

Anaphora and Distributivity

A study of *same*, *different*, reciprocals and *others*

Published by

LOT

Janskerkhof 13

3512 JK Utrecht

the Netherlands

phone: +31 30 253 6006

fax: +31 30 253 6406

e-mail: lot@uu.nl

<http://www.lotschool.nl/>

Cover illustration: Lodge. Photo by Jakub Dotlačil

ISBN: 978-94-6093-018-8

NUR: 616

Copyright © 2010 Jakub Dotlačil. All rights reserved.

Anaphora and Distributivity

A study of *same*, *different*, reciprocals and *others*

Anaforen en distributiviteit

Een studie naar *hetzelfde*, *ander/verschillend*,
reciproken en *de anderen*

(met een samenvatting in het Nederlands)

Proefschrift

Ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. J.C. Stoof,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
vrijdag 12 maart 2010 des ochtends te 10.30 uur

door

Jakub Dotlačil

geboren 14 May 1979
te Teplice v Čechách, Tsjechoslowakije

Promotor: prof.dr E. Reuland

Contents

Acknowledgments	ix
1 Introduction	1
1.1 Introduction	1
1.2 The plan of the thesis	3
2 Properties of pluralities and E(n)Is	5
2.1 Introduction	5
2.2 Experiments on sentences with plural arguments	7
2.2.1 Introduction	7
2.2.2 Gil (1982)	9
2.2.3 Brooks and Braine (1996) and Kaup et al. (2002)	12
2.2.4 Frazier et al. (1999)	15
2.2.5 Summary	16
2.3 Questionnaire testing readings of E(n)Is	17
2.3.1 Method	18
2.3.2 Results	22
2.3.3 Discussion	28
2.4 Summary	33
3 The language of pluralities	35
3.1 Introduction	35
3.2 Syntax and semantics of pluralities	36
3.2.1 Language model	36
3.2.2 The domain of singular and plural individuals and groups . . .	38
3.2.3 Event semantics	44
3.2.4 Compositional interpretation of LF	46
3.2.5 Modification of Landman (2000): Adding covers	60
3.2.6 Distributive quantifiers in the language of events and pluralities	75

3.3	Explaining the differences in distributive readings	81
3.3.1	Introduction	81
3.3.2	Bi-OT account	85
3.3.3	Game-theoretical analysis	91
3.4	Alternative explanations	108
3.4.1	Explanation in terms of scope preferences	109
3.5	Conclusion	115
4	<i>The others and reciprocity</i>	117
4.1	Introduction	117
4.2	Two research strategies on reciprocity	121
4.2.1	The R-approach to reciprocals	121
4.2.2	The DA-approach to reciprocals	122
4.2.3	Properties and predictions of the two strategies	125
4.3	Distributivity and the range of readings	129
4.3.1	Distributivity	129
4.3.2	The range of readings	136
4.3.3	Questionnaire testing the range of readings and distributivity	146
4.4	Novel R-approach to reciprocals	150
4.4.1	Dependent subevents: Argument for a third type of distributivity	152
4.4.2	Reciprocals in dependent event types	164
4.4.3	Previous R-approaches to reciprocals	176
4.4.4	Conclusion	187
4.5	Antecedents of <i>each other</i> and <i>the others</i>	187
4.5.1	Conclusion	189
4.6	A null theory of long-distance reciprocity	190
4.6.1	Introduction	190
4.6.2	Semantics of Attitude Verbs (Cresswell and von Stechow, 1982)	193
4.6.3	Cumulation in <i>de re</i> Attitudes – Dependent Readings	195
4.6.4	Cumulation in <i>de re</i> Attitudes – LDR Readings	200
4.6.5	Consequences of the analysis	203
4.6.6	Conclusion	206
4.6.7	Previous accounts of LDR	207
4.7	Conclusion	211
4.8	Appendix – cross-linguistic data	213
5	<i>Different and same</i>	217
5.1	Introduction	217
5.2	<i>Different and same</i>	218
5.2.1	Introduction	218
5.2.2	<i>Different</i>	218
5.2.3	<i>Same</i>	222
5.2.4	Conclusion	223
5.3	Two strategies	223
5.3.1	Introduction	223

5.3.2	R-strategy	223
5.3.3	A-strategy	232
5.3.4	Consequences of the analysis	233
5.4	Compositional DRT with pluralities (PCDRT)	234
5.4.1	Introduction	234
5.4.2	Compositional DRT with pluralities	235
5.5	<i>Same</i> and <i>different</i> ₁ in PCDRT	240
5.5.1	Discourse-anaphoric reading in PCDRT	240
5.5.2	The sentence-internal reading of <i>same</i> and <i>different</i> ₁ in PCDRT	241
5.5.3	Open problems	249
5.6	Conclusion	251
6	Summary and conclusion	253
	Bibliography	257
	Bibliography	257
	Appendices	269
	Appendices	269
	Inequalities in Games discussed in Chapter 3	269
	Questionnaires	274
	Samenvatting in het Nederlands	293
	Curriculum Vitae	297

Acknowledgments

There were many moments during the writing of my thesis that I was saying to myself: ‘Oh, I am really looking forward to having this slow and difficult process behind me, when I finally will effortlessly and happily be typing acknowledgments.’ Well, now the moment is here and I might reconsider. For one thing, writing the thesis turned out to be much more fun than I expected. For another, writing acknowledgments might be a much more serious job than I would have thought. I never saw any statistics on this, but it seems to me that Acknowledgments is the most read part of the thesis. For my part, I definitely read more Acknowledgments in theses than theses themselves. Forget the covers and summaries! It is Acknowledgments that get attention. I will try not to get it wrong.

First of all, I would like to thank Eric Reuland, my supervisor. When I started my PhD in Utrecht, I had a vague idea what I wanted to do and not so vague ideas what I did not want to do. I was tired of doing just syntax and wanted to move more towards interfaces, and especially semantics. For someone like me, Eric’s supervision fit very well. He always let me explore things but was also always ready to ask me how all these things that I found fascinating at the moment would connect to my to-be thesis. Eric is a rare combination of someone who lets you do whatever crazy things you would like to do but is ready to remind you that your final goal is not to just have fun for a couple of years but to have fun *and* write a thesis. He is also the kind of person who is never tired. Especially in the last part of my writing, it happened often that I sent him a part of chapter late in the evening, and Eric met me the next day at noon with the part already read and a list of comments. I am very thankful to him.

When I started my PhD, Eric was not my only supervisor. My other supervisor was Tanya Reinhart. To be honest, I was a bit scared before my first meeting with her but it turned out to be a very good meeting. And so did the meetings after that one. I worked under Tanya’s supervision on restrictions on inverse scope readings. In the end, I moved away from that topic because the analysis we were pursuing was running into problems that seemed quite daunting. I did not know how to solve them then and I would not know now. However, by studying this topic I learned a lot about scope and

semantics. When Tanya died... it is strange. I did not realize until then how close she was to me. I miss her a lot. I hope she would not be disappointed with the way my thesis turned out.

During my studies, I worked closely with Øystein Nilsen. This should be obvious to anyone who makes it beyond the Acknowledgments since two papers of ours are also embedded in one chapter of my thesis. But our joint work went way beyond these two papers. Øystein always had time to discuss linguistic issues with me, to show me that I am wrong and why I am wrong and, on my lucky days, that I am right and why. Øystein is always extremely careful with linguistic argumentations and well-aware of every development in the field, from mainstream issues to completely obscure ones. I learned a lot from him.

In my third year, ZAS Berlin invited me to be their visiting PhD student for half a year. I am very happy I got this chance. It was during this time that most parts of my thesis took shape, in my head, at least. I am extremely thankful to Manfred Krifka for helping me during my stay in Berlin. I do not know how he does that but despite his thousands of job obligations he somehow always found time to meet me whenever I asked him to. It was always a lot of fun to discuss my research with him. Manfred could always see immediately through the analysis and bring forth many relevant problems and straight away suggest possible solutions to those problems and problems of these solutions and after that mention hundreds of other phenomena that I never realized connect to my study and I should look at. Apart from Manfred, I was very happy I got a chance to discuss linguistics with Uli Sauerland (before he left for Stanford), and Malte Zimmermann, who also arranged a talk for me at his department. Thanks! I liked being a part of the ZAS research institute for a while. It is a very vibrant and active place and it is nice that there is a lot of concern for connecting people's research. I am only sorry that my German is somewhere between Dutch and non-existent, so I could not participate in other activities of the institute.

This is also a nice opportunity, I think, to thank my teachers in Czechia and Norway who steered me towards formal linguistics. These include (in Czechia) Jarmila Panevová, Vladimír Petkevič and Oldřich Uličný. In Norway, I learned a lot about syntax and semantics from Øystein Nilsen, Gillian Ramchand, Peter Svenonius and Knut Tarald Taraldsen. But in fact, the whole linguistic department there was just amazing!

There were many other people who helped me with my research. I would especially like to mention Rick Nouwen, who was always ready to discuss things with me and could always tell me straight away that he does not understand what I am saying and I should be more precise. Somehow, he managed to say it in a way that I never felt stupid. Other people discussing various aspects of my research and helping me understand the topic better were Chris Barker, Janneke ter Beek, Adrian Brasoveanu, Lucas Champollion, Anna Chernilovskaya, Gennaro Chierchia, Alexis Dimitriadis, Gaetano Fiorin, Berit Gehrke, Nino Grillo, Emar Maier, Marijana Marelj, Reinhard Muskens, Marieke Schouwstra, Giorgos Spathas, Roger Schwarzschild, Káča Součková, Radek Šimík, Anna Szabolcsi, Jakub Szymanik, Henriëtte de Swart, Yoad Winter, Arjen Zondervan. Apart from meeting people, I also learned a lot from the reading groups on

semantics in Utrecht and the semantics-pragmatics research group at ZAS Berlin, as well as audiences at conferences where I presented my research.

I would also like to say thank you to all my Utrecht office mates, which were (chronologically): Joost Zwarts, Kriszta Szendrői, Nicole Grégoire, Rick Nouwen, Nino Grillo, Ivana Brasileiro, Mirjam Rigterink, Roberta Tedeschi, Arnout Koornneef, Natalie Boll-Avetisyan, Natalia Slioussar, Joke de Lange, Marieke Schouwstra, Marijana Marelj, Anna Chernilovskaya, Jose Barcena, Dagmar Schadler, Merijn de Dreu and Cem Keskin. My UiL-OTS and LOT school friends and colleagues: Frans Adriaans, Ana Aguilar, Diana Apoussidou, Boban Arsenijević, Natalie Boll-Avetisyan, Alexis Dimitriadis, Xiali Dong, Gaetano Fiorin, Berit Gehrke, Nino Grillo, Gianluca Giorgolo, Nadya Goldberg, Andrea Gualmini, Sander van der Harst, Jutta Hartmann, Veronika Hegedüs, Naomi Kamoen, Cem Keskin, Marlies Kluck, Marjo van Koppen, Huib Kranendonk, Bert Le Bruyn, Tom Lentz, Shakuntala Mahanta, Luisa Meroni, Krzysztof Migdalski, Nataša Miličević, Iris Mulders, Øystein Nilsen, Rick Nouwen, Andreas Pankau, Min Que, Gyuri Rakosi, Mirjam Rigterink, Martin Salzmann, Erik Schoorlemmer, Marieke Schouwstra, Anca Sevcenco, Natalia Slioussar, Giorgos Spathas, Káča Součková, Radek Šimík, Roberta Tedeschi, Christina Unger, Sharon Unsworth, Luis Vicente, Anna Volkova, Hedde Zeijlstra, Arjen Zondervan.

I would like to thank my parents, for putting up with the fact that I moved first to Norway and then to the Netherlands, even though sometimes based on the frequency of my visits it might have looked like I moved to Papua New Guinea. Slibuju, že se polepším a budu doma častěji! And I would like to say thank you to my siblings, my big economist brother Marek and my future big economist sister Petra. It is nice to see that not everyone in our family is stupid, some people actually know what career to pursue.

Finally, I would like to thank Janneke. For helping me with my thesis. For always having time for me. For not minding watching arty and horror and trashy and whatever other movies I came across. For reading books aloud with me in the evening. For making up stories about mrňous. And for much more.

CHAPTER 1

Introduction

1.1 Introduction

This thesis has two goals. The first one is to study the interpretation of sentences with plural arguments. Consider the following example:

- (1) Three boys baked five cakes.

(1) can be interpreted in various ways. One interpretation says that three boys worked as a group and they together baked five cakes. This could be true if, for example, they were a team working in a restaurant. One was preparing dough, the other one made icing and the last one baked the cakes. As a team, they managed to bake five cakes. Another way one can understand (1) is that three boys split the chore of baking cakes: one baked one cake and the other two baked two cakes each. Following the semantic literature, I will call the first reading *a collective reading* and the second reading *a cumulative reading*. Finally, one can interpret (1) as saying that each of three boys baked five cakes so in total fifteen cakes were baked by the boys. I will call this reading *a distributive reading with covariation* since the referent expressed by the object, five cakes, varies with respect to the boys.

It is commonly acknowledged in the semantic literature that all these readings are possible. The situation changes if we consider the following sentence:

- (2) Each of the three boys baked five cakes.

(2) lacks the first two readings that (1) has. The lack of particular readings has led to a better understanding of how we should analyze quantifiers like *each*, as well as a better understanding of collective, cumulative and distributive interpretations. In this thesis, I want to focus on the opposite issue: the status of the various readings that

sentences like (1) should have. Contrary to common assumptions in semantics I am going to argue that (1) does not have all the readings discussed above or at least the readings are not acceptable to the same degree. Various experiments and questionnaire studies point to this conclusion. In particular, the distributive reading with covariation of sentences like (1) is less acceptable than the collective reading (and presumably, it is also less acceptable than the cumulative reading even though this has been tested to a considerably smaller degree). I believe that accounting for the degraded status of the distributive reading with covariation in cases like (1) and other cases reveals some facts about semantics and pragmatics of numeral indefinites and other expressions that disprefer this reading. It also tells us more about the status of the particular reading as compared to other readings which might be preferred in (1). The consequences of these issues have not been explored so far. This thesis is intended to fill this gap.

The second goal of this thesis is the study of anaphoricity. Standardly, most studies of anaphoricity focus on pronouns and reflexives. I am going to focus on the expressions *same*, *different*, *others* and *each other*, which I will call *expressions of (non-)identity*. What these expressions have in common is that they can be anaphoric to a referent in a clause if the clause includes a semantically plural argument. Consider (3a). This sentence can mean that each boy read the same book as the other boys did. Thus, *same* is anaphoric to the books read by the other boys. Obviously, this reading is possible only thanks to the plural argument *the boys* in the clause. Similarly, *different* can be anaphoric to the books that the other boys read in (3b). *Others* and *each other* are anaphoric to the plurality of the boys in (3c) and (3d).

- (3) a. The boys read the same book.
- b. Each boy read a different book.
- c. Each boy talked to the others.
- d. The boys talked to each other.

None of these expressions could be anaphoric to a referent in a clause if the clause does not include a semantically plural argument. Consider (4a) to (4d) which lack a semantically plural argument and consequently, the sentence is either ungrammatical, as in (4d), or we can only interpret the clause if we assume there is some preceding discourse which introduces an element to which the expressions of non-identity are anaphoric ((4a)-(4c)).

- (4) a. Morris read the same book.
- b. Morris read a different book.
- c. Morris talked to the others.
- d. * Morris talked to each other.

The fact that the expressions of (non-)identity require a plural argument connects the study of anaphoricity to the study of the interpretation of sentences with plural arguments. In particular, there is a connection between the expressions of (non-)identity and the ability of plural arguments to license distributive readings with covariation, or collective and cumulative readings. As we will see some expressions of (non-)identity

require the plural argument that licenses collective and cumulative readings (but might not license distributive readings with covariation). Other expressions of (non-)identity require the plural argument that licenses distributive readings with covariation (but does not license collective and cumulative readings). Finally, still other expressions are insensitive to the type of plural argument.

The question is, why do we observe these differences? I am going to argue that they tell us something about the semantics and pragmatics of sentences with plural arguments. In particular, they improve our understanding why the distributive reading with covariation in (1) (and similar examples) is degraded and why this is not true, for instance, in (2). The differences also tell us something about the way anaphoricity works in natural language. In particular, I am going to argue that we can see on the behavior of expressions of (non-)identity that anaphoricity is achieved in two ways. In the first one, anaphoric expressions are quantificational elements that create scope dependencies. This is identical to the way quantifiers are interpreted. In the second one, they require an external mechanism which triggers their anaphoricity. The mechanism can be predicate abstraction and binding by lambda-operators in the approach where anaphors are interpreted as variables, as, for instance, in Heim and Kratzer (1998). Crucially, I am going to argue that if we want to account for the behavior of expressions of (non-)identity we should not assume anything else. For example, we should not postulate anaphors whose variables have to be bound obligatorily, as is often assumed (incorrectly, I believe) for the interpretation of reflexives.

The rest of this chapter shortly summarizes each chapter in the thesis.

1.2 The plan of the thesis

In Chapter 2 I present questionnaires and experiments which show that the distributive reading with covariation is dispreferred with various plural arguments. I show that based on the status of the acceptability of this reading, plural arguments fall into at least three groups. In the second part of Chapter 2 I present a questionnaire that I prepared with Øystein Nilsen, which focuses on expressions of (non-)identity. It shows that some of these can be anaphoric to a referent in a clause if the plural argument present in the clause is of the type that licenses the distributive reading with covariation. Other expressions of (non-)identity require the plural argument to be able to license collective readings or they are insensitive to the type of the plural argument.

Chapter 3 introduces the language of pluralities which is built on Landman (2000). Furthermore, I account for the first fact discussed in Chapter 2: why the distributive reading with covariation is dispreferred with some plural arguments. The analysis is based on the assumption that when interpreting a sentence with plural arguments one also considers alternative ways of saying the same, with different expressions expressing plurality. I consider two formalizations of this analysis: one within the bi-directional Optimality Theory, the other one within a Game Theoretic analysis. I show in detail how, using the approach of Parikh (2000) we explain why the distributive reading with covariation is dispreferred with various plural expressions.

Chapter 4 focuses on two expressions of non-identity: *each other* and the noun *others*. I discern two types of analyses that have been proposed for *each other* and show that only one of them should be used for *each other* while the other one should be kept for *the others*. Afterwards, I propose a novel analysis for *each other*, which however follows the tradition of treating it as a polyadic quantifier, in line of Dalrymple et al. (1998), Sabato and Winter (2005b). My account extends these because it can also assign the correct interpretation to reciprocal sentences with collective predicates, negation and more than two arguments. Furthermore, as I show, the analysis can also deal with long-distance reciprocity, studied in Heim et al. (1991a) and Dimitriadis (2000), among others, and the variety of readings that reciprocal sentences can give rise to.

The division of the analyses into the ones that can account for *each other* and the ones that can account for *the others* follows two different viewpoints on anaphoricity. In one case, an expression is a quantifier. In the other case, an expression includes variables that need to find their referent by some external binding mechanism. I argue that making use of the two strategies leaves no space for an account of anaphoricity in which anaphoric expressions have variables but the variables have to be obligatorily bound in the syntax. Furthermore, I show that *the others*, but not *each other* can be anaphoric to a particular referent in a clause if the plural argument present in the clause can license the distributive reading with covariation. This difference between *the others* and *each other* follows from the account proposed in Chapter 3.

In Chapter 5 I study one expression of identity, the adjective *same*, and one expression of non-identity, the adjective *different*. It will be shown that in languages other than English two expressions are used for *different*, called *different₁* and *different₂* here. I argue that each expression makes use of one strategy of anaphoricity: *different₁* uses the strategy that *the others* makes use of, while *different₂* uses the strategy that *each other* makes use of. Thus, *different₂* is a quantifier, while *different₁* is an adjective with free variables which find their referent in a discourse. One distinction between the two expressions is their behavior in a clause with a plural argument. *Different₁* can be anaphoric to a particular referent in a clause if the plural argument present can license the distributive reading with covariation. On the other hand, *different₂* can be anaphoric to a particular referent in a clause if the plural argument present can license the collective reading. This difference follows from the account proposed in Chapter 3 when combined with the fact that *different₁* and *different₂* select two different strategies for resolving anaphoricity. Finally, *same* differs from both *different₁* and *different₂* since it is insensitive to the type of plural argument. This behavior of *same* also follows from the account of why the distributive reading with covariation is degraded in some cases, as proposed in Chapter 3.

Chapter 6 shortly concludes the thesis.

Properties of pluralities and expressions of (non-)identity

2.1 Introduction

This thesis focuses on the semantics of the lexical items *the others*, *each other*, *same*, *different* and its cognates in languages other than English. Throughout, I will call these expressions *expressions of (non-)identity* (E(n)Is for short). What E(n)Is have in common is their ability to license an additional reading when they appear in a clause with a plural argument. Take, for example, the adjective *same*. *Same* can express that one object is identical to another object mentioned in the previous discourse. This use is shown in (1a) where *the same boy* means ‘the boy identical to the boy introduced previously, which is the boy that I watched’. In this use *same* seems to play a very similar role to pronouns, as one can see when comparing (1a) to (1b).

- (1) a. I watched a boy play flute. Angelica watched the same boy.
- b. I watched a boy play flute. Angelica watched him, too.

Same can play another role once it is in the scope of a plural argument. This use is shown in (2a) where *the same girl* does not mean ‘the girl identical to the girl introduced previously in the discourse’. Rather, the sentence means ‘I dreamed about the same girl that Morris dreamed about’. Following Carlson (1987), I call this reading *sentence-internal*. Notice that (2b) sounds odd which excludes the possibility of treating (2a) as a case identical to the discourse use of *same* where discourse salience of some girl was first accommodated. If the accommodation was possible in (2a) there is no reason it should not be possible in (2b).

- (2) a. Morris and I dreamed about the same girl.
- b. # Morris dreamed about the same girl.

Other expressions apart from *same* that can give rise to the sentence-internal reading in scope of plural arguments are the noun *others* and the adjective *different*. The sentence-internal reading of *the others* is shown in (3a) which can mean ‘each boy hated the boys other than himself’. In this case again, the sentence-internal reading is only possible if *the others* is in scope of a plural argument. Finally, the sentence-internal reading of *different* is shown in (3b) which can mean ‘Morris dreamed about a different girl than I did.’

- (3) a. Each boy hated the others.
- b. Morris and I each dreamed about a different girl.

As we will see *different* is commonly realized as more than one lexical item in languages other than English. For example, it appears as *andere* and *verschieden* in German (Beck, 2000), *ander* and *verschillend* in Dutch, *différent* and *autre* in French (Tovena and van Peteghem, 2002), *jiny* and *různý* in Czech.

Finally, to the group of expressions of non-identity, so far including *same*, *other*, *different* and its counterparts in other languages, I also add reciprocals given their ability to license the sentence-internal reading in a clause with a plural argument. Of course, in the case of reciprocals, this is the only interpretation they can give rise to so they *have* to be in the scope of plural arguments to be interpretable (cf. (4a) and (4b)).

- (4) a. Philip and Hilary like each other.
- b. * Philip likes each other.

However, requiring a plurality to be interpretable is not an exclusive property of reciprocals. It is shared by some of the expressions of *different* in languages other than English, as we will see later.

Even though E(n)Is comprise only a handful of lexical items they played quite an important role in formal semantics, as is obvious from the high number of analyses presented in literature. I will discuss the previous literature in subsequent chapters. The main focus of the thesis is an issue which has usually been put aside. The sentence-internal reading of an E(n)I is possible only when the E(n)I is in the scope of a plural argument. Therefore, our analysis of these lexical items should connect them to our understanding of pluralities. Take, as an example, a simple transitive sentence with plural arguments like (5).

- (5) Three boys invited four girls.

It is commonly assumed that sentences like (5) come with a variety of readings: distributive, cumulative, branching, collective and maybe others. What is less acknowledged in the semantic literature is that these readings do not have the same status according to speakers’ intuitions. For (5), the collective and branching readings are preferred over the distributive reading. As I am going to show this difference is directly connected to the type of plural arguments. Indefinite numerals are only marginally accepted with a distributive reading, unlike some other plural DPs. Clearly, this fact should be either encoded in our analysis of pluralities or should somehow follow from

independent principles of language. Either way, it might have and does have consequences for the analysis of E(n)Is. If we see that the sentence-internal reading of an E(n)I is degraded with plural arguments with which the distributive reading is degraded, and accepted with plural arguments with which the distributive reading is accepted we should make sure that the sentence-internal reading of the E(n)I and distributivity are tightly connected. On the other hand, if the sentence-internal reading of another E(n)I is not degraded in contexts where distributive readings are degraded, the two readings should be disassociated. Furthermore, studying E(n)Is also helps us understand pluralities better. For example, if the sentence-internal reading of E(n)Is is degraded in contexts where distributive readings are degraded this suggests that both degradation have the same cause and thus, for instance, can help us understand why one reading in (5) and other examples is less preferred than others.

This chapter introduces the data that will be analyzed throughout the rest of the thesis. Two kinds of data are discussed. We start with simple transitive sentences like (5). I distinguish two types of readings, distributive and non-distributive, which are commonly recognized and further subdivided in the literature on the semantics of pluralities. In Section 2.2 I discuss psycholinguistic evidence which shows that the acceptability of both distributive and non-distributive readings depends on the type of determiner phrase (DP). This will form the background for the discussion of the next kind of data, discussed in Section 2.3: transitive sentences in which one argument is an E(n)I, i.e., sentences like (2a). The crucial issue is going to be whether the acceptability of the sentence-internal reading of E(n)Is depends on the type of plural DP and whether the difference in the acceptability matches or mismatches the acceptability of distributive readings. These points are summarized in Section 2.4. The analysis of these data will be the main topic of the next chapters.

I do not expect all readers to be interested in every detail of the experiments. Some probably want to move to the analysis as quickly as possible. If you are such a reader I suggest you only read the introduction of the next section (Section 2.2.1), and move straight into the summary of the results, Section 2.3.3.

