Weak Reciprocity without the Cumulative Operator

Jia Ren

University of Massachusetts Amherst; jiaren@umass.edu
* Correspondence: jiaren@umass.edu

Abstract: A proposal in studies on Weak Reciprocity is that sentences with WR interpretation have a cumulative syntax and semantics. (Sternfeld, 1998; Beck, 2001, etc.) In Beck (2001), the author derives the WR interpretation using the cumulative operator. In this paper, we challenge the view that WR sentences involve cumulative operators by discussing some problems with Beck's (2001) proposal on WR sentences. We will outline a new proposal on the structure of WR sentences using Skolemized covers. We derive the WR interpretation with pragmatic weakening instead of quantificational weakness. By the end, we revisit the debate between Winter (2000) and Beck and Sauerland (2000). We suggest that co-distributivity is involved in WR sentences, while cumulativity is not.

Keywords: Reciprocity, Cumulativity, Non-maximality

1. Introduction

In this paper we will focus on reciprocal sentences as below.

(1) The children kicked each other.

(1) is true in a scenario where every child kicked some other child, and every child was kicked by some other child. The interpretation is referred to as Weak Reciprocity (WR). It can be formalized as below.

(2) Weak Reciprocity (Langendoen, 1978)

$$\forall x[x \leq A \rightarrow \exists y \leq A[xRy \wedge x \neq y]] \ \& \ \forall y[y \leq A \rightarrow \exists x \leq A[xRy \wedge x \neq y]]$$

In (1), the children is the antecedent A of the reciprocal pronoun, kicked is the relation R involved. When the sentence is interpreted as WR, every individual in A must be the subject of kick and the object of kick. Moreover, when an individual is the subject of kick, the object must be some other individual in A, and conversely when an individual is the object of kick, some other individual in A must be the subject.

It has long been noticed that reciprocal sentences can have a range of interpretations. (Darymple et al., 1998) The interpretations differ in the strengths. For instance,

(3) The three men knew each other.

For (3), the most salient reading is that for each of the three men, he or she knew every other one of them. The interpretation is stronger than the one formalized in (2). Previous studies refer to this interpretation as Strong Reciprocity (SR), which can be formalized as below.

(4) Strong Reciprocity (SR)

$\forall x \leq A : \forall y \leq A [xRy \wedge x \neq y]$

When a sentence is interpreted as SR, it is assumed that the relation holds between every two different individuals in the antecedent.

The most salient interpretation of (5) is neither WR nor SR. The interpretation is that for each of the pirates, he or she stared at some other pirate(s). The interpretation is weaker than SR and WR.

(5) The pirates stared at each other.

This paper mainly discusses WR. A successful account of the meanings of sentences with reciprocals should offer a way of understanding why sometimes the interpretation is strengthened to SR or weakened to readings like in (5).

An important feature of reciprocal sentences is that the antecedents must be plural. Reciprocal relations cannot hold with a singular item, as shown by the infelicity of (6).

(6) # The child kicked each other.

Given the close relation between reciprocal sentences and plurals, an account of reciprocals necessarily involves studies on plurals. As for WR sentences, there are two main proposals which borrow existing studies in plurals to explain the WR interpretation. One proposal is that WR sentences involve the cumulative operator (Sternefeld, 1998). The cumulative operator is originally proposed for plural sentences with the cumulative inference (Krifka, 1986). The other proposal is that pragmatic restriction derives the WR interpretation (Schwarzschild, 1996). Beck (2001) uses the cumulative operator as the main device to derive WR, and she has some pragmatic restriction, namely ill-fitting covers (see Section 3), for cases like (5), which require further weakening.

In this paper, we challenge Sternefeld and Beck's view that the cumulative operator is needed to give rise to the WR interpretation. We provide evidence showing the absence of the cumulative operator. We propose a simplification of Beck's account. We provide new empirical evidence showing the need for skolemized covers. This innovation improves the expressive power of cover and helps derive the WR interpretation.

The paper will be as below. In section 2, we introduce the cumulative reading of plural sentences and the cumulative operator proposal. After that, we show how Beck derives WR with the cumulative operator. In section 3, we introduce how pragmatics plays a role in interpreting plural sentences. We will focus on Brisson's proposal on non-maximality. We will show how Beck incorporates non-maximality into her proposal. In section 4, We provide two pieces of evidence challenging the view that the cumulative operator derives WR. The first is the locality condition. The second is the compatibility with quantifiers like all. In section 5, we outline a new proposal on the structure of WR sentences. The proposal is simpler than Beck (2001) with no cumulative operator involved, thus avoiding the problems. WR sentences only make use of the distributive operator which is needed for independent reasons. WR sentences do not involve quantificational weakness, but non-maximality effect. We extend Brisson's account of non-maximality to certain new kinds of cases and suggest that cover takes a domain restriction variable. We derive the WR interpretation with the updated cover. In section 6, we test some predictions of the proposal. In section 7, we discuss some potential problems to the current proposal and future directions.

2. Deriving WR with the cumulative operator

In this section, we discuss the proposal of deriving WR sentences with cumulative operators. We explain cumulative predication and the cumulative operator first before we discuss how Beck (2001) derives WR with the cumulative operator.

2.1. The cumulative problem and the cumulative operator solution

Plural predications can have different logical properties.

(7) The men left.

In (7), for the sentence to be true, it's necessary for each of the men to leave. This reading is distributive, as what applies to the whole applies to each of the individuals distributively. Not all plural predications are distributive.

(8) The men met.

In (8), what applies to the whole does not apply to each individual distributively. For the sentence to be true, it is necessary for the men to form a group of meeters. This reading is collective, as the individuals mentioned collectively satisfy the predicate, while no individual satisfies the predicate by himself.

Schwarzschild (1996) gives an elegant analysis to the distributive-collective distinction. Following previous studies, Schwarzschild makes use of the distributive operator. It applies the predicate to each part of the argument. A primary definition of the distributive operator is given below.

(9) $[\![D]\!]=\lambda P_{\langle e,t\rangle}.\lambda x_e.\forall y[y\leq x \rightarrow P(y)]$

The distributive operator takes a one-place predicate as its first argument, and an individual as its second argument. A sentence is true if and only if for each part of the individual, the predicate applies to the parts. While the semantics of the distributive operator involves the parts of the argument, it remains silent about how the mereological structure of the argument is determined. Schwarzschild represents the mereological structure of the argument in the semantics of the distributive operator, thus explaining the distributive and collective distinction. The distinction lies in how the mereological structure of the argument is understood.

Schwarzschild proposes that the distributive operator distributes over a certain universe of discourse. As intuitively, there are always singular entities and plural entities, with the latter having the former as subparts, the universe of discourse is not flat. It has internal structures, where each item is in a cell of the universe. Cover, a concept in topology, is a good model of the above-mentioned universe of discourse, thus it is introduced to natural language semantics by Schwarzschild. The definition of cover is given below.

(10) C is a cover of a set A if and only if:

a. C is a set of subsets of A

b. Every member of A belongs to some set in C

c. ∅ is not in C

The distributive operator is co-indexed with a contextually assigned cover, with the cover restricting the domain of the distributive operator. The updated definition of the distributive operator is as below.

(11) $[\![D_i]\!]^g = \lambda P\_\langle e,t\rangle . \lambda x\_e. \forall y[y \leq x \wedge y \in cov_i \rightarrow P(y)]$

As before, the distributive operator takes a predicate and an individual as its arguments. The difference is that the predicate applies not to every part of the argument, but to every part of the argument which is also a member of the contextually given cover. The distributive-collective distinction can now be captured by different possible covers.

(12) Scenario: there are three men, A, B and C.

The men danced.

LF: [[The men] [D [danced]]

$\forall x[x \leq \text{the men} \wedge x \in cov \rightarrow x \text{ danced}]$

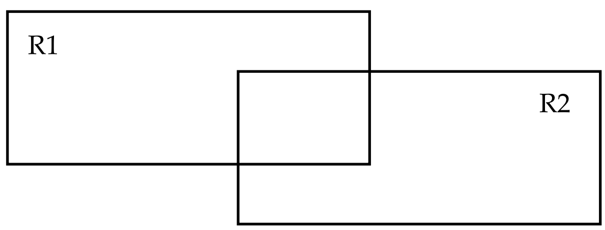
$cov_i = \{\{A\}, \{B\}, \{C\}\ldots\}$

$cov_j = \{\{A, B, C\}\ \ldots\}$

Given the LF, (12) means for all x such that x is a part of the men and x is in the cover, x danced. When the distributive operator is assigned covi, the sentence is true in a scenario where A danced, B danced, and C danced. A picture exemplifying the scenario will involve three solo dances. This is the distributive reading of the sentence. When the distributive operator is assigned covj, the sentence is true in a scenario where A, B and C danced together. A picture exemplifying the scenario will involve a joint dance by three performers. This is the collective reading of the sentence. For the rest of the section, we suppose by default that all covers are like covi in (12), where for the arguments the operator takes, each individual of the argument occupies an independent cell in the cover, so that each of the individual is distributed over by the predicate. Covers of other kind will be mentioned in Sec 3.

The proposal works well so far, yet only one-place predicates are considered. When we think about two-place predicates, there is a reading which cannot be captured by existing tools.

(13) The sides of R1 run parallel to the sides of R2. (Schwarzschild, 1998)



(13) is true in a scenario where each side of R1 runs parallel to some sides of R2, each side of R2 is parallel to some sides of R1. Given the existing proposal on capturing the distributive and collective reading, we predict there are four possible readings of the sentences, i.e., the subject interpreted collectively, and the object interpreted collectively; the subject interpreted distributively, and the object interpreted collectively; the subject

interpreted collectively, and the object interpreted distributively; the subject interpreted collectively, and the object interpreted collectively. The LFs of the four readings are as represented below.

