Cumulative readings of modified numerals: A plural projection approach[1]

Nina HASLINGER — *Georg-August-Universität Göttingen*
Viola SCHMITT — *Karl-Franzens-Universität Graz*

Abstract. This paper addresses cumulative readings of modified-numeral DPs (MNs) like *exactly two boys*. Based on new German data, we argue that MNs are interpreted *in situ* in cumulative sentences, while the maximality condition contributed by the numeral modifier can take wider scope. We present an analysis that combines the Plural Projection system (Schmitt, 2019 a.o.), a surface-compositional approach to cumulativity, with a two-dimensional semantics for the numeral modifiers (Krifka, 1999 a.o.) and derives widest scope for the maximality condition. We then discuss 'non-maximal' readings of MNs (Buccola and Spector, 2016). We show that the availability of such readings depends on the syntactic positions of the MNs, which supports the idea that they are interpreted *in situ*, and sketch a way of deriving this fact.

Keywords: cumulativity, modified numerals, plurality, scope, plural projection.

1. Introduction

1.1. Background

Like other sentences with multiple plural expressions, sentences with plural DPs headed by modified numerals – we call such DPs **MN**s – can exhibit cumulative truth-conditions (Scha, 1981 a.o.): (1a) is true in the 'cumulative' scenario (1b), where exactly two students read books, but neither read exactly two books. (Throughout, Ada, Bea and Carl will be the students.)

(1) a. *Exactly two students read exactly two books.*
b. SCENARIO Ada read book 1. Bea read book 2. No other students read books.

As many have noted (e.g., Krifka, 1999; Landman, 2000; Brasoveanu, 2013; Buccola and Spector, 2016), the behavior of MNs in cumulative sentences raises a problem for the semantic analysis of such expressions. First, the cumulative reading of (1a) is clearly incompatible with the traditional generalized quantifier treatment of MNs in (2) (Barwise and Cooper, 1981).

(2) $[\![\textit{exactly two}\ \text{NP}]\!] = \lambda P_{\langle e,t\rangle}.|[\![\text{NP}]\!] \cap P| = 2$

But if we assume that MNs quantify existentially over pluralities as in (3) – following the intuition that cumulativity is a hallmark of plurality – we run into what is known as 'van Benthem's problem' (van Benthem, 1986) for MNs like *exactly two* that impose a **maximality condition** (i.e., an upper bound). Since (3) only requires the *existence* of a plurality of a certain size, it falsely predicts (4) to be true in a scenario where three students failed the exam.

(3) $[\![\textit{exactly two}\ \text{NP}]\!]$: $\lambda P_{\langle e,t\rangle}.\exists x_e[x \text{ is an } [\![\text{NP}]\!]\text{-plurality} \wedge |x| = 2 \wedge P(x)]$

(4) *Exactly two students failed the exam.*

The maximality condition is also observable in cumulative sentences like (1a): (1a) is false in all the scenarios in (5). We therefore cannot assume that cumulative sentences require special,

[1]Thanks to Jan Köpping, Izabela Jordanoska, Martin Prinzhorn, Benjamin Spector, the reviewers of *Sinn und Bedeutung* 24 and the audiences at the University of Düsseldorf, RTANJ 2, the 20th Szklarska Poręba Workshop and *Sinn und Bedeutung* 24. This research was supported by the Austrian Science Fund (FWF) project P-29240.

Proceedings of Sinn und Bedeutung 24, vol. 1, pp.323–340. Osnabrück University.

'non-maximal' denotations for MNs along the lines of (3). Instead, we need a semantics for MN that encodes the maximality condition both in distributive and in cumulative sentences.

(5) a. SCENARIO: Ada read book 1. Bea read book 2. Carl read book 3.
b. SCENARIO: Ada read book 1. Bea read book 2. Carl read book 2.
c. SCENARIO: Ada read book 1. Bea read book 2. Bea read book 3.

1.2. Claims

Landman (2000) argues that in cumulative sentences, the maximality condition is 'global', i.e., it has scope over all the plurals in the sentence. For (1a), this means that we cannot let each MN introduce a separate maximality condition as in (6a), while interpreting one MN in the scope of the other as in (6b). This would falsely predict (1a) to be true in scenario (5a) (cf. (6c)).

(6) a. $[\![$*exactly two* NP$]\!] = \lambda P_{\langle e,t\rangle}.\exists x_e[x$ is an $[\![\text{NP}]\!]$-plurality $\wedge |x| = 2 \wedge P(x) \wedge \neg\exists y_e[y$ is an $[\![\text{NP}]\!]$-plurality$\wedge x < y \wedge P(y)]]$
b. [*exactly two students* [1 [*exactly two books* [2 [t_1 ***read* t_2]]
c. 'There is a plurality x of two students that cumulatively read exactly two books and there is no plurality $x' > x$ of students that cumulatively read exactly two books.'

So, scope asymmetries between the MNs appear to be disregarded for the purposes of their maximality conditions. This seems to fit well with the traditional 'scopeless' conception of cumulative readings (see Scha, 1981; Sher, 1990; Sternefeld, 1998): The basic idea is that a cumulative reading involving n plurals always involves a cumulation operation that applies to an n-ary relation. In (1a), for instance, a 'cumulative version' $**[\![\textit{read}]\!]$ of the binary relation $[\![\textit{read}]\!]$ is derived by simultaneously summing up readers and objects they read (i.e., if A read book 1, and B read book 2, then $**[\![\textit{read}]\!] = \{\langle \mathbf{a}, \mathbf{b1}\rangle, \langle \mathbf{b}, \mathbf{b2}\rangle, \langle \mathbf{a}+\mathbf{b}, \mathbf{b1}+\mathbf{b2}\rangle\}$). On this approach, if the cumulative relation is not lexicalized, as in (7), the syntax has to derive an LF constituent denoting a suitable relation, e.g. the relation $[\lambda x.\lambda y.y$ wanted to marry $x]$ for (7).[2]

(7) *The two women wanted to marry the two men.* Beck and Sauerland (2000)

Yet, we will present two phenomena that strongly suggest that syntactic asymmetries matter for the interpretation of MNs and therefore should not be 'flattened out' in the way just sketched. Based on this, we will provide a novel treatment of MNs.

Point 1: We first present data showing that the 'plural component' of MNs must be interpreted *in situ*. This leaves us with a 'split-scope' puzzle: The plural component of a MN must be severed from its maximality condition, since the latter must be able to have widest scope. We account for this by combining the **plural projection** mechanism (Haslinger and Schmitt, 2018; Schmitt, 2019) – an analysis of cumulativity in which all plurals are interpreted *in situ* – with a **two-dimensional semantics** that ensures that the maximality condition is 'read off' at the root level and therefore gets widest scope, following Krifka (1999), Landman (2000) and Brasoveanu (2013). The system will derive cumulative readings in a surface-compositional manner. However, their derivation will involve two separate meaning components that basically specify pluralities 'of the right size' and pluralities that are 'too big', respectively. Since the two levels are combined in the truth definition, maximality will come out as a global condition.

[2] A cumulative scenario is e.g. one where woman 1 wanted to marry man 1 and woman 2 wanted to marry man 2.

Point 2: Our second empirical observation fits in well with our general goal of maintaining syntactic asymmetries between MNs in cumulative sentences: We show that the maximality condition of MNs is itself affected by such syntactic asymmetries. MNs permit a **non-maximal reading** (see also Buccola and Spector, 2016), but only if the MN has another plural expression in its scope. We will connect this effect to an independently motivated presupposition associated with such non-maximal readings, which can only be met if the MN has a plurality in its scope. Combined with our surface-compositional semantics for MNs in cumulative sentences, this yields the right distribution of non-maximal readings.

2. 'Split scope' of modified numerals in cumulative sentences

Building on Schmitt's (2019) analogous argument for plural definites, we first show that MNs participating in cumulativity may be interpreted *in situ*. Consequently, the position where the plural component of the MN is interpreted and the scope site of its maximality condition do not always coincide.[3] Consider (8a), which is true in scenario (8b) on one of its readings.

(8) a. *Die zwei Tutoren haben vom Carl verlangt*, [$_P$ ***genau zwei Artikel*** *zu lesen*]
the two TAs have of Carl required exactly two papers PRT read.INF
und [$_P$ *ein Buch zusammenzufassen*].
and a book summarize.INF
'The two TAs required Carl to read exactly two papers and summarize a book.'

b. SCENARIO: Last week, Carl had to miss two classes. He asked the TAs what he should do to catch up. The TA for class 1 told him to read paper 1. The TA for class 2 told him to read paper 2 and summarize one of the books on the syllabus.

The relevant reading involves a cumulative interpretation of the predicate conjunction relative to the two TAs: Since the TA for class 1 did not tell Carl to summarize a book, a distributive reading of the predicate conjunction does not capture scenario (8). But the MN 'exactly two papers' also has a cumulative interpretation relative to the two TAs: Neither TA asked Carl to read more than one paper. Thus, the predicate conjunction and the MN both have plural denotations that are relevant for the cumulative reading of the sentence, but the predicate conjunction contains the MN. This configuration is a problem for any analysis which, as sketched above, derives cumulative readings by syntactically deriving the relevant cumulative relation: Moving only the plural subject and the predicate conjunction would create the relation in (9a) and thus fail to account for the cumulative reading of the MN. Moving the MN and the plural subject yields the relation in (9b), leaving the cumulative reading of the predicate conjunction unexplained. And moving first the MN and then the remnant of the predicate conjunction yields a configuration with an unbound trace that has the wrong meaning (see Schmitt, 2019 for details).