2.2 Experiments on sentences with plural arguments

2.2.1 Introduction

Transitive sentences with plural arguments like (5) are underspecified or many-way ambiguous, depending on one's viewpoint on what counts as an independent reading. For the purposes of this chapter, two reading types are going to be relevant. In the first one, one argument scopes and distributes over the other argument. Consider (5). Two cases of this reading can be distinguished. First, the subject can distribute over the object, in which case each boy invites four girls and the girls are different for each boy, so up to twelve girls were invited. This reading could be paraphrased as 'three boys each invited four girls'. Second, the object can distribute over the subject. This reading could be paraphrased as 'each of the four girls was invited by three boys', and in this case up to twelve boys in total invited four girls. I am going to call this reading

a *distributive reading with covariation*, or, shortly, a *distributive reading*, in case no confusion is likely to arise.

There is another reading of sentences with plural arguments in which neither of the arguments distributes over any other argument(s). This reading is often taken to comprise various sub-readings, called a *branching reading*, a *cumulative reading*, a *collective reading* etc. Since at this point further distinctions of this reading are irrelevant, I am going to label it a *non-distributive reading*. Under the non-distributive reading (5) says that ‘there are three boys, there are four girls, and the three boys invited the four girls’.

It is commonly assumed that any plural argument, with the exception of DPs with distributive quantifiers like *every* and possibly other quantifiers like *most*, can give rise to both distributive reading with covariation and non-distributive readings. This is at least the impression one gets from reading, for example, Kamp and Reyle (1993); Lasersohn (1995); Landman (1995, 2000); Winter (2001a), among many others. But this viewpoint is by no means shared without exception. Schwarzschild (1992) claims that sentences like (6) do not allow the distributive reading with covariation (John and Mary each own a different car), or allow it only very marginally. In Schwarzschild (1996), he modifies this stance and says that the distributive reading is always possible, but requires the right context.

(6) John and Mary own a car.

Beghelli and Stowell (1997) claim that subjects headed by *each* or *every*, and definite or indefinite plurals can all give rise to the distributive reading even though the first two subjects ‘seem to favor a distributive construal over the indefinite object somewhat more strongly than the other subjects [...] do but this does not appear to be an absolute requirement’. Scha (1981) claims that definite plurals cannot give rise to the distributive reading. For Link (1991), definite plurals can give rise to the distributive reading but this reading is dispreferred (the preferred reading is the collective reading, which is a subtype of the non-distributive reading). Williams (1991) marks the distributive reading with definite plurals as unacceptable, in which he is followed by Moltmann (1992) even though Moltmann adds that ‘there are counterexamples to this view’, and specifically mentions a handful of papers that disagree with this judgement. A somewhat contradicting picture is given in Roberts (1990) who says that the distributive reading is hard to get with definite plurals but it is possible in the right context, which makes definite plurals different for example, from numeral indefinites (like *three men*) since the latter can easily get the distributive reading. Later on, however, she claims that the distributive reading is hard to get with numeral indefinites as well.

The point of mentioning all this literature is to show that contrary to what one might think there is in fact very little agreement on the acceptability of the distributive reading. We have seen that while some authors claim that the distributive reading is fully possible with any DP, others believe that the distributive reading has a special status with some DPs. The question is what the status is. The answer to that ranges from marginal to ungrammatical. The second question is which DPs are only marginally acceptable or even ungrammatical with the distributive reading. The answer to that ranges from DP coordination (Schwarzschild, 1992), definite plurals (Link, 1991;

Williams, 1991; Moltmann, 1992), to indefinite numerals (Roberts, 1990), and to any non-distributive DP (Beghelli and Stowell, 1997). Obviously, this is a domain where one might want to use more a careful way of studying the data to see when the distributive reading is possible and how degraded it is. As far as I know, there are, in fact, four experiments that concentrated on readings of the transitive sentences with plural arguments. As we are going to see, the results of these experiments show that the distributive reading is degraded with many DPs. These include universal quantifiers headed by *all*, indefinite numerals, definite plurals and DP coordinations. Furthermore, it seems that the degradation is only mild - commonly, still the majority of people find the sentences grammatical. The data from these experiments are summarized in the following subsections.

2.2.2 Gil (1982)

David Gil in his paper discussed what readings two sentence forms, active and passive, of sentences like (7) can get in three languages. The languages he studied are Dutch, Hebrew and Bengali.

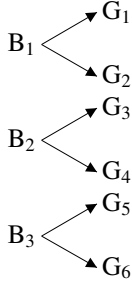
- (7) Three boys saw two girls.

He translated the sentences (7) into the respective languages using both active and passive verb forms. Here, I concentrate only on the active form of those sentences since the passive form brings in unnecessary complications irrelevant for the point I am making. The translated sentences for (7) are presented in (8a)-(8c).

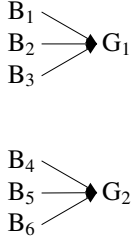
- (8) a. Drie jongens zagen twee meisjes
three boys saw-pl two girls
- b. šloša banim raʔu štey banot
three-m boys saw-3pl two-f girls
- c. tinṭi ċ^hēlē duṭi mēyēkē dēk^hēċ^hil
three-class boy-nom two-class girl-acc saw-3

He designed four possible situations using simple drawings. In these drawings the letters *B* represented boys and the letters *G* girls. The relation *seeing* was represented by an arrow. Thus, $B_1 \longrightarrow G_1$ means ‘the boy 1 saw the girl 1’. Along with the drawings a simple verbal explanation was added specifying who saw whom, as shown also here. In the actual questionnaire, each boy and girl got a name in each language.

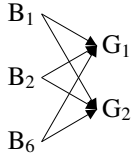
- (9) B_1 saw G_1 and G_2 , B_2 saw G_3 and G_4 , B_3 saw G_5 and G_6 .



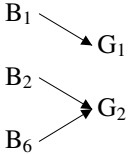
- (10) B_1, B_2 and B_3 saw G_1 , B_4, B_5 and B_6 saw G_2 .



- (11) B_1, B_2 and B_3 each saw G_1 and G_2 .



- (12) B_1 saw G_1 , B_2 and B_3 saw G_2 .



As one can see the four pictures differentiate the readings discussed in the introduction. The first two pictures represent distributive readings. In particular, in (9) the subject distributes over the object (=surface scope distributive reading) while in (10) the object takes scope over the subject (=inverse scope distributive reading). The other two pictures represent non-distributive readings. (11) represents a branching reading.¹ The picture in (12) represents what goes under the name of cumulative or co-distributive reading (see Beck and Sauerland 2000 for various definitions and names of this reading in the semantic literature).

¹Its truth-condition can be captured without switching to any logic higher than the first-order logic if we use branching quantification, hence the name (Hintikka, 1974; Sher, 1990; Schein, 1993). Nowadays, the

Gil distributed the Dutch, Hebrew, and Bengali sentences along with the drawn scenarios which were accompanied by verbal explanations to the speakers of the languages (49 Dutch speakers, 141 Hebrew speakers, and 26 Bengali speakers) and asked them to judge in which scenarios the sentence is true. For each scenario (9)–(12) the speakers were given three options: TRUE (=the sentence is true in this scenario), FALSE (=the sentence is false in this scenario), TRUEFALSE (=the sentence could be viewed as true or false given the scenario). The last option was designed for those subjects who judged a test sentence to be interpretable in two (or more) ways one of which would render it true, and the other false. It was also used by subjects who could not make a clear decision between the two extreme choices (true and false), or who viewed the state of affairs to be a possible but not very likely interpretation of the sentence.²

Notice that if the sentences (8a)–(8c) were simply ambiguous between the four readings the sentences should be judged as TRUE or TRUEFALSE for each scenario depicted in the four pictures. However, this is not what Gil found. There were clear asymmetries between the scenarios. The non-distributive readings were judged as TRUE more often than the distributive readings, and the difference has turned out to be significant for each language. In fact, the four readings could be ordered with respect to how many speakers found them true: *branching reading* \geq *cumulative reading* \geq *surface scope distributive reading* \geq *inverse scope distributive reading*, where the first reading is accepted by almost everyone, and the last reading is accepted by almost no one. If we take $A \geq B$ to mean ‘the proportion of speakers that accept A as true is bigger than the proportion of speakers that accept B and the difference is significant *or* the proportions of speakers that accept A and that accept B are not significantly different’, then the hierarchy holds for each of the three languages that Gil tested.

The results show that a predicate whose subject is an indefinite numeral has the distributive reading as a dispreferred interpretation, and the preferred interpretation is the non-distributive reading. However, distributive readings are not ungrammatical for every speaker. Apart from Bengali, where distributive readings were fully rejected probably for independent reasons (see discussion of that in Gil 1982, pages 461–462), the picture depicting the surface scope distributive reading was accepted by more than half of the speakers as TRUE or TRUEFALSE. Preferably, the semantic theory of distributivity should then not only be able to explain why distributive readings are degraded with numerals but also why there are many speakers who still accept distributive readings in these cases. I will come back to this point in the next chapter.

However, one should also be cautious with drawing conclusions from the questionnaire of Gil (1982). First, Gil tested only one verb. It is possible that the marginal status of distributive readings and the preference for non-distributive readings is not

status of branching quantification is questioned. Beghelli and Stowell (1997) show that the branching reading can be derived without use of branching quantification and the empirical coverage is better. Schlenker (2006) discusses other cases of branching readings (originally discussed in Hintikka, 1974) and shows that branching quantification encounters problems here as well.

²This is a strange collection of responses into one category but since not many people used TRUEFALSE as their response this problem is irrelevant.

a general phenomenon, it is only an effect that one finds with this particular verb. Second, the task that people got is rather abstract. They received a sentence and a diagram and have to say if the sentence is true or false given the diagram. This leaves a lot to people's imagination. It could be, for example, that distributive readings are as accepted as other readings, they just require more creativity on the side of the tested subjects. In general, we would prefer a questionnaire whose results depend on issues outside of individual language abilities (like people's creativity) as little as possible. Third, it is not clear from the paper if the items were randomized. This leaves one wondering whether we are not just observing an effect of the order in which the tested reading were presented (distributive readings first, non-distributive readings last). The next experiments that I am going to discuss avoid these problems.

2.2.3 Brooks and Braine (1996) and Kaup et al. (2002)

Brooks and Braine's experiments (Brooks and Braine, 1996) were designed to test children's comprehension of the universal determiners *all* and *each*, as well as numerical adjectives like *three*, as compared to comprehension of the same lexical items by adults. The issue relevant for the discussion here is their Experiment 2, which was partly built on Ioup (1975). They tested which reading is preferred with which determiner and whether this preference depends on the syntactic position of the determiner phrase.

Six sentence types were used in the experiment. In three of them the matrix verb was in active voice, in the other three the matrix verb was in passive voice. The sentence types were further distinguished by the type of determiner in the DP that was assigned the agent theta role. The determiner was either *each*, *all* *the*, or, in case the DP contained a numerical adjective, it was null. This gives us six sentence types.

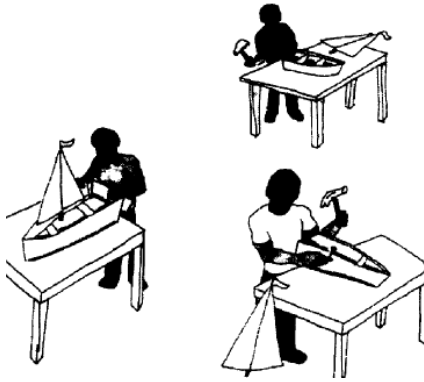
- (13) a. Three NPs are V-ing an NP
 b. All the NPs are V-ing an NP
 c. Each NP is V-ing an NP
 d. An NP was being V-ed by three NPs
 e. An NP was being V-ed by all the NPs
 f. An NP was being V-ed by each NPs

The *V* in each sentence type was in each test item substituted by one of the six verbs tested (*build*, *climb*, *lift*, *wash*, *pull*, or *load*), the *NPs* were substituted by appropriate noun phrases. Each test item was accompanied by two pictures. One picture represented a distributive reading. The other picture depicted the agents engaged in a collective activity. Notice that this depiction represents a non-distributive reading. For example, one test item was:

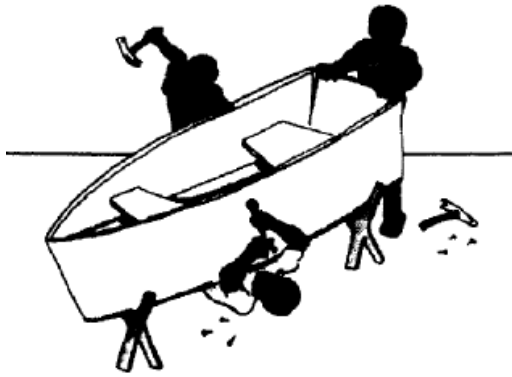
- (14) Three boys are building a boat.

This test item was accompanied by the following two pictures.

- (15) a. Picture depicting distributive reading



- b. Picture depicting collective activity



The subjects' task was to choose a picture that they thought went best with the presented sentence. Afterwards, they were asked if the sentence could also fit the other picture.

Since I am mainly interested in adults' readings of active transitive sentences, I put aside the results of sentences in passive, and I am not going to discuss how children judged the presented sentences. This leaves us with 20 adults tested in Brooks and Braine's Experiment 2 and their picture selection for sentence types (13a)-(13c). If each sentence was ambiguous between a distributive and non-distributive reading we would expect that each reading would be selected in 50% of cases, or close to that number (since there would be no preference the participants would randomly choose a picture). But Brooks and Braine (1996) found quite a different distribution, as we can see in Table 2.1. Table 2.1 shows that people preferred the picture depicting collective activity (=non-distributive reading) when sentences had the subject *three NPs*, as well as *all the NPs*. When the subject was *each NP*, the picture representing the distributive reading was strongly preferred. In case of a random choice between the two pictures, we would expect a binomial distribution with probability $p = 0.5$ and number of cases $n = 60$ (20 adults each of which got three test items of one sentence type). We can then calculate how likely it is that the numbers in (2.1) arise by random choice. The

Table 2.1: Percentage of answers in which subjects chose distributive or collective interpretations

	Distributive reading (picture 15a)	Collective reading (picture 15b)
Three NPs are V-ing an NP	2.5%	97.5%
All the NPs are V-ing an NP	16.7%	83.3%
Each NP is V-ing an NP	99.2%	0.8%

probability that 97.5% of responses favouring the non-distributive reading in *Three NPs are V-ing an NP* is due to the random choice is 1.6×10^{-13} . The probability that 83.3% of responses favouring the non-distributive reading in *All the NPs are V-ing an NP* is due to the random choice is 1.7×10^{-4} , i.e., 0.0002%. The probability that 0.8% of responses favouring the non-distributive reading in *Each NP is V-ing an NP* is due to the random choice is 5×10^{-15} . Thus, assuming that the three sentences (13a)-(13c) are ambiguous and the results in Table 2.1 are due to random choice is highly unlikely. This is in fact supported by the second task in which people had to say whether both pictures were an acceptable description of the sentence. As Table 2.2 shows most people in fact rejected a collective interpretation with *each*, and some people rejected a distributive interpretation with *all* (Brooks and Braine 1996 did not test sentences with numerals in this task). Based on the two experiments, I conclude the following: DPs

Table 2.2: Percentage of participants who accepted collective or distributive interpretations

	Distributive reading (picture 15a)	Collective reading (picture 15b)
<i>each NP</i>	100%	25%
<i>all NPs</i>	70%	100%

headed by numerals lead to a preference for non-distributive readings over distributive readings. The same, albeit less significantly, holds for DPs headed by *all the*. In case of DPs headed by *each* the situation is reversed: distributive readings are preferred in this case. The result for indefinite numerals confirms the result of Gil (1982) discussed in the previous section. Notice that the results here are more robust than in Gil (1982) because various verbs were tested. Furthermore, the task was somewhat more natural (selection of pictures which depicted the situation, rather than judging an abstract graph) and test items were given in randomized order.

Another experiment on the availability of distributive and non-distributive readings has been reported in Kaup et al. (2002), who tested whether the German pronouns *sie* ‘they’ and *beide* ‘both’ favor distributive or non-distributive readings. A test item was a sentence like (16a) where the clausal subject was either ‘they’ or ‘both’. After reading the sentence the participants (60 students) were asked the question (16b).

- (16) a. Sie / Beide brachten ein Geschenk mit.
They / Both brought a gift with

‘They / Both brought a gift.’

- b. Wie viele Geschenke wurden mitgebracht?
 How many gifts were with-brought?
 ‘How many gifts were brought?’

The possible answers to the question ‘how many gifts were brought’ was ‘one’ or ‘two’. The first answer corresponds to the non-distributive reading of the previous sentence, the second answer corresponds to the distributive reading. The participants were instructed to indicate when they thought that both answers are possible responses to the question, or when only one answer is possible, or when one answer is more likely. The answers were coded as (1)=‘one’ is the only possible answer, (2)=‘one’, or ‘two’, but ‘one’ is more likely, (3)=‘one’ or ‘two’, (4)=‘one’, or ‘two’, but ‘two’ is more likely, (5)=‘two’ is the only possible answer. The mean interpretation score was 1.40 when the subject was ‘they’, which indicates a strong preference for the non-distributive reading. When the subject was ‘both’ the mean interpretation score was 3.50, indicating a preference for distributive readings. The difference between the two scores was significant. This suggests that to indefinite numerals that Gil (1982) and Brooks and Braine (1996) discussed we can add another plural DP that makes us prefer the non-distributive reading over the distributive reading: the pronoun ‘they’. On the other hand, ‘both’ lies somewhere between ‘they’ and DPs that strongly prefer distributive readings, like quantifiers headed by *each* as shown in Brooks and Braine’s experiment.

2.2.4 Frazier et al. (1999)

The last experiment to be discussed here that tested the availability of non-distributive and distributive readings was an on-line task. Consider sentences like (17a)-(17d).

- (17) a. Billy and Francis **each** got one ice cream cone the last time we came to Barts.
 b. Billy and Francis **together** got one ice cream cone the last time we came to Barts.
 c. Billy and Francis got one ice cream cone **each** the last time we came to Barts.
 d. Billy and Francis got one ice cream cone **together** the last time we came to Barts.

The examples differ in two respects: first, whether the floating quantifier *each* or the adverb *together* is present in the clause ((17a) and (17c) contra (17b) and (17d)), second, whether *each/together* are in the preverbal position ((17a)-(17b) or follow the object ((17c)-(17d)). The differences are highlighted. Recall that transitive sentences with plural arguments are ambiguous between distributive and non-distributive readings. However, because of the presence of *each/together* in (17a)-(17d) the presented sentences are disambiguated. In particular, floating *each* forces the distributive interpretation while *together* forces the non-distributive interpretation. Crucially,

each/together are introduced right after the plural DP in (17a) and (17b) but only after the object in (17c) and (17d). This means the latter pair is locally ambiguous (between distributive and non-distributive readings). It only becomes disambiguated after *each/together* is encountered. Frazier et al. (1999) tested reading difficulties connected to the quadruple of sentences like these by examining participants' eye movements in a reading task. When the disambiguator between distributive and non-distributive readings (*each* or *together*) appeared preverbally it had no effect on the speed of the reading. However, when the disambiguator followed the object then sentences with *each* (as in (17c)) showed a slowdown as compared to sentences with *together* (as in (17d)). The slowdown was detected at the region following the disambiguator (*the last time* in the example above) and was significant in both first pass reading times, as well as total reading times. Also, there was a significant increase of regression out of this region when the late disambiguator was *each*.

The asymmetry between *each* and *together* in locally ambiguous sentences should be familiar to anyone acquainted with the work on garden path effects. Garden path effects show up in locally ambiguous sentences in which the preferred parsing of the local ambiguity turns out to be incorrect at the point of disambiguation. Consider (18a) and (18b).

- (18) a. Since Jay always jogs a mile this seems like a short distance to him.
- b. Since Jay always jogs a mile seems like a very short distance to him.

When *a mile* is encountered it is parsed as the object of the sentence. This way of parsing the DP has various explanations in the psycholinguistic literature. Frazier (1978) assumes that the human parser always prefers to parse a new material into the clause currently processed, unless it violates grammatical principles. Pritchett (1992) makes use of the principle of Theta Attachment, and states that Theta Criterion must be satisfied as much as possible at every stage of processing. Whatever the reason for this preference in parsing, it distinguishes between (18a) where it turns out to be correct, and (18b) where it turns out to be wrong. In the latter case, a different parsing must therefore be considered. This should then cause slowdown in the reading of (18b), as compared to (18a), after the disambiguator is encountered. In fact such a slowdown was found in the eye-tracking experiment of Frazier and Rayner (1982).

Similarly, the slowdown in (17c) is most likely caused by the readers' parsing preference that is at odds with the interpretation that *each* triggers. Since *each* triggers the distributive reading of the sentence this means that the part prior to *each* (i.e., *Billy and Francis got one ice cream cone*) is not parsed as the distributive reading. Rather, it is parsed as the non-distributive reading, which is compatible with the interpretation of *together*. This suggests that a DP coordination is another phrasal constituent that makes us prefer the non-distributive over distributive reading.

2.2.5 Summary

We have seen four experiments in this section. They showed that the non-distributive reading of predicates is preferred if the subject is a numeral DP like *three boys* (Gil,

1982; Brooks and Braine, 1996), the pronoun *they* (Kaup et al., 2002), or a coordination of proper names (Frazier et al., 1999). On the other hand, if the subject is a universal distributive quantifier like *each boy*, the distributive reading of the sentence is strongly preferred over the collective reading (Brooks and Braine, 1996). Furthermore, the experiments of Brooks and Braine (1996) and Kaup et al. (2002) showed that quantifiers like *all the boys* or *both* lie somewhere in between the two extremes.

The DPs can thus be split into (at least) three groups depending on the level of acceptability of distributive readings. The groups that arise correspond to groups well-known from other work on semantics of quantifiers (see, for example, Reinhart 1997, Beghelli and Stowell (1997), Winter 2001a). Here, I follow the terminology of Beghelli and Stowell (1997). We have (going from the best subjects of predicates interpreted distributively to the worst ones):

- (19) a. **DPs with universal distributive quantifiers (DQ-DPs)**, which are morphologically singular and with which the distributive reading of predicates is strongly preferred (*each boy*)
- b. **DPs with counting quantifiers (CQ-DPs)**, which have quantifiers that determine cardinality of expressions, are morphologically plural (*all boys*, *both boys*, modified numeral DPs), can occur with some collective predicates (as we will see in Chapter 3) and with which the distributive reading is slightly dispreferred
- c. **Group-denoting DPs (G-DPs)**, which consist of numeral DPs, definite DPs, coordinations of proper names (*three NP...*, *the NP*, *DP* and *DP*), and which can support collective interpretations in cases in which counting quantifiers are impossible (as we will see in Chapter 3) and with which the distributive reading is more strongly dispreferred

In the next section I am going to turn to the questionnaire I conducted with Øystein Nilsen, which tested how sentence-internal readings of expressions of non-identity (E(n)Is) are sensitive to the type of DP.

2.3 Questionnaire testing readings of E(n)Is

In this questionnaire, Øystein Nilsen and I tested the sentence-internal reading of E(n)Is. We have seen in the introduction (Section 2.1) that the sentence-internal reading of E(n)Is requires a plural DP antecedent. The questionnaire we conducted was supposed to further differentiate between plural DP antecedents. More specifically, we were interested in the following issues:

1. Are there E(n)Is whose sentence-internal reading is dependent on the type of DP in the same way as distributive readings?
2. Are there E(n)Is whose sentence-internal reading also shows the sensitivity to the type of DP but a different one than distributive readings?

- c. DP [_{VP} V ... een verschillend NP]
DP [_{VP} V ... a different NP]
- d. DP [_{VP} V ... dezelfde NP]
DP [_{VP} V ... thesame NP]
- e. DP [_{VP} V ... een NP]
DP [_{VP} V ... a NP]

The DP in the sentence forms could be anything from the following: universal distributive quantifiers *elk NP* ‘each/every NP’, *ieder NP* ‘each/every NP’, other universal quantifiers *alle NP* ‘all the NPs’, definite plurals *de NP* ‘the NPs’, coordination *DP en DP* ‘DP and DP’. With respect to the grouping in (19) the first two DPs are representatives of universal distributive quantifiers, the third is a representative of counting quantifiers, and the last two are representatives of group-denoting DPs. In addition to these we also tested the negative quantifier *geen NP* ‘no NP’, which is commonly included in the group of distributive quantifiers.

The object positions were filled in by E(n)Is, with the exception of (21e), which tested distributive readings. This was added to the questionnaire mainly to make sure its validity. If we confirm the results of the previous studies we can be more secure that the questionnaire is setup correctly. Unlike in previous studies, we did not use only indefinites and numerals in the object position when testing distributive readings. We also used modified numerals. The fact that these, as we are going to see, do not covary with respect to group-denoting DPs will play a role when we are going to explain why distributive readings are degraded. This will be discussed more in Chapter 3.

The NP and V were filled in according to the presented scenario. For example, for Scenario 1 these were some of the test items that tested distributive readings and sentence-internal readings of *de anderen* ‘the others’ with various quantifiers (for the full list, see Appendix):

- (22) a. Iedere matroos was jaloers op de anderen.
every sailor was jealous of the others
- b. De matrozen waren jaloers op de anderen.
the sailors were jealous of the others
- c. Jip, Jaap, en Joop waren jaloers op de anderen.
Jip, Jaap, and Joop were jealous of the others
- d. Iedere matroos vond een onderkomen.
every sailor found a dwelling
- e. De matrozen vonden een onderkomen.
the sailors found a dwelling
- f. Jip, Jaap, en Joop vonden een onderkomen.
Jip, Jaap, and Joop found a dwelling

Notice that distributive readings of (22d-f) are forced given the scenario. Thus, if someone responds that he can say, for example, (22f) in the scenario we know that he accepts the distributive reading of that sentence. Similarly, (22a-c) are interpretable

within the scenario only as sentence-internal readings (paraphrased as ‘every sailor was jealous of the other sailors’, or ‘the sailors were jealous of each other’).