(14) a. [[The sides of R1] [run parallel to [the sides of R2]]]

b. [[The sides of R1] Di [run parallel to [the sides of R2]]]

$\forall x(x \leq \text{the sides of R1} \wedge x \in cov_i \rightarrow x \text{ run parallel to the sides of R2})$

c. [[The sides of R1] [[the sides of R2] [D_i [1[run parallel to t_1]]]]]

$\forall x(x \leq \text{the sides of R2} \wedge x \in cov_i \rightarrow \text{the sides of R1 run parallel to } x)$

d. [[The sides of R1] [D_i [[the sides of R2] [D_j [1[run parallel to t_1]]]]]]

$\forall x(x \leq \text{the sides of R1} \wedge x \in cov_i \rightarrow \forall y(y \leq \text{the sides of R2} \wedge y \in cov_j \rightarrow x \text{ run parallel to } y)$

None of the above LFs captures the intended reading. The intended reading is that (13) is true in a scenario where each of the sides of R1 and each of the sides of R2 are involved in the parallel relation. The reading is weaker than the distributive interpretations. The distributive interpretations of the sentence will require each of the sides of R1 to be parallel to each of the sides of R2. The reading is also different from the collective interpretations. There is a grain of the distributive flavor in the intended reading which is necessary but absent in the collective interpretations. The intended interpretation wants each part of the subject to be the agent, and each part of the object to be the theme. The meaning is lost in collectivity. The cumulative inference poses a challenge, as it has a mixed flavor of both collectivity and distributivity.

To formalize the interpretation of data like (13), Krifka (1986) introduces the cumulative operator, which gives a cumulative interpretation to two place predicates. A cumulative interpretation involves, as summarized by Champollion (2020), two entities in a symmetric cross-product-like relation.

(15) $[\![**]\!] = \lambda P_{\langle e,\langle e,t\rangle\rangle} .\lambda x_e.\lambda y_e.\forall x' [x' \leq x \rightarrow \exists y' [y' \leq y \wedge P(x')(y')]] \wedge \forall y' [y' \leq y \rightarrow \exists x' [x' \leq x \wedge P(x')(y')]]$.

Beck (2000) gives the following updated definition of the cumulative operator by restricting the domain of the cumulative operator with covers, following Schwarzschild (1996).

(16) $[\![**_i]\!]^g = \lambda P_{\langle e,\langle e,t\rangle\rangle} .\lambda x_e.\lambda y_e.\forall x' [x' \leq x \wedge x' \in cov_i \rightarrow \exists y' [y' \leq y \wedge y' \in cov_i \wedge P(x')(y')]] \wedge \forall y' [y' \leq y \wedge y' \in cov_i \rightarrow \exists x' [x' \leq x \wedge x' \in cov_i \wedge P(x')(y')]]$.

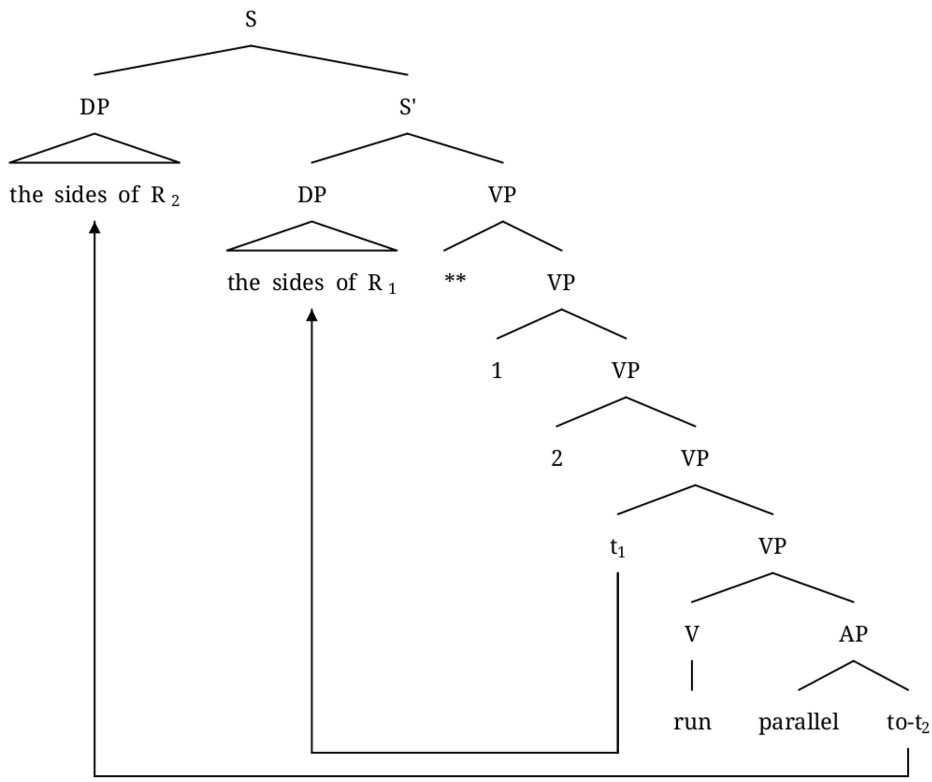
The cumulative operator takes a two-place predicate and two individuals, X and Y, as its arguments. For the sentence to be true, every part of the argument X which is also in the cover must be the subject of the relation, every part of the argument Y which is also in the cover must be the object of the relation. Moreover, when an individual in X is the subject of the relation, the object must be some individual in Y. Conversely, when an individual in Y is the object of the relation, some other individual in X must be the subject.

We will show how the operator gives the cumulative interpretation to (13), repeated below as (17).

(17) The sides of R1 run parallel to the sides of R2.

LF: [[The sides of R2] [the sides of R1] [$**_i$ [1[2[t_1 run parallel to t_2]]]]]

(18) $[\![S]\!]^g = \forall x'\ [x' \leq \text{the sides of R1} \wedge x' \in cov_i \rightarrow \exists y'\ [y' \leq \text{the sides of R2} \wedge y' \in cov_i \wedge \text{PARALLEL}(x')(y')]] \wedge \forall y'\ [y' \leq \text{the sides of R2} \wedge y' \in cov_i \rightarrow \exists x'\ [x' \leq \text{the sides of R1} \wedge x' \in cov_i \wedge \text{PARALLEL}(x')(y')]]$.



As defined in (16), the cumulative operator always takes a two-place predicate and adds a cumulative meaning to the predicate. To have a two-place predicate, some special syntactic movement is necessary. In (17), the subject the sides of R1 and the object the sides of R2 QR to sentence-initial positions, leaving behind a two-place predicate, as shown by the tree. The cumulative operator takes the predicate and gives the predicate a cumulative interpretation. The predicate thus means two arguments are in the run-parallel relation cumulatively. (17) is thus true when each of the sides of R1 runs parallel to some side of R2, and each of the sides of R2 runs parallel to some side of R1.

There is more than one theory on cumulativity, yet this cumulative operator proposal by Krifka (1986), later extended by Beck and Sauerland (2000), is successful in many ways. The proposal derives the cumulative inference in a straightforward manner. Besides, the proposal gives a good explanation to the locality condition of cumulative sentences.

The cumulative inference is local, as shown by the following pair of examples.

(19) Beck and Sauerland (2000)

a. Scenario: Max wants to marry a dentist named Ann; Peter wants to marry a dentist named Amy.

Max and Peter want to marry two dentists.

b. Scenario: Max said that Bill married a dentist named Ann; Peter said that Bill married a dentist named Amy.

\# Max and Peter said that Bill married two dentists.

The scenarios given in (19) force a cumulative reading of the sentences. (19)a is felicitous, while (19)b is not. (19)b cannot have a cumulative reading of Max and Peter and two dentists, it can only mean Max and Peter both said that Bill married two dentists. Under the cumulative operator proposal, the locality is explained. As in (17), QR is necessary for cumulativity. The cumulative operator takes two-place predicates, and two-place predicates are derived through QR. The locality of cumulativity comes from the locality of QR. As for (19), given that infinite clauses are not islands for QR, while finite clauses are, the cumulative inference is available in the former but not the latter.

2.2. Beck (2001)

We will show how Beck extends the cumulative operator proposal to WR sentences.

Beck proposes that each other is a definite noun. Following Heim et al. (1993), reciprocals denote a group containing all members of the antecedent, minus the individuals which are distributed over. The definition is as below.

(20) $[\![\text{each other}]\!]^g = \iota x[x \leq g(2) \wedge x \neq g(1)]$

LF: [the[other$_1$ of Pro$_2$]]

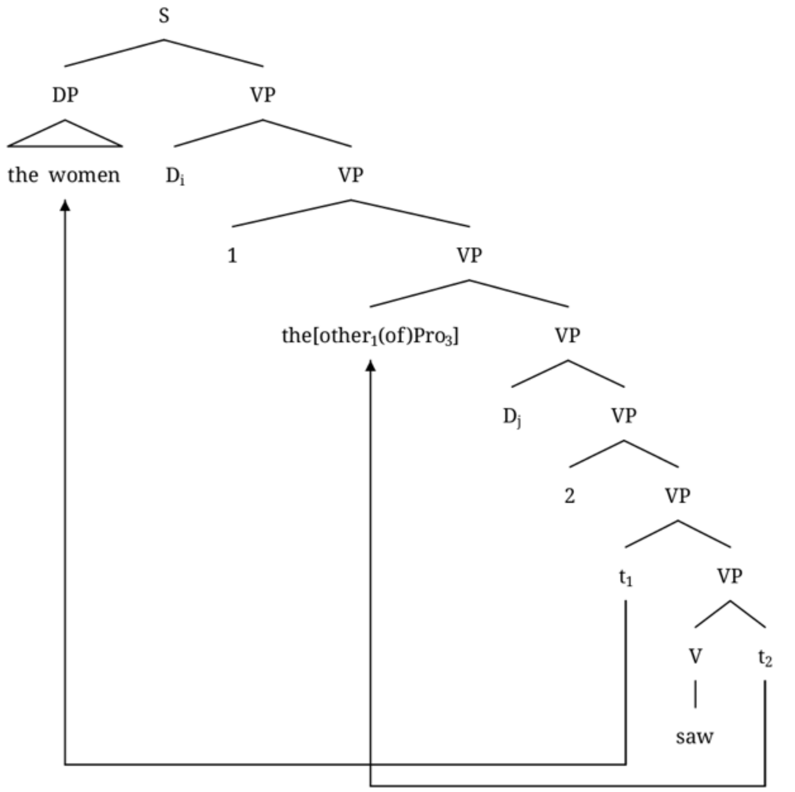
Each other takes two variables. One of them is a variable bound by the antecedent, the other is a pronoun which is co-referent to the antecedent. The reciprocal pronoun is a definite plural.