(9) a. $\lambda P_{\langle e,t\rangle}.\lambda x_e.x$ required Carl to do P
b. $\lambda y_e.\lambda x_e.x$ required Carl to read y and summarize a book

We obtain a more adequate paraphrase of the cumulative reading of (8a) if we take both the part structure introduced by the MN and the part structure introduced by the predicate conjunction

[3] Schein (1993) makes a related point within a very different theory of plurality. In our terms, he shows that the part structure introduced by a MN in the scope of a distributive operator interacts with the cumulative interpretation of plurals outside the scope of that operator. It is therefore impossible to interpret all the plurals 'participating' in cumulativity at the same scope site. (8a) extends this conclusion to examples without distributive operators.

into account. Ignoring the maximality condition for a moment, (8a) can be paraphrased as saying that the TAs stand in a cumulative relation to a sum of three predicates like (10).

(10) $[\lambda x.x$ reads paper $1] + [\lambda x.x$ reads paper $2] + [\lambda x.x$ summarizes a book$]$

To derive a meaning along the lines of (10), the MN in (8a) (or at least the part of this DP that contributes pluralities of papers) must be interpreted in the scope of the predicate conjunction. What are the consequences of this result for the analysis of the maximality condition? Several analyses of cumulative readings of MN (Krifka, 1999; Landman, 2000; Brasoveanu, 2013) compute the maximality condition in a separate semantic dimension: Starting from an 'ordinary', non-maximal cumulative interpretation, the maximality condition is added at the sentence level and ends up taking scope over all plurals participating in the cumulative relation. For (8a), this would mean that the maximality condition takes scope above the matrix subject. This matches our intuitions about its interaction with the intensional verb *verlangt* 'required': The potential paraphrase (11a), which quantifies over numbers higher than 2 in the scope of 'required', is false in (8b), since it entails that reading more than two papers was forbidden. In the more adequate paraphrase (11b), the maximality condition outscopes 'required'.

(11) a. 'For every world w in which the TAs' requirements (taken together) are satisfied, there is no number $n > 2$ such that Max reads n books in w.'
b. 'There is no number $n > 2$ such that for every world w in which the TAs' requirements (taken together) are satisfied, Max reads n books in w.'

In sum, (8a) shows that existing arguments against a 'symmetric' analysis of cumulative sentences, where all relevant plurals move to the same scope site, extend to MNs: The MN in (8a) must remain within the predicate conjunction, which is in the scope of 'required'. Yet, the maximality condition appears to outscope 'required', hence for this meaning component of MNs, a 'scopeless' analysis could still be correct: (8a) does not falsify an analysis in which maximality comes in at the level of the root clause. We will now describe a system that combines the two-dimensional approach to maximality (Krifka, 1999; Landman, 2000; Brasoveanu, 2013) with the Plural Projection approach to cumulativity, which was motivated by data analogous to (8a) and will allow us to derive cumulative readings while interpreting all plurals *in situ*.

3. Plural Projection

The Plural Projection system has two crucial traits.[4] First, by default, an expression containing a semantically plural subexpression will itself count as semantically plural, regardless of its type. This requires a notion of plurality that extends to abstract model-theoretic objects like functions and truth values. For instance, as DP1 in (12) contributes pluralities of individuals, VP1 contributes pluralities of predicates, S1 pluralities of propositions and VP2 pluralities of predicates. This 'projection' mechanism for semantic plurality gives us a way of interpreting plurals in cumulative sentences *in situ*, which we will need to model the 'split scope' of MNs.

(12) $[_{\text{S2}}$ $[_{\text{DP2}}$ *the two TAs*$]$ $[_{\text{VP2}}$ *made* $[_{\text{S1}}$ *Carl* $[_{\text{VP1}}$ *read* $[_{\text{DP1}}$ *two papers*$]]]]]$

If DP1 is interpreted *in situ*, we cannot derive the cumulative reading of (12) via a cumulative relation between people and things they made Carl read, as there is no constituent denoting

[4]For reasons of space, we cannot spell out all the formal details of the system here; for a more precise discussion and independent support for the components of the analysis, see Schmitt (2019) and Haslinger and Schmitt (2018).

this relation. This motivates the second trait of the system: cumulative readings are due to a semantic rule applying whenever a node immediately dominates two plural expressions such that one of them denotes a plurality of functions of some type $\langle a,b\rangle$ and the other a plurality of matching arguments of type a. Since the 'projection' mechanism makes VP2 in (12) a plural expression, this rule will kick in at node S2 and derive a cumulative sentence meaning.

3.1. Cross-categorial plurality and plural sets

The idea that functional and propositional expressions can be plural requires a non-standard ontology. First, we posit a cross-categorial plurality-forming operation, which is primitive even for functional types. For example, one-place predicates can be summed up via an operation $+_{\langle e,t\rangle}$ which is not reducible to the sum operation on D_e or to conjunction in D_t.[5] We associate every semantic type a with an **atomic domain** A_a and a **full domain** D_a which also includes pluralities formed from the elements of D_a. The full domain D_a is closed under an operation $+_a$ which maps any subset of D_a to its sum. We stipulate that the structures $(\mathscr{P}(A_a)\backslash\{\emptyset\},\bigcup)$ and $(D_a,+_a)$ must be isomorphic. Thus, the elements of D_a are in a one-to-one correspondence with non-empty subsets of A_a. This is illustrated for individuals in (13a) and for one-place predicates in (13b). (We suppress type subscripts on $+$ and write $a+b$ for $+(\{a,b\})$.)

(13) a. If $\mathbf{a, b, c} \in A_e$, then $\{\mathbf{a},\mathbf{b},\mathbf{c},\mathbf{a}+\mathbf{b},\mathbf{b}+\mathbf{c},\mathbf{a}+\mathbf{c},\mathbf{a}+\mathbf{b}+\mathbf{c}\} \subseteq D_e$.
b. If $\mathbf{P},\mathbf{Q},[\lambda x_e.\mathbf{P}(x)\vee\mathbf{Q}(x)] \in A_{\langle e,t\rangle}$, then $\{\mathbf{P},\mathbf{Q},[\lambda x_e.\mathbf{P}(x)\vee\mathbf{Q}(x)],\mathbf{P}+\mathbf{Q},$ $\mathbf{P}+[\lambda x_e.\mathbf{P}(x)\vee\mathbf{Q}(x)],\mathbf{Q}+[\lambda x_e.\mathbf{P}(x)\vee\mathbf{Q}(x)],\mathbf{P}+\mathbf{Q}+[\lambda x_e.\mathbf{P}(x)\vee\mathbf{Q}(x)]\} \subseteq D_{\langle e,t\rangle}$.

Furthermore, semantically plural expressions do not simply denote pluralities from some domain D_a, but actually alternative sets containing such pluralities – what we call **plural sets**. Thus, besides the enriched domain D_a, there will be a type a^* of plural sets with elements from D_a. Since we do not want plurals with 'parts' of type a (say, conjunctions with conjuncts of type a) to be treated like predicates of type $\langle a,t\rangle$ by the composition rules, we assume that the domain A_{a^*} of plural sets and the domain $A_{\langle a,t\rangle}$ of one-place predicates are disjoint, but have the same algebraic structure – the operations $\cup,\cap$ and $\backslash$ are defined on both. We will write plural sets in square brackets **[ ]** instead of the usual set brackets. Some examples are given in (14).

(14) a. Some elements of A_{e^*}: $[\mathbf{a},\mathbf{b}]$, $[\mathbf{a}+\mathbf{b},\mathbf{b}+\mathbf{c}]$, $[\mathbf{a}+\mathbf{b}+\mathbf{c}]$
b. Some elements of $A_{\langle e,t\rangle^*}$: $[\mathbf{smoke},\mathbf{drink}]$, $[\mathbf{smoke}+\mathbf{drink},\mathbf{drink}+[\lambda x_e.\mathbf{drink}(x)$ $\vee\,\mathbf{smoke}(x)]]$, $[\mathbf{smoke}+\mathbf{drink}+[\lambda x_e.\mathbf{drink}(x)\vee\mathbf{smoke}(x)]]$

Our use of an extra level of plural sets is closely related to Alternative Semantics treatments of indefinites (Kratzer and Shimoyama, 2002). In particular, at the level of the root clause, we will obtain a set of pluralities of propositions and get to our truth conditions by existentially quantifying over this set. Before discussing the compositional system, we give some examples of expressions whose meanings involve plural sets in a non-trivial way.

3.2. Expressions that introduce plurality

Consider first some sample denotations for plural expressions, given the toy model (15a). Plural definites denote singleton plural sets containing the maximal element of the NP denotation (15b). In the interpretation of indefinites, on the other hand, plural sets play a non-trivial role:

[5] See Schmitt, 2019 for motivation of this move.

two papers denotes the set of all pluralities of two papers (15c).