Within each list, two scenarios tested sentence-internal readings of *de anderen* ‘the others’ and distributive readings, and the other two scenarios tested sentence-internal readings of *een ander* ‘different₁’, *een verschillend* ‘different₂’, and *dezelfde* ‘same’. To see how the sentence-internal readings of the latter *E(n)Is* were tested take a look at the following scenario:

Scenario 2. Het land Ukhbar is verdeeld in twee provincies: het noorden, en het zuiden. Het noorden heeft drie steden: Appel, Sinaasappel en Peer. Het zuiden heeft ook drie steden: Den, Berk en Palm. Er zijn geen andere steden of dorpen in het land Ukhbar.

Vroeger had elke stad in Ukhbar zijn eigen eenheid voor lengte. In het noorden konden ze het niet eens worden welke eenheid de beste was, dus zelfs vandaag de dag gebruikt elke stad nog zijn eigen lengte-eenheid. In het zuiden konden de steden het daarentegen wel eens worden over een lengte-eenheid. Alle steden daar gebruiken die eenheid, en de andere eenheden zijn afgeschaft.

The country Ukhbar is divided into two provinces: the north one and the south one. The north province has three towns: Appel, Sinaasappel and Peer. The south province also has three towns: Den, Berk and Palm. There are no other towns in Ukhbar.

In the past, each town in Ukhbar had its own length unit. In the north, they could not agree which of these units is the best one. Therefore, each city kept its own length unit even nowadays. In the south province the cities agreed on one length unit. All the towns there use only that length unit, and the other ones are forgotten.

Again, after reading the scenario the participants were presented with a sentence and asked whether they can say it, given the scenario, and, as before, they had six buttons to choose from.

Given this scenario, these were some of the sentences that tested sentence-internal readings of *ander* ‘different₁’, *verschillend* ‘different₂’, and *dezelfde* ‘same’ (for the full list, see Appendix):

- (23) a. Iedere stad in het noorden heeft een andere lengte-eenheid.
every town in the north has a different length-unit
- b. De steden in het noorden hebben een andere lengte-eenheid.
the towns in the north have a different length-unit
- c. Iedere stad in het noorden heeft een verschillende lengte-eenheid.
every town in the north has a different length-unit
- d. De steden in het noorden hebben een verschillende lengte-eenheid.
the towns in the north have a different length-unit
- e. De steden in het zuiden hebben dezelfde lengte-eenheid.
the towns in the south have the-same length-unit

Notice that if a participant says that he, for example, can say (23a) in the given context we know that he can use that sentence to express the sentence-internal reading of ‘different₁’ (paraphrased as ‘every town in the north has a length unit different from the length units that the other towns in the north have’). *Een ander* can also have a discourse reading but this could hardly be available in this case given that no other length unit is mentioned in the sentence. It is unlikely that the participants, when judging the sentence, could recall some of the length units mentioned in the background scenario and understand ‘different₁’ as ‘different from this unit length’ (i.e., make use of the discourse reading of ‘different’). Furthermore, if they could do this they should be able to use the same strategy regardless of the type of DP argument, and thus we would not expect any effect of DP, which, as we will see, is incorrect. Secondly, we used controls to see if the participants really cannot make use of the discourse reading. One of the controls we used in this scenario was the following sentence:

- (24) Iedere stad in het zuiden heeft een andere lengte-eenheid.
 every town in the south has a different length-unit

(24) is false under the sentence-internal reading (because all towns in the south have the same length unit) but would be true if the participants were able to consider some other length unit(s) from the background scenario and use the discourse reading of *een ander*. For instance, if the participants could consider the length units of the towns in the north and treat *ander* ‘different₁’ as anaphoric to these length units the sentence could be judged as true. However, this control was rejected by more than 98% participants, which suggests that they could not construct the discourse reading.

To sum up, the test items differed with respect to two parameters: first, what expression appeared in sentence-internal readings (*de anderen* ‘the others’, *een ander* ‘different₁’, *verschillend* ‘different₂’, *dezelfde* ‘same’, or no expression in case of distributive readings); second, what type of subject was used (headed by the determiner *elk* ‘each/every’, *ieder* ‘each/every’, *alle* ‘all’, *geen* ‘no’, *de* ‘the’, or the subject consisted of singular DPs coordinated by *en* ‘and’). To facilitate the upcoming discussion, I use the following coding: *elk/ieder/alle/geen/de/en* signals what type of subject was used. *de anderen/een ander/een verschillend/dezelfde* signals what type of expression was tested in its sentence-internal reading. Distributivity is coded as *dist*. For example, if we later see *En-dist* that means that we are talking about test items in which the subject was a coordination of DPs, and in which the predicate was tested for distributive readings.

Apart from the test items, each scenario had two controls. Some of the controls, as I have already mentioned, tested if the participants could not access other readings than the sentence-internal reading with *een ander* ‘different₁’. The remaining controls tested if the participant read the scenario by presenting sentences that were either clearly false or true given the scenario. For example, for Scenario 1, one of the controls was the following:

- (25) Er waren drie matrozen op het eiland.
 there were three sailors on the island.

If a participant failed to correctly answer three or more controls his responses were excluded. This affected about 10 participants in total, so less than 1% of all the participants of the questionnaire.

2.3.2 Results

For each test sentence, we got a distribution of judgments which was strongly non-normal: the extreme values 1 and 6 were selected more often than middle values. The speakers tended to select extreme values 1 or 6 probably because we had used relatively “mild” labels for the extreme values (‘can say it/cannot say it’). Had we used labels like “perfect” and “completely impossible” instead, the results might have been different. Another reason might be the question we asked. We specifically asked if the participants can say the presented sentence given the scenario. This might invite yes/no answers and thus choosing only 1 or 6 and much less the values in between (even though, I should stress that almost every participant did use the middle values, albeit usually only in a few cases). If we phrased the question differently (for example, *How natural does the sentence sound given the scenario?*), the participants might be more likely to grade their judgements more (Featherson, 2007).

Given the bimodal distribution, we could not rely on parametric statistical tests. We used two non-parametric test, namely *the Mann-Whitney test* and *the Wilcoxon signed-rank tests* to overcome this problem. For the same reason, I generally don’t report *mean scores* since the bimodality of the score distributions would render the means misleading. Instead, I report percentages of responses for each value from 1 to 6.

With *the Mann-Whitney tests*, one ranks all the individual scores obtained in two independent groups for a variable, and tests whether the *sum of ranks* is significantly different between the two groups. If the sum of ranks in the first group is significantly lower than the sum of ranks for the second, then the first group had a *higher acceptance* for the test sentences than the second group.

With a Wilcoxon’s *signed-rank test*, one tests the *same* group of speakers whether they liked one sentence better than another. One ranks their judgments separately for the two cases and checks for each speaker, whether his rank in one case is higher/lower than in the other. If the number of speakers who went *down* in rank from the first to the second case is *significantly larger* than the number of speakers who went *up* in rank, then the first sentence is judged *worse* than the second.

Because we carried out a number of tests for each lexical item whose sentence-internal readings were tested we were in danger of rejecting the null hypothesis and finding significant difference between scores where there is none (Type I error). To avoid this problem we divided the level of significance ($p = .05$) by 10 which is a number of tests carried out per *E(n)I*. This is a conservative way of avoiding Type I error and might lead to the opposite problem, Type II error, i.e., accepting the null hypothesis where it should be rejected. However, as we are going to see almost every test either has very high level of confidence, with $p < .0005$, or would not be significant even at $p = .05$ level so in reality Type II error is not going to affect us.

Distributive readings

We compared speakers’ reactions to four subjects of predicates interpreted distributively, *geen* (‘no’), *ieder* (‘every’), *de NP* (definite plural) and *en* (coordination of DPs). I start with the last three because the results of distributive readings with the negative quantifier are somewhat more complicated.

Table 2.3 shows how people accepted the distributive reading with the three subjects.³

The subject headed by *ieder* ‘every’ was accepted better than the coordination of proper names or definite plurals. The difference between the distributive quantifier and the coordination was significant (Mann-Whitney’s test, $U = 53118.50$, effect size= .25, $p < .0005$), and the same is true for the difference between the distributive quantifier and the definite plural (Mann-Whitney’s test, $U = 47054$, effect size= .36, $p < .0005$).

Table 2.3: Percentage of responses for distributive reading with *ieder* (‘every’), *de NP* (definite plural) and *en* (coordination of DPs) (1=best, 6=worst)

	1	2	3	4	5	6
<i>Ieder</i> ‘every’	91.21%	2.75%	0.55%	3.3%	0.55%	1.65%
<i>En</i> ‘and’	71.2%	10.13%	9.07%	6.13%	1.33%	2.13%
<i>De</i> ‘the’	60.85%	12.17%	7.67%	9.26%	3.97%	6.08%

When testing *geen* we encountered some problems. According to Table 2.4 it looks like the distributive reading was judged the worst when the subject was headed by *geen* ‘no’. And in fact, the lower scores of the distributive reading with *geen* ‘no’ as compared to the scores of the distributive reading with definite plurals and coordinations were significant (in both cases using Mann-Whitney’s test, $p < .001$). However, that

Table 2.4: Percentage of responses for distributive reading with *geen* ‘no’ (1=best, 6=worst)

	1	2	3	4	5	6
<i>Geen</i> ‘no’	49.87%	10.24%	12.3%	9.29%	4.37%	13.93%

conclusion is probably wrong. It turned out that one scenario was flawed. In that scenario three cities were presented each of which ‘has a church which is very popular among the citizens’. The test item to test the distributive readings was:

- (26)
- Geen stad heeft meer dan één kerk.
no city has more than one church

However, the fact that a city has a famous church does not in any way mean that a city has no more than one church. This caused most people to judge the test item

³In this table and all the subsequent tables the following convention is used: in the highest row the best accepted test type is placed (in this case, the distributive reading with a distributive quantifier as the subject), and as one goes down the acceptance decreases.

(26) as 6 (the lowest score) which skewed the data. Once we recognized the flaw and excluded this scenario, the picture changed. We obtained the results in Table 2.5, where, as one can see, the distributive reading with the negative quantifier is judged much better than before. In fact, now the distributive reading with *geen* ‘no’ is judged significantly better than the distributive reading where the subject is a coordination of proper names (Mann-Whitney’s test, $U = 13486$, effect size= .22, $p < .0005$). It is also judged significantly better than the distributive reading where the subject is a definite plural (Mann-Whitney’s test, $U = 14312.5$, effect size= .20, $p < .0005$). The difference between the distributive readings with *geen* ‘no’ and *ieder* ‘every’ is not significant (Wilcoxon signed rank test, $p = .03$).

Table 2.5: Percentage of the responses to the distributive reading with *geen* ‘no’ excluding the flawed scenario (1=best, 6=worst)

	1	2	3	4	5	6
<i>Geen</i> ‘no’	83.52%	3.85%	3.85%	2.2%		6.59%

To conclude coordination of DPs and definite plurals are less accepted with *Dist*, i.e., the distributive reading of predicates, than the subjects headed by *ieder* ‘each/every’. Furthermore, once we excluded the flawed scenario we saw that the coordination of DPs and the definite plurals are also less accepted with *Dist* than the subjects headed by *geen*.

The sentence-internal reading of *de anderen* ‘the others’

We compared how the sentence-internal reading of *de anderen* (‘the others’) was accepted depending on whether the antecedent of *de anderen* was *de NP* (definite plural) and *en DP* (coordination of DPs), or the DPs headed by *elk* ‘each/every’, *ieder* ‘each/every’ *geen* ‘no’, or *alle* ‘all’. As with the distributive reading in the previous section , I will put aside *geen* and come back to it later.

Table 2.6 summarizes how people accepted the sentence-internal reading of *de anderen* with all types of subjects (apart from the one headed by *geen* ‘no’). The subjects headed by *elk* were accepted significantly better than definite plural (Mann-Whitney’s test, $U = 111394$, effect size= .25, $p < .0005$) and than coordinations of proper names (Mann-Whitney’s test, $U = 20002$, effect size= .35, $p < .0005$). The same holds for the subjects headed by *ieder* and definite plurals (Mann-Whitney’s test, $U = 113037$, effect size= .24, $p < .0005$) and the subjects headed by *ieder* and coordinations (Mann-Whitney’s test, $U = 17876$, effect size= .41, $p < .0005$). Furthermore, the subject headed by *alle* ‘all’ was preferred for the sentence-internal reading of *de anderen* over the coordination of proper names (Mann-Whitney’s test, $U = 45447$, effect size= .28, $p < .0005$). *Alle* ‘all’ was also preferred in this case over the definite plural. The difference would be significant but because of the Bonferonni correction (see the discussion on Type I error above) it ended up as non-significant (Mann-Whitney’s test, $U = 29946$, effect size= .06, $p = .03$). We also compared the various quantificational antecedents of *de anderen* to each other. *Ieder* and *elk*

Table 2.6: Percentage of responses for the sentence-internal reading of *de anderen* ‘the others’ with *elk*, *ieder*, *alle*, *de* and *en* (1=best, 6=worst)

	1	2	3	4	5	6
<i>Elk</i> ‘every’	64.22%	11.01%	5.5%	8.62%	2.39%	8.26%
<i>Ieder</i> ‘every’	62.43%	11.97%	6.08%	8.1%	2.03%	9.4%
<i>Alle</i> ‘all’	45.92%	9.3%	8.17%	17.46%	3.66%	15.49%
<i>De</i> ‘the’	43.09%	7.9%	11.67%	11.13%	7.18%	19.03%
<i>En</i> ‘and’	23.73%	6.4%	14.93%	13.33%	10.13%	31.47%

are not significantly different (Wilcoxon signed-rank test, $p = .7$), but they are both significantly better than *alle* (comparing *elk* and *alle*: Wilcoxon signed-rank test, $T = 7602$, effect size= .25, $p < .0005$; comparing *ieder* and *alle*: Wilcoxon signed-rank test, $T = 5094$, effect size= .39, $p < .0005$).

Let us come back to *geen* ‘no’. Table 2.7 shows the percentages of responses to the sentence-internal reading of *de anderen* with the negative quantifier. It looks like the negative quantifier was one of the worst antecedents. However, once more this is because of a flaw in one scenario. The problematic scenario specified that three sailors were *jealous of* each other, and the test sentences were *every sailor is jealous of the others* vs. *no sailor cared for* (‘hield van’) *the others*, etc. but of course, the last sentence doesn’t follow from the context, while the first one does. If we separate out the

Table 2.7: Percentage of the responses to the sentence-internal reading of *de anderen* ‘the others with *geen* ‘no’ (1=best, 6=worst)

	1	2	3	4	5	6
<i>geen</i> ‘no’	36.58%	9.19%	10.85%	15.99%	6.06%	21.32%

problematic scenario, we get the percentages of responses summarized in Table 2.8. Definite plurals and coordinations are significantly worse than the negative quantifier (comparing the negative quantifier and the definite plural: Mann-Whitney’s test, $U = 74448$, effect size= .22, $< .0005$; comparing the negative quantifier and the coordination of proper names: Mann-Whitney’s test, $U = 36264$, effect size= .41, $p < .0005$). Furthermore, keeping the flawed scenario excluded from the consideration, the acceptance of the sentence-internal reading of *de anderen* shows no significant difference between the negative quantifier and distributive quantifiers, i.e., a DP either headed by *elk* ‘each/every’ (Wilcoxon signed-rank test, $p = .6$) or a DP headed by *ieder* ‘each/every’ (Wilcoxon signed-rank test, $p = .8$). There is a slight difference between the negative quantifier and the quantifier headed by *alle* ‘all’ but this turns out to be non-significant because of the Bonferroni correction (Wilcoxon signed-rank test, $T = 4051$, effect size= .19, $p = .01$).

To conclude we saw that definite plurals and coordinations are the worst antecedents of sentence-internal readings of *de anderen* (table 2.6). The DPs headed by *ieder* and *elk* are the best antecedents. We are also adding to this group the DPs headed

Table 2.8: Percentage of the responses to the sentence-internal reading of *de anderen* ‘the others’ with *geen* ‘no’ excluding the flawed scenario (1=best, 6=worst)

	1	2	3	4	5	6
<i>geen</i> ‘no’	67.43%	6.29%	4.0	5.71%	1.71%	14.86%

by *geen* with the proviso that it is correct to exclude one scenario (table 2.8). The DPs headed by *alle* are somewhere in between the first and the second group with respect to their ability of licensing sentence-internal readings of *de anderen* (table 2.6).

The sentence-internal reading of *een ander* ‘different₁’

The sentence-internal reading of *een ander* ‘different₁’ was tested with the quantified DPs headed by *ieder* ‘each/every’, *geen* ‘no’, and *alle* ‘all’. Furthermore, they were tested with definite plurals (*de NP*) and coordination of DPs (*En*).

Table 2.9 summarizes how people accepted the sentence-internal reading of *een ander* with all types of subjects. First, we compare how the definite plurals and the conjoined DPs differed from the quantified DPs. Since definite plurals appeared in different lists than quantified DPs, the Mann-Whitney test, which compares independent conditions, was used. The coordination was used in the same lists as quantified DPs which warrants the application of the Wilcoxon signed-rank test. The coordination was a worse antecedent of *een ander* than the quantified subjects headed by *ieder* ‘every’ (Wilcoxon signed-rank test, $T = 1578.5$, $p < .0005$) and *alle* ‘all’ (Wilcoxon signed-rank test, $T = 2716.5$, $p < .0005$). Definite plurals were also worse antecedents of *een ander* than quantified subjects headed by *ieder* (Mann-Whitney’s test, $U = 26968$, $p < .0005$) and *alle* (Mann-Whitney’s test, $U = 36354.5$, $p < .0005$). On the other hand, the coordination was a better antecedent than negative quantifiers (Wilcoxon signed-rank test, $T = 9510.5$, $p < .0005$). The definite plural was a better antecedent than the negative quantifier, even though the difference was non-significant after the Bonferroni correction (Mann-Whitney’s test, $U = 60745$, $p = .01$). Finally, as in

Table 2.9: Percentage of responses for the sentence-internal reading of *een ander* ‘different₁’ with *ieder*, *alle*, *de*, *en* and *geen* (1=best, 6=worst)

	1	2	3	4	5	6
<i>Ieder</i> ‘every’	82.99%	6.1%	3.88%	2.59%	0.74%	3.7%
<i>Alle</i> ‘all’	70.89%	10.49%	5.24%	3.98%	1.81%	7.59%
<i>En</i> ‘and’	49.82%	10.81%	11.9%	8.79%	4.58%	14.1%
<i>De</i> ‘the’	31.9%	8.24%	13.8%	9.5%	9.14%	27.42%
<i>Geen</i> ‘no’	28.24%	6.29%	10.25%	11.33%	9.35%	34.53%

previous cases we compared how quantified DPs differed from each other. *Geen* was judged as the worst one, which is unsurprising since it was judged worse than both definite plurals and conjoined DPs, to begin with. More interestingly, the comparison

between *alle* and *ieder* revealed that the former was significantly worse than the latter (using the Wilcoxon signed-rank test, $T = 513.50$, $p < .005$). This means that the antecedents of *een ander* are ordered with respect to their ability to license the sentence-internal reading on a scale similar to the antecedents of *de anderen*. *Ieder* is the best antecedent, followed by *alle* which is further followed by *en* and *de*. Unlike with *de anderen*, *geen* is the worst antecedent of all. Unlike in the previous cases here it is not the effect of a scenario, which becomes clear when we compare scenarios in which *geen-een ander* was judged - the difference between the scenarios is far from significant ($p = .3$). This shows that the degraded status of *geen* as the antecedent of *een ander* is genuine.

The sentence-internal reading of *een verschillend* ‘different₂’

The sentence-internal reading of *een verschillend* ‘different₂’ was tested with the same antecedents as *een ander*: quantified DPs headed by *ieder* ‘each/every’, *geen* ‘no’, and *alle* ‘all’; definite plurals (*de NP*) and coordination of DPs. We proceed as in the previous cases.

Table 2.10 summarizes how people accepted the sentence-internal reading of *een verschillend* with all the DP types. As before, we first compare how the definite plurals and the conjoined DPs differed from the quantified DPs. The coordination of proper names was tested in different lists than the quantified DPs. The same holds for the definite plurals and the quantified DPs, with the exception of *ieder* ‘every’ which was tested in the same lists as definite plurals. Thus, we use the Mann-Whitney test in all cases apart from the last one. Coordinations of proper names were significantly bet-

Table 2.10: Percentage of responses for the sentence-internal reading of *een verschillend* ‘different₂’ with *ieder*, *alle*, *de*, *en* and *geen* (1=best, 6=worst)

	1	2	3	4	5	6
<i>Alle</i> ‘all’	79.65%	8.67%	4.25%	2.83%	1.77%	2.83%
<i>En</i> ‘and’	75.41%	8.74%	6.56%	2.37%	3.28%	3.64%
<i>De</i> ‘the’	60.8%	11.07%	8.71%	5.26%	4.90%	9.26%
<i>Ieder</i> ‘every’	62.77%	7.3%	8.58%	5.84%	4.38%	11.13%
<i>Geen</i> ‘no’	42.67%	8.61%	9.71%	7.51%	7.14%	24.36%

ter antecedents of *een verschillend* than quantifiers headed by *ieder* (Mann-Whitney’s test, $U = 54479$, $p < .0005$) and quantifiers headed by *geen* (Mann-Whitney’s test, $U = 43756$, $p < .0005$). They were not significantly different from quantified DPs headed by *alle* (Mann-Whitney’s test, $p = .12$). On the other hand, definite plurals were significantly worse antecedents of *een verschillend* than quantifiers headed by *alle* (Mann-Whitney’s test, $U = 49219$, $p < .0005$), and they were not significantly different from quantifiers headed by *ieder* (Wilcoxon signed-rank test, $p = .5$). Like coordinations definite plurals were significantly better than negative quantifiers (Mann-Whitney’s test, $U = 48898$, $p < .0005$).

Finally, as in the previous cases we compared how quantified DPs differed from each other. *Alle* was judged as the best one. It was significantly better than *ieder* (Mann-Whitney $U = 55271, p < .0005$). In turn, *ieder* was significantly better than *geen* (Mann-Whitney $U = 42818, p < .0005$).

The sentence-internal reading of *dezelfde* ‘the same’

We assumed that it is uncontroversial that the sentence-internal reading of ‘the same’ is possible with quantified DPs. We only tested how it is accepted with definite plurals. This is summarized in Table 2.11. As one can see, the sentence-internal reading of *dezelfde* is accepted almost by everyone. In fact, its score does not differ from the acceptability of a distributive reading when the subject is a distributive quantifier, i.e., headed by *every* (Mann-Whitney’s test, $p = .4$), and it is significantly better than the distributive reading with definite plurals (Mann-Whitney’s test, $U = 49210.5, p < .0005$).

Table 2.11: Percentage of responses for the sentence-internal reading of *dezelfde* ‘the same’ with *de* (1=best, 6=worst)

	1	2	3	4	5	6
<i>De</i> ‘the’	88.35%	4.88%	2.17%	0.27%	1.36%	2.98%

2.3.3 Discussion

In Section 2.2 we have seen various evidence that the distributive reading of predicates depends on the DP type of the subject. In particular, DPs can be split into three groups (see (19)) that can be ordered on a scale of how acceptable they are with the distributive reading of predicates. The scale is given in Table 2.12.⁴

Table 2.12: Subjects of distributive readings

DQ-DPs	<	CQ-DPs	<	G-DPs
--------	---	--------	---	-------

In the presented questionnaire we have tested the following assumption: are the antecedents of the sentence-internal reading of E(n)Is ordered similarly as the subjects of the distributive reading of predicates? The answer is: some are and some are not.

First, let me go back to the distributive reading tested in the questionnaire. We have seen that universal distributive quantifiers are better subjects of predicates interpreted distributively than coordinations of proper names and definite plurals. The latter two belong to group-denoting DPs. Thus, these results confirm the results of the previous

⁴‘ $\alpha < \beta$ ’ should be read as α is preferred over β as the subject of a predicate that is interpreted distributively; recall that DQ-DPs=DPs with universal distributive quantifiers, CQ-DPs=DPs with counting quantifiers, G-DPs=group-denoting DPs, see(19).

studies on distributivity, which is encouraging. They also confirm the intuition that theoretical linguists had concerning the degraded status of the distributive reading of predicates whose subjects are definite plurals (see Section 2.2.1). Furthermore, the anti-additive quantifier ‘no NP’ did not differ from universal distributive quantifiers. This is in accordance with most research which considers the negative quantifier a distributive quantifier.

We can now order the antecedents of the sentence-internal reading of *de anderen* ‘the others’. The tested universal distributive quantifiers did not differ but they were significantly better antecedents than *all NP*, i.e., a counting quantifier, which in turn was a significantly better antecedent than group-denoting DPs.⁵ This gives us the following hierarchy which is equal to Table 2.12.

Table 2.13: Antecedents of the sentence-internal reading of *de anderen* ‘the others’

DQ-DPs	<	CQ-DPs	<	G-DPs
--------	---	--------	---	-------

The sentence-internal reading of *een ander* ‘different₁’ also closely matches distributive readings. The quantified DPs headed by *ieder* ‘every/each’ were the best antecedents followed by the quantified DPs headed by *alle* ‘all’, which was further followed by definite DPs and conjoined DPs. Up to this point, this was again a perfect match with distributive readings. However, the fact that the negative quantifier *geen NP* ‘no NP’ was judged as the worst antecedent is surprising in this respect (Table 2.14).

Table 2.14: Antecedents of the sentence-internal reading of *een ander* ‘different₁’

<i>ieder NP</i> ‘every NP’	<	CQ-DPs	<	G-DPs	<	<i>geen NP</i> ‘no NP’
----------------------------	---	--------	---	-------	---	------------------------

The sentence-internal reading of *een verschillend* ‘different₂’ differed much more from the distributive reading. The best antecedents were counting quantifiers and group-denoting DPs, followed by significantly worse universal distributive quantifiers, followed by the negative quantifier. In fact the situation is slightly more complicated because in the group of group-denoting DPs coordinations were better than definite plurals. The former was accepted as much as the counting quantifier *alle NP* while the second was accepted only slightly better than the universal distributive quantified DP ‘every NP’ (Table 2.15). I will come back to the difference between G-DPs in a second.