With the semantics of each other and the distributive operator, one can derive the interpretation of reciprocal sentences. The LF is as below.

(21) The women saw each other.

LF: [[The women] [D$_i$ [1 [[the [other$_1$ (of) Pro$_3$]] [D$_j$ [2 [t$_1$ saw t$_2$]]]]]]

$[\![S]\!]^g = \forall z' [z' \leq \text{the women} \wedge z' \in cov_i \rightarrow \forall y[y \leq \iota x[x \leq \text{the women} \wedge x \neq z'] \wedge y \in cov_j \rightarrow z' \text{ saw } y]]$



As mentioned in Sec 1, WR is weaker than the interpretation we derived above. As people introduce the cumulative operator to explain plural sentences which cannot be formalized using the distributive operator, people try to use the cumulative operator to derive WR. Based on previous study by Sternfeld (1998), Beck (2001) gives the following analysis to WR sentences.

(22) The children kicked each other.

LF: [Pro$_2$ [the children]$_1$ $**_i$ [1[2[t$_1$ kicked [the [[other x$_1$] (of) t$_2$]] …]

$[\![S]\!]^g = \forall z' [z' \leq \text{the children} \wedge z' \in cov_i \rightarrow \exists y' [y' \leq \text{the children} \wedge y' \in cov_i \wedge KICK(z')(\iota x[x \leq y' \wedge x \neq z'])]] \wedge \forall y' [y' \leq \text{the children} \wedge y' \in cov_i \rightarrow \exists z' [z' \leq \text{the children} \wedge z' \in cov_i \wedge KICK(z')(\iota x[x \leq y' \wedge x \neq z'])]]$.

The sentence means for every child, there is some other child that he or she kicked. Besides, for every child, he or she is kicked by some other child.

In this section, we review the cumulative operator approach to WR. We explain the cumulative inference and the cumulative operator. We also introduce how Beck extends the cumulative operator analysis to WR sentences. In the following section, we introduce another ingredient in Beck's proposal, namely weakening through non-maximality.

3. Non-maximality and WR

An immediate problem with the cumulative operator analysis to WR is that there are many reciprocal sentences whose interpretation are not captured by the cumulative proposal.

(23) The pirates stared at each other.

(23) is true in a scenario where every pirate stared at some other pirate, while not all pirates were stared at. This is weaker than the cumulative interpretation. Beck suggests that this is also an example of the WR. The meaning difference comes from pragmatic slackness, or the non-maximality effect. We will explain what is non-maximality before we introduce how Beck (2001) formalizes it.

3.1. *The non-maximality effect*

Plural definites can allow for exceptions.

(24) Bar-Lev (2021)
Context: There was a clown at my kid's birthday party. Someone asks me if they gave a funny performance. I reply:
The kids laughed.

The sentence above is judged true even if there were a few children who didn't laugh. The plural noun thus has a non-maximal reading in the sentence.

There are many proposals trying to explain the non-maximality effect. As their differences will not influence our discussion, we will only mention Brisson's proposal here. Brisson (2003) formalizes the non-maximality effect based on Schwarzschild's (1996) proposal on cover.

The definitions of cover and cover-based distributive operator are repeated below.

(25) C is a cover of a set A if and only if:
C is a set of subsets of A
Every member of A belongs to some set in C
∅ is not in C

(26) $[\![D_i]\!]^g = \lambda P_{\langle e,t\rangle}.\lambda x_e.\forall y\,(y \leq x \wedge y \in cov_i \rightarrow P(y))$

In Schwarzschild (1996), the cover of the distributive operator is a cover of the universe of discourse. In (12) we mention covers in which all subparts of the arguments taken by the distributive operators are in a cell with only other subparts of the same argument. We will show them again with (27).

(27) Context: there are six children, a, b, c, d, e and f. m is not a child.
The children danced.
LF: [The children [D [danced]]]
$[\![S]\!]^g = \forall y[y \in Cov \wedge y \leq \text{the children} \rightarrow y \text{ danced}]$

Cov_i = {{a}, {b}, {c}, {d}, {e}, {f}, {m…} …}
Cov_j = {{a, b, c, d, e, f}, {m…} …}

(27) means for all y such that y is a subpart of the children and y is in the cover, y danced. Suppose the cover is Cov_i. Each of a, b, c, d, e and f is an of the children, and is in the cover. Thus, the sentence means a, b, c, d, e and f danced, or each of the children danced. Similarly, suppose the cover is Cov_j. The set containing a, b, c, d, e and f is in the cover. The sentence thus means that the group containing a, b, c, d, e and f danced. A commonality of Cov_i and Cov_j is that every individual of the children is included in a cell with nothing else or only some other children. Thus, when the distributive operator takes the children as its argument, the predicate applies to all parts of the children. Covers like Cov_i and Cov_j are the so-called good-fit cover. The definition is given below.

(28) Good fit (Brisson, 2003)
For some cover of the universe of discourse Cov and some DP denotation X, Cov is a good fit with respect to X iff $\forall y[y \in X \rightarrow \exists Z[Z \in Cov \wedge y \in Z \wedge Z \subseteq X]]$.

Good-fit cover is a relative notion. A distributive operator has a co-indexed cover as its domain restrictor, and it takes an entity and a one-place predicate as its arguments. A cover is good fit with respect to the entity taken by the distributive operator, if all parts of the entity is in a cell with nothing else or only some other part of the entity in the cover, thus the predicate applies to all parts of the children collectively or distributively. Covers which are not good fit are ill-fitting. We give an example of ill-fitting covers below.

(29) Context: there are six children, a, b, c, d, e and f. m is not a child.
The children danced.
LF: [The children [D [danced]]]
$[\![S]\!]^g = \forall y[y \in Cov \wedge y \leq \text{the children} \rightarrow y \text{ danced}]$
$Cov_k = \{\{a\}, \{b\}, \{c\}, \{d\}, \{e\}, \{f, m\ldots\} \ldots\}$

As mentioned above, (27) means for all y such that y is a subpart of the children and y is in the cover, y danced. Besides, as given in the context, there are six children, a, b, c, d, e and f. m is not a child. In Cov_k, f, which is one of the children, is in a cell with a non-children m. Thus, given the semantics, the predicate does not apply to f. The children excluding f danced.

Ill-fitting covers give rise to the pragmatic weakening of plural definites. As mentioned in Laserson (1999), people speak loosely. People often say things which are not precisely true but are close enough to truth for practical purposes. The pragmatic slackness is captured by ill-fitting covers. When the distributive operator takes an entity, and a part of the entity is in a cell with non-participants in the cover, the part of the entity is in a pragmatic junkpile, which is silently ignorable at a particular moment. For instance, in (29), when Cov_k is used, f is silently ignorable. As long as almost all of the children danced, it's enough to say the children danced.

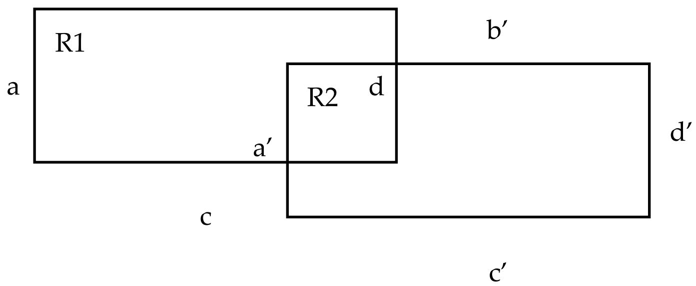
All, according to Brisson (2003), does not contribute any quantificational force. All eliminates from the set of all available covers any covers that are not a good fit with respect to the DP it is construed with. For the example above, all the children danced filtered out Cov_k. Thus, the sentences only have the maximal reading.

3.2. Non-maximality and reciprocals

In 3.1, we reviewed the non-maximal interpretation of plural definites and Brisson's (1999) formalization of non-maximality in distributive sentences. As mentioned in (16) in 2.1, Beck (2000) gives an updated definition of the cumulative operator based on cover. It is predicted that the non-maximality effect also exists in cumulative sentences and WR sentences. In this section, we show how Beck makes use of Brisson's ill-fitting cover proposal to derive the interpretations of cumulative sentences and certain WR sentences.

We will first show how Brisson's ill-fitting proposal can be extended to certain cumulative sentences. Scha (1984) has a series of good examples of cumulative sentences with imprecise non-maximal interpretations. We borrow one of them here.

(30) The sides of rectangle 1 cross the sides of rectangle 2. (Scha, 1984)



(30) is a plural sentence with the cumulative inference. In the picture, two sides of R1 cross two sides of R2. The scenario verifies (30), despite the fact that the cumulative inference requires each side of R1 to cross some side of R2, and each side of R2 to cross some side of R1. The plural definites in (30) have non-maximal interpretations in this scenario.

The non-maximal reading of the cumulative sentence can be explained with Beck's cumulative operator and Brisson's ill-fitting covers. We repeat Beck's definition of the cumulative operator below.