(15) a. $[\![paper]\!] = \{\mathbf{p1}, \mathbf{p2}, \mathbf{p3}\}$, $[\![book]\!] = \{\mathbf{b1}, \mathbf{b2}\}$
b. $[\![\textit{the papers}]\!] = \mathbf{[p1 + p2 + p3]}$
c. $[\![\textit{two papers}]\!] = \mathbf{[p1 + p2, p1 + p3, p2 + p3]}$

We take conjunction to denote an operation $\bigoplus_a$ which maps any nonempty set $S \subseteq D_a$ to a single element of D_a. If the elements of S are not plural sets, they are simply summed up using $+$, but for plural sets, the semantic contribution of $\bigoplus$ is more complex: It forms the set of all pluralities obtained by selecting an element from each set in S and summing up the selected elements. This operation is formalized in (16) and illustrated by (17). As (17) shows, it applies to plural sets of functions and of individuals in exactly the same way.

(16) a. For any type a, the operation $\bigoplus_a : \mathscr{P}(D_a)\backslash\{\emptyset\} \to D_a$ is defined as follows:
For any nonempty $S \subseteq D_a$:
(i) If a is a non-plural type (i.e. a is not of the form b^*): $\bigoplus_a S = +_a S$.
(ii) If $a = b^*$ for some type b: $\bigoplus_a S = [\bigoplus_b(\{f(X^*) \mid X^* \in S\}) \mid$
f is a function from S to $D_b \wedge \forall X^* \in S : f(X^*) \in X^*]$
b. $[\![and_{\langle a^*,\langle a^*,a^*\rangle\rangle}]\!] = \lambda x_{a^*}.\lambda y_{a^*}.x \oplus_{a^*} y$ for any type a

(17) a. $[\![\textit{paper 1, paper 1 and paper 3}]\!] = \mathbf{[p1]} \oplus \mathbf{[p2]} \oplus \mathbf{[p3]} = \mathbf{[p1 + p2 + p3]}$
b. $[\![\textit{two papers and a book}]\!] = \mathbf{[p1 + p2, p1 + p3, p2 + p3]} \oplus \mathbf{[b1, b2]} = \mathbf{[p1 + p2 +}$ $\mathbf{b1, p1 + p3 + b1, p2 + p3 + b1, p1 + p2 + b2, p1 + p3 + b2, p2 + p3 + b2]}$
c. $[\![\textit{read and summarize}]\!] = \mathbf{[}(\lambda x.\lambda y.\mathbf{read}(x)(y)) + (\lambda x.\lambda y.\mathbf{summarize}(x)(y))\mathbf{]}$

(17c) illustrates one application of higher-type pluralities: analyzing plural-based readings of conjunctions of any type (Schmitt, 2019). More importantly, we can now implement the two assumptions sketched above – that cumulativity arises in local function-argument configurations, and that an expression with a semantically plural subexpression is itself plural.

3.3. Projection and Cumulative Composition

We first illustrate the results the composition mechanism should deliver for example (12). The embedded VP1 receives the denotation in (18a), which preserves the structure of the plural set $[\![\textit{two papers}]\!]$ in (15c). Informally speaking, (18a) is the result of applying the function **read** (abbreviated as **R**) 'pointwise' to the atoms of the pluralities in (15c). Similarly, for the embedded clause S1 in (12), we derive a plural set of propositions (18b) by pointwise application of the atoms in (18a) to Carl. Finally, we apply the matrix verb **make** to each atomic proposition in (18b), resulting in another plural set of predicates, (18c).[6]

(18) a. $[\![\textit{read two papers}]\!] = \mathbf{[R(p1) + R(p2), R(p1) + R(p3), R(p2) + R(p3)]}$
b. $[\![\textit{Carl read two papers}]\!] = \mathbf{[R(p1)(c) + R(p2)(c), R(p1)(c) + R(p3)(c),}$ $\mathbf{R(p2)(c) + R(p3)(c)]}$
c. $[\![\textit{make Carl read two papers}]\!] = \mathbf{[make(R(p1)(c)) + make(R(p2)(c)),}$ $\mathbf{make(R(p1)(c)) + make(R(p3)(c)), make(R(p2)(c)) + make(R(p3)(c))]}$

To interpret the matrix clause S2, we must compose (18c) with the plural set contributed by the

[6]We here assume that D_t is the set of all propositions (sets of possible worlds), not the set of truth values: *make* is thus a type $\langle t, \langle e,t\rangle\rangle$ function that combines with its argument via ordinary functional application.

subject DP, (19a), in a way that encodes cumulativity. The required cumulative truth conditions can be stated as follows: There is a predicate plurality P in (18c) such that each atomic part of P is satisfied by at least one atomic part of $\mathbf{t1}+\mathbf{t2}$ and each atomic part of $\mathbf{t1}+\mathbf{t2}$ satisfies at least one atomic part of P. Importantly, instead of directly mapping (18b) and (19a) to a proposition expressing this condition, we now introduce a rule that maps them to a set of pluralities of propositions, each of which essentially corresponds to a different cumulative scenario. Some sample elements of this set are given in (19b). Each element will be the result of combining the atoms of some predicate plurality from (18c) with the atoms of some individual plurality from (19a), such that each atom of the predicate plurality is used at least once and each atom of the individual plurality is used at least once. This is formalized by the notion of a **cover** (20): A cover is a relation between predicate parts and matching argument parts such that each atomic part of the predicate, as well as each atomic part of the argument, occurs in the relation.

(19) a. $[\![\textit{the two TAs}]\!] = [\mathbf{t1}+\mathbf{t2}]$

b. $[\![(12)]\!] = [\mathbf{make}(\mathbf{R}(\mathbf{p1})(\mathbf{c}))(\mathbf{t1}) + \mathbf{make}(\mathbf{R}(\mathbf{p2})(\mathbf{c}))(\mathbf{t2}), \mathbf{make}(\mathbf{R}(\mathbf{p1})(\mathbf{c}))(\mathbf{t2}) + \mathbf{make}(\mathbf{R}(\mathbf{p2})(\mathbf{c}))(\mathbf{t1}), \mathbf{make}(\mathbf{R}(\mathbf{p1})(\mathbf{c}))(\mathbf{t1}) + \mathbf{make}(\mathbf{R}(\mathbf{p3})(\mathbf{c}))(\mathbf{t2}), \mathbf{make}(\mathbf{R}(\mathbf{p1})(\mathbf{c}))(\mathbf{t2}) + \mathbf{make}(\mathbf{R}(\mathbf{p3})(\mathbf{c}))(\mathbf{t1}), \ldots]$

(20) a. Let $P \in D_a, x \in D_b$. A relation $R \subseteq A_a \times A_b$ is a **cover** of (P,x) iff $\boldsymbol{+}(\{P' \mid \exists x' : (P',x') \in R\}) = P$ and $\boldsymbol{+}(\{x' \mid \exists P' : (P',x') \in R\}) = x$.

b. Example: $\{(\mathbf{P},\mathbf{a}),(\mathbf{Q},\mathbf{b})\}$, $\{(\mathbf{P},\mathbf{b}),(\mathbf{Q},\mathbf{a})\}$ and $\{(\mathbf{P},\mathbf{a}),(\mathbf{Q},\mathbf{a}),(\mathbf{Q},\mathbf{b})\}$ are among the possible covers of $(\mathbf{P}+\mathbf{Q},\mathbf{a}+\mathbf{b})$.

The new composition rule, **Cumulative Composition (CC)** (21), proceeds as follows: It takes a plural set of type $\langle a,b\rangle^*$ (the 'function set') and a plural set of type a^* (the 'argument set'). For every relation R that is a cover of some plurality in the function set and some plurality in the argument set, it performs functional application for each function-argument pair in R and sums up the results, creating a plurality of type b. The type b pluralities obtained for the different covers are finally collected into a plural set of type b^*. The reader can check that this rule yields the plural set indicated in (19b) when applied to (18c) and (19a). This plural set, which is of type t^*, can then be mapped to a truth value via definition (22) – which states that a plural set of propositions is true iff it contains at least one plurality all atomic parts of which are true.

(21) Cumulative Composition (CC)

a. For any $P^* \in D_{\langle a,b\rangle^*}$ and $x^* \in D_{a^*}$: $\mathscr{C}(P^*,x^*) = [\bigoplus(\{P'(x') \mid (P',x') \in R\}) \mid \exists P \in pl^{*-1}(P^*), x \in pl^{*-1}(x^*) : R \text{ is a cover of } (P,x)]$

b. For any meaningful expressions ϕ of type $\langle a,b\rangle^*$ and ψ of type a^*, $[\phi\ \psi]$ is a meaningful expression of type b^*, and $[\![\phi\ \psi]\!] = \mathscr{C}([\![\phi]\!], [\![\psi]\!])$.

(22) The truth value of an expression α of type t^* in a world w is 1 iff $\exists p \in [\![\alpha]\!].\forall p' \leq_a p.p'(w) = 1$ and 0 iff $\forall p \in [\![\alpha]\!].\exists p' \leq_a p.p'(w) = 0$.

Since a and b in (21) may be arbitrary types, this rule generalizes immediately to 'subsentential' instances of cumulativity. For example, the embedded VP in (23a) is assigned the denotation in (23b) (**S** stands for **summarize**) – a plural set of type $\langle e,t\rangle^*$.