Simple plots might help visualize the situation. Figure 2.1 shows two graphs. The first one plots mean values for the subjects of distributive predicates. We see that the mean score is the lowest with universal distributive quantifiers and goes higher as we

⁵This is a slight oversimplification. We have seen that *alle NP* was better than definite plurals but the significance only reached the level of significance $p = .03$ which is not significant given the Bonferroni correction. But I will keep the distinction here. We will see in Chapter 4 where another questionnaire on the sentence-internal reading of *the others* is discussed that the distinction is warranted.

Table 2.15: Antecedents of the sentence-internal reading of *een verschillend* ‘different₂’

<i>alle NP</i> ‘all NP’	<	<i>de NP</i> ‘the NP’	<	<i>geen NP</i> ‘no NP’
<i>DP en DP</i> ‘DP and DP’		<i>ieder NP</i> ‘every NP’		

go along the x-axis, and drops again with the negative quantifier. The second graph plots the distributive reading and the sentence-internal reading of *de anderen* ‘the others’ together. The sentence-internal reading of *de anderen* is in general worse than the distributive reading but it is sensitive in the same way to the type of antecedents as the distributive reading is sensitive to the type of subjects. The only difference is between coordinations and definite plurals. The distributive reading prefers the former but the sentence-internal reading of *de anderen* prefers the latter.

I believe that the preference for coordinations over the definite plural in the first case has nothing to do with distributivity per se but rather with a general strategy on referring. When the name of an individual is known one prefers to use that over using a definite description. Thus, if one knows that my name is *Jakub*, one normally prefers using that name over using a definite description like *the man* or *the linguist*. Thus holds especially when the name is short (see Ariel 1990 on the preference along this line). Given that the scenarios always introduced (short first) names it is likely that speakers preferred using them over using definite descriptions and this showed in the higher acceptance of the former. If this is correct then the preference for using proper names instead of definite descriptions has nothing to do with distributivity. This leaves the opposite pattern in case of *de anderen* unexplained. I will not be able to offer an account why this opposite pattern holds in case of *de anderen* ‘the others’. However, it is possible that this has to do with the fact that *de anderen* is anaphoric to the plurality and the anaphoric link is more easily accessible when the plurality is presented as sharing some property, which happens when we use one definite description, and it is less easily accessible when no such property in common is presented. The latter is true when we create a plurality simply by listing names.

Figure 2.2 plots the mean scores of distributive readings and the sentence-internal reading of *een ander* ‘different₁’ and *een verschillend* ‘different₂’. The sentence-internal reading of ‘different₁’ shows the same sensitivity to the type of DPs as distributive readings do. The only difference is that the sentence-internal reading of ‘different₁’ is most degraded with negative quantifiers and this is not true for distributive readings.

Finally, in case of the sentence-internal reading of *een verschillend* ‘different₂’, the only match between the two readings is in people’s reactions to the coordination and the definite plural, as can be seen in Figure 2.2. As I argued above, the difference between the coordination and the definite plural quite likely has to do with the general preference for using proper names when they are familiar (which they were in the scenarios) and thus is no support for the parallelism between *een verschillend* and the distributive reading.

Finally, we also briefly studied the sentence-internal reading of *dezelfde* ‘the same’.

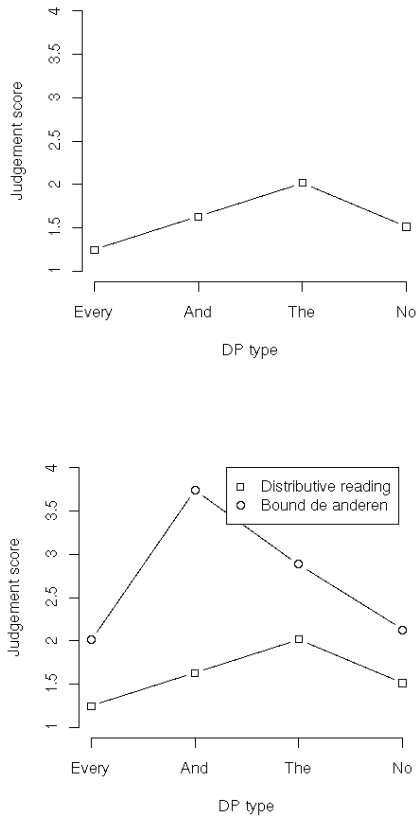


Figure 2.1: Mean scores on distributive readings and the sentence-internal reading of *de anderen* conditionalized by DP type (1=best, 6=worst)

We have seen that the acceptability of *dezelfde* with definite plurals was comparable to the acceptability of the distributive reading with distributive quantifiers, and was significantly better than the acceptability of the distributive reading with definite plurals. While one should test the sentence-internal rading of ‘the same’ further, I take it that these data suggest that the sentence-internal reading of ‘the same’ is not sensitive to the antecedent in the way the distributive reading is.

To conclude, the distributive reading is sensitive to the type of subject. The same sensitivity applies to the sentence-internal reading of *de anderen* (but something extra needs to be said about the preference for definite plurals over coordinations of proper names in the latter case). The sentence-internal reading of *een ander* shows also similar sensitivity to its antecedents (but something extra must be said about the degraded status of negative quantifiers). Finally, *een verschillend* is preferred with

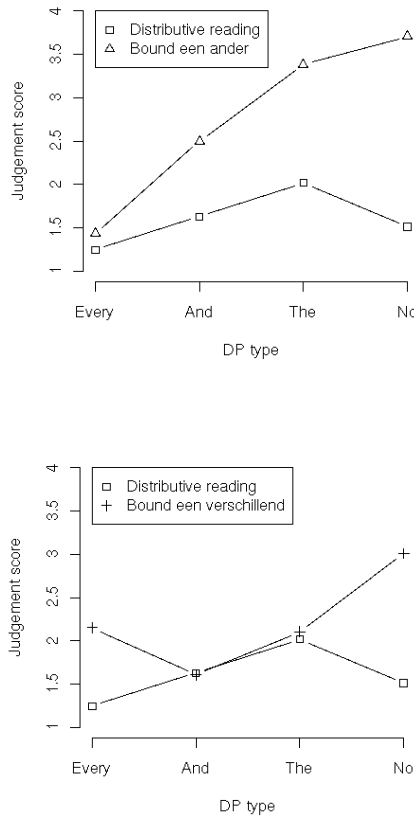


Figure 2.2: Mean scores on distributive readings and the sentence-internal reading of *een ander* and *een verschillend* conditionalized by DP type (1=best, 6=worst)

non-distributive quantifiers, and *dezelfde* is not degraded with non-distributive quantifiers (and, presumably, accepted with distributive quantifiers, as well, even though the last point was not tested but the data there seem uncontroversial).

These facts are in need of explanation. In Chapter 3 I will take on the issue concerning distributivity. I present a theory that explains why the type of DPs plays a role in licensing distributive readings. In Chapter 4 I turn my attention to the sentence-internal *de anderen* ‘the others’ and show how its sensitivity to the type of antecedent can be explained. This, furthermore, will have consequences for the analysis of reciprocity, which I am going to discuss in the same chapter. In Chapter 5 I am going to discuss the semantics of *een ander* ‘different₁’, *een verschillend* ‘different₂’ and *dezelfde* ‘the same’ and how we can derive their differences with respect to type of antecedents.

2.4 Summary

This chapter focused on data from various experiments and questionnaires and it formed the empirical basis for the upcoming analysis. We have distinguished two types of readings for sentences with plural arguments: distributive and non-distributive readings. We have seen that plural subjects are not equal with respect to their ability to license these readings. They are ordered as summarized in table 2.16 ($\alpha < \beta$ should be read as α is preferred over β as the subject of a predicate that is interpreted distributively).

Table 2.16: Subjects of distributive readings

DQ-DPs	<	CQ-DPs	<	G-DPs
--------	---	--------	---	-------

As I have shown, there is quite a lot of evidence for this hierarchy. In the second part of this chapter, I discussed the questionnaire that tested similar DPs and their ability to license sentence-internal readings of E(n)Is in Dutch. The questionnaire gave us the following results:

1. The sentence-internal reading of *de anderen* ‘the others’ is preferred with the DPs that are preferred for the distributive reading and dispreferred with the DPs that are dispreferred for the distributive reading (but something extra needs to be said about the preference for definite plurals over coordinations, which is reversed in case of distributive readings).
2. The sentence-internal reading of *een ander* ‘different₁’ is preferred with the DPs that are preferred for the distributive reading and dispreferred with the DPs that are dispreferred for the distributive reading (but something extra must be said about the degraded status of negative quantifiers as antecedents of *een ander*).
3. The sentence-internal reading of *een verschillend* is preferred with non-distributive quantifiers.
4. The sentence-internal reading of *dezelfde* ‘the same’ is not degraded with DPs that are dispreferred for the distributive reading.

An adequate analysis should capture the relation between distributivity and sentence-internal readings of E(n)Is should be captured. On top of it, we should also capture why only some E(n)Is show this relation, and others lack it. All this is the topic of the next chapters.

CHAPTER 3

The language of pluralities

3.1 Introduction

The previous chapter discussed empirical findings which show that the acceptability of the distributive reading with covariation depends on the DP type. The DPs can be split into (at least) three groups. The scale was discussed in Chapter 2 and is repeated here in Table 3.1.¹

Table 3.1: Acceptability of distributive readings with covariation

DQ-DPs	<	CQ-DPs	<	G-DPs
--------	---	--------	---	-------

Furthermore, we have seen that a similar scale holds for antecedents of the bound readings of some E(n)Is, like Dutch *de anderen* ‘the others’, and *een ander* ‘different’.

In this chapter, I offer an analysis of plurality and distributivity that captures the scale in Table 3.1. I am going to argue that the scale in Table 3.1 can be derived if we assume that G-DPs are in competition with CQ-DPs and DQ-DPs when searching for their interpretation. The competition viewpoint accounts for the fact that distributive readings are dispreferred in case of G-DPs and CQ-DPs, as well as for why distributive readings are more dispreferred with G-DPs than with CQ-DPs.

The chapter is organized as follows. In Section 3.2 I introduce a language model as assumed in the framework of Government and Binding and more recently, Minimalism, and the semantic framework for events and pluralities as developed in Landman

¹‘ $\alpha < \beta$ ’ should be read as α is preferred over β as the argument in whose scope a distributive reading with covariation can take place. DQ-DPs=DPs with universal distributive quantifiers, CQ-DPs=DPs with counting quantifiers, G-DPs=group-denoting DPs, see (19) in Chapter 2.

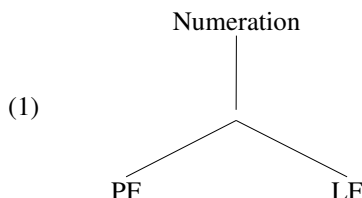
(2000). After that I discuss a few examples and modifications to the original system. This formal semantic framework will also be used in the following chapters until we switch to discourse semantics with pluralities. In Section 3.3 I discuss two systems that assume competition in the grammar. The two systems are Optimality Theory and game-theoretic analysis of Parikh (2000). I argue that the latter is to be preferred in accounting for the data concerning the distributivity since it avoids some stipulations of the OT framework that are quite ad hoc. In Section 3.4 I compare this account to possible alternatives.

3.2 Syntax and semantics of pluralities

3.2.1 Language model

I am going to assume the language model developed in the Government and Binding Theory (Chomsky, 1981, 1986), which later evolved into Minimalism (Chomsky, 1995, 2001).

Within the Minimalist Program, the computational system of the human language faculty is the T-model. The computational system connects the phonetic form, or PF (which is fed into phonological and phonetic processes), the logical form, or LF (which is fed into semantic and pragmatic processes) and the numeration (the selection from the lexicon). We can picture the model as shown in (1), where the lines represent connections between the interfaces established by the computational system.



The two operations available in the computational system are Merge and Move. Merge takes two elements from the numeration and merges them into one new object (a phrase marker). Move dislocates an object within the phrase marker. Move can apply on any of the branches in the model drawn in (1). If it applies in the upper branch the dislocation of a constituent has an effect on both PF and LF. If it applies at the leftmost branch it affects only the PF. If it applies on the rightmost branch it affects the LF. Quantifier raising, for example, is movement of the last sort.

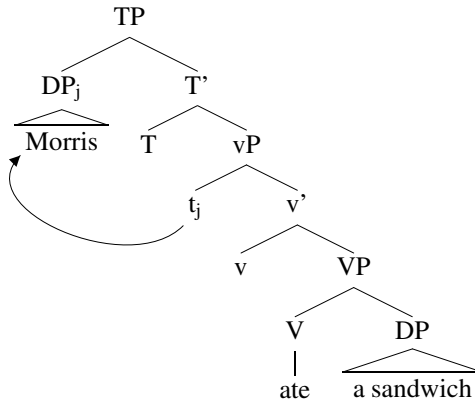
I assume a standard clause structure, in which the verbal domain is below the TP (tense phrase), which in turn is below the CP (complementizer phrase). The TP is just an abbreviation, in fact, it probably consists of many functional projections (Split-IP Hypothesis, on this see especially Cinque, 1999). The same holds for the CP (Rizzi, 1997). Since I am not really concerned here with the syntax of functional projections I will ignore these issues. Furthermore, I follow Marantz, Kratzer and others (Kratzer, 2003) and assume that in the verb phrase a separate head introduces the external argument. Finally, in English, the highest argument undergoes movement to the Specifier

of TP. In Minimalism it is common to assume that the moved phrase leaves a copy of itself in the original position. For expository purposes I will assume that the moved element leaves behind a coindexed trace. Treating these as traces simplifies the analysis of movement in semantics. Ultimately everything I am going to say could be reinterpreted within the Copy Theory of Movement (see Fox 2002 for a proposal that achieves this).

All this means that a sentence like (2) gets the syntactic structure in (2a).

(2) Morris ate a sandwich.

a.



This syntactic structure is linearized as (2). The linearization could be done for example using Kayne's LCA, in which the phrase α precedes β iff α asymmetrically c-commands β . Alternatively, we could use linearization as developed in Chomsky (1995), or recent modifications of the LCA. However, I will not dwell on this issue here given the PF side of computation is not the main concern of this thesis.²

In the next part I move to the interpretation of language. I start with modelling the domain of singular and plural individuals and groups (Section 3.2.2). In Section 3.2.3 I combine singular and plural individuals with event semantics. In Section 3.2.4 I show how LF constituents are interpreted in event semantics with pluralities, and go through a few examples of derivations. In Section 3.2.5 I modify the system of pluralities to include covers, which narrows down the ambiguity in sentences with plural arguments. Finally, in Section 3.2.6 I show how distributive quantifiers are to be analyzed.

²Later on I will assume multiple adjunction at LF structures. This is incompatible with the original LCA which disallows multiple adjuncts but it is compatible with some versions of linearization in which the LCA holds only at the PF side of the computation. If we wanted to adhere to the original LCA, we could assume more than one projections to avoid multiple adjunction at the LF side. I will leave it open what approach is better.

3.2.2 The domain of singular and plural individuals and groups

Introduction

I follow Landman (2000) and take the domain of individuals and groups to be his language of events and plurality, with some slight modifications. In this section, I show how the domain of individuals is modelled. In the next section, I discuss event semantics.

Pluralities

Take the name *Morris*. What does it refer to? A straightforward answer is that *Morris* refers to the individual named Morris, and the sentence *Morris walks* can be formalized in first-order logic:

- (3) $\text{walk}(\text{Morris})$

What about *Morris and Philip* in *Morris and Philip walk*? If we do not want to model anything beyond individuals we could say that the conjunction really conjoins propositions which were phonologically reduced. We could then capture the meaning of sentences like *Morris and Philip walk* by the following formalization:

- (4) $\text{walk}(\text{Morris}) \wedge \text{walk}(\text{Philip})$

There is a well-known problem to this solution: it is not general enough. It works in cases like the one above because the predicate *walk* is distributive. That means that the following entailment is valid: if several individuals walk, it is true that each of these individuals walk. However, many predicates in natural language lack this entailment.³ Take, for example, the collective predicate *be a couple*. Following the same strategy as above, we would get (5b) for the sentence (5a). But while this sentence is true in case Philip and Hilary form one couple, the formula in first-order logic, which postulates nothing beyond individuals, is clearly false.

- (5) a. Philip and Hilary are a couple.
 b. $\text{couple}(\text{Hilary}) \wedge \text{couple}(\text{Philip})$

The problem is that collective predicates do not license the entailment pattern that we saw with *walk*. We need something like (6) but for that to make sense, one needs to see what the plurality Hilary and Philip refer to.

- (6) $\text{couple}(\text{Philip and Hilary})$

There are three answers to the question of what plurals refer to.

In set-based theories, plurals are taken to denote sets (Landman, 1989a,b; Schwarzschild, 1996; Winter, 2001a). Plurals are a single entity, a set, which of course can have many

³In his historical overview, Lasersohn 1995 shows that already Aristotle was aware of the challenges that non-distributive predicates pose to the analysis of conjunctions. It is known that the problem these predicates raise does not concern only coordinations. It affects many other language expressions that intuitively seem to refer to plural objects.

elements as its members. Collective predicates (like *be a couple*) have sets in their extension. (5a) then can be expressed as follows (using the standard notation of set theory):

$$(7) \quad \text{couple}(\{\text{Hilary}, \text{Philip}\})$$

This would be true if the plural entity denoted by *Hilary and Philip* (which we model as a set consisting of two individuals, Hilary and Philip) is a couple. This is reasonably close to the meaning of the original sentence.

In mereological approaches, it is assumed that plurals are individuals (Link, 1998; Landman, 2000). They differ from what we normally conceive of as an individual, i.e., a singular entity, only in the following respect: individuals in mereological approaches can have other individuals as their parts. We postulate an operation that takes two individuals and merges them into one plural individual. Let us notate the operation as \oplus . Now we can denote (5a) as:

$$(8) \quad \text{couple}(\text{Hilary} \oplus \text{Philip})$$

Finally, proponents of plural logic assume that plurals do not denote one entity as in the previous theories (a set, a plural individual) but rather refer to many individuals directly. This is formalized in first-order logic with variables which are sorted into singulars and plurals where the plurals range over plural entities directly (Boolos, 1984; Schein, 1993; Linnebo, 2008).

The three viewpoints on what plurals refer to differ first and foremost in their ontological commitments. Thus, for example, it is taken as problematic that treating plurals as sets means that once I utter (5a) I committed myself to the existence of sets, and furthermore, I believe that the property of *being a couple* is a property that holds of sets rather than (pairs of) individuals directly. While ontological commitments are important for linguistic theories⁴ they lie outside of the scope of this thesis. Because of that I do not feel forced to use one approach and disregard the others.

To model the interpretation of singular and plural DPs I am going to follow mereological approaches, in particular Landman (2000). Everything I am going to say could be translated into other approaches to pluralities if one wishes so. Readers familiar with Landman (2000) are encouraged to skim the following part.

Pluralities in Landman (2000)

The interpretation domain of individuals is a set D_e . Both singular and plural individuals are in D_e . D_e is a structure ordered by 'sum', \oplus . Any plural individual is a sum of singular individuals, while any singular individual is a sum of itself and nothing else. For example, $\text{Philip} \oplus \text{Hilary}$ is a plural individual, the sum of Philip and Hilary. On the other hand, Philip is only the sum of itself and nothing else. We can furthermore define the part-of relation, \leq , where $a \oplus b = b$ iff $a \leq b$. Thus, $\text{Philip} \oplus \text{Hilary}$ have

⁴Boolos (1984) constructs examples in which the set theory used to model pluralities might run into Russell's paradox. However, Linnebo (2008) discusses various objections one could consider in order to counter Boolos' arguments.

the singular individual Philip and the singular individual Hilary as its part. Philip has only itself as its part.

The structure $\langle D_e, \oplus \rangle$ is a part-of structure. A part-of structure is an i-join semilattice, which means that $\langle D_e, \oplus \rangle$ satisfies the following properties:

- (9) a. D_e is a non-empty set
- b. \oplus is a function that assigns to every non-empty subset X of D_e an element of D_e , $\oplus X$
- c. the relation \leq , defined above, is a partial order on D_e (it is reflexive, transitive, antisymmetric)
- d. for every non-empty subset X of D_e : $\oplus X$ is the join of X

The join of X mentioned in (9d) is a member of D_e that satisfies two conditions: (10a): every member of X must be part of the join of X , (10b): the join of X is the smallest element that satisfies the first condition.

- (10) a. $(\forall a)(a \text{ is the join of } X \rightarrow (\forall x \in X)(x \leq a))$
- b. $(\forall a)(a \text{ is the join of } X \rightarrow (\forall b \in D_e)(\forall x \in X)(x \leq b \rightarrow a \leq b))$

When creating the join of two individuals, I will usually write $a \oplus b$ instead of the more elaborate $\oplus\{x : x = a \vee x = b\}$.

Apart from being an i-join semilattice, a part-of structure satisfies a few further conditions. First, it has no minimal element (11). If we take singular individuals, for example, Philip, and Hilary, there is no element that is part of each of them and they share.

- (11) No minimal element: $\neg(\exists a \in D_e)(\forall x \in D_e)(a \leq x)$

Second, part-of structures are distributive. If we take some plural individuals, b and c , whose join is another plural individual, $b \oplus c$ and we look at some part of $b \oplus c$, this part must satisfy the following: either it also is a part of b or it is a part of c , or it is the sum of some part of b and some part of c (12).

- (12) Distributivity: if $a \leq b \oplus c$ then: either $a \leq b$ or $a \leq c$
or
 $(\exists b' \leq b)(\exists c' \leq c)(a = b' \oplus c')$

Finally, part-of structures have a witness. First, we must define a proper part and overlap. A proper part is defined in (13a). Overlap is defined in (13b). Intuitively, two elements overlap if they share something that is part of both of them.

- (13) a. $a < b$ iff $a \leq b \wedge a \neq b$
- b. $a \circ b$ iff $(\exists c)(c \leq a \wedge c \leq b)$

The witness condition requires that if some element of D_e has a proper part, then there is some other part of this element that does not overlap with it (14).

- (14) Witness: if $a < b$ then $(\exists c \leq b)\neg(a \circ c)$

To sum up, part-of structures are witnessed distributive i-join semilattices. It can be shown that they are isomorphic to set theoretic lattices with the 0 (the minimal element) cut off (Landman, 1991). In other words, the part-of structure $\langle D_e, \oplus \rangle$ is isomorphic to $\langle \wp(D_e) - \{\emptyset\}, \cup \rangle$. It is this isomorphism that enables such an easy switch between mereological approaches and set-theoretic approaches to pluralities (Landman, 1989a, 1991).

In case of singular and plural individuals we want to restrict our attention to atomic part-of structures. Atoms are elements that have no proper parts. Atomic part-of structures are the structures which for every element have an atom below it (15).

$$(15) \quad \text{Atomic: } (\forall x \in D_e)(\exists a \in D_e)(a \text{ is an atom} \wedge a \leq x)$$

Thus, the interpretation domain for singular and plural individuals is an atomic part-of structure $\langle D_e, \oplus, AT \rangle$, where AT is a set of atoms.

This domain has properties that seem intuitively necessary for the domain of singular and plural individuals. In particular:

1. $\langle D_e, \oplus, AT \rangle$ is *atomistic*: every individual is the join of the atoms that are part of that individual. Imagine what would happen if this would not hold. Then we would have a plural individual, *blik*, which we would never arrive at by joining singular individuals. This clearly is against our intuition about plurals, where every plural individual *is* the join of some singular individuals. While we could have semilattices that are not atomistic, $\langle D_e, \oplus, AT \rangle$ cannot be like that.
2. $\langle D_e, \oplus, AT \rangle$ satisfies *distinctness*: for every two sets of atoms, if the sets are not identical their joins are not identical either. Again, imagine what would happen if this was not satisfied. Then we would have two sets of singular individuals, let us say {Philip, Hilary}, and {Philip, Morris}. The join of the first set would give us a plural individual, let us say *blok*, and *blok* would also be the join of the second set. Now, interpreting *blok* would give us troubles: we would know that this is a plural individual, which could be the join of two different sets of singular individuals, and we would have no idea which one it is. Clearly, there are no such plural individuals in natural language. And again, while we could have a semilattice that does not satisfy distinctness, $\langle D_e, \oplus, AT \rangle$ must satisfy it.

I will not dwell here more on the properties of the domain of singular and plural individuals. For proofs of the two properties above, as well as discussion of other properties, readers are encouraged to consult chapters six and seven of Landman (1991). I will only add one more definition. For every non-empty subset X of D_e , $[X]$, the i-join semilattice generated by X , is the closure of X under \oplus , i.e.:

$$(16) \quad [X] = \{y \in D_e : (\exists Y \subseteq X)(y = \oplus Y)\}$$

$[X]$ is the smallest sub-i-join semilattice of D that contains $\oplus X$. If X is a set of non-overlapping elements then $[X]$ is a part-of structure with X the set of atoms.

Now, we have a domain of singular and plural individuals. We further modify it by adding groups. We sort the set of atoms into individual atoms and group atoms.

Furthermore, we add two operations, *group formation*, notated as \uparrow , and *membership specification*, notated as \downarrow . When the group formation is applied to a plural individual its value is a group atom. When the group formation is applied to an individual atom its value is the same individual atom. The membership specification maps a group atom onto a plural individual that is a member of the group, and an individual atom onto itself. The domain of singular, plural individuals and groups is defined as follows:

- (17) A domain of singular, plural individuals and groups is a structure $\langle D_e, \oplus, \text{AT}, \text{IND}, \text{GROUP}, \uparrow, \downarrow \rangle$, where:
- a. $\langle D_e, \oplus \rangle$ is a domain of singular and plural individuals
 - b. AT is a set of atoms, defined above and repeated here:
 $\text{AT} = \{a \in D_e : \neg(\exists x \in D_e)(x < a)\}$
 - c. $\text{AT} = \text{IND} \cup \text{GROUP}$, and $\text{IND} \cap \text{GROUP} = \emptyset$.
 - d. $[\text{IND}]$ is a sublattice generated by IND, the set of individual atoms.
 - e. \uparrow is a one-to-one function from $[\text{IND}]$ into AT such that:
 - $\forall x \in [\text{IND}] - \text{IND} (\uparrow x \in \text{GROUP})$
 - $\forall x \in \text{IND} (\uparrow x = x)$
 - f. \downarrow is a function from AT onto $[\text{IND}]$ such that:
 - $\forall x \in [\text{IND}] (\downarrow(\uparrow x) = x)$
 - $\forall x \in \text{IND} (\downarrow x = x)$

Before providing two examples of this domain, let me shortly go through the summary of the system in (17).