(31) $[\![**\_i]\!]=\lambda P\_\langle e,\langle e,t\rangle\rangle .\lambda x\_e.\lambda y\_e.\forall x\hat{}'\ [x\hat{}'\leq x\wedge x\hat{}'\in cov\_i\rightarrow \exists y\hat{}'\ [y\hat{}'\leq y\wedge y\hat{}'\in cov\_i\wedge P(x\hat{}')(y\hat{}')]]\wedge \forall y\hat{}'\ [y\hat{}'\leq y\wedge y\hat{}'\in cov\_i\rightarrow \exists x\hat{}'\ [x\hat{}'\leq x\wedge x\hat{}'\in cov\_i\wedge P(x\hat{}')(y\hat{}')]]$.

The LF and semantics of (30) are given below.

(32) The sides of rectangle 1 cross the sides of rectangle 2.

LF: [[The sides of rectangle 1] [the sides of rectangle 2] [$**_i$[2[1[t_1 cross t_2]]]]]

$[\![S]\!]^g=\forall x'\ [x'\leq \text{the sides of } R1\wedge x'\in cov_i\rightarrow \exists y'\ [y'\leq \text{the sides of } R2\wedge y'\in cov_i\wedge CROSS(x')(y')]]\wedge \forall y'\ [y'\leq \text{the sides of } R2\wedge y\hat{}'\in cov_i\rightarrow \exists x'\ [x'\leq \text{the sides of } R1\wedge x'\in cov_i\wedge CROSS(x')(y')]]$.

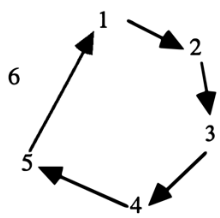
The non-maximal interpretation can be derived with the following ill-fitting cover.

(33) $Cov_i=\{\{d\},\{c\},\{a'\},\{b'\},\{a,b,c',d',t\ldots\}\ldots\}$

The sides of R1 and the sides of R2 are included in Cov_i. d and c of R1 are in independent cells, a′ and b′ of R2 are in independent cells. a and b of R1, c′ and d′ of R2 are in a cell with t, which is neither sides or R1 nor sides of R2. They are thus in the pragmatic junkpile in the scenario, being silently ignored in the utterance.

As non-maximality effect and ill-fitting covers exist in sentences with definite plurals, Beck (2001) suggests that the same holds for reciprocal sentences. It is observed that (34) is true in the scenario described in the picture below.

(34) The pirates stared at each other.



In the scenario, there are six pirates. Five pirates stared at some other pirate and was stared at by some other pirate. One pirate, namely pirate 6, was neither staring at other pirate nor being stared at by other pirates. The scenario verifies (34) according to Beck, and it can be derived with the cumulative operator and the ill-fitting cover. The LF and the semantics of (34) are given below.

(35) The pirates stared at each other.

LF: [Pro2 [the pirates]1 $**_i$[1 [Cov [2 [Cov [t_1 [stared at [the [other x1] (of) t_2]]…]

$[\![S]\!]^g=\forall z'\ [z'\leq \text{the pirates}\wedge z'\in cov_i\rightarrow \exists y'\ [y'\leq \text{the pirates}\wedge y'\in cov_i\wedge STAREAT(z')(\iota x[x\neq z'\wedge x\leq y'])]]\wedge \forall y'\ [y'\leq \text{the pirates}\wedge y'\in cov_i\rightarrow \exists z'\ [z'\leq \text{the pirates}\wedge z'\in cov_i\wedge STAREAT(z')(\iota x[x\neq z'\wedge x\leq y'])]]$.

According to Beck, (34) is an example of WR. The scenario which verifies the sentence is weaker than WR because of non-maximality or the ill-fitting cover. The ill-fitting cover chosen by the context in (34) is as below.

(36) $Cov_i=\{\{p_1\},\{p_2\},\{p_3\},\{p_4\},\{p_5\},\{p_6,q\ldots\}\ldots\}$

In the cover, pirates 1, 2, 3, 4 and 5 occupy individual cells in the cover. Pirate 6 is in a cell with some non-pirates, thus being silently ignored in the scenario. Thus, (34) is true in a scenario where five of the six pirates stared at some other pirate, and were stared at by some other pirate, while one of the pirates neither stared at some other pirate nor was stared at by some other pirate.

In this section we reviewed the notion of non-maximality and Brisson's proposal on formalizing it with ill-fitting covers. Together with Section 2, we show how Beck derives WR. Beck (2001) uses the cumulative operator as the main device in deriving WR, and she has some pragmatic restriction, namely ill-fitting covers, for cases which require further weakening. In the coming section, we will point out some problems with Beck's proposal.

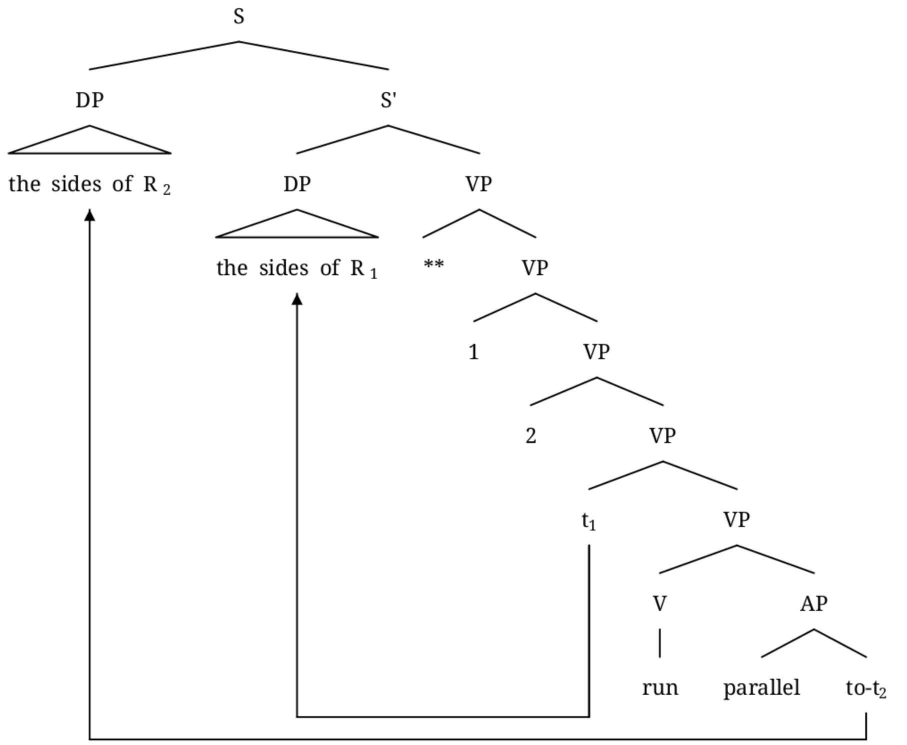
4. Problems with Beck (2001)

To derive the WR interpretation, Beck uses two weakening mechanisms, i.e., weakening by introducing the quantificationally weak cumulative operator, and weakening by pragmatics using ill-fitting covers. Beck uses the former as the main device. In this section, we challenge this choice. We give two pieces of evidence showing the absence of the cumulative operator.

4.1. Syntactic arguments

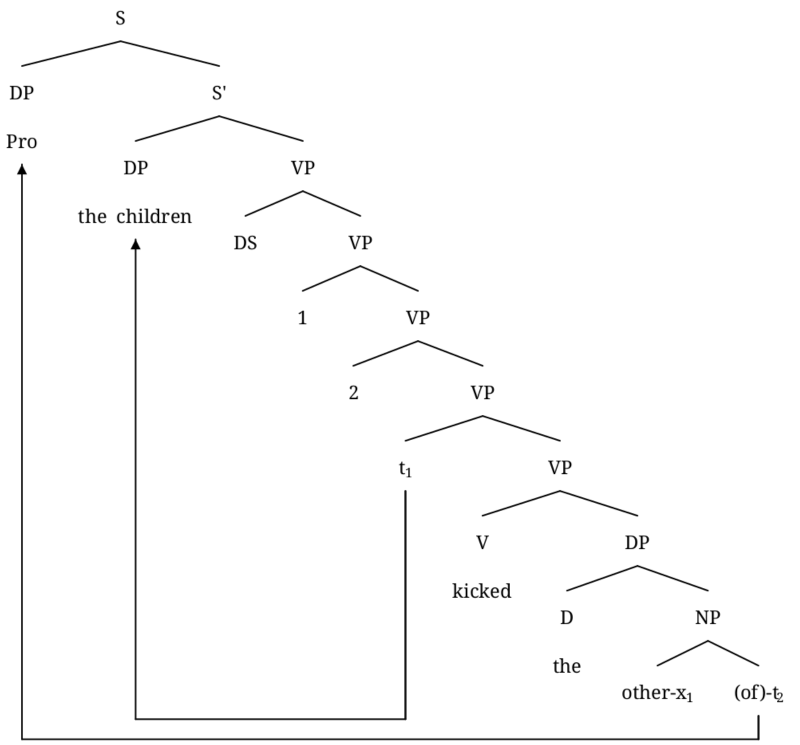
In Sec 2.1, we showed that the cumulative operator requires QR of the cumulated arguments, as shown in (17), repeated below as (37).

(37) The sides of R1 run parallel to the sides of R2.



In Sec 2.2, we showed that Beck assumes that in WR sentences the cumulative operator helps give the WR interpretation. Similar to the cumulative sentence above, the antecedent of the reciprocal and a pronoun which is co-indexed with the antecedent inside each other need to QR, as shown in (22) above, repeated below as (38).

(38) The children kicked each other.



The cumulative operator gives the predicate a cumulative interpretation, and the cumulated arguments need to QR. If we suppose that WR involves the cumulative operator, as proposed by Beck (2001), we predict that WR interpretation is possible only when QR of the antecedent and the pronoun are possible. We show one case where the prediction is contradicted below.