(23) a. *The TAs made Carl* ***read and summarize two papers***.

b. $\mathscr{C}([\mathbf{R}+\mathbf{S}],[\mathbf{p1}+\mathbf{p2},\mathbf{p1}+\mathbf{p3},\mathbf{p2}+\mathbf{p3}])$
$= [\mathbf{R}(\mathbf{p1})+\mathbf{S}(\mathbf{p2}), \mathbf{R}(\mathbf{p2})+\mathbf{S}(\mathbf{p1}), \mathbf{R}(\mathbf{p1})+\mathbf{S}(\mathbf{p3}), \mathbf{R}(\mathbf{p3})+\mathbf{S}(\mathbf{p1})\ldots]$

We have now specified the semantic operation behind cumulativity, but we still seem to lack a formal account of the 'projection' behavior exemplified in (18). But closer inspection reveals that we get this behavior almost for free with our composition rule in (21). Consider node VP1, where the non-plural predicate **read** combines with the plural set of individuals from (15c). Given what we have said so far, this configuration should lead to a type mismatch: To combine with a plural set of type e^*, the predicate would need to have type $\langle e,\langle e,t\rangle\rangle^*$, not just $\langle e,\langle e,t\rangle\rangle$. To resolve this mismatch, we stipulate that any meaning can be shifted to a singleton plural set via the operation in (24a).[7] This allows us to apply the CC rule in (21) to $[\![\textit{two papers}]\!]$ and the singleton plural set **[read]**. Since **read** is atomic, there is exactly one cover for each plurality of two papers – e.g., $\{(\mathbf{read},\mathbf{p1}),(\mathbf{read},\mathbf{p2})\}$ for the plurality $\mathbf{p1}+\mathbf{p2}$. The plural set returned by the CC rule, given in (24b), thus mirrors the structure of $[\![\textit{two papers}]\!]$, with **read** applied 'pointwise' to each plurality of papers. The results in (18b) and (18c) are derived analogously.

(24) a. For any type a: $[\![\Uparrow_a]\!] = \lambda x_a.[x]$
b. $\mathscr{C}([\mathbf{read}],[\![\text{two papers}]\!]) = \mathscr{C}([\mathbf{R}],[\mathbf{p1}+\mathbf{p2},\mathbf{p1}+\mathbf{p3},\mathbf{p2}+\mathbf{p3}])$
$= [\mathbf{R}(\mathbf{p1})+\mathbf{R}(\mathbf{p2}),\mathbf{R}(\mathbf{p1})+\mathbf{R}(\mathbf{p3}),\mathbf{R}(\mathbf{p2})+\mathbf{R}(\mathbf{p3})]$

To summarize, we introduced a new approach to semantic composition in plural sentences, which permits plural expressions of any category, including predicates and propositions. The default means of combining a plural expression with its sister is the CC rule which, when applied to a node whose daughters have types $\langle a,b\rangle^*$ and a^*, returns a plural set of type b^*. Consequently, expressions containing a semantically plural subexpression will themselves have plural meanings. This lets us reduce non-lexical cumulative relations, as in (12), to a combination of local cumulation operations, without having to assume a special LF syntax for plurals.

3.4. Back to our motivating example

We now return to our original example (8a). In the Plural Projection framework, there is no need to move the plurals from the embedded clause to a scope site in the matrix clause. This solves our dilemma concerning the cumulative reading of the predicate conjunction. To see this, consider Figure 1, a simplified version of (8a) without the numeral modifier. The structure of the plural set $[\![\textit{two papers}]\!]$ projects to $[\![\textit{read two papers}]\!]$: **read** being a non-plural predicate, there is only one cover for each paper-plurality. The predicate conjunction then combines with two plural sets of predicates, which are combined via $\bigoplus$. The output of $\bigoplus$ is a set of pluralities consisting of three properties each: two properties that each correspond to reading some particular paper and one that corresponds to summarizing a particular book. Due to further 'projection' steps, the matrix VP ends up denoting a set of pluralities of three properties, where two properties correspond to making Carl read some particular paper and the third corresponds to making Carl summarize a particular book. This set cumulates with the matrix subject, yielding a set of pluralities of propositions that is mapped to a truth value via definition (22).

In sum, we are now closer to an analysis of (8a), since the compositionality puzzle posed by the cumulative VP conjunction is solved. Yet, we still cannot model the maximality conditions of MNs and, in particular, the 'split scope' effect observed in Section 2. Although *zwei Artikel* 'two papers' denotes a set of pluralities of *exactly* two papers, the numeral in Figure 1 gets an 'at

[7] This type-shift is independently motivated within our system by the behavior of operations like conjunction: It behaves analogously for plural and non-plural conjuncts and should thus have a single lexical entry for both cases.

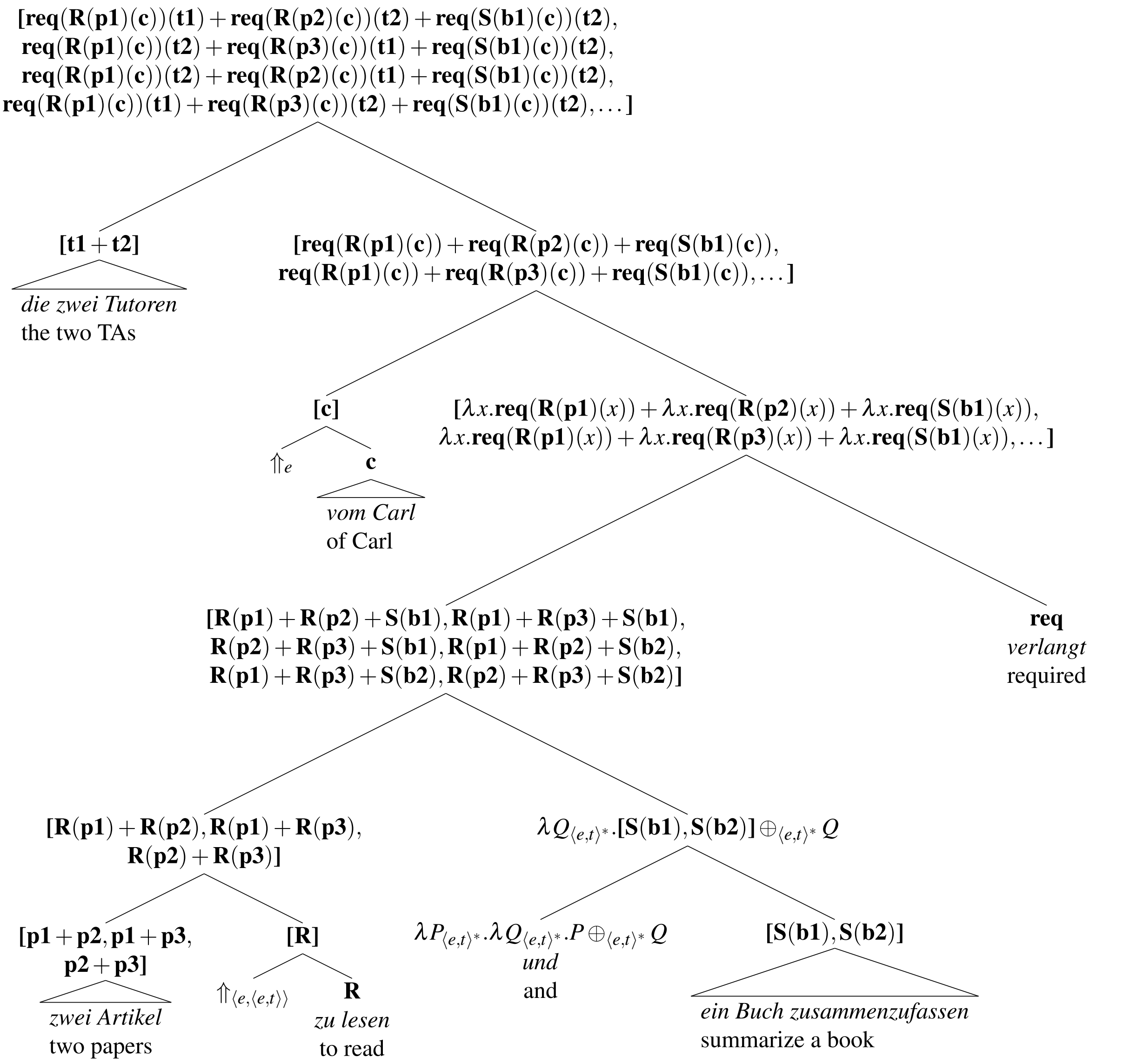

Figure 1: Derivation of a sentence analogous to (8a), without the numeral modifier

least' reading: As our truth definition in (22) quantifies *existentially* over sums of propositions which correspond to sums of two papers, we predict that the sentence can be true even if the TAs, between them, made Carl read *three* papers and summarize a book. So the present analysis of unmodified numerals does not generalize to MNs since it runs into the same problem as analyses based on classical existential quantification over plural individuals (cf. Section 1).