The domain for singular, plural individuals and groups is, as said before, a part-of structure. I discussed the most important properties of this structure above.

AT, atoms in this structure, are of two sorts: first, IND, which is a set of individual atoms. Philip or Hilary are members of IND. Second, GROUP, which is a set of group atoms. We can generate the sublattice which consists of individuals (singular and plural) and excludes groups. This sublattice is $[\text{IND}]$. Notice that $[\text{IND}]$ includes individual atoms. If we want to arrive at the sublattice that consists only of plural individuals, we have to subtract IND, i.e., $[\text{IND}] - \text{IND}$.

Furthermore, we have two operations in our domain. First, the group formation, \uparrow . \uparrow can be approximated to the appositive use of *as a group*. It takes a plural individual and returns a group atom. In this way, it is like the interpretation of the DP *Philip and Morris, as a group*. However, it is not completely the same. The reason for its difference is its behavior with singular individuals. Apposition *as a group* is ungrammatical here:

- (18) *Philip, as a group, conjured a mouse.

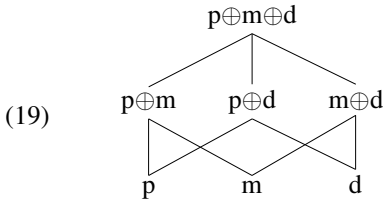
On the other hand, \uparrow is defined for individuals (in that case, \uparrow behaves as the identity function). Notice that \uparrow is a *one-to-one* function and *into*. It is *into* (and not *onto*) because there can be other group atoms in the domain which are not values of \uparrow . Take *Philip and Morris*. They can be a group in many ways. To make a group, people need

to satisfy various properties. I will say more about this in Section 3.2.5. Crucially, by satisfying these properties, one and the same people can create many groups (Landman, 1989b). It is only one of these groups that is accessed by \uparrow . Furthermore, \uparrow is *one-to-one* because if individuals that make up groups are different then their groups are different, too.

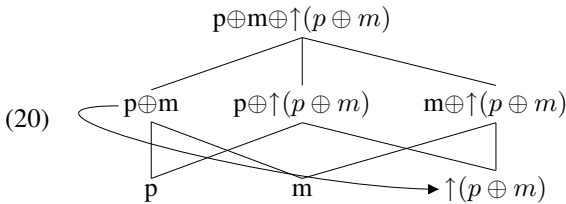
The second operation is the membership specification, \downarrow . \downarrow gives us members of groups. Like \uparrow , it can be applied to individuals as well, in which case it is the identity function. Notice that unlike \uparrow , the membership specification is *onto*. Thus, every individual (plural or singular) can make a group.

The way the system is set up, iteration of groups is not possible since \uparrow has [IND] as its domain. In this way, the system, as we have it here, which follows Landman (2000), differs from Landman (1989a) and others. Reasons for no iteration of groups were given in Schwarzschild (1996).

Let me conclude by going through two examples. Assume that our domain of individuals, D_e has only three individuals: Philip, Morris, and Desire. The interpretation domain for singular and plural DPs is then the structure $\langle \{\text{Philip, Morris, Desire, } \oplus\} \rangle$. We can draw this in a Hasse diagram, which are commonly used for lattices, as in (19). The bottom elements are atoms. Notice that apart from the fact that (19) is the domain for singulars and plurals, it is also [IND] (i.e., a sublattice generated by all non-group atoms, which are Philip, Morris, and Desire).



To have the full domain with groups we must also add group atoms. Since that would make (19) quite unreadable, I will do it below for a simpler domain. In this case, we have only two individuals: Philip and Morris. They together form a group (result of application of \uparrow , which is notated in the graph by an extra arrow $k \oplus m \rightarrow \uparrow(k \oplus m)$). This group atom can then form joins with other members of [IND], which gives us the full picture in (20).



3.2.3 Event semantics

Landman's theory follows the tradition of event semantics, which finds its origin in Davidson (1967). In his paper, Davidson offered an account of why a sentence like (21) entails any of its alternatives in which one or more adjuncts are dropped. The reason is simple. The sentence asserts the existence of an event. The adjuncts are related to the event in particular relations, expressed by prepositions (or cases in other languages than English), and they are added as separate conjuncts (see (21a)). The entailment pattern then falls out simply from axioms of first-order logic.

(21) Jones buttered the toast in the bathroom with a knife at midnight.

$$\text{a. } (\exists e) \left(\begin{array}{l} \text{BUTTER}(e)(\text{JONES})(\text{THE TOAST}) \wedge \text{IN}(e)(\text{THE BATHROOM}) \\ \wedge \text{WITH}(e)(\text{A KNIFE}) \wedge \text{AT}(e)(\text{MIDNIGHT}) \end{array} \right)$$

Since the work of Parsons (1990), event semantics is commonly set up in a neo-Davidsonian framework. Here, the verb itself is just a predicate of events, and its arguments are introduced as separate conjuncts, through thematic roles (agent, theme, etc.). Thematic roles are functions from events to individuals. The neo-Davidsonian approach captures people's intuitions about entailment patterns that were still not accounted for in the original approach by Davidson (1967). Namely, if we hear a sentence like (21) we also know that buttering took place, the toast was buttered, etc. In the neo-Davidsonian framework, (21) looks as follows:

$$(22) \quad (\exists e) \left(\begin{array}{l} \text{BUTTER}(e) \wedge \text{Agent}(e)(\text{JONES}) \wedge \text{Theme}(e)(\text{THE TOAST}) \\ \wedge \text{IN}(e)(\text{THE BATHROOM}) \wedge \text{WITH}(e)(\text{A KNIFE}) \\ \wedge \text{AT}(e)(\text{MIDNIGHT}) \end{array} \right)$$

As is commonly assumed, events are organized in the structure that is isomorphic to the structure imposed on individuals (Bach, 1986; Krifka, 1989). This means that D_v , the domain of events, is a part-of structure. The part-of structure is atomic for quantized events, and non-atomic for cumulative events. This is again similar to the domain of individuals, where mass nouns form non-atomic part-of structures, while count nouns form atomic part-of structures.

In the framework here, both stative and eventive predicates have some elements of D_v in its domain. Thus, what I call events corresponds to what Bach (1986) calls eventualities.

There is one potential problem in introducing arguments and modifiers separately from the verb denotation, discussed in Carlson (1984), Parsons (1990), Landman (2000) and others. Consider the following example:

- (23) a. Morris bought Ulysses from Philip.
 b. Philip sold Ulysses to Morris.
 c. Only one transaction took place.
 d. Hence, Morris sold Ulysses to himself.

Obviously, this inference is invalid. We might derive this inference as valid, incorrectly, if we analyze the first two sentences as (24a) and (24b), respectively and assume that the events are identical. In that case (24c) is a valid inference.

- (24) a. $(\exists e) \left(\begin{array}{l} \text{BUY}(e) \wedge \text{Agent}(e)(\text{MORRIS}) \wedge \\ \wedge \text{Theme}(e)(\text{ULYSSES}) \wedge \text{Source}(e)(\text{PHILIP}) \end{array} \right)$
 b. $(\exists e) \left(\begin{array}{l} \text{SELL}(e) \wedge \text{Agent}(e)(\text{PHILIP}) \wedge \\ \wedge \text{Theme}(e)(\text{ULYSSES}) \wedge \text{Goal}(e)(\text{MORRIS}) \end{array} \right)$
 c. $(\exists e) \left(\begin{array}{l} \text{SELL}(e) \wedge \text{Agent}(e)(\text{MORRIS}) \wedge \\ \wedge \text{Theme}(e)(\text{ULYSSES}) \wedge \text{Goal}(e)(\text{MORRIS}) \end{array} \right)$

The problem in event semantics is that arguments and modifiers are related to the verb only indirectly, through the event variable. If we let two verbs describe the same event then we can no longer ensure that arguments of one verb must always be tied only to that verb. This results in deriving the invalid inference.

As standardly assumed, this leads us to the conclusion that events are extremely fine-grained, so (23a) and (23b) are not true of the same event after all. One common way to achieve this fine-grainedness is to postulate the unique role requirement (Carlson 1984; Landman 2000 and others):

(25) **The Unique Role Requirement:**

If a thematic role is specified for an event, it is uniquely specified.

The Unique Role Requirement blocks the inference in (23). If Agent is specified for the event of buying and for the event of selling, these must be two different events otherwise the unique role requirement would be violated. Since they are two different events, the inference does not go through.

Apart from assuming the Unique Role Requirement, I am also going to assume that thematic roles apply only to atoms, that is, if a thematic role is true of some individual, then the individual is an atom.⁵ The two conditions together mean that for sentences with plural arguments like (26a) we cannot simply say that Philip and Morris were agents of the event of kissing and Hilary and Desire were themes of the same event (26b). Due to the two requirements on the thematic roles (26b) would always be false.

- (26) a. Philip and Morris kissed Hilary and Desire.
 b. $(\exists e) \left(\begin{array}{l} \text{KISS}(e) \wedge \text{Agent}(\text{PHILIP AND MORRIS})(e) \wedge \\ \wedge \text{Theme}(\text{HILARY AND DESIRE})(e) \end{array} \right)$

To deal with these cases, I assume that predicates and thematic roles are cumulative (Landman, 2000; Kratzer, 2003) (summative in the terminology of Krifka (1989)). Cumulativity of predicates and roles can be captured if we pluralize them by *. Speaking somewhat loosely, we want the pluralized predicates and roles to hold of singular individuals *and* all their joins, i.e., plural individuals built up from the singular individuals. Thus, * should do the same work as [], which generates the i-join semilattice from a set (see (16)). For example, if we have the set of events $X = \{e_1, e_2, e_3\}$ its pluralized version should be the same as $[X]$, which is the set $\{e_1, e_2, e_3, e_1 \oplus e_2, e_1 \oplus e_3, e_2 \oplus e_3, e_1 \oplus e_2 \oplus e_3\}$. In case of pairs, I assume, as is standard, that \oplus is defined pointwise. This definition is important for thematic roles.

⁵This assumption helps deal with mixed readings, in which a predicate is true neither of atomic individuals nor of a group but of subgroups of a group (Schwarzschild, 1996). For details see Landman (1989a). I will discuss these readings in Section 3.2.5.

$$(27) \quad \langle \alpha, \beta \rangle \oplus \langle \gamma, \delta \rangle = \langle \alpha \oplus \gamma, \beta \oplus \delta \rangle$$

For example, if we have $R = \{\langle a_1, e_1 \rangle, \langle a_2, e_2 \rangle\}$ then $[R] = \{\langle a_1, e_1 \rangle, \langle a_2, e_2 \rangle, \langle a_1 \oplus a_2, e_1 \oplus e_2 \rangle\}$.

Using the star operator, we can deal with examples like (26a) by following Landman (2000), Kratzer (2003) and others and using the pluralized versions of roles and predicates. Thus, (26a) is in event semantics:

$$(28) \quad (\exists e)(*kiss(e) \wedge *Agent(e)(p \oplus m) \wedge *Theme(e)(h \oplus d))$$

(28) is true if e is a plural event which has subevents of kissing, Morris and Philip are the plural agent of e , so their parts are agents of the subevents of e (agents of kissing). Similarly, Desire and Hilary are the plural theme of e , so their parts are themes of the subevents of e (themes of kissing). Thus, (28) is true if, for example, e has subevents e_1, e_2 , where Morris is the agent of kissing in e_1 and Desire is the theme, and Philip is the agent of kissing in e_2 and Hilary is the theme. This correctly captures the intuition that (29a) and (29b) together entail (26a).

- (29) a. Morris kissed Desire.
b. Philip kissed Hilary.

Following Landman (2000); Kratzer (2003) and others, I assume that thematic roles and verbal predicates are always pluralized. This works correctly for plural individuals as we have just seen and does no harm in case arguments are atomic individuals because in that case the pluralized and non-pluralized thematic roles collapse to mean the same.

3.2.4 Compositional interpretation of LF

In this section I discuss in detail how LF tree structures are systematically interpreted in the language of plurals and events that I summarized in the previous sections. The interpretation proceeds in two steps. First, every LF structure is translated into an expression in a formal language \mathcal{L} . Then the expressions of \mathcal{L} are interpreted in a model \mathcal{M} . \mathcal{L} and \mathcal{M} are a language and a model of type theory with events and individuals as basic types.

I will first start with the types that are allowed. Then, I discuss the properties of the language \mathcal{L} and the model \mathcal{M} , and after that I show how LF expressions are translated into the expressions of \mathcal{L} .

Types

In type theory, every expression is of a particular type, which determines whether expressions are individuals, relations, predicates etc. The types also determine whether expressions can combine with each other. Combinations of expressions correspond to functional application at a semantic level.

Commonly the set of possible types is given in two steps. First, we define basic types. Second, we add a rule which can recursively form complex types. The two rules are given in (30).

- (30) The set of types TYPE is the smallest set of strings such that:
- a. $e, v, t \in \text{TYPE}$, where
 - e is the type of entities
 - v is the type of events
 - t is the type of truth values
 - b. If $\sigma, \tau \in \text{TYPE}$ then $\langle \sigma, \tau \rangle \in \text{TYPE}$

Language \mathcal{L}

We want to define the set of expressions that form \mathcal{L} . We proceed in the same manner as with types, first by defining basic expressions and then complex expressions:

- (31) The set of EXP, expressions of \mathcal{L} consists of:
- a. The set of basic expressions BAS_EXP_τ , for τ being any type, such that $\text{BAS_EXP}_\tau = \text{Con}_\tau \cup \text{Var}_\tau$, where
 - Con_τ is the (possibly empty) set of constants of type τ
 - Var_τ is the set of variables of type τ
 - b. The set of complex expressions:
 - If $\phi, \varphi \in \text{EXP}_t$, then $\neg\phi, (\phi \wedge \varphi) \in \text{EXP}_t$
 - If $\alpha, \beta \in \text{EXP}_\tau$, then $(\alpha = \beta) \in \text{EXP}_t$
 - If $\varphi \in \text{EXP}_t$ and $x \in \text{Var}_\tau$, then $(\exists x)(\varphi) \in \text{EXP}_t$
 - If $\varphi \in \text{EXP}_\sigma$ and $x \in \text{Var}_\tau$, then $\lambda x.\varphi \in \text{EXP}_{\langle \tau, \sigma \rangle}$
 - If $\varphi \in \text{EXP}_{\langle \tau, \sigma \rangle}$ and $\phi \in \text{EXP}_\tau$, then $\varphi(\phi) \in \text{EXP}_\sigma$

Other logical operators are abbreviations of the ones that we already have:

- $(\forall x)(\varphi) = \neg(\exists x)(\neg\varphi)$
- $(\phi \vee \varphi) = \neg(\neg\phi \wedge \neg\varphi)$
- $(\phi \rightarrow \varphi) = \neg(\phi \wedge \neg\varphi)$

Table 3.2 gives examples of basic expressions with their types and names. Notice some conventions for variables that will be used throughout. Small letters from the end of the alphabet ($x, y \dots$) are used for variables of type e . The letter e is commonly used for variables of type v . Capital letters are used for variables of type $\langle e, t \rangle, \langle v, t \rangle$ and higher types, even though R is used exclusively for variables of type $\langle e, \langle v, t \rangle \rangle$. In unclear cases I will subscript variables to express which type they belong to. For example, $P_{\langle \langle e, t \rangle, t \rangle}$ is a set of nominal predicates, i.e., a generalized quantifier. Thematic roles are commonly differentiated by various names (agent, theme, recipient etc.) Since the thematic role label is not going to be important in the thesis, I will use Θ with subscripts.

So far, EXP consists of expressions that are commonly assumed in the study of natural language semantics. Apart from these, we also need to specify expressions that are related to the study of pluralities. These are given here:

Table 3.2: Some constants and variables of \mathcal{L}

Expression of \mathcal{L}	Type	Name
PHILIP, HILARY	e	entity
x, y, \dots	e	entity variable
e	e	event variable
BOY	$\langle e, t \rangle$	nominal predicate
P, Q	$\langle e, t \rangle, \dots$	variables of higher types than e
SLEEP, KISS	$\langle v, t \rangle$	verbal predicate
Θ_1, Θ_2	$\langle e, \langle v, t \rangle \rangle$	thematic role
R	$\langle e, \langle v, t \rangle \rangle$	role variable

- (32) a. If $\alpha, \beta \in \text{EXP}_e$, then $\alpha \leq \beta \in \text{EXP}_t$
b. If $\alpha, \beta \in \text{EXP}_e$, then $\downarrow\alpha, \uparrow\alpha, \alpha \oplus \beta \in \text{EXP}_e$
c. If $P \in \text{EXP}_{\langle e, t \rangle}$, then $\sigma(P), \oplus(P) \in \text{EXP}_e$
d. If $\phi \in \text{EXP}_\tau$, where $\tau = \langle e, t \rangle$ or $\tau = \langle v, t \rangle$ or $\tau = \langle e, \langle v, t \rangle \rangle$, then $*\phi \in \text{EXP}_\tau$

In the examples below, I will use two other abbreviations:

- $(\forall x)(x \leq z \rightarrow \varphi) = (\forall x \leq z)(\varphi)$
- $(\exists x)(x \leq z \wedge \varphi) = (\exists x \leq z)(\varphi)$

Model \mathcal{M}

A model contains denotations for every type of expression. The model \mathcal{M} is a pair $\langle D, \mathbb{I} \rangle$, where D is an interpretation domain and \mathbb{I} is an interpretation function. To define D we first need to define the domains of individuals, events, numbers and truth values and complex domains formed out of those.

- (33) a. D_e is the domain of singular and plural individuals with groups discussed in Section 3.2.2
b. D_v is the domain of events discussed in Section 3.2.3
c. $D_t = \{0, 1\}$, the set of truth values
d. $D_{\langle \tau, \sigma \rangle} = \{f \mid f \text{ is a function from } D_\tau \text{ to } D_\sigma\}$ for every complex type $\langle \tau, \sigma \rangle$

We can define D as:

$$(34) \quad D = \bigcup_{\tau \in \text{TYPE}} D_\tau$$

The interpretation function maps every constant in \mathcal{L} to an object in D , such that for every type τ and every constant c of type τ which is in \mathcal{L} we have $\mathbb{I}(c) \in D_\tau$:

$$(35) \quad \mathbb{I} = f : \text{Con} \rightarrow D \text{ s.t. for every } c \in \text{Con}_\tau : \mathbb{I}(c) \in D_\tau$$

Interpretation of \mathcal{L}

Interpretation of \mathcal{L} is given relative to g , a variable assignment. We first recursively define a set of free variables, FV :

- (36) a. If $\xi \in Var$ then $FV(\xi) = \xi$
 b. $FV(\neg\alpha) = FV(\alpha)$
 c. $FV(\alpha \wedge \beta) = FV(\alpha \vee \beta) = FV(\alpha(\beta)) = FV(\alpha) \cup FV(\beta)$
 d. $FV((\exists\xi)(\alpha)) = FV(\lambda\xi.\alpha) = FV(\alpha) \setminus \{\xi\}$

g is a function $Var \rightarrow D$. g is called a proper assignment for α iff all free variables in α are in the domain of g . Assuming a proper assignment for formulas below, we can give recursive interpretation of expressions of \mathcal{L} as in (37)-(38). In (38) I use the notation $g[d/x]$. $g[d/x]$ is a variable assignment g' that assigns d to x and otherwise is identical to g .

(37) Simple expressions:

- a. $|c|^{\mathcal{M},g} = \mathbb{I}(c)$ if c is a constant
 b. $|\xi|^{\mathcal{M},g} = g(\xi)$ if ξ is a variable

(38) Complex expressions:

- a. $|\neg\phi|^{\mathcal{M},g} = 1$ iff $|\phi|^{\mathcal{M},g} = 0$
 b. $|\phi \wedge \varphi|^{\mathcal{M},g} = 1$ iff $|\phi|^{\mathcal{M},g} = 1$ and $|\varphi|^{\mathcal{M},g} = 1$
 c. $|\phi = \varphi|^{\mathcal{M},g} = 1$ iff $|\phi|^{\mathcal{M},g} = |\varphi|^{\mathcal{M},g}$
 d. $|(\exists\xi_\tau)(\phi)|^{\mathcal{M},g} = 1$ iff for some $d \in D_\tau$ $|\phi|^{\mathcal{M},g[d/\xi]} = 1$
 e. $|\lambda\xi_\tau.\phi_\sigma|^{\mathcal{M},g}$ = the function f with the domain D_τ and range D_σ such that for every $d \in D_\tau$: $f(d) = |\phi|^{\mathcal{M},g[d/\xi]}$
 f. $|\alpha(\beta)|^{\mathcal{M},g} = |\alpha|^{\mathcal{M},g}(|\beta|^{\mathcal{M},g})$

The interpretations of expressions related to the study of pluralities are given below (where the interpretations of $\oplus, \leq, \downarrow, \uparrow, *$ are model-independent).

- (39) a. $|\alpha \leq \beta|^{\mathcal{M},g} = 1$ iff $|\alpha|^{\mathcal{M},g} \leq |\beta|^{\mathcal{M},g}$
 b. $|\alpha \oplus \beta|^{\mathcal{M},g} = |\alpha|^{\mathcal{M},g} \oplus |\beta|^{\mathcal{M},g}$
 c. $|\oplus(P)|^{\mathcal{M},g} = \oplus(|P|^{\mathcal{M},g})$
 d. $|\downarrow\alpha|^{\mathcal{M},g} = \downarrow(|\alpha|^{\mathcal{M},g})$, $|\uparrow\alpha|^{\mathcal{M},g} = \uparrow(|\alpha|^{\mathcal{M},g})$
 e. $|\ast P|^{\mathcal{M},g} = \ast(|P|^{\mathcal{M},g})$

I have discussed the interpretation of $\oplus, \leq, \downarrow, \uparrow$ in Section 3.2.2, and have informally shown in Section 3.2.3 what interpretation the \ast operator should get. There, we have seen that \ast should be assimilated to the function that forms sub-i-join semilattices, $[\]$. Strictly speaking, in the semantic framework we are going to use we cannot just equate the star operator with $[\]$. The reason is that predicates and thematic roles are not sets, but (characteristic) functions, as we will see below. To be more precise, \ast should then be defined as follows:

(40) Definition of pluralization

- a. $*P = \lambda x.x \in [\{a : Pa\}]$
- b. $*R = \lambda x\lambda y.\langle x, y \rangle \in [\{\langle a, b \rangle : R(a)(b)\}]$

This is a definition of $*$ as applying to types $\langle v, t \rangle$ and $\langle e, t \rangle$ (40a), $\langle e, \langle v, t \rangle \rangle$ (40b). If we wanted we could further generalize $*$ to apply to any higher type that ends in t (for this, see Lasersohn 1998) but since we will not need it I will not do that here.

Expressions of \mathcal{L} can be converted into other expressions of \mathcal{L} by α -conversion and β -reduction. α -conversion works as re-naming of bound variables, β -reduction as the application of λ -functions to arguments. If φ can be converted to ϕ by application of α -conversions and β -reductions then we say that φ can be reduced to ϕ . I symbolize “ $N[M/y]$ ” as the result of substituting the free occurrences of y in N by M . The unpleasant thing is that M , as an expression, can have free variables which end up being bound in N (for discussion, see for example, Andrews 1986). To avoid this, I define the substitution recursively as follows:

- (41) a. If N is a simple expression then $N[M/y]$ results in substituting y by M in N .
- b. If $N = \neg N', \sigma N', \oplus N', \downarrow N', \uparrow N', *N'$ for N' an arbitrary expression then $N[M/y] = N'[M/y]$
- c. If $N = N' \wedge N'', N' = N'', N'(N'')N' \leq N'', N' \oplus N''$ for N', N'' as arbitrary expressions then $N[M/y] = N'[M/y]$ and $N''[M/y]$
- d. If $N = (\exists \xi)(N'), \lambda \xi.N'$ then $N[M/y] = N'[M/y]$ if $\xi \notin FV(M)$

We can now define α -conversion and β -reduction:

- (42) a. α -conversion:
 $\lambda \xi.\alpha = \lambda \zeta.\alpha[\zeta/\xi]$
- b. β -reduction:
 $\lambda \xi.\alpha(\beta) = \alpha[\beta/\xi]$

Translation of LF constituents into expressions of \mathcal{L}

Now, we are ready to go through the last step, the translation of LF constituents to \mathcal{L} expressions. As in Heim and Kratzer (1998), the translation function $[[\]]$ is given first for terminal nodes, whose translation is specified in the lexicon, and then for non-terminal nodes, whose translations are driven by general principles.