First, we will show that quantifiers cannot take scope out of the possessor of a complex DP. (39) interprets as there is no such a president x such that x's biography discusses x's death. The quantifier no president can QR from the specifier position of the complex DP to scope over the bound variable his.

(39) No president's biography discusses his death.

$\neg\exists x[\text{president}(x) \wedge x'\text{ s biography discusses } x'\text{ s death}]$

When the quantifier is inside the possessor, however, the quantifier can no longer take scope out of the possessor. The intended reading of (40) is that there is no president such that the final chapter of some biography about him discusses his death. The interpretation is not available, showing that the quantifier nobody cannot QR from a complement embedded in the specifier of the complex DP to scope over the bound variable him.

(40) Some biography about no president's final chapter discusses his death.

$\neq\neg\exists x[\text{president}(x) \wedge \exists y[\text{biogrphy}(y) \wedge \text{about}(y)(x) \wedge y'\text{ s final chapter discusses } x'\text{ s death}]$

The intended reading of (40) is that there is no president such that some biography about him's final chapter discusses his death. The interpretation is not available, showing that the quantifier nobody cannot QR from the inside of the specifier of the complex DP to scope over the bound variable him.

The contrast above shows that the specifier position of a complex DP is an island for QR. Quantifiers can QR from the specifier of a DP, but quantifiers cannot QR from the inside of the specifier of a DP.

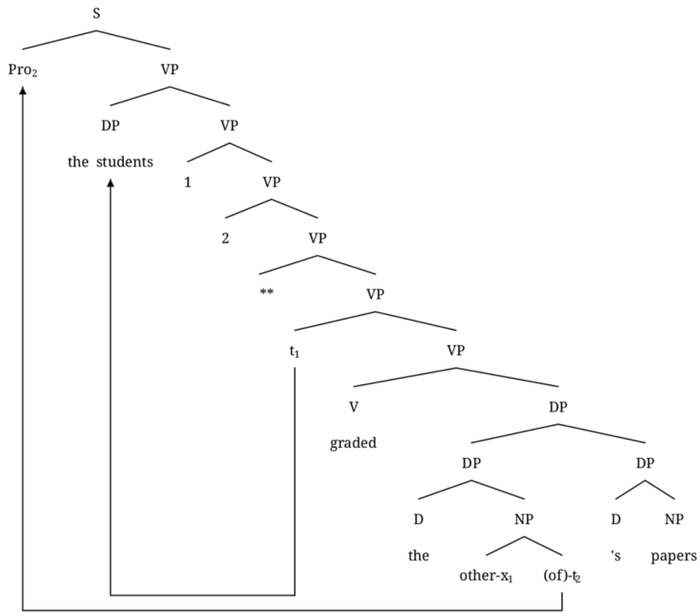
As mentioned above, the cumulative operator requires the cumulated arguments to QR. We show with (40) that the specifier of a DP is an island for QR. Beck assumes that in WR sentences, the antecedent and a pronoun which is co-indexed with the antecedent QR. We thus predict them to show up in places where QR can take place. It is not possible for the pronoun which is co-indexed with the antecedent to be inside the specifier of a DP. The prediction is not borne out, as shown by the example below.

(41) The students graded each other's papers.

(41) is true in a scenario where every student graded the paper of some other student, and every student's paper got graded by some other student. It has a WR interpretation. Following Beck (2001), the LF and syntactic structure of the sentence will be as below.

(42) The students graded each other's papers.

LF: [Pro_2 [the students]$_1$ **[1[2[t_1 graded [the [other x_1 (of) t_2]]'s papers]]]]]



In (42), as shown in the tree, each other is the specifier of a complex DP. The pronoun which is co-indexed with the antecedent in each other is inside the specifier of the complex DP. As mentioned above, the specifier of a complex DP is an island for QR. Thus, the pronoun cannot QR from the position it occupies in the sentence. If the QR cannot take place, we predict that it is no longer possible to have a two-place predicate for the cumulative operator to take, thus the cumulative interpretation and WR should be unavailable. This is not correct, as Ex (42) still has a WR interpretation.

A natural question to ask is whether it is possible for the cumulative operator to take a two-place predicate without quantifier raising the arguments. It is possible, but the semantics cannot compose, as we will show below.

(43) The students graded each other's papers.

LF: [[the students$_2$] [[** [graded]] [[the [other x_1 (of) Pro$_2$]]'s papers]]]

In Ex (43), we have a two-place predicate graded. If we let the cumulative operator to take the predicate as its argument, and give it a cumulative inference, there will be an unbounded variable in each other, that is x1. One may QR the students to get the free variable bound, as in the LF below.

(44) [[the students] [1 [[t_1 ** [graded]] [[the [other x_1 (of) Pro$_2$]]'s papers]]]]

This LF is syntactically grammatical but semantically invalid. Its semantics is given in (45).

(45) $[\![(44)]\!] = \forall x[x \leq \iota z' \text{ s papers}[z \leq \text{the students} \wedge z \neq \text{the students}] \rightarrow \exists y[y \leq \text{the students} \wedge y \text{ graded } x]] \wedge \forall y[y \leq \text{the students} \rightarrow \exists x[x \leq \iota z' \text{ s papers}[z \leq \text{the students} \wedge z \neq \text{the students}] \wedge y \text{ graded } x]]$

According to (45), (44) interprets as for all x such that x is the papers of the unique z who is a part of the students but not identical to the students, there exists y who is a part of the students, such that y graded x; besides, for all y who is a part of the students, there exists x who is the papers of the unique z who is a part of the students but not identical to the students, such that y graded x. When no distributive operators are involved, the meaning of each other is not as intended. There is no such unique z which is a part of the students and not identical to the students. Thus, the interpretation of each other and the LF in (44) are wrong.

As shown in (43) and (44), when no QR is involved, one gets false predictions. This emphasizes the necessity of QR for Beck's proposal. The unavailability of QR and the cumulative reading and the availability of WR in the same context challenge the view that WR involves the cumulative operator.

4.2. Universal quantifiers and cumulativity

The availability of the cumulative interpretation is restricted. An early generalization is that cumulative quantification appears with non-upward monotone quantifiers. (Landman, 2000; Winter, 2001) Recent works try to give a more fine-grained description and explanation. Attempts include Zweig (2008, 2009), Champollion (2020) and Haslinger and Schmitt (2018), etc. We believe a complete description and explanation of the phenomenon requires a separate paper. Here, we only present one empirical generalization and show its implications to WR. It challenges the view that WR sentences involve cumulativity.

It is observed that when certain quantifiers appear in the sentences, they block a cumulative interpretation of plural expressions in lower syntactic positions. All is one such quantifier.

(57) The students read ten books.

(58) # All the students read ten books.

(57) has more than one interpretation, one of which is a cumulative interpretation such that each student read some of the ten books, and each of the ten books was read by some student(s). The reading is available in (57), but not available in (58). (58) only has the interpretation in which each of the student read ten possibly different books.

The contrast does not show up when all is in the object, as shown below.

(59) Ten students read the books.

(60) Ten students read all the books.

If reciprocal sentences are interpreted as WR because of cumulativity, we predict that when all shows up in the subject, the WR reading is not available for reciprocal sentences. The prediction is not borne out, as shown by the following pairs of examples.

(61) The pirates stared at each other.

(62) All the pirates stared at each other.

(61) has a WR interpretation, such that each pirate stared at some other pirates, and each pirate was stared at by some other pirate. If WR comes from cumulativity, we predicted that when a cumulative interpretation is blocked, WR should not be available. In Ex (62), we have all in the subject of the sentence. It will block the cumulative interpretation, as we showed above, but Ex (62) still has a WR interpretation. It is true in the same scenarios where (61) is true.

(63) also has a WR interpretation, such that each student graded some other student's paper, and each student's paper was graded by some other students. The reading is still available in Ex (64).

(63) The students graded each other's papers.

(64) All the students graded each other's papers.

There are other quantifiers which also block the cumulative readings, like every, most, etc. They can also be used with WR examples freely, although we will not go into details here. These examples challenge Beck's (2001) proposal which explains the weak readings of reciprocal sentences with the cumulativity readings observed in plural sentences.

5. Proposal

Beck derives WR with the cumulative operator and ill-fitting covers, as explained in Section 2.2 and Section 3.2. In Section 4, we gave two pieces of evidence against the proposal that the cumulative operator is involved in deriving WR. In this section we will show that ill-fitting covers alone can derive WR without the cumulative operator. The new proposal will be theoretically simpler, and it avoids the above-mentioned problems.

Before delving into the new proposal, a detour into the structure of covers is needed. We will show that the current proposal of covers is inadequate. It has a problem of under-generation. To fix the problem, a new proposal on covers is needed. We thus Skolemized covers, giving covers a more complicated structure and a stronger explanatory power.

5.1. A Skolemized cover

In this section we point out the problem with existing definition of covers and give an updated definition to them. First, we review the existing definition of covers.

The distributive operator contains a universal quantifier in its semantics. Covers restrict the domain of the universal quantifier in the distributive operator, as shown below.

(65) $[\![D_i]\!]^g = \lambda P_{\langle e,t\rangle}.\lambda x_e.\forall y[y \leq x \wedge y \in cov_i \rightarrow P(y)]$

The distributive operator has an index. The index corresponds to a variable which is assigned a value from the context by the assignment function g. The variable is named cover. The distributive operator takes a predicate and an entity as its arguments. The predicate applies to the parts of the entity which are also members of the cover.

Covers are variables over sets of sets. We repeat the definition of cover below.

(66) C is a cover of a set A if and only if:

a. C is a set of subsets of A;

b. Every member of A belongs to some set in C;

c. ∅ is not in C.

The covers of distributive operators are sets of subsets of the universe of discourse. Covers provide partitions of the universe of discourse for the distributive operator to distribute over.