4. Adding the maximality condition to the system

Following most existing work on cumulative readings of non-upward-monotonic quantifiers (e.g. Krifka, 1999; Landman, 2000; Brasoveanu, 2013), we will assume a two-dimensional semantics for such quantifiers, resulting in two meaning components at the sentence level that are connected via the truth definition. In order not to prejudge the question how our system relates to other applications of two-dimensional semantics (such as scalar alternatives), we will call the two meaning components of an expression α the **'lower set'** $[\![\alpha]\!]^{\Uparrow}$ and the **'upper set'** $[\![\alpha]\!]^{\Downarrow}$.

At the level of the root clause, these will be two plural sets of propositions. The basic intuition will be that we existentially quantify over $[\![\alpha]\!]^{\Uparrow}$ as before – at least one plurality in $[\![\alpha]\!]^{\Uparrow}$ must consist of true propositions – but $[\![\alpha]\!]^{\Downarrow}$ additionally tells us which propositional pluralities are 'too big' to make α true. Consider first the two-dimensional meaning of the modifier *exactly n* in (25). While $[\![\textit{exactly n}]\!]^{\Uparrow}$ maps a predicate to the set of all sums of *n* individuals satisfying that predicate, $[\![\textit{exactly n}]\!]^{\Downarrow}$ yields the set of all sums of *n or more* such individuals.[8] For most non-quantificational expressions, the two meaning components are identical, (26).

(25) a. $[\![\textit{exactly n}]\!]^{\Uparrow} = \lambda P_{\langle e,t\rangle}.[x \in D_e \mid |x| = n \wedge P(x)]$
b. $[\![\textit{exactly n}]\!]^{\Downarrow} = \lambda P_{\langle e,t\rangle}.[x \in D_e \mid |x| \geq n \wedge P(x)]$

(26) If, for some expression α, $[\![\alpha]\!]$ is specified in the lexicon without specifying the dimension ($\Uparrow$ or $\Downarrow$), then $[\![\alpha]\!]^{\Uparrow} = [\![\alpha]\!]^{\Downarrow} = [\![\alpha]\!]$.

For a complex expression, composition proceeds for both dimensions in parallel: Its lower set is obtained by combining the lower sets of its daughter nodes, using whichever semantic rule is applicable – the CC rule or regular functional application. Its upper set is obtained in the same way from the upper sets of its daughter nodes. For instance, the two meaning components of *exactly two students* are computed by applying the functions in (25) to $[\![\textit{student}]\!]$:

(27) a. $[\![\textit{exactly two students}]\!]^{\Uparrow} = [\![\textit{exactly two}]\!]^{\Uparrow}([\![\textit{students}]\!]^{\Uparrow}) = [x \in D_e \mid |x| = 2 \wedge \mathbf{students}(x)] = [\mathbf{a}+\mathbf{b}, \mathbf{b}+\mathbf{c}, \mathbf{a}+\mathbf{c}]$ for $[\![\textit{student}]\!] = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$
b. $[\![\textit{exactly two students}]\!]^{\Downarrow} = [\![\textit{exactly two}]\!]^{\Downarrow}([\![\textit{students}]\!]^{\Downarrow}) = [x \in D_e \mid |x| \geq 2 \wedge \mathbf{students}(x)] = [\mathbf{a}+\mathbf{b}, \mathbf{b}+\mathbf{c}, \mathbf{a}+\mathbf{c}, \mathbf{a}+\mathbf{b}+\mathbf{c}]$ for $[\![\textit{student}]\!] = [\mathbf{a}, \mathbf{b}, \mathbf{c}]$

Let us say we want to derive a denotation for (4) by composing (27) with the non-plural predicate *failed the exam*. To combine with plural sets, both components of the latter must be shifted to the singleton plural set **[fail]**. We then apply the CC rule for each of the two dimensions:

(28) a. $[\![\textit{exactly two students failed}]\!]^{\Uparrow} = \mathscr{C}([\mathbf{fail}], [\![\textit{exactly two students}]\!]^{\Uparrow}) = [\mathbf{fail}(\mathbf{a}) + \mathbf{fail}(\mathbf{b}), \mathbf{fail}(\mathbf{b}) + \mathbf{fail}(\mathbf{c}), \mathbf{fail}(\mathbf{a}) + \mathbf{fail}(\mathbf{c})]$
b. $[\![\textit{exactly two students failed}]\!]^{\Downarrow} = \mathscr{C}([\mathbf{fail}], [\![\textit{exactly two students}]\!]^{\Downarrow}) = [\mathbf{fail}(\mathbf{a}) + \mathbf{fail}(\mathbf{b}), \mathbf{fail}(\mathbf{b}) + \mathbf{fail}(\mathbf{c}), \mathbf{fail}(\mathbf{a}) + \mathbf{fail}(\mathbf{c}), \mathbf{fail}(\mathbf{a}) + \mathbf{fail}(\mathbf{b}) + \mathbf{fail}(\mathbf{c})]$

How does such a pair of plural sets of propositions map to a truth value? Intuitively, a sentence

[8] Corresponding lexical entries for two other numeral modifiers are given in (i) and (ii). Note that for *at least n*, the two components are identical, which corresponds to the intuition that it lacks a maximality condition. Lexical entries like (i), (ii) and (25) can be derived systematically from the classical non-plural generalized-quantifier meaning of numeral modifiers (cf. (iii)). This suggests that (i), (ii) and (25) do not reflect the *basic* meaning of numeral modifiers. There is one potential reason to assume that the non-plural generalized-quantifier meanings of MNs play a role in the grammar: *at most n* lacks an existential entailment when combined with distributive predicates. However, we think it also lacks this entailment in some cumulative sentences, indicating that instead of positing a separate non-plural meaning, one should revise the plural meaning in (i).

(i) a. $[\![\textit{at most n}]\!]^{\Uparrow} = \lambda P_{\langle e,t\rangle}.[x \in D_e \mid |x| \leq n \wedge P(x)]$
b. $[\![\textit{at most n}]\!]^{\Downarrow} = \lambda P_{\langle e,t\rangle}.[x \in D_e \mid P(x)]$

(ii) $[\![\textit{at least n}]\!] = \lambda P_{\langle e,t\rangle}.[x \in D_e \mid |x| \geq 3 \wedge P(x)]$

(iii) For a determiner Q of type $\langle\langle e,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$, we can derive meanings $[\![\text{Q'}]\!]^{\Uparrow}$ and $[\![\text{Q'}]\!]^{\Downarrow}$ as follows:
a. $[\![\text{Q'}]\!]^{\Uparrow} = \lambda P_{\langle e,t\rangle}.[x \in D_e \mid P(x) \wedge [\![\text{Q}]\!](\lambda y_e.1)(\lambda y_e.y \leq_a x)]$
b. $[\![\text{Q'}]\!]^{\Downarrow} = \lambda P_{\langle e,t\rangle}.[x \in D_e \mid P(x) \wedge \exists y[y \leq x \wedge y \in [\![\text{Q'}]\!]^{\Uparrow}(\lambda y_e.1)]]$

will count as true if there is a plurality p in its lower set that consists only of true propositions, *and there is no plurality q in the upper set that consists only of true propositions and properly contains p as a part.* We formalize this by introducing an operator $\mathscr{W}$ that maps the two meaning components of an expression of type t^* to a single proposition:

(29) a. For two plural sets $L, U \in D_{t^*}$: $\mathscr{W}(L,U) = \lambda w.\exists p \in L.\forall p' \leq_a p.p'(w) = 1 \wedge \neg\exists q \in U.p < q \wedge \forall q' \leq_a q.q'(w) = 1$

b. The truth value of an expression ϕ of type t^* in a world w is $\mathscr{W}([\![\phi]\!]^{\Uparrow}, [\![\phi]\!]^{\Downarrow})(w)$.

According to (29), (4) is true if the following holds: First there must be a propositional plurality in (28a) that consists only of true parts. This is the case whenever there are two students who failed the exam. Second, there must be at least one such plurality that is not a proper part of a plurality in (28b) that only has true parts. This second condition is met whenever exactly two students failed. But if there is a sum of three students who failed, it will correspond to a plurality of three true propositions in (28b) which has proper parts in (28a), making the sentence false.

Now consider the maximality conditions of the MNs in the cumulative sentence (1a). The lower set of *exactly two books* contains all sums of two books. The structure of this set 'projects' when combined with **[read]** via CC (30a). Another application of CC yields the lower set at the clausal level (30b), which contains the propositional pluralities that 'cover' two books and two students. In contrast, the upper set of the VP also contains predicate pluralities that amount to reading more than two books (30c) – and at the sentence level, the upper set contains propositional pluralities which only count as true if more than two students read books, and propositional pluralities which only count as true if more than two books were read by students (30d). Recall that all scenarios in (5) above make (1a) false. Each corresponds to a propositional plurality that is in (30d), but not in (30b), and has proper parts in (30b). In each scenario, there will be some pluralities in (30b) that consist of only true propositions, but definition (29) predicts (1a) to be false as all such pluralities are proper parts of an element of (30d).