Translation of terminal nodes Table 3.3 gives examples of translations that are specified in the lexicon. *Sleep* and *love* are two examples of verbs. As one can see their translations are of the same type. This might come as a surprise given that the first is intransitive and the second is transitive. But recall that we operate in event semantics, where arguments are introduced separately, by thematic roles (v, V_{Θ_2} etc.), and therefore, the two verbs do not differ in their types but in their selection restrictions (the first one does not select V_{Θ_2} while the second one does). Alternatively, we could

Table 3.3: Translations of some terminal LF nodes into \mathcal{L}

Expression of LF	Translation	Type of \mathcal{L} expression
sleep	$\lambda e. * \text{SLEEP}(e)$	$\langle v, t \rangle$
love	$\lambda e. * \text{LOVE}(e)$	$\langle v, t \rangle$
v	$\lambda x. \lambda e. * \Theta_1(x)(e)$	$\langle e, \langle v, t \rangle \rangle$
V_{Θ_2}	$\lambda x. \lambda e. * \Theta_1(x)(e)$	$\langle e, \langle v, t \rangle \rangle$
Philip	$\lambda P. P(\text{PHILIP})$	$\langle \langle e, t \rangle, t \rangle$
boy	$\lambda x. \text{BOY}(x)$	$\langle e, t \rangle$
boys	$\lambda x. * \text{BOY}(x)$	$\langle e, t \rangle$
three	$\lambda P. \lambda x_e. AT(x) = 3 \wedge P(x)$	$\langle \langle e, t \rangle, \langle e, t \rangle \rangle$
the	$\lambda P \lambda Q. Q(\sigma P)$	$\langle \langle e, t \rangle, \langle \langle e, t \rangle, t \rangle \rangle$
t_n	x_n	e
he_n	x_n	e
*	*	$\langle \langle e, \langle v, t \rangle \rangle, \langle e, \langle v, t \rangle \rangle \rangle$ $\langle \langle e, t \rangle, \langle e, t \rangle \rangle, \langle \langle v, t \rangle, \langle v, t \rangle \rangle$

make thematic roles part of the lexical semantics of the verbs. Nothing I say below hinges on this choice, but the first option gives a simpler definition of A-LIFT, to be defined below.

Table 3.3 also gives examples of translations of lexical items in nominal domains. I opt for treating DPs in a uniform manner so even proper names like Philip are interpreted as generalized quantifiers, i.e., of type $\langle \langle e, t \rangle, t \rangle$. Similarly, the definite article is taken to be a quantifier. *The* picks the maximal element of its complement, given that the maximal element is in the set of the things that the complement of *the* refers to. This is achieved by using σ -operator, known from Link's work (Link, 1998), which is defined as:

$$(43) \quad |\sigma P|^{\mathcal{M},g} = |\oplus P|^{\mathcal{M},g} \text{ if } |\oplus P|^{\mathcal{M},g} \in |P|^{\mathcal{M},g}, \text{ undefined otherwise}$$

To be sure, to make the undefinedness condition fully work we would need to amend rules to allow for presupposition projections. Since this is orthogonal to the issues relevant here, I will ignore presupposition of *the* in later discussion.

Apart from definites and proper names, the table also shows the translation of numerals. They are taken to be NP modifiers (type $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$). Thus, NPs with indefinite numerals are predicates. In order to be able to interpret them as quantifiers, we need to postulate existential raising (Partee, 1987). I do that below. The table also shows the translation of traces (t_n). Their translation is identical to that of pronouns. Finally, $*$ is simply translated as $*$ (whose interpretation we have seen above).

Plurals (see example *boys* in the table) are interpreted as adding $*$ to nominal predicates. Recall that $*$ is interpreted as creating a sublattice from the set to which the original predicate refers. Since $[[\text{boy}]]$ is a set of singular boys, $[[\text{boys}]]$, then, is a set of singular and plural boys. Thus, both plural *and* singular individuals are in the domain of plural DPs. This might seem counterintuitive. Such semantics would suggest that (44) is true when I read one book, which is wrong.

- (44) I read books.

There are good reasons to believe that plural DPs do include singular and plural entities in its extension. One of them is that in downward entailing contexts plural DPs in fact refer to both singular and plural individuals. For example, (45) means that I did not read any single book.

- (45) I did not read books.

One common way to reconcile the paradoxical behavior of plurals is to keep the assumption that plurals refer to both singular and plural entities, and derive their ‘more than one’ reading in non-downward entailing contexts through scalar implicatures. Scalar implicatures, introduced into linguistics by Horn (1972), are conversational implicatures which are based on Grice’s maxim of quantity. The maxim states that speakers should be maximally informative and can be summarized as follows: If two sentences can be ordered on some specified scale of informativeness, one should utter the more informative statement provided one has evidence for that statement to be true. Thus, if someone utters a less informative statement, this is taken to imply that the speaker has no evidence that the stronger statement is true. This is further strengthened into the implication that the speaker believes that the stronger statement is not true. For example, (46b) is more informative than (46a) (since the former entails the latter). If a speaker utters (46a) the hearer can conclude that the speaker believes that (46b) is not true. Thus, (46a) ends up meaning ‘some, but not all, students arrived’. Scalar implicatures modify this viewpoint by adding that when we make the comparison between what was said and what could have been said, the alternatives that are to be considered can only differ by substituting a scalar item of one type for another scalar item of the same type, where what is a scalar item is stipulated. Quantifiers like *some* and *all* are scalar items of the same type.

- (46) a. Some students arrived.
b. All the students arrived.

Plurals and singulars might then be another case of scalar items. Since plurals have no number requirement they are informationally weaker than singulars. Thus, if a speaker uses a plural in her utterance the hearer can conclude that the speaker believes that the same utterance with singular is not true. In conclusion, while plurals are number neutral in semantics, they end up meaning ‘more than one’ when scalar implicatures are calculated. In downward entailing contexts the scale of informativeness is reversed, and therefore, plurals are not informationally weaker than singulars in these contexts, which explains the disappearance of the ‘more than one’ reading in cases like (45). For a detailed analysis of plurals along this line, see especially Zweig (2008) and literature therein (but see Farkas and de Swart 2007 for a differing viewpoint, in which plurals are interpreted already in semantics as referring to more than one object, and singulars are underspecified for number).

I mentioned before that we need existential raising in some cases. I show its translation together with one other type-shifting operator and one sort-shifting operator in Table 3.4. First, let me discuss the existential raising. Nowadays, the existential raising

Table 3.4: Translations of shifting/sorting operators into \mathcal{L}

\mathbb{E}	$\lambda P_{\langle e, t \rangle} . \lambda Q_{\langle e, t \rangle} . (\exists x)(Px \wedge Qx)$
$A - LIFT$	$\lambda Q_{\langle e, \langle v, t \rangle \rangle} . \lambda T_{\langle \langle e, t \rangle, t \rangle} . \lambda e . T(\lambda x . Q(x)(e))$
$\uparrow - LIFT$	$\lambda Q_{\langle \langle e, t \rangle, t \rangle} . \lambda P_{\langle e, t \rangle} . Q(\lambda x . P \uparrow x)$

is commonly taken to be polymorphic, so it can existentially close not only predicates but higher types as well. The generalized version of existential raising can be used when one analyzes indefinites using choice functions (Reinhart, 1997) which lift the predicate to the type of quantifier (Winter, 1997, 2001a) and which are existentially closed at some higher level. Since these issues are not really important for me, I will just use Existential raising as originally proposed in Partee (1987) (notated here as \mathbb{E}). $A - LIFT$ is another type-lifting operation. It is necessary in event semantics if we want to enter DPs in surface positions, and is parallel to Landman's LIFT operation. Finally, $\uparrow - LIFT$ lifts a DP to the group interpretation. Again, $\uparrow - LIFT$ is also assumed in Landman's system. Later on, we will see that we can dispense with the last sort-shifting operator.

Translation of non-terminal nodes The translation of non-terminal nodes is specified by the composition rules in Table 3.5. Most of them follow Heim and Kratzer (1998), but there are a few differences. First, I make the rule of predicate modification apply to event predicates and since it is now similar to Kratzer's event identification (which in turn is similar to theta identification of Higginbotham 1985), I use that term. Event identification is defined as a polymorphic operation so it applies to any two arguments as long as their type ends in $\langle v, t \rangle$. The definition is given here:

(47) Definition of EI :

- a. $[[Y]]_{\tau_1} EI[[Z]]_{\tau_2} = \lambda e . [[Y]](e) \wedge [[Z]](e)$ if $\tau_1 = \tau_2 = \langle v, t \rangle$
- b. $[[Y]]_{\tau_1} EI[[Z]]_{\tau_2} = \lambda P_{\sigma_1} . [[Y]](P) EI[[Z]]$ if $\tau_1 = \langle \sigma_1, \sigma_2 \rangle$ and σ_2 ends in $\langle v, t \rangle$ and $\tau_2 = \langle v, t \rangle$
- c. $[[Y]]_{\tau_1} EI[[Z]]_{\tau_2} = \lambda P_{\sigma_1} . \lambda Q_{\sigma_3} . [[Y]](P) EI[[Z]](Q)$ if $\tau_1 = \langle \sigma_1, \sigma_2 \rangle$ and $\tau_2 = \langle \sigma_3, \sigma_4 \rangle$ and σ_2, σ_4 end in $\langle v, t \rangle$

The second modification of Heim and Kratzer (1998) consistsoff adding Existential Closure (EC), which existentially quantifies over events.⁶ Finally, I add function application with LIFT which makes sure that arguments can be entered into event types.

⁶Alternatively, we could avoid postulating EC and assume instead that the event argument is bound when it is introduced into the derivation, and modifiers can access this event argument through Existential Disclosure (Dekker, 1993).

Table 3.5: Rules for the translation of LF non-terminal nodes

Non-branching nodes (NN)

If X is a non-branching node and Y its daughter, then:

$$[[X]] = [[Y]]$$

Functional application (FA)

If X is a branching node with two daughters, Y and Z , then:

$$[[X]] = [[Y]]([Z])$$

Function application with LIFT (FAL)

If X is a branching node with two daughters, Y and Z , and $[[Y]]$ is of type $\langle\langle e, t \rangle, t\rangle$ and $[[Z]]$ is of type $\langle e, \langle v, t \rangle \rangle$ then:

$$[[X]] = [[A-LIFT]]([Z])([[Y]])$$

Predicate abstraction (PA)

If X is a branching node with two daughters, Y and Z , where Y dominates only a numerical index i , then:

$$[[X]] = \lambda x. [[Z]][x/x_e]$$

Event identification (EI)

If X is a branching node with two daughters, Y and Z and $[[Y]]$ is of type $\langle \tau_1 \rangle$ and $[[Z]]$ is of type $\langle \tau_2 \rangle$ and τ_1 and τ_2 end in $\langle v, t \rangle$ then:

$$[[X]] = [[Y]]EI[[Z]]$$

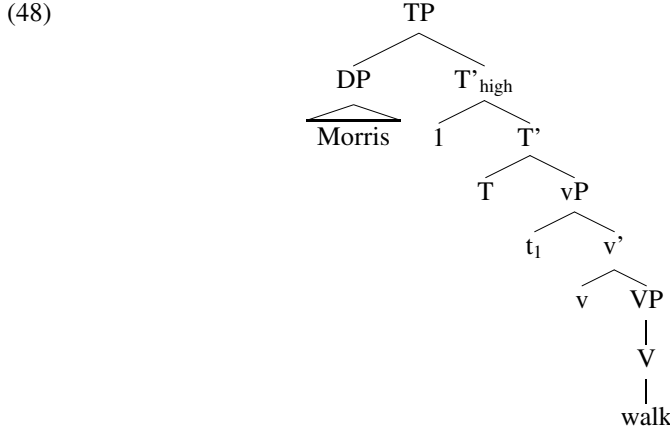
Existential closure (EC)

If X is a node whose translation into \mathcal{L} is of type $\langle v, t \rangle$ then it can be translated as $EC([X])$ where:

$$EC([X]) = (\exists e)([X](e))$$

Examples

A simple sentence like *Morris walks* has the tree structure in (48). I assume that the DP is interpreted in its surface position. In reality, the situation is more complicated because the DP can also be interpreted in its base thematic position (total reconstruction, see Sauerland and Elbourne 2005 and references therein), but that would complicate the discussion on the semantics below and is irrelevant for our purposes here. In the tree I subscript one T' node as T'_{high} since this makes it easy to refer to it later. Node subscripting has no theoretical significance.



The translation of the LF nodes, from bottom to top, is given in (49). Throughout, I ignore the semantic contribution of T.

- (49)
- a. $[[VP]]_{(NN)} [[V]] = \lambda e(*WALK(e))$
 - b. $[[v']]_{(EI)} [[v]] EI[[VP]] = \lambda x. \lambda e(*\Theta_1(x)(e) \wedge *WALK(e))$
 - c. $[[vP]]_{(FA)} [[v']](t_1) = \lambda x. \lambda e(*\Theta_1(x)(e) \wedge *WALK(e))(x_1) = (\beta - reduction) \lambda e(*\Theta_1(x_1)(e) \wedge *WALK(e))$
 - d. $[[T'_{high}]]_{(PA)} \lambda x. (\lambda e. * \Theta_1(x_1)(e) \wedge *WALK(e))[x/x_1] = \lambda x. \lambda e(*\Theta_1(x)(e) \wedge *WALK(e))$
 - e. $[[TP]]_{(FAL)} [[LIFT]]([T'])[[DP]] = \lambda T. \lambda e(T(\lambda x. * \Theta_1(x)(e) \wedge *WALK(e)))(\lambda P. P(MORRIS)) = (\beta - reduction) \lambda e(*\Theta_1(MORRIS)(e) \wedge *WALK(e))$
 - f. $[[TP]]_{(EC)} (\exists e)(* \Theta_1(MORRIS)(e) \wedge *WALK(e))$

Given the interpretation rules for our language \mathcal{L} $|49f|^{\mathcal{M},g}$ is true iff there is an event of Morris walking which obviously corresponds to the meaning of the sentence *Morris walks*. Notice that it is the event of Morris walking, not an event in which Morris walks, or an event which includes Morris walking etc. because we want (49f) to be true of events which include Morris walking and nothing else, i.e., we want minimal

events. If we did not restrict the events to minimal ones, we would not be able to correctly capture the way they can be measured, temporally modified etc. For example, the sentence *Morris walked for seven hours* would be true if Morris walked for thirty minutes, a game of soccer took place for one hour and half and Mary baked a cake for five hours. Obviously, this is not what we want. On the other hand, if we restricted the events that can satisfy (49f) to the ones that include Morris and walking and nothing else, the sentence *Morris walked for seven hours* would be accounted for correctly. To formalize the minimality requirement, I am going to assume the exemplification of events, as in Kratzer's work (see Kratzer 2008 and references therein).

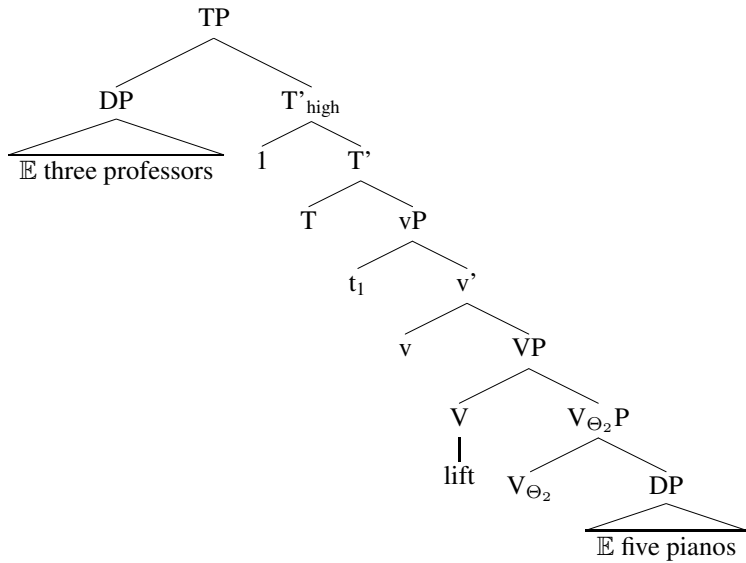
(50) **Exemplification**

An event e exemplifies a proposition p iff whenever there is a part of e in which p is not true then e is a minimal event in which p is true.

Propositions are sets of events. The event e which would exemplify the proposition ' λe . Morris walked in e for seven hours' includes Morris and his walking for seven hours and nothing else.

Let me move to an example with plural arguments, whose LF structure is given below.

- (51) a. Three professors lifted five pianos.
b.



Again, going from bottom up, we get the following translation of the nodes of (51b) (here, I present translations after β – reduction):

- (52) a. $[[E \text{ five pianos}]] = \lambda P((\exists y_e)(|AT(y)| = 5 \wedge *PIANO(y) \wedge P(y)))$
 b. $[[V_{\Theta_2}P]] =_{(FAL)} [[LIFT]]([V_{\Theta_2}])[[E \text{ 5 pianos}]] =$
 $\lambda e((\exists y_e)(|AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(y)(e)))$
 c. $[[VP]] =_{(EI)} \lambda e((\exists y_e)(*LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(y)(e)))$

- d. $[[v']] =_{(EI)} \lambda x. \lambda e ((\exists y_e) (*\Theta_1(x)(e) \wedge *LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(y)(e)))$
- e. $[[vP]] =_{(FA)} \lambda e ((\exists y_e) (*\Theta_1(x_1)(e) \wedge *LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(y)(e)))$
- f. $[[T'_{high}]] =_{(PA)} \lambda x. \lambda e ((\exists y_e) (*\Theta_1(x)(e) \wedge *LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(y)(e)))$
- g. $[[TP]] =_{(FAL+EC)} (\exists e, x_e, y_e) (|AT(x)| = 3 \wedge *PROFESSOR(x) \wedge *\Theta_1(x)(e) \wedge *LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(y)(e))$

The formula in (52g) says that there is an event of three professors lifting five pianos between them. This is true if we have one event which has subevents e_1, e_2, e_3 such that in e_1 one professor lifted two pianos, in e_2 another professor lifted one piano, and in e_3 the third professor lifted two pianos. This is a cumulative reading (Scha, 1981; Krifka, 1989; Schein, 1993).⁷ To derive other readings, it suffices that $\uparrow -LIFT$ applies to the arguments before they combine with thematic roles. This gives us three further readings:

- (53) a. $\uparrow -LIFT(\mathbb{E} 3 \text{ PROFESSORS}) :$
 $(\exists e, x_e, y_e) (|AT(x)| = 3 \wedge *PROFESSOR(x) \wedge *\Theta_1(\uparrow x)(e) \wedge *LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(y)(e))$
- b. $\uparrow -LIFT(\mathbb{E} 5 \text{ PIANOS}) :$
 $(\exists e, x_e, y_e) (|AT(x)| = 3 \wedge *PROFESSOR(x) \wedge *\Theta_1(x)(e) \wedge *LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(\uparrow y)(e))$
- c. $\uparrow -LIFT(\mathbb{E} 3 \text{ PROFESSORS}), \uparrow -LIFT(E 5 \text{ PIANOS}) :$
 $(\exists e, x_e, y_e) (|AT(x)| = 3 \wedge *PROFESSOR(x) \wedge *\Theta_1(\uparrow x)(e) \wedge *LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge *\Theta_2(\uparrow y)(e))$

(53a) says that three professors, as a group, lifted five pianos one by one. (53b) says that three professors lifted one by one five pianos. Finally, (53c) says that three professors, as a group, lifted five pianos, as one group. These are collective readings.

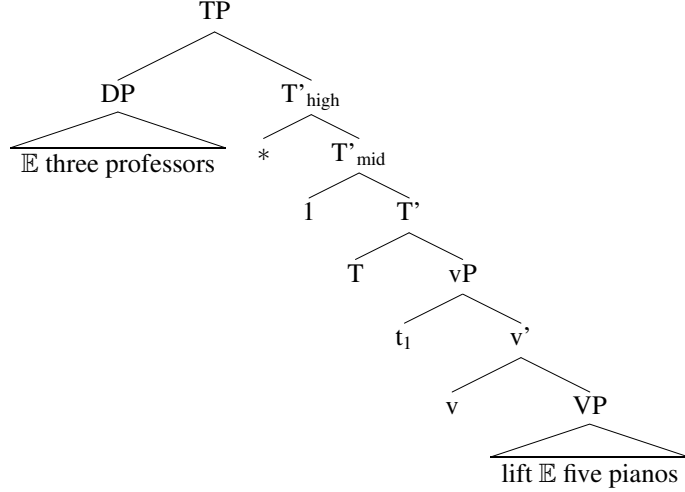
Right now, these are all the readings we get. In particular we cannot get distributive readings with covariation. (51a) cannot in our system mean ‘three professors each lifted five pianos and fifteen pianos were lifted in total’ (surface scope distributive reading with covariation). Landman (2000) gets this reading by adding scope mechanisms into his system. He assumes two scope mechanisms. One of them is similar to Montagovian Quantifying-In and it is responsible for scope readings of quantifiers like *every NP*. I will discuss that later when I turn to the semantics of distributive quantifiers. The other scope mechanism has a built-in distributivity over the whole predicate. As Landman notes, the second scope mechanism is simply a combination of Quantifying-In and insertion of $*$.

⁷(52g) does not exclude the situation where, for example, five professors lifted ten pianos. To have (52g) describe the maximal situation we would have to take scalar implicatures introduced by numerals into account. This is done in Chapter 7 of Landman (2000). Since it is orthogonal to the point of this thesis, I will ignore this issue here.

To simplify Landman's system, I make use of this observation. The distributive reading, then, is a case of scoping over the existential closure where $*$ applies to the predicate created by Predicate Abstraction.⁸ Thus, the distributive reading has a structure similar to the one above, the only difference is the introduction of $*$:

- (54) a. Three professors lifted five pianos.

b.



The only requirement on EC (Existential Closure) is that it applies to a node of type $\langle v, t \rangle$. Thus, EC does not need to wait till the subject in Spec,TP applies to the lifted event predicate. EC can apply at T' after which PA takes place. Thus, we have:

- (55) a. $[[T'_{mid}]] = (PA)$
 $\lambda x. (\exists e) (*\Theta_1(x)(e) \wedge *LIFT(e) \wedge *\Theta_2(5 \text{ PIANOS})(e))$
- b. $[[T'_{high}]] = (FA)$
 $\lambda x. x \in [\{a : (\exists e) (*\Theta_1(a)(e) \wedge *LIFT(e) \wedge *\Theta_2(5 \text{ PIANOS})(e))\}]$

$[[T'_{high}]]$ is true if we can split x into parts and each part lifted five pianos. Thus, we have the following equivalence:

$$(56) \quad \lambda x. x \in [a : Pa] \Leftrightarrow \lambda x. (\exists a_1, \dots, a_n) (a_1 \oplus \dots \oplus a_n = x \wedge Pa_1 \wedge \dots \wedge Pa_n)$$

Suppose we consider the case in which x is split into individual parts. In that case we have the following equivalence:

$$(57) \quad \lambda x. x \in [a : a \in AT \wedge Pa] \Leftrightarrow \lambda x. (\forall a \leq x \wedge a \in AT) (Pa)$$

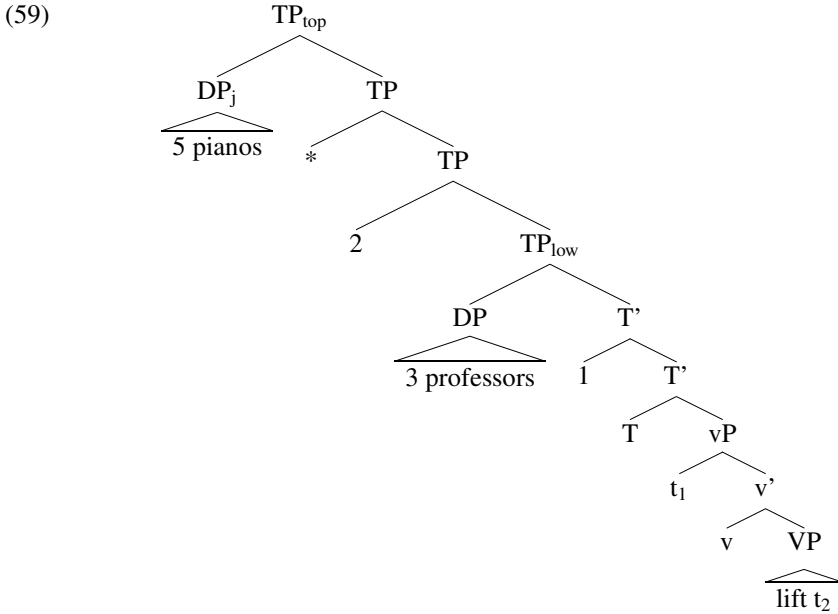
Thus, if we assume that x is split into individuals in (55b) we get the reading that each individual in x lifted five pianos. If we apply the subject DP to (55b) we get:

⁸See Sauerland (1998), Beck and Sauerland (2000), Barker (2007), among others, for various arguments that the semantic argument created by predicate abstraction can be available as an argument of various scoping elements, one of which is the $*$ operator.

- (58) $[[TP]] =_{(FA)}$
 $(\exists x)(|AT(x)| = 3 \wedge *PROFESSOR(x) \wedge (\forall a \leq x \wedge a \in AT)(\exists e, y)(* \Theta_1(a)(e) \wedge$
 $*LIFT(e) \wedge |AT(y)| = 5 \wedge *PIANO(y) \wedge * \Theta_2(y)(e))$

This is true if each professor lifted five pianos, and furthermore, the lifted pianos can differ between the professors. Thus, (58) can be interpreted as the distributive reading with covariation where the subject distributes over the object.

To get the other distributive reading ('five pianos were each lifted by three professors', which would be true if up to fifteen professors did lifting) I assume quantifier raising (QR) of the object (Heim and Kratzer, 1998). QR targets DPs. The DPs that undergo QR are dislocated at LF. In case of the inverse scope reading here the object moves from its base position and is adjoined to TP, after Predicate Abstraction and insertion of *:



The tree structure in (59) is interpreted as shown in (60). I present only two steps in the derivation. First, TP_{low} at which point EC applies. After that Predicate abstraction and insertion of * makes sure that the QR-ed object is going to distribute over the rest of the clause, (60b), rewritten as (60c). In the formulas to come, I will sometimes use universal quantification rather than sets and sublattices since I find the latter formula harder to read. In other words, when talking about distributive readings, I sometimes use (60c) rather than (60b).