There are cases where covers are quantificationally bound, and this requires an extension of the definition.

(67) Scenario: Teacher Ta and Tb co-taught a class with 50 students, s1 to s50. They met with the students on two different occasions. When Ta was meeting with the students, s2 and s3 were missing. When Tb was meeting with the students, s4 and s5 were missing. m is not a student.

The teachers each met their students.

LF: [[The teachers] [Di [[their students] [Dj [t1 met t2]]]]

∀x(x≤the teachers∧x∈cov_i→∀y(y≤x^' s student∧y∈cov_j→x met y))

In (67), the cover of the lower distributive operator is bound by the teachers. For each teacher, a different cover of the universe of discourse is chosen, so that different students are silently ignored in the context, leading to different non-maximal interpretations of their students. We will show why the current definition of covers does not give us the intended interpretation below.

(68) $[\![cov_i]\!] = \{\{T_a\}, \{T_b\}, \ldots\}$

$[\![cov_j]\!] = \{\{s_1, s_4, s_5, s_6, s_7 \ldots s_{50}\}, \{s_2, s_3, m\ldots\} \ldots\}$

For (67), we suppose covi and covj have the values as above. Ta and Tb each occupy an individual cell in covi, thus the predicate applies to each of Ta and Tb. In covj, all the students except for s2 and s3 are in a cell, s2, s3 and a non-student m are in a cell, thus being silently ignored in the scenario. (67) thus means both Ta and Tb met every student except for s2 and s3. This is not the intended interpretation, where only Ta met every student except for s2 and s3.

Note that other choices of covers will run into similar problems.

(69) $[\![cov_i]\!] = \{\{T_a\}, \{T_b\}, \ldots\}$

$[\![cov_j]\!] = \{\{s_1, s_2, s_3, s_6, s_7 \ldots s_{50}\}, \{s_2, s_3, m\ldots\} \ldots\}$

The covers in (69)give us the interpretation both Ta and Tb met every student except for s4 and s5. This is not the intended interpretation, where only Tb met every student except for s4 and s5.

(70) $[\![cov_i]\!] = \{\{T_a\}, \{T_b\}, \ldots\}$

$[\![cov_j]\!] = \{\{s_1, s_6, s_7 \ldots s_{50}\}, \{ s_2, s_3, s_4, s_5, m\ldots\} \ldots\}$

The covers in (70) give us the interpretation both T_a and T_b met every student except for s_2, s_3, s_4 and s_5. This reading is weaker than the intended reading. Although it is verified in the scenario in (67), this cannot work as a general solution to the problem. Imagine an extreme scenario where for every student, there is a teacher who didn't meet him. In that case, one needs a cover in which all the students are ignored. This does not fit with people's intuition and understanding of the scenario. We conclude that under the current definition of covers, there's no way for them to be quantificationally bound. We thus give the following updates to covers.

We propose that covers can have a more complicated structure than previously proposed. There are two types of covers. The first type of covers are as proposed in Schwarzschild (1996) and later used in Brisson and Beck (2001). The second type of covers are Skolem functions which take individual variables. The values of the Skolem functions and the individual variables are determined by the assignment function. The Skolem function maps the individuals to the covers. The new semantics of covers are as below.

(71) $[\![cov_{i1}]\!]^g = g(i)$ iff $g(i)$ is a cover of the universe of discourse.

$[\![cov_{j2}]\!]^g = g(j)$ only if $g(j)$ is a function from entities to covers of the universe of discourse.

We will show the explanatory power of the new semantics of covers with (67).

(72) $[\![cov_{i1}]\!]^g = g(i) = \{\{T_a\}, \{T_b\}\}$

$[\![cov_{j2}\ pro_1]\!]^{g[1\to T_a]} = g(j)(T_a) = \{\{s_1, s_4, s_5, s_6, s_7 \ldots s_{50}\}, \{s_2, s_3, m\ldots\} \ldots\}$

$[\![cov_{j2}\ pro_1]\!]^{g[1\to T_b]} = g(j)(T_b) = \{\{s_1, s_2, s_3, s_6, s_7 \ldots s_{50}\}, \{s_4, s_5, m\ldots\} \ldots\}$

Covi is a Type 1 cover. It is contextually assigned a value. Each of T_a and T_b are an element of the cover, thus the predicate distributes to T_a and T_b. Cov_j takes a variable which is bound by the teachers. Its value co-varies with each teacher. When the variable equals T_a, then s_2 and s_3 are in a subset with non-students, thus being silently excludable in the context; when the variable equals T_b, then s_4 and s_5 is in a subset with non-students, thus being silently excludable in the context. For each teacher, the non-maximal interpretation of the students is different. This gives us the intended interpretation.

It should be noted that the proposal we give to covers is reminiscent of previous studies on quantifier domain restrictions, as shown below.

(73) Adapted from von Fintel (1994)

Sweden is a funny place. Every tennis player looks like Bjorn Borg, and more men than women watch tennis on TV. But everyone dislikes foreign tennis players.

(74) Adapted from Stanley and Szabo (2000)

In some of my classes, every student failed.

In (73), the expression everyone is most naturally interpreted as every Swedes. The domain of the quantifier every is people in Sweden, although in Sweden is not explicitly stated. In (74), the domain of the lower quantifier every student is quantificationally bound by the higher quantifier some of my classes. The meaning of the sentence is that in some of my classes, every student failed in that class. von Fintel (1994) proposes that quantifiers are indexed. The index corresponds to some Skolem function variables. The functions take individual variables and map the individuals to sets. The sets restrict the domain of the quantifiers. We leave an analysis on the relations between the two similar phenomena to future studies. As the current paper is concerned, it should be noted that empirically, quantifier domain restrictions in (73) or (74), and covers are independently needed. The former applies to nominal arguments, the latter applies to distributive operators. The former is usually a set, the latter contains a set of subsets of the universe of discourse.

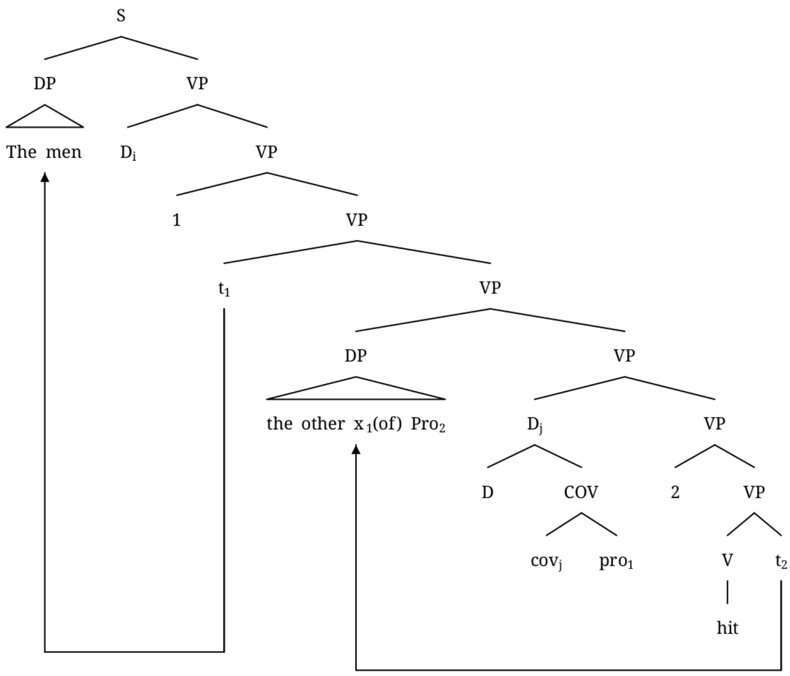
5.2. Deriving WR

In this section, we will show how the updated covers derive WR without the cumulative operator. We will use the distributive operator with covers instead.

(75) The men hit each other.

LF: [[The men] [[D cov_i] [1 [the [other x_1 (of) Pro_4]] [[D [cov_j pro_1]] [2 [t_1 [hit t_2]]]]]

$\forall x(x \leq \text{the men} \wedge x \in g(i) \rightarrow \forall y(y \leq \iota z[z \in g(4) \wedge z \neq g(1)] \wedge y \in g(j)(pro_1) \rightarrow x \text{ hit } y))$



(75) is true in a scenario where there were four men, namely m1, m2, m3 and m4. n is not a man. m1 hit m2, m2 hit m1, m3 hit m4, m4 and m3. Each man hit some other man, and each man was hit by some other man. The following covers derive the intended interpretation.

(76) $[\![cov_{i1}]\!]^g = g(i) = \{\{m_1\},\{m_2\},\{m_3\},\{m_4\}\ldots\}$

(77) $[\![cov_{j2}\ pro_1]\!]^{g[1\text{->}m1]} = g(i)(m_1) = \{\{m_2\}, \{m_1, m_3, m_4, n\ldots\}\ \ldots\}$

$[\![cov_{j2}\ pro_1]\!]^{g[1\text{->}m2]} = g(i)(m_2) = \{\{m_1\}, \{m_2, m_3, m_4, n\ldots\}\ldots\}$

$[\![cov_{j2}\ pro_1]\!]^{g[1\text{->}m3]} = g(i)(m_3) = \{\{m_4\}, \{m_1, m_2, m_3, n\ldots\}\ldots\}$

$[\![cov_{j2}\ pro_1]\!]^{g[1\text{->}m4]} = g(i)(m_4) = \{\{m_3\}, \{m_1, m_2, m_4, n\ldots\}\ldots\}$

Cov_i has a value as in (76). In the cover, each of the men occupies an individual cell. Thus, the predicate applies to each of the men. Cov_j takes a variable bound by the men. When assigned different values, the function maps the individual to different covers. When the variable equals m_1, then m_1, m_3 and m_4 are in a subset with non-men, thus being silently excludable in the context. This gives us m_1 hit m_2. When the variable equals m_2, then m_2, m_3 and m4 are in a subset with non-men, thus being silently excludable in the context. This gives us m_2 hit m_1. When the variable equals m_3, then m_1, m_2 and m_3 are in a subset with non-men, thus being silently excludable in the context. This gives us m_3 hit m_4. When the variable equals m_4, then m_1, m_2 and m_3 are in a subset with non-men, thus being silently excludable in the context. This gives us m_4 hit m_3. The covers above give us the WR interpretation.

