and ass terms: A lattice-theoretical approach. In R. Baeuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use and interpretation of language* (pp. 302–323). Berlin: Mouton de Gruyter.

Link, G. (1987). Generalized quantifiers and plurals. In P. Gärdenfors (Ed.), *Generalized quantifiers: Linguistic and logical approaches* (pp. 151–180). Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-3381-1_6.

Minor, S. (2017). *Dependent plurals and the semantics of distributivity*. PhD dissertation, UiT - The Arctic University of Norway.

Muskens, R. (1996). Combining Montague semantics and discourse representation. *Linguistics and Philosophy*, *19*, 143–186.

Nouwen, R. (2003). *Plural pronominal anaphora in context: Dynamic aspects of quantification*. Utrecht: LOT.

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

Partee, B. (1987). Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh, & M. Stokhof (Eds.), *Studies in discourse representation theory and the theory of generalized quantifiers* (pp. 115–141). Dordrecht: Foris.

Partee, B. H. (1975). Comments on C. J. Fillmore's and N. Chomsky's papers. In R. Austerlitz (Ed.), *Scope of American linguistics* (pp. 197–209). Lisse: The Peter De Ridder Press.

Partee, B. H. (1985). "Dependent plurals" are distinct from bare plurals. Ms., University of Massachusetts at Amherst.

Partee, B. H., & Rooth, M. (1983). Generalized conjunction and type ambiguity. In R. Baeuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, use and interpretation of language* (pp. 361–383). Berlin: Mouton de Gruyter.

Roberts, C. (1990). *Modal subordination, anaphora, and distributivity*. New York: Garland.

Sauerland, U. (1994). Coditributivity and reciprocals. In V. Samiian, & J. Schaeffer (Eds.), *Proceedings of 24th Western conference on linguistics (WECOL 94)* (pp. 224–237), Fresno, CA: California State University.

Sauerland, U. (1998). Plurals, derived predicates and reciprocals. In U. Sauerland, & O. Percus (Eds.), *The interpretive tract, MIT working papers in linguistics 25*. (pp. 177–204). Cambridge, MA: MITWPL.

Sauerland, U. (2003). A new semantics for number. In R. B. Young, & Y. Zhou (Eds.), *Proceedings of SALT 13* (pp. 258–275). https://doi.org/10.3765/salt.v13i0.2898.

Sauerland, U., Anderssen, J. & Yatsushiro, K. (2005). The plural is semantically unmarked. In S. Kepser, & M. Reis (Eds.), *Linguistic evidence* (pp. 409–430). Berlin: Mouton de Gruyter.

Scha, R. (1984). Distributive, collective and cumulative quantification. In J. Groenendijk, T. M. V. Janssen, & M. Stokhof (Eds.), *Truth, interpretation, and information* (pp. 131–158). Dordrecht: Foris.

Schein, B. (1993). *Plural and events*. Cambridge, MA: MIT Press.

Schwarzschild, R.(1996). *Pluralities*. Heidelberg: Springer.

Spector, B. (2003). Plural indefinite DPs as PLURAL-polarity items. In J. Quer, J. Shroten, M. Scorretti, P. Sleeman, & E. Verheugd (Eds.), *Romance languages and linguistic theory 2001: Selected papers from 'Going Romance'* (pp. 295–313). Amsterdam: John Benjamins.

Spector, B. (2007). Aspects of the pragmatics of plural morphology: On higher-order implicatures. In U. Sauerland, & P. Stateva (Eds.), *Presupposition and implicature in compositional semantics* (pp. 243–281). London: Palgrave-Macmillan.

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

Swart, H. de (2006). Aspectual implications of the semantics of plural indefinites. In S. Vogeleer & L. Tasmowski (Eds.), *Non-definiteness and plurality* (pp.169–189). Amsterdam: John Benjamins.

van den Berg, M. (1990). A dynamic predicate logic for plurals. In M. Stokhof, & L. Torenvliet (Eds.), *Proceedings of the 7th Amsterdam Colloquium* (pp. 29–51). ILLC, University of Amsterdam, Amsterdam.

van den Berg, M. (1993). Full dynamic plural logic. In K. Bimbo, & A. Mate (Eds.), *Proceedings of the Fourth Symposium on Logic and Language*. Budapest: Aron Publishers.

van den Berg, M. (1994). A direct definition of generalized dynamic quantifiers. In P. Dekker, & M. Stokhof (Eds.), *Proceedings of the 9th Amsterdam Colloquium* (pp. 121–140). ILLC, University of Amsterdam.

van den Berg, M. (1996). *Some aspects of the internal structure of discourse. The dynamics of nominal anaphora*. PhD thesis, University of Amsterdam.

Winter, Y. (2000). Distributivity and dependency. *Natural Language Semantics*, 8, 27–69.

Winter, Y. (2001). *Flexibility principles in Boolean semantics*. Cambridge, MA: MIT Press.

Zweig, E. (2008). *Dependent plurals and plural meaning*. PhD thesis, New York University.

Zweig, E. (2009). Number-neutral bare plurals and the multiplicity implicature. *Linguistics and Philosophy*, 32, 353–407.