- (60) a. $[[TP_{low}]] = (\exists e, x)(|AT(x)| = 3 \wedge *PROFESSOR(x) \wedge * \Theta_1(x)(e) \wedge$
 $*LIFT(e) \wedge * \Theta_2(x_2)(e))$
 b. $[[TP_{top}]] = (\exists y)(y = 5 \text{ PIANOS} \wedge$
 $y \in [\{b : (\exists e, x)(|AT(x)| = 3 \wedge *PROFESSOR(x) \wedge * \Theta_1(x)(e) \wedge$
 $*LIFT(e) \wedge * \Theta_2(b)(e))\}])$

- c. $[[\text{TP}_{\text{top}}]] = (\exists y)(y = 5 \text{ PIANOS} \wedge (\forall b \leq y \wedge b \in AT)(\exists e, x)(|AT(x)| = 3 \wedge * \text{PROFESSOR}(x) \wedge * \Theta_1(x)(e) \wedge * \text{LIFT}(e) \wedge * \Theta_2(b)(e)))$

It has been noted that inverse scope distributive readings are very marginal (Gil (1982), and my discussion in Section 2.2.2). If we want to block this reading we could stipulate in our grammar that QR is impossible in this case. Alternatively, its marginal status might follow from processing considerations discussed in VanLehn (1978), Tunstall (1998) and Pykkänen and McElree (2006), who assume that the reading is dispreferred because it requires an extra QR operation (see also Reinhart 2006 for arguments why QR operations resulting in inverse scope are costly).

If we do not block the inverse scope distributive reading, we end up with the total number of eight LF representations for transitive sentences with two plural arguments. These are summarized in (61): four representations in which the arguments are not scopally ordered (61a-d), two representations in which the subject distributes over the rest of the sentence, including the object (61e-f), and two representations in which the object distributes over the rest of the sentence, including the subject (61g-h). The last four representations can, but do not have to, give rise to distributive readings with covariation.

- (61) a. $\text{DP}_{\text{subj}} - \text{DP}_{\text{obj}}$
 b. $\uparrow(\text{DP}_{\text{subj}}) - \text{DP}_{\text{obj}}$
 c. $\text{DP}_{\text{subj}} - \uparrow(\text{DP}_{\text{obj}})$
 d. $\uparrow(\text{DP}_{\text{subj}}) - \uparrow(\text{DP}_{\text{obj}})$
 e. $\text{DP}_{\text{subj}} \in [\{x : \dots \text{DP}_{\text{obj}} \dots\}]$
 f. $\text{DP}_{\text{subj}} \in [\{x : \dots \uparrow \text{DP}_{\text{obj}} \dots\}]$
 g. $\text{DP}_{\text{obj}} \in [\{x : \dots \text{DP}_{\text{subj}} \dots\}]$
 h. $\text{DP}_{\text{obj}} \in [\{x : \dots \uparrow \text{DP}_{\text{subj}} \dots\}]$

I will now modify Landman's system by combining it with covers, which will cut down the number of representations.

3.2.5 Modification of Landman (2000): Adding covers

As the language of pluralities and events is at this point, it follows Landman's system very closely. I will modify it now by adding covers, following Schwarzschild's work (Schwarzschild, 1996). Covers are also considered in Landman's book (Chapter six) but mainly to compare his system to Schwarzschild's, and are not used otherwise. Using covers is not an exclusive property of Schwarzschild (1996), they are commonly considered in connection with the semantics of pluralities (see, for instance, Gillon 1987, Verkuyl and van der Does 1996, Beck 1999, among others), even though my motivation of using covers will mainly draw on the discussion in Schwarzschild (1996).

Consider the following situation (from Schwarzschild 1996, p. 67):

- (62) Two merchants are attempting to price some vegetables. The vegetables are sitting before the merchant, piled up in several baskets. To determine their price, the vegetables need to be weighed. Unfortunately, our merchants do not have an appropriate scale. Their grey retail scale is very fine and is meant to weigh only a few vegetables in a time. Their black wholesale scale is coarse, meant to weigh small truckloads. Realizing this, one of the merchants truthfully says (...):
 The vegetables are too heavy for the grey scale.
 The vegetables are too light for the black scale.

The closest we can get to replicate the meaning of the two utterances in (62) is to analyze these sentences by pluralizing on the event predicate and theta roles.

- (63) a. $(\exists e)(*\Theta_1(\sigma(*VEGETABLE)(e)) \wedge *too\ heavy\ for\ the\ grey\ scale(e))$
 b. $(\exists e)(*\Theta_1(\sigma(*VEGETABLE)(e)) \wedge *too\ light\ for\ the\ black\ scale(e))$

As I said before basic (=non-pluralized) predicates are true only of individuals. But because of that (63a) and (63b) is true only if individual pieces of vegetables are light or heavy. In our scenario above this would make (63a) false and (63b) true. But intuitively both utterances are true. If we, on the other hand, lifted the roles' arguments to groups we would get collective readings:

- (64) a. $(\exists e)(*\Theta_1(\uparrow\sigma(*VEGETABLE)(e)) \wedge *too\ heavy\ for\ the\ grey\ scale(e))$
 b. $(\exists e)(*\Theta_1(\uparrow\sigma(*VEGETABLE)(e)) \wedge *too\ light\ for\ the\ black\ scale(e))$

But now, (64a) is true and (64b) is false (because all the vegetables summed up together are not too light for the black scale). This shows that we are in need of intermediate readings in which some subparts of *the vegetables* are relevant.

A common way to improve on this situation is to introduce covers. A cover of X is a set of sets such that X is a subset of the union of the sets. In our part-of structure, we can mimic the work of covers by introducing *Cov*, a set of group and individual atoms and require it to cover plural individuals:

- (65) *Cov* covers a plural individual x iff every atom of x is part of some $\downarrow a$, such that $a \in Cov$

Cov restricts *. We define a new pluralizer, C_* on predicates and thematic roles as follows.

- (66) Definition of pluralization on predicates and thematic roles with Covers
 a. $C_*P = \lambda x.x \in [\{\downarrow a : Pa \wedge a \in Cov\}]$, or, alternatively,
 $C_*P = \lambda x.(\forall a \in Cov)(\downarrow a \leq x \rightarrow Pa)$
 b. $C_*R = \lambda x\lambda e.\langle x, e \rangle \in [\{\langle \downarrow a, e' \rangle : R(a)(e') \wedge a \in Cov\}]$

Notice that the pluralizer on predicates is defined twice, shown in 66a). The two definitions are not completely equivalent. For instance, suppose that x would consist of four parts, a_1, \dots, a_4 , and that every pair's group is in the cover (that is, $\uparrow(a_1 \oplus a_2)$ is in the cover, $\uparrow(a_1 \oplus a_3)$ is in the cover and so on). In this case, the first definition of C_*

makes C_P true if, for example, the predicate P is true for $\uparrow(a_1 \oplus a_2)$ and $\uparrow(a_3 \oplus a_4)$ but the second definition makes C_P false. That is, the second definition is stronger than the first one. In the cases I am going to discuss this difference will be irrelevant so I will use the two definitions interchangeably.

Let us see how this works on the previous example, which should now make use of pluralizers restricted to covers:

- (67) a. $(\exists e)(C_*\Theta_1(\sigma(*\text{VEGETABLE}))(e)) \wedge *_{\text{too heavy for the grey scale}}(e)$
 b. $(\exists e)(C_*\Theta_1(\sigma(*\text{VEGETABLE}))(e)) \wedge *_{\text{too light for the black scale}}(e)$

Assume the merchants have four vegetables: an aubergine, a carrot, a paprika, and a zucchini, and the first two are in one basket, and the last two are in another basket. The plural individual $V = a \oplus c \oplus p \oplus z$ is covered by Cov iff every atom of V is part of some $\downarrow a$, such that $a \in Cov$. This is true if, for example $Cov = \{a, c, p, z\}$. Employing this Cov would lead to the same interpretation as we get with $*$. That is, each vegetable is too light/heavy. Assume instead $Cov = \{\uparrow(a \oplus c), \uparrow(p \oplus z)\}$. Now, (67a) and (67b) say that the aubergine and carrot together, and the paprika and the zucchini together, are heavy/light. Since this is the distribution of vegetables into baskets, we correctly get the interpretation we wanted.

I assume that Cov restricts pluralization only in the verbal domain. Pluralization of nouns is not restricted by covers. Besides, I do not assume that the verb itself (set of events) is, when pluralized, restricted by covers. We could introduce Cov in this case but since the domain of events has no groups, the introduction of covers would be vacuous. For the same reason, only the individual argument in thematic roles is restricted by covers.

Notice that using covers simplifies constructions with $*$ operators. In particular, we now have the following equivalences:

- (68) a. $C_*C_*\Theta(x)(e) \Leftrightarrow C_*\Theta(x)(e)$
 b. $C_*(\lambda x.(\exists e) \dots C_*\Theta(x)(e) \dots) \Leftrightarrow C_*(\lambda x.(\exists e) \dots \Theta(x)(e) \dots)$

This is because C_* gives us a set of atomic entities, either individual or group atoms because Cov is a set of atoms. If some pluralized thematic role or a predicate applies to an atomic entity, pluralization is vacuous.

Apart from covers, Schwarzschild (1996) also postulates paired covers for relations. These capture properties of cumulative readings (see (52g)). For example, (69) is judged true in the situation depicted in Figure 3.1. Scha (1981) who originally discusses this case, derives this meaning by adding a meaning postulate to the grammar, which states that ‘As are parallel to Bs’ iff every atom of A is parallel to some atom of B, and every atom of B is parallel to some atom of A.

- (69) The sides of the upper rectangle are parallel to the sides of the lower rectangle.

In the system that I developed here, we do not need meaning postulates to derive this reading because it falls out from pluralization. The cumulative reading is captured in this formula:

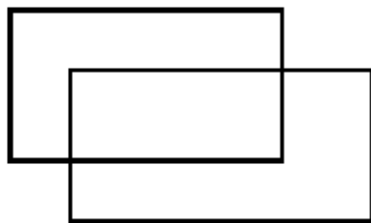


Figure 3.1: Situation in which (69) is judged as true

$$(70) \quad (\exists e) \left(\begin{array}{l} *PARALLEL(e) \wedge \\ \wedge^{C_1} * \Theta_1(\sigma\{x : x \text{ ONE SIDE OF THE UPPER RECTANGLE}\})(e) \wedge \\ \wedge^{C_2} * \Theta_2(\sigma\{x : x \text{ ONE SIDE OF THE LOWER RECTANGLE}\})(e) \end{array} \right)$$

(70) is true if *parallel* holds of some subevents of e , where in each subevent one of the lines of the upper rectangle is the agent, and one of the lines of the lower rectangle is the theme, i.e., one of the lines in the upper rectangle is parallel to one of the lines in the lower rectangle. This is true in Figure 3.1. More generally, we only require that every relevant subpart (every subpart that is in *Cov*) of some plural DP be an argument of its non-pluralized thematic role. In case of transitive relations like *parallel* this means that every relevant subpart of one argument is related by the relation to some relevant subpart of the other argument. Therefore, the truth conditions are the same in this case as the ones achieved by Scha's meaning postulates.

Schwarzschild (1996) shows that Scha's meaning postulates overgenerate. (71) seems false in (3.2), even though it is true in this situation that every dashed line is parallel to some continuous line, and every continuous line is parallel to some dashed line. Similar problems have been discussed in Beck (1999), Landman (2000) and others.

(71) The dashed lines run parallel to the continuous lines.

Schwarzschild accounts for the difference between (71) and (69) by postulating paired covers. These pair together subparts of the two arguments, and the relation must hold between the members of each pair. Schwarzschild suggests that pairing lines which are parallel is more natural in Figure 3.1 than in Figure 3.2, which explains the difference in judgements.

I believe that paired covers are not needed in the system here. The reason is that their work might be taken over by events. In (69) we need subevents in which a line of one rectangle is parallel to a line of the other rectangle (70). The split of the event that achieves this is quite natural. The subevents are convex spaces in which the two closest lines are put together. This should be obvious from the first picture in Figure 3.3. On the other hand, in the case of (71) the split of the event that puts together



Figure 3.2: Situation in which (71) is judged as false

dashed and continuous lines which are parallel is much more complicated. Subevents here are not convex spaces and they put together lines which are not closest to each other. This should be clear from the second picture in Figure 3.3. In particular, e_1 is a subevent in which a dashed and a continuous line are put together to the exclusion of a line in between them.

This explanation, similar to Kratzer (2008), assumes that events are constrained in various ways. The constraints should only hold for atomic events because ideally, we want to leave the structure intact, i.e., keep treating it as a part-of structure. This makes the domain of events quite similar to the domain of entities. In the latter case, what is an atom (be it an individual or a group) is restricted in various ways. But any atoms can be joined together, and thus, there are no restrictions on *plural* elements apart from the fact that it must be built up from atoms. Similar restrictions then should hold of events. While the discussion of what these restrictions are goes way beyond the topic of my thesis, I would like to point out at least some reasonable properties, borrowing insights from psychology.

In 1923, Max Wertheimer brought forth the question of the organization of perceptual information. Why is it that we perceive the visual world surrounding us as being occupied by discrete objects, when the retinal image is just a juxtaposition of elementary physical sensations? The study of the perceptual organization fell well within the agenda of the Gestalt theory, which assumed that many phenomena, and visual perception in particular, are not just summation of parts, but are determined by the intrinsic nature of the whole (Ellis, 1938, Part 1). In his seminal paper (Wertheimer, 1923), Wertheimer discovered various principles that organize the visual stimulation into groups and objects based on properties in the image. This leads to the perception of categories that are nowhere present in the original. He studied lines of dots organized in various ways and showed that our perception of how they are grouped follows some simple principles. Figure 3.4 demonstrates three of those principles. Line A shows a line of dots. There is no perceived grouping within these dots, except the line of dots as a whole. In Line B the spacing between the dots is changed and this leads to a change in the perception. The dots that are close to each other are perceived

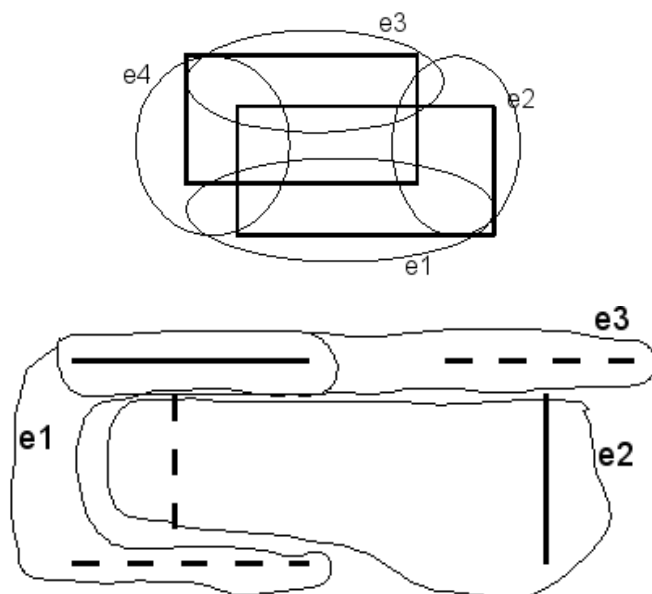


Figure 3.3: Split of events for (69) and (71)

as one unit. This shows the principle of *proximity* at work: all else being equal, the closest elements tend to be grouped together. Apart from the principle of proximity, Wertheimer discovered other principles, of which I represent here only the principle of *similarity*: all else being equal, the most similar elements tend to be grouped together. This is shown in Lines C and D where the elements that share color (Line C) or shape (Line D) are perceived as one unit. Other properties, like the same orientation, or size, can be equally well used by the principle of similarity. Line E shows another principle, discussed by Palmer (1992). Here, dots that are in an ellipse are perceived as one unit. This is dubbed as the principle of *common region*: all else being equal, elements that are located within the same closed region of space will be grouped together.

There are further principles, uncovered by Wertheimer (1923) and others. A very good overview of these can be found in Chapter 6 of Palmer (1999) (see also Palmer and Rock 1994).

The principles of grouping have been applied outside of the study of vision, for example, in auditory perception, or in social psychology, in studies of social entities (Campbell, 1958; Hamilton and Sherman, 1996). The last application is especially relevant here since it shows that our perception of social grouping follows the principles discussed above. It is quite reasonable to assume that groups in the interpretation domain of pluralities are not really different from groups studied in social psychology. In fact, it could be shown that in example after example which Schwarzschild discusses to motivate the existence of paired covers, the principles of Gestalt psychology for

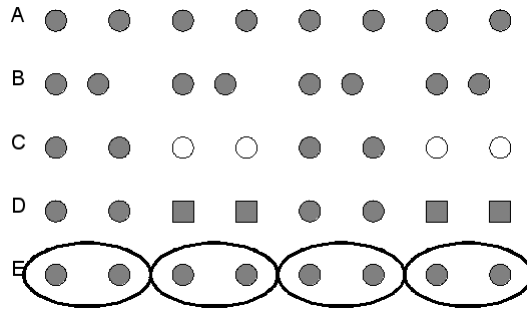


Figure 3.4: Demonstration of grouping principles

grouping are exploited. For instance, in (62) the context forces some vegetables to be closer to each other than other vegetables (the principle of proximity), and some are located within the same region (basket), separately from the others (the principle of common region).

Atomic events, I assume, should be similar to groups in the following respect: they include entities that are, according to the Gestalt principles, perceived as one group. Again, these principles might be more clearly seen at work when we draw pictures (like we did in Figures 3.1 and 3.2) but they apply outside the spatial domain. Consider the following situation, from Kratzer (2003). Suppose you and I each have a donkey and a cat, and these are all the animals we have. My donkey looks just like your donkey, and my cat looks just like your cat. In this situation, (72a) is true and (72b) is false.

- (72) a. My animals look like your animals.
 b. My animals look different from your animals.

At this point it should be clear why the two sentences differ. In (72a) we have a plural event which can be split into two subevents which *look alike* has in its extension. In one of the subevents we compare my cat to your cat, in the other subevent we compare my donkey to your donkey, and since they look alike the sentence is true. In (72b) we again need two subevents. But this time, in one of them we have my donkey and your cat, and in the other one we have my cat and your donkey. This latter split violates the principle of similarity, discussed above.

Notice that Scha (1981) can account for the difference between (72a) and (72b) too. In his analysis, it suffices to say that a meaning postulate which derives cumulative readings can be bundled with *look like* but cannot be bundled with *look different*. But the approach here goes one step further. Not only does it capture the difference between the two examples, it also explains why they should differ, based on assumptions that are independent of these two examples. In this respect, this approach might also improve on Schwarzschild's work, who never offers any principles which guide selection of covers. To be sure, what I am saying here about group atoms and atomic

events is rather a suggestion for future research than a serious analysis, and there are many issues one would have to look at to fully establish that this approach is right. I only want to clear up one thing that might seem problematic at this point. Consider the following example:

(73) My cat is different from your donkey.

(73) is true. But how come it is true? Does not (73) violate the principle of similarity in the precisely same way as (72b)? The answer is: It does but that does not matter in evaluating the sentence. To see why let us go back to Figure (3.4) and compare Lines A and C. In Line A there is no grouping perceived. In Line C, we perceive grouping by colour. But why do we not perceive grouping in Line A? For instance, the first two dots are exactly the same in both lines, and still, they are perceived as one group only in Line C. The reason for this difference is that the Gestalt principles are not absolute principles that each group must satisfy. Rather, they are constraints that guide the parsing of the spatial domain. If one ends up with a unique parse of the image that satisfies most Gestalt principles then the perception of groups arise. If there is no unique parse, no groups arise. So, perception of groups in Line A is missing not because we cannot find objects that satisfy the Gestalt principles but because we find too many of them. In Line C, on the other hand, the principle of similarity starts playing a role only because some of the parses do *not* satisfy it. Thus, we are looking for the optimal parse, which tries to maximize the satisfaction of the principles, and if we find one we can perceive a grouping. This is the reasoning usually implemented in formalizations of the parsing of images. I will not go into any such formal framework here, but the interested reader might take a look at Feldman (1999), and especially Feldman (1997) and literature therein for one particular formalization, called Minimal Model theory.

Why then do (72b) and (73) differ? In the first case, we have a plural event and are looking for atomic events which together would make the sentence true. But we cannot pick e_1 consisting of my donkey and your cat and e_2 consisting of my cat and your donkey. The reason is not that the objects in e_1 and e_2 violate the principle of similarity but that there is a different parse of the plural event which does not violate that principle. And therefore, the subevents must be members of the latter, optimal, parse. In (73) my cat and your donkey can be grouped together because this is the only parse that can be considered.⁹

By now, it should be clear why (71) is false in picture 3.2. We cannot split the whole situation into e_1, e_2, e_3 . In particular, e_1 cannot be considered because it violates the principle of proximity, and there is a different parse which does not violate

⁹Thus, this approach avoids a pitfall into which Kratzer's (2003) analysis falls, I believe. In her approach atomic events must consist of substantive pluralities. What is a substantive plurality can change from one context to another. In the context given for (72a) and (72b), Kratzer assumes that substantive pluralities correspond to kinds, and this correctly captures why (72a), but not (72b), is true. But then, in the same context, it is not clear how we can ever evaluate (73) as true. Here, the event consists of a donkey and a cat. These are two different kinds, and therefore, by stipulation, not a substantive plurality in the context. Therefore, we should not even be able to consider an event which consists of them. Notice that we cannot blame the difference between the two cases on context since it is kept exactly the same.

this principle. But of course, the latter parse puts together two lines that are perpendicular, and this makes the sentence false. Furthermore, the original parse is not convex. Convexity is commonly assumed to be important in the perception of objects (Palmer, 1999) but it probably plays a role in the perception of groups too (see, for example, its importance in the study of attention, which is closely related to grouping, as discussed in Scholl 2001, Cave and Bichot 1999).

Number of representations, branching and cumulative readings and the Strongest Meaning Hypothesis

In Landman's system modified by the use of covers, we have only three LF representations for transitive sentences with two plural arguments.¹⁰ These are summarized in (74). In the first one the arguments are not scopally ordered. In the next one the subject distributes over the rest of the sentence, including the object (74b), and in the last one the object distributes over the rest of the sentence, including the subject (74c). Notice that this simplifies the original system of Landman, where we had eight representations for the same sentence type.

- (74) a. $DP_{\text{subj}}-DP_{\text{obj}}$
 b. $DP_{\text{subj}} \in \{ \{x : \dots DP_{\text{obj}} \dots \wedge x \in Cov \} \}$
 c. $DP_{\text{obj}} \in \{ \{x : \dots DP_{\text{subj}} \dots \wedge x \in Cov \} \}$

The main simplification comes from the fact that collective readings, derived by using $\uparrow -LIFT$ in the original system, are now just extreme cases of (74a). (74a) can be interpreted in various ways, depending on the values of *Cov* in thematic roles:

- *Cov* consists of groups. This derives the collective reading.
- *Cov* consists of atomic individuals, which derives a cumulative reading.
- *Cov* consists of subgroups of plurals. This derives cumulative readings where subgroups, not individuals, are related to each other. This reading appears when the relevance of subgroups is supported in the context. To this date it is most extensively discussed in Schwarzschild (1996), and I discussed it in the previous subsection.

To the readings that we have so far, I want to add one more that I believe (74a) derives. This is a branching reading. For example, in a sentence *Morris and Philip saw two girls*, the branching reading could be paraphrased as 'there are two girls and each of Morris and Philip saw each of the two girls'. In Gil (1982) this reading was accepted by everyone (for the sentence *Three boys saw two girls*), and in fact, was preferred over the cumulative readings. I am going to show that the branching reading follows from (74a) and is preferred over the cumulative reading because of the Strongest Meaning

¹⁰Thus, using covers has a nice side-effect of cutting down the number of meanings in sentences with plural DPs. In fact, Verkuyt and van der Does (1996) discuss a system which, thanks to using covers (or, to be more precise, partitions), narrows down the number of readings in sentences like 'five boys lifted three tables' to two (one if we disregard the possibility of inverse scope).

Hypothesis, SMH, see Dalrymple et al. (1998). This is similar to the analysis of Beck and Sauerland (2000); Beck (2001); Winter (2001b).

The preference for the strongest possible interpretation has been studied in the theory of plurals and especially in the domain of reciprocity. The investigation in the latter domain led to the postulation of various interpretations of reciprocals which are related by entailment, and the SMH which states that one chooses the strongest meaning of the reciprocal that is compatible with the context.

Beck (2001) argues that the SMH causes one to prefer distributive readings over cumulative readings whenever the distributive reading is compatible with the context and world knowledge. Her argument is based on the following contrast:

- (75) a. The two women know the two men.
b. The two women married the two men.

As she notes, the first sentence is likely to be judged false if one of the women does not know one of the men. On the other hand, the second sentence is fine if each woman married only one man. This situation is argued to follow from the SMH. The distributive reading here entails the cumulative reading. Because of the SMH the first reading is chosen whenever it is compatible with our world knowledge, which it is with the predicate *to know*. On the other hand, since one usually marries only one man, the second sentence falls back on the cumulative reading. A similar argument is made in Winter (2001b) for the pair (76). While the first sentence seems false or odd if Mary saw only John and Sue saw Bill and George, for example, the second sentence is completely natural if Mary gave birth to John, and Sue gave birth to Bill and George (cumulative reading). The only difference between Beck (2001) and Winter (2001b) is that Winter claims that the SMH does not let us choose between distributive and cumulative readings but it can weaken distributive readings in cases where this is forced by our world knowledge (as is the case with the relation *giving birth*).

- (76) a. Mary and Sue saw John, Bill and George.
b. Mary and Sue gave birth to John, Bill and George.

When all these accounts try to show that the SMH makes us choose a distributive reading over a cumulative reading (or a weakened distributive reading, as is the name used in Winter 2001b) they use data where one cannot differentiate between distributive and branching readings because both plural arguments in their example are definite. In other words, even if one argument scopes and distributes over the other, we cannot notice it in the interpretation since it cannot induce any variation on it. To see this consider (76a). The distributive interpretation of the sentence is given in (77). The object is in the scope of the subject. But because the object consists of a conjunction of proper names, the conjunction is the same for every x . I will call this kind of interpretation *distributive reading with no variation*.