The proposal can also derive readings weaker than WR.

(78) The students followed each other.

(78) is true in a scenario where there are three students, s_1, s_2, and s_3. Suppose t is not a student. s_1 follows s_2, and s_2 follows s_3. No other student followed any other student. The interpretation is weaker than WR, as not every student follows some other student, consider s_3, who is not following any other student. Besides, not every student is followed by some other student, consider s_1, who is the last one in the line.

Dalrymple et al. (1998) and Beck (2001) propose that asymmetric relations like follow mean follow or precede in sentence like (78), thus the meaning can be derived in the same way WR sentences are derived. This proposal has some issues. For instance, if we suppose follow means follow or precede in sentences like (78), it is unexplained why the following sentence is less felicitous than (78).

(79) # The two students followed each other.

Under our proposal, examples like (78) can be derived as below.

(80) The students followed each other.

LF: [[The students] [[D cov_i] [1 [t_1 [the [other x_1 (of) Pro_2]] [[D [cov_j pro_1]] [2 [followed t_2]]]]]

$\forall x(x \leq \text{the students} \wedge x \in g(i) \rightarrow \forall y(y \leq \iota z[z \in \text{the students} \wedge z \neq x] \wedge y \in g(j)(pro_1) \rightarrow x \text{ followed } y))$

(81) $[\![cov_{i1}]\!]^{g} = \{\{s_1\},\{s_2\},\{s_3,t\},\ldots\}$

(82) $[\![cov_{j2}\ pro_1]\!]^{g[1\text{->}s1]} = \{\{s_2\},\{s_1, s_3, t\ldots\}\ldots\}$

$[\![cov_{j2}\ pro_1]\!]^{g[1\text{->}s2]} = \{\{s_3\},\{s_1, s_2, t\ldots\}\ldots\}$

$[\![cov_{j2}\ pro_1]\!]^{g[1\text{->}s3]} = \{\{s_1, s_2, s_3, t\ldots\}\ldots\}$

Covi has a value as in (81). In the cover, s_1 and s_2 occupy individual cells. s_3 is in a cell with non-students. Thus, the predicate applies to s_1 and s_2 but not s_3. Cov_j takes a variable bound by the students. When assigned different values, the function maps the individual to different covers. When the variable equals s_1, then s_1 and s_3 are in a subset with non-student, thus being silently excludable in the context. This gives us s_1 followed s_2. When the variable equals s_2, then s_1 and s_2 are in a subset with non-student, thus being silently excludable in the context. This gives us s_2 followed s_3. When the variable equals s_3, then all students are in a subset with non-men, thus s_3 did not follow anyone. The covers above give us the intended interpretation.

5.3. Comparison with Beck (2001)

Beck (2001) uses both the cumulative operator and ill-fitting covers to derive the WR interpretation. We outline a proposal which derives WR using only ill-fitting covers without the cumulative operator. Thus, our proposal is theoretically simpler. In Section 4, we discuss two problems with Beck's proposal. In this section, we show how our proposal avoids those problems.

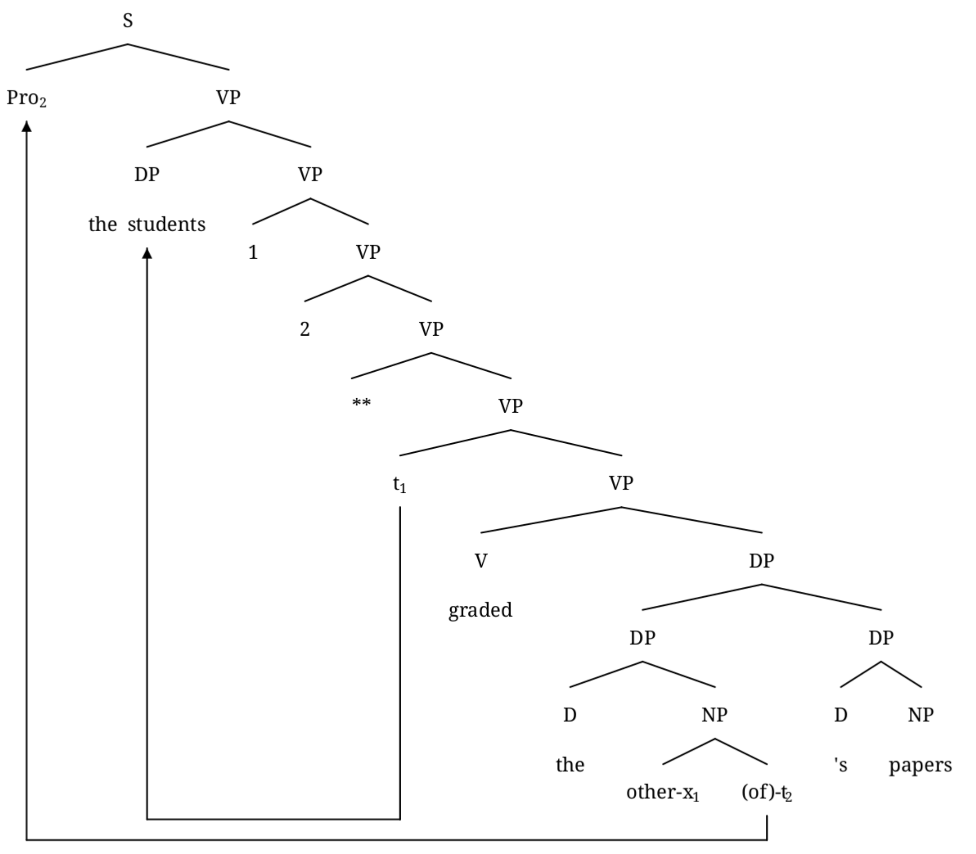
The quantifier raisings required in Beck (2001) and our proposal are different. Beck requires a pronoun inside each other to QR, our proposal requires each other to QR. We use (42) as an example, repeated below as (83).

(83) The students graded each other's papers.

As mentioned in 4.1, Beck's analysis for (83) is as below. We show that it is not possible to QR out from the inside of the specifier of a DP. (84), however, requires this illegal QR. Thus, (84) is not the correct analysis for (83).

(84) Beck (2001)

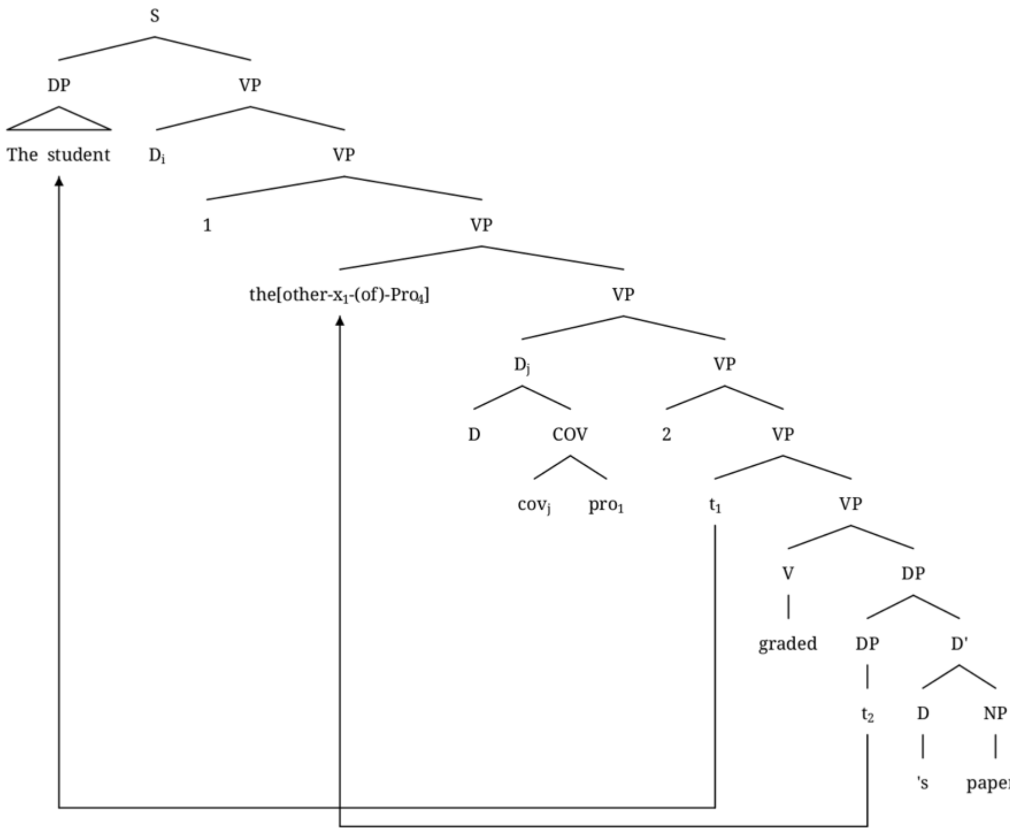
LF: [Pro$_2$ [the students]$_1$ **[1[2[t$_1$ graded [the [other x$_1$ (of) t$_2$]]'s papers]]]]]



Following our proposal, (83) is analyzed as (85). (83) is true in a scenario where for each part of the students which is also in covi, they graded the papers written by the parts of students who are in the cover which covj maps pro1 to.