(30) $[\![\textit{student}]\!] = [\mathbf{a},\mathbf{b},\mathbf{c}]$. $[\![\textit{book}]\!] = [\mathbf{b1},\mathbf{b2},\mathbf{b3}]$

a. $[\![\textit{read exactly two books}]\!]^{\Uparrow} = \mathscr{C}([\mathbf{R}], [\![\textit{exactly two books}]\!]^{\Uparrow})$
$= [\mathbf{R}(\mathbf{b1}) + \mathbf{R}(\mathbf{b2}), \mathbf{R}(\mathbf{b2}) + \mathbf{R}(\mathbf{b3}), \mathbf{R}(\mathbf{b1}) + \mathbf{R}(\mathbf{b3})]$

b. $[\![\textit{exactly two students read exactly two books}]\!]^{\Uparrow} = [\mathbf{R}(\mathbf{b1})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{b}),$
$\mathbf{R}(\mathbf{b1})(\mathbf{b}) + \mathbf{R}(\mathbf{b2})(\mathbf{a}), \mathbf{R}(\mathbf{b1})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{b}), \ldots]$

c. $[\![\textit{read exactly two books}]\!]^{\Downarrow} = \mathscr{C}([\mathbf{R}], [\![\textit{exactly two books}]\!]^{\Downarrow})$
$= [\mathbf{R}(\mathbf{b1}) + \mathbf{R}(\mathbf{b2}), \mathbf{R}(\mathbf{b2}) + \mathbf{R}(\mathbf{b3}), \mathbf{R}(\mathbf{b1}) + \mathbf{R}(\mathbf{b3}), \mathbf{R}(\mathbf{b1}) + \mathbf{R}(\mathbf{b2}) + \mathbf{R}(\mathbf{b3})]$

d. $[\![\textit{exactly two students read exactly two books}]\!]^{\Downarrow} = [\mathbf{R}(\mathbf{b1})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{b}),$
$\mathbf{R}(\mathbf{b1})(\mathbf{b}) + \mathbf{R}(\mathbf{b2})(\mathbf{a}), \mathbf{R}(\mathbf{b1})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{b}), \ldots,$
$\mathbf{R}(\mathbf{b1})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{b}) + \mathbf{R}(\mathbf{b3})(\mathbf{b}), \mathbf{R}(\mathbf{b1})(\mathbf{a}) + \mathbf{R}(\mathbf{b2})(\mathbf{b}) + \mathbf{R}(\mathbf{b2})(\mathbf{c}), \ldots]$

We now return to our original puzzle concerning example (8a), which contains the MN *exactly two papers* embedded in a VP conjunction. The lower set of this VP conjunction is the same as the denotation we computed in Figure 1 for the corresponding VP without a numeral modifier (31a). The upper set additionally contains predicate pluralities that amount to reading more than two papers and summarizing a book, like $\mathbf{R}(\mathbf{p1}) + \mathbf{R}(\mathbf{p2}) + \mathbf{R}(\mathbf{p3}) + \mathbf{S}(\mathbf{b1})$ in (31b). (32a) gives some examples of propositional pluralities that are in the lower set at the sentence level (which is the plural set computed in Figure 1). Crucially, the structure of (31b) will project to the upper set of (8a) in a completely analogous manner (32b). Unlike (32a), (32b) will also contain sums of propositions which intuitively express that the TAs, between them, required Carl to read a certain sum of *more than two* papers and summarize a certain book.

(31) a. $[\![\textit{read exactly two papers and summarize a book}]\!]^{\Uparrow}$
$= \mathbf{[R(p1)+R(p2)+S(b1), R(p1)+R(p3)+S(b2), \ldots]}$

b. $[\![\textit{read exactly two papers and summarize a book}]\!]^{\Downarrow}$
$= \mathbf{[R(p1)+R(p2)+S(b1), R(p1)+R(p3)+S(b2), \ldots, R(p1)+R(p2)+R(p3)+ S(b1), R(p1)+R(p2)+R(p3)+S(b2), \ldots]}$

(32) a. $[\![(8a)]\!]^{\Uparrow} = \mathbf{[req(R(p1)(c))(t1)+req(R(p2)(c))(t2)+req(S(b1)(c))(t2)}$,
$\mathbf{req(R(p1)(c))(t2)+req(R(p3)(c))(t1)+req(S(b2)(c))(t2), \ldots]}$

b. $[\![(8a)]\!]^{\Downarrow} = \mathbf{[req(R(p1)(c))(t1)+req(R(p2)(c))(t2)+req(S(b1)(c))(t2)}$,
$\mathbf{req(R(p1)(c))(t2)+req(R(p3)(c))(t1)+req(S(b2)(c))(t2), \ldots}$,
$\mathbf{req(R(p1)(c))(t1)+req(R(p2)(c))(t2)+req(R(p3)(c))(t2)}$
$\mathbf{+\,req(S(b1)(c))(t2), \ldots]}$

In a scenario where the TAs, between them, assigned *three* papers to Carl and one of them made him summarize a book, (32b) will contain a sum that consists only of true propositions. As this plurality will have proper parts in (32a), but is not itself in (32a), (8a) is correctly predicted to be false. But if there are exactly two papers Carl was required to read by the TAs, and one book he had to summarize, this will translate into a propositional plurality in (32a) that is *not* a proper part of any true element of (32b), and the sentence is predicted to be true. Thus, we now have an account of the seemingly conflicting properties of cumulative sentences with MNs: We can interpret the plurals *in situ*, while the maximality condition is read off at the sentence level.

5. Non-maximal readings of MNs and syntactic asymmetries

We have modelled two properties of cumulative sentences: The fact that the 'plural component' of MNs in cumulative sentences is interpreted *in situ*, and the maximality condition. We will now show that there is an interesting relation between these seemingly independent components: MN can have non-maximal readings, but only in some syntactic configurations.[9] This interaction between the maximality condition and the syntactic position of a MN supports our asymmetric treatment of cumulativity. Yet, the non-maximal reading only surfaces under a certain *pragmatic* condition that resembles the presuppositions of scalar particles like *even*. We will formulate this condition in such a way that it can only be met if the MN combines with a set of non-trivial pluralities via CC, which will derive the distribution of non-maximal readings.

5.1. Syntactic constraints on the non-maximal reading: Sensitivity to scope

Like our example in (1a), the German sentence in (33a) contains two MNs and has a cumulative reading. Yet, it can be be true in the 'non-maximal cumulative' scenario (33b), where more than 2 students read books (cf. Buccola and Spector, 2016 for similar English examples).

(33) a. *Genau zwei Studenten haben genau 100 Bücher gelesen.*
exactly two students have exactly 100 books read
'Exactly two students read exactly 100 books.'

b. SCENARIO: A read 60 and B 40 books. C also read some of these books.

[9] This is not predicted by previous analyses, which either only derive the maximal reading (Krifka, 1999; Landman, 2000; Brasoveanu, 2013) or derive non-maximal readings 'across the board' (Buccola and Spector, 2016).

Importantly, the availability of such non-maximal readings is tied to the *syntactic position* of the MN (Schmitt, 2015): They are only possible if the MN c-commands (or, simplifying, outscopes) another plural expression.[10] This is illustrated in (34). (34b), where the MN subject *genau drei Wiener* ('exactly three Viennese') c-commands the plural object PP, is true in the 'non-maximal' scenario in (34a). But (34c), where the plural object PP has been scrambled across the subject – and the latter thus no longer c-commands a plural expression – lacks the non-maximal reading for the MN – the sentence is false in the scenario.[11]

(34) a. SCENARIO: A remote village has 50 inhabitants. Three Viennese people have many friends there: A is friends with 15 villagers, B with another 15, C with the remaining 20. Some other Viennese people have one or two friends there.

b. *Anscheinend sind genau drei Wiener mit allen Dorfbewohnern befreundet.*
apparently are exactly three Viennese with all villagers friendly
'Apparently, exactly three Viennese people are friends with all the villagers.'

c. *Anscheinend sind* [*mit allen Dorfbewohnern*]$_1$ *genau drei Wiener* t_1 *befreundet.*

5.2. Pragmatic constraints on the non-maximal reading: The 'surprise condition'

The availability of non-maximal readings in cumulative sentences is not just constrained by the syntactic position of the MNs. In (35a), *genau zwei Studenten* c-commands another plural, but cannot get a non-maximal reading: (35a) is false in the 'non-maximal' scenario in (35b).

(35) a. *Genau zwei Studenten haben genau drei Bücher gelesen.*
'Exactly two students read exactly three books.'

b. SCENARIO: A read book 1. B read book 2 and book 3. C also read book 3.

The contrast between (35a) and (33a) must be tied to the number of books read, since the sentences are otherwise identical. We submit that it is connected to an intuitive notion of 'surprise' or 'unexpectedness'. Assuming that people usually read 1-2 books, it is significantly more surprising that two students (so few students!) cumulatively read 100 books than if some larger plurality of students (say, 50) had done so. But since it is expected, and thus not surprising, for two students to cumulatively read three books, this is not significantly more unexpected than if some larger plurality of students had done so. We will now try to derive the distribution of non-maximal MNs from this 'pragmatic' constraint. Not only is it related to the restriction that the MN must have a plural in its scope in cumulative sentences – it will also explain why MNs with distributive predicates never permit non-maximal readings (as pointed out by Buccola and Spector, 2016 for English), regardless of the material in their scope.

The basic idea is the following. Non-maximal readings of MNs can only occur under a certain condition, which involves a 'significantly more surprising' (henceforth 'more surprising') relation that is probably related to the semantics of scalar particles like *even*. We will not attempt to formalize this relation here, but note that 1) it must be context-dependent and 2) a proposition

[10] A different kind of interaction between cumulativity and c-command or scope is found with *every*-DPs (Champollion, 2010) and related expressions in German and other languages (Haslinger and Schmitt, 2018, 2019), cf. the data in Schein, 1993; Kratzer, 2003. Yet, Haslinger and Schmitt (2019) argue that a direct correlation with scope does not capture the German data and appeal to additional constraints on c-command relations at different derivational stages. We have not yet investigated whether the data involving MNs also warrant such a qualification.

[11] The MN has the same thematic role in (34b) and (34c), hence the constraint cannot be captured in terms of thematic roles (as assumed by Buccola and Spector, 2016 for non-maximal readings more generally).

p clearly cannot be more surprising than a proposition q if q entails p.[12] (However, we will not identify the relation with 'less probable', given the issues Greenberg, 2016 raises for the case of *even*.) For a given MN like *exactly two students*, the condition requires that it is more surprising that some plurality in the lower set of the MN satisfies the predicate expressed by the sister of the MN (e.g., some plurality of exactly two students cumulatively read exactly 100 books) than that some larger plurality in the upper set has this property (e.g, some plurality of more than two students cumulatively read exactly 100 books). If the condition holds, the MN may be interpreted as non-maximal, while the maximality conditions of other plurals are unaffected.

Given this informal description, let us see how this condition interacts with our previous assumptions to derive the distribution of non-maximal readings. For a distributive predicate **P**, it cannot be more surprising that some plurality of exactly two students is **P** than that some plurality of more than two students is **P**, since the latter statement entails the former. We thus predict distributive sentences to lack non-maximal readings. To account for the structural asymmetry in cumulative sentences like (33a) or (35a), it is crucial that the surprise condition makes reference to the sister of the MN in question. Since our plural projection semantics from Section 3 derives cumulativity in several steps rather than forming a cumulative relation right away, there is a crucial difference between the sister of the syntactically lower MN and the sister of the syntactically higher MN: Only the latter will denote a non-trivial plural set. This property of our semantics is behind the observation in (34) that only MNs with a plural in their scope have non-maximal readings. If we combine the predicate in the scope of the lower plural – **[read]** in (33a) and (35a) – with a non-plural subject **x**, it cannot be more surprising that **x** read a plurality of exactly three books than that **x** read a plurality of more than three books, since the latter condition entails the former. So the surprise condition can never be met for the lowest MN in a cumulative sentence, which is thus expected to behave like MNs in distributive sentences and disallow non-maximal readings. In contrast, the higher MN combines with a non-trivial plural set (created via projection of the lower plural). Hence, we do not get the 'automatic' entailment relation that conflicts with the surprise condition in the other cases: It is possible that more than two students cumulatively read 100 books while no plurality of exactly two students cumulatively read 100 books. In sum, the surprise condition can only be met if the MN has another plural in its scope. Assuming that students usually read 1-2 books each, it will hold in (33a) (it is more surprising that some plurality of exactly 2 students cumulatively read exactly 100 books than that some plurality of more than 2 students did so), but not in (35a).

5.3. A preliminary analysis sketch

We now show informally how this idea can be added to the system developed in the previous sections. We posit an operator $\mathscr{S}$, defined in (36): It attaches to DPs (ψ) and considers certain results of combining their denotations with the denotations of their scope (ϕ). The presupposition of $\mathscr{S}$ encodes the surprise condition: It relates two propositions p, q, requiring that p is significantly more surprising than q. Each of p and q is derived by applying our truth-definition

[12]The role of entailment here closely resembles Buccola and Spector's (2016) pragmatic account for the lack of non-maximal readings in distributive sentences. On both accounts, it is crucial that distributive predication gives rise to a certain entailment pattern, but cumulative predication does not. Their proposal differs from ours in that our relation is meant to be stronger than entailment (based on the contrast between (33a) and (35a)) and in the way it interacts with the compositional semantics.

(29) to a pair of plural sets derived via CC: p states that there is a plurality in the lower set of the MN, which is such that one of the predicate pluralities in the lower set of ϕ – but no larger predicate plurality in the upper set of ϕ – cumulatively holds of it. q states that there is a plurality in the upper set of the MN which is not in its lower set and which is such that one of the predicate pluralities in the lower set of ϕ – but no larger predicate plurality in the upper set of ϕ – cumulatively holds of it. So, if our MN is *exactly two students*, p will relate to pluralities of two students, and q will relate to pluralities of more than two students, and both p and q will 'preserve' maximality conditions imposed by elements in the scope of the MN. If defined, the operator yields a new pair of plural sets: The lower set is obtained by CC in the usual way, but the upper set results from CC combining again the MN's *lower* set with the upper set of ϕ. Thus, $[\![\mathscr{S}]\!]$ removes the effect of the maximality condition from the MN denotation.

(36) $[\![[\mathscr{S}\ \psi]\ \phi]\!]^{\Uparrow}$ and $[\![[\mathscr{S}\ \psi]\ \phi]\!]^{\Downarrow}$ are defined only if $\mathscr{W}(\mathscr{C}([\![\phi]\!]^{\Uparrow},[\![\psi]\!]^{\Uparrow}),\mathscr{C}([\![\phi]\!]^{\Downarrow},[\![\psi]\!]^{\Uparrow}))$ is significantly more surprising than $\mathscr{W}(\mathscr{C}([\![\phi]\!]^{\Uparrow},[\![\psi]\!]^{\Downarrow}\backslash[\![\psi]\!]^{\Uparrow}),\mathscr{C}([\![\phi]\!]^{\Downarrow},[\![\psi]\!]^{\Downarrow}\backslash[\![\psi]\!]^{\Uparrow}))$.
  a. If the definedness condition is met: $[\![[\mathscr{S}\ \psi]\ \phi]\!]^{\Uparrow} = \mathscr{C}([\![\phi]\!]^{\Uparrow},[\![\psi]\!]^{\Uparrow})$
  b. If the definedness condition is met: $[\![[\mathscr{S}\ \psi]\ \phi]\!]^{\Downarrow} = \mathscr{C}([\![\phi]\!]^{\Downarrow},[\![\psi]\!]^{\Uparrow})$

Let us first see what happens when we insert this operator into the distributive sentence (4), as in (37a). The presupposition requires that (37b) is more surprising than (37c) (we use our toy model for illustration) – but this is impossible since the latter entails the former. Hence, when the operator attaches to a MN without a plural in its scope, it gives rise to a presupposition failure and we never get non-maximality. We assume that the operator is optional, but cannot be inserted if it leads to automatic presupposition failure (see e.g. Abrusán, 2011).

(37) a. [[$\mathscr{S}$ *exactly two students*] *failed the exam*]
  b. $\mathscr{W}(\mathscr{C}([\![\textit{fail}]\!]^{\Uparrow},[\![\textit{exactly two students}]\!]^{\Uparrow}),\mathscr{C}([\![\textit{fail}]\!]^{\Downarrow},[\![\textit{exactly two students}]\!]^{\Uparrow}))$
  $= \mathscr{W}(\mathscr{C}(\mathbf{[F]},\mathbf{[a+b,b+c,a+c]}),\mathscr{C}(\mathbf{[F]},\mathbf{[a+b,b+c,a+c]}))$
  = that there is a plurality consisting of two students that failed
  c. $\mathscr{W}(\mathscr{C}([\![\textit{fail}]\!]^{\Uparrow},[\![\textit{exactly two students}]\!]^{\Downarrow}\setminus[\![\textit{exactly two students}]\!]^{\Uparrow}),$
  $\mathscr{C}([\![\textit{fail}]\!]^{\Downarrow},[\![\textit{exactly two students}]\!]^{\Downarrow}\setminus[\![\textit{exactly two students}]\!]^{\Uparrow}))$
  $= \mathscr{W}(\mathscr{C}(\mathbf{[F]},\mathbf{[a+b+c]}),\mathscr{C}(\mathbf{[F]},\mathbf{[a+b+c]}))$
  = that there is a plurality consisting of more than two students that failed

Next, we consider cumulative sentences with two plurals like (33a) or (35a). We first check what happens when $\mathscr{S}$ attaches to the lower MN in (35a), as shown in (38a). As in (37), both meaning components of this MN's sister yield the 'trivial' plural set **[read]**, which has no interesting part structure. But as (38a) will require the operator to compare plural sets of one-place predicates, rather than propositions, a full analysis would require a cross-categorial version of (36), which space limitations prevent us from addressing. We thus make the simplifying assumption that the definedness condition checks if (36) holds whenever the predicate combines with an atomic individual. This yields the requirement in (38b) – which cannot be met. So any MN without a plural in its scope behaves like an MN in a distributive sentence: $\mathscr{S}$ cannot adjoin to this MN, as the presupposition would never be met, and we never derive a non-maximal reading.[13]

[13] We here omit the discussion of collective predicates, which Buccola and Spector (2016) cite as another context where MNs exhibit non-maximal readings. First, our system for cumulativity does not yet extend to collective predicates, so the predictions we make are unclear. More importantly, the empirical situation is also unclear to us.