(85) LF: [[The students] [[D cov$_i$] [1 [t$_1$ [[the [other x$_1$ (of) Pro$_4$]] [[D [cov$_j$ pro$_1$]] [2 [graded [t$_2$'s papers]] …]

$[\![(83)]\!]=\forall x[x\leq \text{the students}\wedge x\in g(i)\rightarrow \forall y[y\leq \iota z[z\leq \text{the students}\wedge z\neq x]\wedge y\in g(j)(pro_1)\rightarrow x$ graded y' s papers]]



In 4.1, we show that although it is not possible for quantifiers to scope out of the specifier of a DP, it is possible if the quantifier is the specifier of the DP. In (85), we have each other being the specifier of a DP. It QR out of the DP. As that is not an island for QR, the movement is possible. Thus, our proposal obeys the island effect of QR.

In 4.2, we show that certain quantifiers like all can block the cumulative reading. If we predict WR involves cumulativity, we predict WR to be unavailable in contexts where the cumulative reading is unavailable. We showed that the prediction was not borne out. Our proposal only makes use of the distributive operators. If a distributive sentence can be used with these quantifiers, WR sentences should allow these quantifiers, too. In contexts where a cumulative reading is unavailable, a distributive reading is always available, as shown below.

(86) All the students wear the uniforms.

√ Each student wears more than one uniform.

From the discussions above, we see that our proposals avoid the problems mentioned in Section 4. It is thus not only theoretically simpler than Beck (2001), but also fits better with the facts.

6. Forcing Maximality

In the previous section, we outline a proposal which derives WR with ill-fitting covers. Thus, we predict the availability of WR to be determined by the availability of ill-fitting covers. We will discuss two cases here. In 6.1, we discuss cases where ill-fitting covers are not preferred. We show that WR readings are also not preferred in these scenarios. In 6.2, we discuss cases will ill-fitting covers are highly infelicitous. We show that WR readings are also infelicitous in these scenarios. We thus prove the close relations between WR and ill-fitting covers.

6.1. Small numbers and statives

The availability of non-maximality effect is influenced by lexical factors. For instance, as observed by Yoon (1996), exceptions are easier to get for episodic predicates than stative predicates. It is reported that while it is very hard to allow exceptions in (89), it is easier to get non-maximal reading for sentences like (90).

(87) Stative predicates (Yoon, 1996)

The children (who are playing in the garden) are eight years old.

(88) Episodic predicates (Yoon, 1996)

The children (who ate pizza here last night) danced in the street.

Now we consider reciprocal sentences in similar contexts. In (91) and (92) different predicates are used. An episodic verb is used in (91), a stative verb is used in (92). Under a WR interpretation, we expect the sentences to be true in a scenario where each child is in relation to some other child, and some individual is in relation to each child. It turns out that while (91) is natural in such a scenario, (92) is less natural.

(89) Scenario: Every child kicked a small proportion of the other child, and each child was kicked by some other child.

The children kicked each other.

(90) Scenario: Every child knew a small proportion of the other child, and each child was known by some other child.

? The children knew each other.

Another observation that is mentioned by many is that when a small number of entities is the reference of the plural definite in the context, it is hard to get the non-maximality effect. For instance, while (94) allows a non-maximal reading, people are very reluctant to allow exceptions under the context in (93).

(91) Bar-Lev (2021)

Context: Someone asks me who among the five adults and the ten kids at the birthday party laughed. I reply:

The kids laughed.

(92) Bar-Lev (2021)

Context: There was a clown at my kid's birthday party. Someone asks me if they gave a funny performance. I reply:

The kids laughed.

In (95) and (96), a same sentence appears under different contexts. For one of the contexts, the children refers to many children, in the other context, the children refers to three children. A WR interpretation predicts both sentences to be true in a scenario where each child kicked some other child, and some child kicked each child. It turns out that while such a reading is natural in (95), it is less natural in (96).

(93) Context: there were many children who were involved in the mass fight. A child did not kick another child in the process.

The children kicked each other.

(94) Context: there are three children, namely John, Bill and Mary who were fighting. Mary did not kick Bill.

? The children kicked each other.

6.2. Forcing maximal interpretations

As observed in Brisson (1998), all can remove the non-maximality effect of the associated nouns. Thus, we predict that when all is used in reciprocal sentences, certain interpretations may become unavailable.

We repeat (78) as (97) below. As mentioned in 5.2, (97) can only be interpreted under ill-fitting covers. Given a finite line of students which fits to world knowledge, there must be a first one, who did not follow anyone, and a last one, who was not followed by anyone.

(95) The students followed each other.

(96) # All the students followed each other.

If we remove the non-maximality of the agent with all, the sentence becomes infelicitous, as shown in (98).

Some consultants judge (98) as natural. We give two arguments showing that it does not contradict our proposal. Following Lasersohn's (1999) theory on non-maximality, operators like all do not strictly prohibit non-maximal readings. Instead, they narrow the interpretation down to be closer to the maximal interpretation. Thus, there's always room for a non-maximal interpretation.

Besides, for people who judge (98) as natural, the sentence becomes unnatural when we add more factors that remove non-maximality, as shown below.

(97) ## All the three students followed each other.

7. Remaining issues

In this section we will discuss some potential problems of this current proposal and outline some potential future directions.

This paper mainly discusses proposals on plural predication by Schwarzschild (1996), Beck and Sauerland (2000) and papers on cumulativity and reciprocals by Beck. We provide empirical evidence showing that reciprocal sentences with WR interpretation differs from characteristic cumulative sentences in their syntactic restrictions and compatibility with quantifiers like all. Many papers haven been written on cumulativity following the studies we discussed here. The nature of cumulativity is still a topic with many discussions. Here, we want to highlight one potential future direction on the relation between reciprocal sentences and cumulativity.

Winter (2000) gives a distributive analysis to examples like below.

(98) Winter (2000)

Scenario: At a shooting range, each soldier was assigned a different set of targets and had to shoot at them. At the end of the shooting we discovered that:

The soldiers hit the targets.

Sentences like (100) have the so-called co-distributive readings. The co-distributive reading is special in that the interpretation of the plural the targets is dependent on the plural the soldiers, so that for each soldier, he hit his targets. Winter (2000) derives the so-called co-distributive reading, which is very similar to the cumulative reading, with the distributive operator. Plurals like the targets are treated as dependent plurals, they contain a covert variable, which can be bound by higher quantifiers. Depending on the value of the variable, the interpretation of the dependent plural varies. Our proposal on Skolemized covers is reminiscent of Winter (2000). Winter (2000) Skolemized the definite plurals, we Skolemized the covers. For both proposals, we let the interpretation of a definite plural to vary with a higher quantifier.

Two questions will follow this similar treatment between WR sentences and co-distributive sentences. A first question is what's the relation between co-distributivity and cumulativity. A second question is what's the relation between WR sentences and co-distributivity. Both questions need further studies, here we only give some tentative thoughts. For the first question, it is stated in the paper that whether the co-distributive reading and the cumulative reading are the same is an open question, although Winter (2000) and a recent paper by Minor (2022) provide evidence supporting the view that they are separate interpretations.

For the second question, we will give a piece of data supporting the view that the co-distributive reading is different from the cumulative reading, while WR sentences behave more similar to the former than the latter. Co-distributive sentences, like WR sentences, can be used with all without a double distributive interpretation.

(99) Scenario: there were a total of three participants. One participant stared at two zebras, another participant stared at five zebras, the last participants stared at three.

All the three participants stared at the ten zebras.

(100) Scenario: there were a total of three participants. One participant stared at one zebra that was exhibited to him, another participant stared at three zebras that were exhibited to him, the last participants stared at two zebras that were exhibited to him.

All the participants stared at the zebras.

(101) Scenario: there were a total of three participants. They each stared at the other participant which was close to him.

All the participants stared at each other.

In (99), we have an example of a cumulative sentence. All is incompatible with the cumulative interpretation. In (100), we have a co-distributive example. With all, the co-distributive interpretation is still available. These example show that reciprocal sentences are more similar to co-distributive sentences than cumulative sentences.

References

1. Bar-Lev, Moshe E. 2021. An Implicature account of Homogeneity and Non-maximality. Linguistics and Philosophy 44(5). 1045–1097.

2. Beck, Sigrid. 2001. Reciprocals are definites. 35.

3. Beck, Sigrid & Uli Sauerland. 2000. Cumulation is needed. Natural Language Semantics 8(4). 349–371.

4. Brisson, C. 2003. Plurals, all and the nonuniformity of collective predication. 56.

5. Champollion, Lucas. 2020. Distributivity, Collectivity, and Cumulativity. en. In Daniel Gutzmann, Lisa Matthewson, C ´ecile Meier, Hotze Rullmann & Thomas Zimmermann (eds.), The Wiley Blackwell Companion to Semantics, 1st edn., 1–38. Wiley.

6. Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam Mchombo & Stanley Peters. 1998. Reciprocal Expressions and the Concept of Reciprocity.

7. Križ, Manuel. 2016. Homogeneity, Non-Maximality, and all. en. Journal of Semantics 33(3). 493–539.

8. Landman, Fred. 2000. Events and Plurality. en. Gennaro Chierchia, Pauline Jacobson & Francis J. Pelletier(eds.). Vol. 76 (Studies in Linguistics and Philosophy). Dordrecht: Springer Netherlands.

9. Langendoen, D. Terence. 1978. The Logic of Reciprocity. Linguistic Inquiry 9(2). 177–197. http://www.jstor.org/stable/4178051.

10. Schmitt, Viola. 2018. Pluralities in intensional contexts: Cumulativity and individuation. 177.

11. Schwarzschild, Roger. 1996. Pluralities.

12. Sternefeld, Wolfgang. 1998. Reciprocity and cumulative predication. 36.

13. Winter, Yoad. 2000. Distributivity and Dependency.

14. Zweig, Eytan. 2009. Number-neutral bare plurals and the multiplicity implicature. en. Linguistics and Philosophy 32(4). 353–407.