(38) a. [*read* [$\mathscr{S}$ *exactly three books*]]
b. For a given individual **x**, **that there is a plurality of exactly three books such that x read each of them** is significantly more surprising than **that there is a plurality of more than three books such that x read each of them**

Finally, let us consider what happens when $\mathscr{S}$ attaches to an MN that *does* have a plural in its scope, as in (39a) (for our example (35a)). Here, $\mathscr{S}$ will require that the proposition in (39b) is significantly more surprising than the one in (39c). Since (39c) does not entail (39b), insertion of $\mathscr{S}$ in this case does not automatically lead to presupposition failure. This means we can now evaluate whether, given our assumptions about reading habits, (39b) is significantly more surprising than (39c). In a context where students are assumed to read 1-2 books on average, it is not, so the non-maximal reading should be blocked. But in the same context, a sentence with a higher numeral in object position, like (33a) above, would meet the surprise condition.

(39) a. [[$\mathscr{S}$ *exactly two students*] [*read exactly three books*]]
b. $\mathscr{W}(\mathscr{C}([\![\textit{read exactly three books}]\!]^{\Uparrow}, [\![\textit{exactly two students}]\!]^{\Uparrow}),$
$\mathscr{C}([\![\textit{read exactly three books}]\!]^{\Downarrow}, [\![\textit{exactly two students}]\!]^{\Uparrow}))$
= that there is a plurality of two students that cumulatively read three books and that did not read more than three books
c. $\mathscr{W}(\mathscr{C}([\![\textit{read exactly three books}]\!]^{\Uparrow}, [\![\textit{exactly two students}]\!]^{\Downarrow} \setminus [\![\textit{exactly two students}]\!]^{\Uparrow}),$
$\mathscr{C}([\![\textit{read exactly three books}]\!]^{\Downarrow}, [\![\textit{exactly two students}]\!]^{\Downarrow} \setminus [\![\textit{exactly two students}]\!]^{\Uparrow}))$
= that there is a plurality of more than two students that cumulatively read three books and that did not read more than three books

Thus, the restrictions on non-maximal readings follow from a particular pragmatic condition on such readings together with our two-dimensional semantics. The interesting empirical question how this condition relates to overt scalar particles must be left unresolved here.[14]

6. Conclusion and open problems

We developed a compositional system with the following properties: 1) Plurals in cumulative sentences are always interpreted *in situ*. 2) The maximality conditions associated with plural quantifiers, particularly MNs, are computed as a separate component of meaning within a two-dimensional semantics, which gives them scope over all plurals in the sentence. 3) Under certain contextual conditions, the maximality condition of MNs can be obviated, but only if they cumulate with another plural in their scope. Each of these properties is shared by some existing accounts of cumulativity for MN, but no previous analysis combines all of them.

We conclude by mentioning several problems and open questions. The most serious problem concerns a class of examples that systematically challenge assumption 2): Sentences with an MN in the scope of a distributivity operator and another MN that outscopes that operator, such as (40a). While we did not discuss distributivity yet, Haslinger and Schmitt (2019) give a Plural Projection analysis of *every*. Extending this approach to *each* is not trivial, but the general predictions it would make are clear: (40b) is taken to denote a set of propositional pluralities that correspond to particular assignments of two books to each student. When we add our

[14]Overt scalar particles like *selbst* (≈ English *just* in the use discussed by Panizza and Sudo, 2019) are not very good in cumulative sentences with non-maximal readings. Yet, their semantic impact in cases without cumulativity seems very similar to the role we attribute to $\mathscr{S}$. For the moment, we leave this issue to further research.

semantics for MNs, we get (40b) as our lower set for (40a), but the upper set would contain all propositional pluralities in which two or more students are each related to two or more books. We then predict (40a) to be false if there is a third student who read ten books, which does not seem correct. Intuitively, the problem is that the maximality condition of *exactly two books* should not 'project' unchanged through the distributivity operator.

(40) a. *Exactly two students (each) read exactly two books.*
b. $[\![$*two students each read two books*$]\!] = \mathbf{[R(b1)(a)+R(b2)(a)+R(b3)(b)}$ $\mathbf{+R(b4)(b), R(b1)(a)+R(b2)(a)+R(b2)(b)+R(b3)(b),\ldots]}$

One might take this to show that distributive sentences do not involve plural sets. But this is not an option in light of the more complex examples discussed by Schein (1993), where the higher MN additionally cumulates with another plural that also outscopes *each*. Schein shows that the truth conditions of such sentences are sensitive to the part structure introduced by the MN embedded under *each*. The problem posed by (40) must therefore be solved within a theory of cumulativity. While the mixed cumulative/distributive configurations discussed by Schein do not pose a general problem for our theory (see Haslinger and Schmitt, 2019), we do have a problem with those of his examples that involve two MNs on different 'sides' of a distributive operator. This suggests that there should be a more complex interaction between distributivity and the two-dimensional semantics, the details of which must be left to future work.[15]

Another issue for further research is the relation between the non-maximal reading of MNs and the special interpretation of plural definites, and plural universals like *all*, in cumulative sentences. Schmitt (2015) and Buccola and Spector (2016) show that it is sometimes sufficient if the cumulative predicate is true of a part of the plurality the DP introduces – e.g., (41) can be true even if some activists did not call any voters. This reading and the non-maximal reading of MNs seem to have a similar distribution (Schmitt, 2015), which suggests they are related.

(41) *The 50 activists called exactly 10 voters.*

Finally, our analysis of the modifiers was tailored for cumulative readings of MNs, but they are in fact cross-categorial and not restricted to numerals (Krifka, 1999). Our notion of 'upper' and 'lower sets' should thus follow from a general two-dimensional semantics for the modifiers. It is unclear how this could be done, especially in the case of scales that are not entailment-based.

References

Abrusán, M. (2011). Presuppositional and negative islands: a semantic account. *Natural Language Semantics 19*, 257–321.

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

Beck, S. and U. Sauerland (2000). Cumulation is needed: a reply to Winter (2000). *Natural Language Semantics 8*(4), 349–371.

[15] Schein's (1993) analysis of MNs does not face this criticism, as it is not two-dimensional. So does it account for cumulative sentences without encoding the maximality conditions in a plural-quantifier meaning for the MN – an idea we argued against in Section 1? To the extent we understand his proposal, it seems to avoid the problem discussed in Section 1 due to a syntactic LF for cumulative sentences where none of the MNs is in the scope of the others. It is, however, unclear to us 1) how maximality conditions are then computed compositionally and 2) how to encode the 'surprise' condition from Section 5, which is sensitive to plurals in the scope of the MN.

Brasoveanu, A. (2013). Modified numerals as post-suppositions. *Journal of Semantics 30*, 155–209.

Buccola, B. and B. Spector (2016). Modified numerals and maximality. *Linguistics and Philosophy 39*, 151–199.

Champollion, L. (2010). Cumulative readings of *every* do not provide evidence for events and thematic roles. In M. Aloni, H. Bastiaanse, T. de Jager, and K. Schulz (Eds.), *Logic, Language and Meaning*, Heidelberg, pp. 213–222. Springer.

Greenberg, Y. (2016). A novel problem for the likelihood-based semantics of *even*. *Semantics and Pragmatics 9*(2), 1–28.

Haslinger, N. and V. Schmitt (2018). Scope-related cumulativity asymmetries and cumulative composition. In S. Maspong, B. Stefánsdóttir, K. Blake, and F. Davis (Eds.), *Proceedings of SALT 28*.

Haslinger, N. and V. Schmitt (2019). Asymmetrically distributive items and plural projection. Ms., University of Vienna.

Kratzer, A. (2003). The event argument and the semantics of verbs. Ms., University of Massachusetts at Amherst.

Kratzer, A. and J. Shimoyama (2002). Indeterminate pronouns: The view from Japanese. In Y. Otsu (Ed.), *The Proceedings of the Third Tokyo Conference on Psycholinguistics*, Tokyo, pp. 1–25. Hituzi Syobo.

Krifka, M. (1999). At least some determiners aren't determiners. In K. Turner (Ed.), *The semantics/pragmatics interface from different points of view*, pp. 257–291. Elsevier.

Landman, F. (2000). *Events and Plurality*. Dordrecht: Kluwer Academic Publishers.

Panizza, D. and Y. Sudo (2019). Minimal sufficiency with covert *even*. To appear in *Proceedings of GLOW 42*.

Scha, R. (1981). Distributive, collective and cumulative quantification. In J. Groenendijk, T. Janssen, and M. Stokhof (Eds.), *Formal Methods in the Study of Language*, Volume 2, pp. 483–512. Amsterdam: Mathematisch Centrum.

Schein, B. (1993). *Plurals and Events*. MIT Press.

Schmitt, V. (2015). Asymmetries in plural predication. Talk at Sinn und Bedeutung 20, Tübingen.

Schmitt, V. (2019). Pluralities across categories and plural projection. *Semantics and Pragmatics 12*(17), 1–49.

Sher, G. (1990). Ways of Branching Quantifiers. *Linguistics and Philosophy 13*(4), 393–422.

Sternefeld, W. (1998). Reciprocity and cumulative predication. *Natural Language Semantics 6*(3), 303–337.

van Benthem, J. (1986). *Essays in Logical Semantics*. Kluwer Academic Publishers.