- (77) a. $[Mary\ and\ Sue] * \lambda x [John, Bill\ and\ George] * \lambda y. x\ saw\ y$
b. $(\forall x \leq m \oplus s)(\forall y \leq j \oplus b \oplus g)(x\ saw\ y)$

We could distinguish the distributive and branching reading if we used indefinite arguments. This is precisely what Gil (1982) did. He used the relation *to see* with two indefinite arguments and asked people to judge the acceptability of surface scope and inverse scope distributive readings, as well as the branching reading and the cumulative reading. It turned out that the distributive readings were dispreferred while the branching reading was fully accepted, and preferred over the cumulative reading. For a moment I will put the distributive readings aside. I will offer an analysis independent of the SMH on why they are degraded. This leaves us with cumulative and branching readings. Obviously, given the SMH, we expect the latter to be preferred over the former since the latter entails the former and thus, it is stronger in sense of Dalrymple et al. (1998).

How do we derive the branching reading? The first option that probably comes to mind is to add a new silent operator or new scope mechanism (Sher, 1990; Landman, 2000) and assume that the SMH makes us choose between this and the operation that derives the cumulative reading. But this approach is problematic because of the data in Beghelli et al. (1997). Their argument goes as follows. The branching reading already falls out in two ways. Either it is the case of a distributive reading with no variation (the dependent argument is definite, specific indefinite etc.) or it is compatible with the LF interpretation that derives cumulative readings. As Beghelli et al. show when we can independently establish that the cumulative reading and the distributive reading with no variation are impossible, the branching reading disappears. Thus, either the branching reading is derived by a special mechanism which is for some independent reason blocked when the cumulative reading and distributive reading with no variation are blocked, or the branching reading is an epiphenomenon.¹¹ I will assume that the branching reading is an epiphenomenon. This still leaves two options. First, the branching reading is a case of a distributive reading with no variation. Alternatively, it appears due to the same process that derives the cumulative reading. If Beghelli et al. (1997) are correct we need both options to derive the branching reading. Thus, the branching reading of (78a) would be one of the interpretations of (78b):

- (78) a. Morris and Philip saw two girls.
 b. $(\exists e)(*see(e) \wedge C_*\Theta_1(m \oplus p)(e) \wedge C_*\Theta_2(TWO\ GIRLS)(e))$

This, however, faces an immediate problem. The event in which the branching reading of (78a) holds is never going to be the event which makes (78b) true. Suppose there is an event e' in which the cumulative reading of (78a) is true, so Morris saw a girl and Philip saw a girl in e' . Furthermore, there is an event e in which the branching reading is true, so Morris saw two girls in e and Philip saw the same two girls in e .

¹¹Landman (2000) claims that the branching reading is needed independently of the cumulative reading because of the interpretation one gets when one considers the interaction of two plural arguments with negative quantifiers. Landman builds up a scope mechanism that can derive this reading. As far as I can see, there is no reason why this mechanism should be blocked when the cumulative reading and the distributive reading with no variation are blocked, so this way of deriving the branching reading fails the observation of Beghelli et al. (1997). I am going to argue in Chapter 4 that the reading he considers as an argument for the (independent derivation of the) branching reading follows from the cumulative reading when one bears in mind that predicates and relations place a requirement of indivisibility on their plural arguments, as extensively discussed in Löbner (2000).

And finally, there are no other events in which the cumulative/branching reading of (78a) holds. Now, recall that events which make a proposition true must exemplify it. The proposition here is the following (characteristic function of the) set of events:

$$(79) \quad \lambda e. *see(e) \wedge C_{*\Theta_1}(m \oplus p)(e) \wedge C_{*\Theta_2}(\text{TWO GIRLS})(e)$$

Exemplification as repeated from above:

(80) **Exemplification**

An event e exemplifies a proposition p iff whenever there is a part of e in which p is not true then e is a minimal event in which p is true.

The problem is that e , the branching reading event, must have e' as its part. Since e' is part of e and e' makes the proposition (79) true and it is a minimal event which makes the proposition true e' exemplifies the event. This means that (81) repeated from above, is satisfied for e' but not for e .

$$(81) \quad (\exists e)(*see(e) \wedge C_{*\Theta_1}(m \oplus p)(e) \wedge C_{*\Theta_2}(\text{TWO GIRLS})(e))$$

In general, branching readings will always have cumulative readings as their parts, and because of that the event in which the branching reading is true will never exemplify a proposition. This makes it impossible to force the branching reading for any proposition since the event representing the branching reading is never going to make any proposition true to begin with. We could give up exemplification but that seems like a bad move since it is necessary if we want to measure and quantify over events (see Kratzer 2008).

However, a similar problem shows up in other environments than those involving branching readings. For example, we can see the same problem with conditionals. It is common to analyze these using situations/events, so a conditional would state that whenever there is an event for which the antecedent holds, this event can be extended to another event in which the consequent holds. Thus, (82) is true if every event e' which makes the proposition 'Hilary kisses Morris' true can be extended to e which makes the proposition 'Morris is happy' true.

$$(82) \quad \text{Whenever Hilary kisses Morris Morris is happy.}$$

Given our assumptions, e' and e must exemplify the propositions. Exemplification does the right job, as one can see in the following example. (83) is true if every event e' which makes the proposition 'snow falls around here' true can be extended to e which makes the proposition 'it takes 10 volunteers to remove the snow' true. What event is e' which exemplifies the proposition in the antecedent? Since every part of snow falling is snow falling then e' is an event in which snow falls and nothing else happens.

$$(83) \quad \text{When snow falls around here, it takes ten volunteers to remove it.}$$

This makes many events exemplify the proposition, from a few snowflakes dropping to huge snowstorms. The events exemplifying the proposition gives us a homogeneous set, and we cannot count such sets or quantify over them. So, some further principle

is necessary to allow quantification. Casati and Varzi (1999) and others argue that in counting we consider maximal spatiotemporally self-connected entities. Applying this to (83) we get that the sentence is true if for every event e' which represents the maximal stretch of time and space in which it consistently snowed here e' can be extended to e which exemplifies the proposition of the consequent. This interpretation seems to correspond to our intuition that (83) does not talk about a few snowflakes but about the cases of consistent snowfalls that happen around here.

So far, the exemplification seems to work correctly. But now consider the following example, from Kratzer (2008):

- (84) Whenever a man rides a donkey, the man gives a treat to the donkey.

We are looking for an event that exemplifies the proposition in the antecedent. Suppose e is an event in which the proposition is true and which consists of a man, a donkey, and the man riding the donkey. This event has parts in which the proposition is not true, for example, an event which lacks the man. Thus, we need to pick up a minimal event in which the proposition is true. But because of that the event which exemplifies the antecedent is one that includes a man, a donkey, and *a minimal ride*. But obviously, (84) does not require that the rides should be minimal, which would only multiply the number of sweets that the man must give. Rather, (84) considers only full rides, where a full ride is a temporally closed unit, that is from the point of the man getting on the donkey to the point of the man stepping off the donkey. The problem is not that we cannot derive the correct reading. This should be possible. It is standard in event and situation semantics to assume that the event/situation in the consequent of conditionals extends the event/situation of the antecedent (Elbourne 2006, Kratzer 2008). Thus, even if we consider the minimal ride in the consequent, it could still be true that the man gives a treat to the donkey only at some bigger event which extends the minimal ride. There is nothing that blocks this bigger event to include a full ride. However, we still derive the reading that the man keeps feeding the donkey after every minimal ride and (84) does not seem to have that interpretation. In general, we face the following problem. Whenever we add an entity to a state or activity event we should stop considering the maximal spatiotemporally self-connected event and start considering only the minimal event. But this does not seem to be true, as shown on (84).

Kratzer (2008) proposes that we can avoid this problem if we let exemplification apply to subevents of e which only include riding. This could be achieved if we let exemplification play a role not only on the whole but in each conjunct as well. This is shown in (85), where EX is a function that takes a proposition and outputs a set of events that exemplify the proposition.

- (85) $(\exists e \in EX(\lambda e'. (\exists x, y)(\text{MAN}(x) \wedge \text{DONKEY}(y) \wedge e' \in EX(\lambda e''. \text{RIDE}(e'')))) \wedge e' \in EX(\lambda e''. \Theta_1(x)(e'')) \wedge e' \in EX(\lambda e''. \Theta_2(y)(e''))))$

Since *to ride* is a homogeneous predicate all parts of e' make the proposition of riding, being Θ_1 and being Θ_2 true. Assuming again that for the purposes of quantifying we consider maximal events stretched in the spatiotemporal domain which are self-connected, we correctly get that e exemplifies the proposition in the antecedent if

there is a man and a donkey in e and one maximal event in which the man rides the donkey.

Letting exemplification play a role in conjuncts helps us with the problem we faced with the branching reading. To derive cumulative and branching readings we need to pluralize thematic roles, where pluralization on thematic roles is repeated here:

$$(86) \quad \begin{aligned} * \Theta(x)(e) = \\ \langle x, e \rangle \in [\{\langle a, b \rangle : \Theta(a)(b)\}] \end{aligned}$$

It suffices now to require that the subevent where the role holds for an atomic entity is exemplified:

$$(87) \quad \begin{aligned} \text{Pluralization of thematic roles with exemplification:} \\ * \Theta(x)(e) = \\ \langle x, e \rangle \in [\{\langle a, e' \rangle : e' \in EX(\lambda e''. \Theta(a)(e''))\}] \end{aligned}$$

To see how this works, consider plural thematic roles with exemplification in (88b) which is the interpretation of (88a).

$$(88) \quad \begin{aligned} \text{a.} \quad & \text{Morris and Philip saw Hilary and Desire.} \\ \text{b.} \quad & (\exists e \in EX(\lambda e'. e' \in [\{\langle e_2 : e_2 \in EX(\lambda e_3. \text{SEE}(e_3))\} \wedge \langle m \oplus p, e' \rangle \in \\ & [\{\langle x, e_2 \rangle : e_2 \in EX(\lambda e_3. \Theta_1(x)(e_3))\} \wedge \langle h \oplus d, e' \rangle \in [\{\langle x, e_2 \rangle : e_2 \in \\ & EX(\lambda e_3. \Theta_1(x)(e_3))\}]])) \end{aligned}$$

Event e in (88b) exemplifies the whole proposition. In turn e' has, as its parts, subevents which exemplify seeing. Furthermore, e' has subevents which exemplify for Morris that he is Θ_1 , and the same for Philip. Suppose Morris saw both Hilary and Desire and this was one event. Then this event exemplifies $\lambda e_3. \Theta_1(m)(e_3)$, as we want. Similarly for Philip. The same reasoning applies to the subevents of e' which exemplify for Hilary that she is Θ_2 and the same for Desire. Thus, (88b) is true of e which is the event of Morris seeing Hilary and Desire and Philip seeing Hilary and Desire.

At this point, an event can exemplify the branching reading. But the discussion above showed us that the branching reading not only can be true, it should be true when possible. This is where the SMH comes to play a role. To be sure, Dalrymple et al.'s Strongest Meaning Hypothesis has been proposed just to deal with reciprocals even though already in the original paper they proposed some possible extensions. These included the interpretation of presuppositions triggered by *manage* which seem to be governed by the Strongest Meaning Hypothesis: in other words, one selects the strongest presupposition compatible with the context and world knowledge. In our example, I assume that the selection between branching and cumulative readings is governed by the Strongest Meaning Hypothesis as well. The proposition interpreting the branching reading entails the proposition interpreting the cumulative reading and following the Strongest Meaning Hypothesis, we require the event to exemplify the strongest proposition.

Notice that nothing enables us to get to a weaker reading than the cumulative one, which makes our account different from Winter's approach (Winter, 2001b) who starts with the strongest reading and allows its weakening. Winter's approach faces

the problem of too weak sentences which should be true in special cases. Consider, for example, (89), from his paper (p. 354).

- (89) # Mary and Sue gave birth to John.

Since one person can be given birth by only one woman, the sentence should be true in Winter's approach if Mary or Sue gave birth to John but in this case the sentence just sounds odd. Compare this to (90), which is fine if Mary gave birth to John and Sue gave birth to Bill and George.

- (90) Mary and Sue gave birth to John, Bill and George.

As Winter notes, '[i]t seems that although the predicate to give birth can allow some weakening effects (cf. (90)), there is a 'lower bound' to the weakening that can take place: each member of the group argument should take part in at least one 'giving birth' relation.' This is precisely what we predict. Since the cumulative reading is the weakest reading possible, it represent the lowest bound of what the weakest reading can be. It seems to me that this is correct. In all the examples discussed in Beck and Sauerland (2000); Beck (2001); Winter (2001b) the weakest reading in sentences with plural arguments is the cumulative reading, while the strengthened reading is the branching reading. There is one single counterexample to this, from Winter (2001b), given here:

- (91) The cats are sitting in the baskets.

Winter claims that (91) is true if there are two cats, three baskets and cat 1 is sitting in basket 1, cat 2 is sitting in basket 2 and no cat is sitting in the third basket. This would be a weaker reading than the cumulative reading since the third basket is present in the situation but no cat is sitting in it. However, I have serious doubts whether when we evaluate the sentence the third basket really plays any role. It is well-known that definites can be understood as being dependent on another DP (Chierchia, 1995; Winter, 2000) in which case (91) expresses that the cats are sitting in their baskets. If the baskets 1 and 2 were the cats' baskets then the sentence is true. But in that case, we once again have the cumulative reading since the third basket is irrelevant for the evaluation of the sentence. If Winter wants to sustain his claim that (91) is true in the situation with two cats and three baskets due to his weakening account then the same weakening should appear if we substituted the definite DPs with other plural arguments. For example:

- (92) a. # Kees and Poes are sitting in the baskets 1, 2 and 3.
b. # The cats are sitting in the three baskets.

Contrary to Winter's expectations, (92a,b) are bad in the situation described above. And we know why: the situation described above is false under the cumulative reading, one cannot get weaker than that, and the dependent definite reading is not available in (92).

To conclude, I argued that the Strongest Meaning Hypothesis plays a role in the interpretation of sentences with plural arguments. In particular, it favors the branching

reading over the cumulative one. The SMH is also going to play a role in the next chapter, where we study reciprocals. Using the Strongest Meaning Hypothesis offers part of the explanation why some readings are preferred over others with group denoting DPs (G-DPs). I should point out, though, that we still did not explain the main task of this chapter: why the distributive reading with covariation is degraded with G-DPs. Finally, I want to make the following convention: even though I assume that events and their subevents are exemplified, I will not write exemplification in formulas. This is because it would make every representation very complex, and exemplification is not going to be relevant for most of the discussion.

This concludes my analysis of G-DPs. In the next section I shortly discuss how other DP types, DQ-DPs and CQ-DPs can be dealt with in this system.

3.2.6 Distributive quantifiers in the language of events and pluralities

So far, I focused on definite plurals and indefinites, which belong to the group of group-denoting DPs, G-DPs, see Table 3.1. But the system that we have should be extendable to other plural DPs. In this section, I discuss two other DP types which, unlike the DPs that we looked at so far, are more limited in the number of readings they license. The DQ-DPs, DPs with universal distributive quantifiers (*every boy*, *each boy*, *no boy*...) are compatible only with distributive readings. The CQ-DPs, DPs with counting quantifiers (*all the boys*, *both boys*, modified numerals) are compatible with distributive readings and a subset of collective and cumulative readings.

To see the varying behavior of the DPs, take, as an example, the predicate *lift five pianos*. When the subject is a G-DP then the sentence can have the cumulative reading ('the boys did lifting of pianos such that in total five pianos were lifted' in (93a)), or the subject can be interpreted collectively ('the boys, as a group, lifted five pianos' in (93a)). Finally, the distributive reading is marginally possible. The same range of readings appears with other G-DPs, for instance, for coordinations of proper names, see (93b). CQ-DPs can be interpreted collectively ((93c) can be read as 'all the boys, as a group, lifted five pianos'). The distributive reading is another possibility (albeit again, it is slightly marginal). The cumulative interpretation is less clear. Its acceptability with CQ-DPs is a matter of discussion in semantic literature and while it might be possible, it is harder than in case of G-DPs. Finally, the same sentence with a DQ-DP as the subject cannot have either cumulative or collective readings (93d).

- (93) a. The boys lifted five pianos.
 b. Philip, Morris, Hilary and Desire lifted five pianos.
 c. All the boys lifted five pianos.
 d. Every boy lifted five pianos.

A similar distinction between three types of DPs can be seen when we consider collective predicates like *gather in the hallway* and *be a rock band*. While definite plurals and coordinations of proper names (=G-DPs) can be the subject of both ((94a) and

(94b)), *all the boys* (CQ-DPs) can be the subject only of the former collective predicate (94c), and *every boy* (=DQ-DPs) cannot be the subject of either of them (94d).

- (94) a. The boys gathered in the hallway/are a rock band.
 b. Philip, Morris, Hilary and Desire gathered in the hallway/are a rock band.
 c. All the boys gathered in the hallway/*are a rock band.
 d. * Every boy gathered in the hallway/is a rock band.

We have already seen how collective, cumulative, and distributive readings are derived. We need to add the lexical semantics of DQ-DPs and CQ-DPs which captures their incompatibility with some of the readings. This is done in the following two subsections. To simplify things, I concentrate here only on the lexical semantics of *each*, as a representative of the DQ-DPs, and *all*, as a representative of CQ-DPs. After that, I derive the generalization on the distributivity hierarchy (Section 3.3).

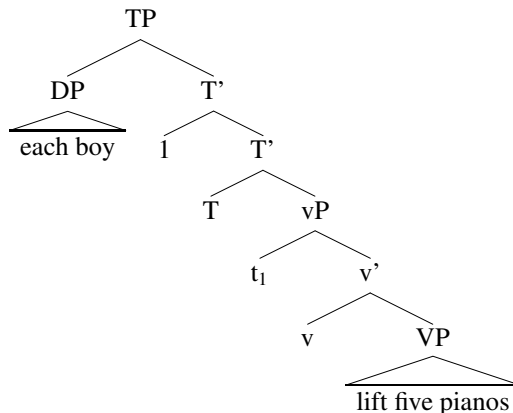
Each NP

Each NP cannot be the subject of any collective predicates or lead to cumulative readings (see Kamp and Reyle 1993; Winter 2001a, among many others). We have seen their incompatibility with collective predicates in (94d), and in (93d) we have seen that they cannot lead to cumulative readings. A common way to derive these data is to force *each NP* to scope above the event closure and distribute down to atomic entities. The meaning of *each* is given below. This combines with an NP to yield a generalized quantifier. The generalized quantifier applies to a property and distributes over it.

- (95) Interpretation of *each*, first version
 $[[\text{each}]] = \lambda P. \lambda Q. (\forall x \in P)(x \in AT \wedge Qx)$

Suppose we have a sentence (96a). After QR applies to *each boy* the resulting structure is the tree in (96b). Furthermore, the event closure must apply in the tree structure below the quantifier, so below the T'-level.

- (96) a. Every boy lifted five pianos.
 b.



The property to which the generalized quantifier applies is ‘ $\lambda x.x$ lifed five pianos’, and the whole sentence is translated as:

$$(97) \quad (\forall x \in \text{boy})(x \in AT \rightarrow (\exists e, y)(\text{*LIFT}(e) \wedge {}^C\text{*Th}(e)(y) \wedge |AT(y)| = 5 \wedge \text{*PIANO}(y) \wedge {}^C\text{*Ag}(e)(x)))$$

This is the only reading available for (96a), and it is the distributive reading. Apart from blocking the cumulative reading, the semantics of *each* also blocks its compatibility with collective predicates. To give one example, (98a) gets the interpretation (98b). Because the universal quantifier outscopes $\exists e$ the sentence is true iff for each single boy, there was an event of gathering. Because *gather* requires multiple agents this is false (or undefined if we take the incompatibility of *gather* with singular agents as presupposition failure).

$$(98) \quad \begin{array}{ll} \text{a.} & \text{* Each boy gathered.} \\ \text{b.} & (\forall x \in \text{boy})(x \in AT \rightarrow (\exists e)(\text{GATHER}(e) \wedge {}^C\text{*Ag}(e)(x))) \end{array}$$

There is one issue that the interpretation of distributive quantifiers does not capture. An indefinite in the scope of a distributive quantifier is commonly taken to introduce referents that differ from each other. Consider the following sentence. It is preferred that more than one boy is involved, that is, boys climbed up different trees (as confirmed in the experiment of Brooks and Braine 1996).

$$(99) \quad \text{Each boy climbed up a tree.}$$

That there is a strong preference to understand indefinites in the scope of distributive quantifiers as introducing several referents differing from each other has been confirmed in questionnaires as well as self-paced reading tasks (apart from Brooks and Braine 1996, see, for example, Tunstall 1998 or Anderson 2004). This is not taken care of by the interpretation of *each* in (95). Nothing requires that the property to which the generalized quantifier applies differs in any way for each atom. To put it differently, we see the following: the distributive quantifier *each* requires not only a distributive reading but it also requires a distributive reading with covariation if there is an indefinite DP in its scope (see also Beghelli and Stowell 1997 for more arguments in this direction; Anderson 2004 shows that the same holds for the quantifier *every*). There are two possible ways to capture this. First, we could assume that the distributive reading with covariation is a case of pragmatic strengthening of the semantic interpretation (95). Second, we could assume a different meaning for distributive quantifiers. I will assume the second option here even though later on I will comment on the first option as well. The interpretation of *each* I am going to assume is the following:

$$(100) \quad \begin{array}{l} \text{Interpretation of } \textit{each}, \text{ final version} \\ [[\textit{each}]] = \lambda P. \lambda Q. (\forall x, y \in P)(x \neq y \wedge x, y \in AT \rightarrow Qx \wedge Qy) \end{array}$$

(100) distributes over pairs of atoms. Suppose there is an indefinite inside the property to which the quantifier applies. An indefinite introduces a referent that should be novel in discourse. But since in the new interpretation of *each* the discourse consists of the application of *Q* to two distinct atoms, the referent introduced by an indefinite should

be different in each case. Thus, the preference for different trees in (99) is identical to one's preference to consider different trees in (101).¹²

- (101) Persse climbed up a tree. Angelica climbed up a tree. Frobisher climbed up a tree.

All NP

All NP differs from *each NP* in being compatible with some collective predicates, as we saw in (93c) and (94c). This behavior of *all NP* has been noticed in Dowty (1987), who suggests the following way of differentiating between the collective predicates. The ones that are compatible with *all NP* as the subject (*gather*, *meet*, *lift a piano together*...) have “distributive subentailments”, unlike the ones which cannot combine with *all NP* (*be a rock group*, *elect a president*). Dowty does not offer a definition of “distributive subentailments” but intuitively, it seems to stand for entailments of collective predicates that hold of atomic entities. For example, *gather* is taken to have the distributive subentailment, roughly, ‘come to be in the same place at the same time as a lot of other people’.

A problem with this account is that we do not have an independent definition of “distributive subentailments”. *Gather*, which can combine with *all NPs*, is taken to have “distributive subentailments”. But we would surely be able to think of similar subentailments for other collective predicates, which cannot combine with *all NP*. (102) is ungrammatical because *be a fraternity* is a collective predicate that cannot combine with *all NP*. Yet it is not hard to think of the “distributive subentailment” of *be a fraternity*, at least in Utrecht, where guys are a fraternity if each of them studies, is a boy, has a beer belly and goes to Janskerkhof 11 every Thursday to get very drunk and produce a lot of noise. Still, these subentailments do not seem to help in licensing *all*.

- (102) * All the boys are a fraternity.

Taub (1989) improves on Dowty's approach by adding a novel observation. I follow Brisson (2003) and call this Taub's generalization, given here:

- (103) Taub's generalization:
The collective predicates that disallow *all* are the collective predicates denoting states and achievements.

Be a group or *be a fraternity* are states. *Gather* is an accomplishment. Very closely related predicates supporting Taub's generalization come from examples (104a), which involves a state predicate, and is ungrammatical with *all*, and (104b), which involves an accomplishment, and is grammatical. (104c) shows that collective achievements

¹²To be sure, in order to fully capture covariation, we also need to represent the condition of ‘novelty’ which indefinites impose. Without that, (100) does not differ from (95). I am not going to represent this condition of indefinites here, but it is commonly assumed in dynamic semantics. In Chapter 5, I consider a dynamic semantic version of *each* in which one distributes over pairs, which captures the fact that indefinites in the scope of distributive quantifiers tend to vary since indefinites introduce new referents in discourse. This account of *each* follows Brasoveanu (2009).

cannot have a DP headed by *all* as the subject either, which further corroborates Taub's generalization. I mark (104c) not as ungrammatical but as inappropriate (#). This is because the sentence is excluded in its most natural reading, namely, 'the students, as one group, elected a president'. However, the sentence has other readings which are grammatical. Suppose we talk about elections at a high school, where each class elects its own president. If we talk about all the students at the high school (104c) is grammatical in this context (Brisson 2003, with p.c. to Veneeta Dayal).

- (104) a. * All the students are a group.
 b. All the students formed a group.
 c. # All the students elected a president.

Taub's generalization has recently been analyzed by Brisson (Brisson, 1998, 2003), who assumes that *all* requires insertion of C_* . The meaning of *all* is in putting specific requirements on C_{ov} that comes with C_* . The requirements are not relevant for the topic of this thesis, so I ignore them here.

To derive the availability of collective readings in activities and accomplishments, Brisson follows the research on aspect and event semantics, and assumes that accomplishments and activities have as a subcomponent the bleached predicate DO. DO is projected as a head that takes VPs as its complement. What counts as DO depends very much on the lexical verb. For example, in case of *build a house*, DO consists of hammering, sawing, putting up walls, etc. In case of *run*, DO is probably much more restricted in its meaning, and consists of moving one's legs very quickly. The crucial assumption is that DO can pick out parts of the main event. Thus, a sentence like (105) states that there is an event which has DO-parts and the agents of the DO-parts are the boys.

- (105) The boys gathered.

Introducing parts of the events enables one to keep the assumption that *all* forces the introduction of C_* even in collective readings. This is capturing Dowty's intuition within event semantics that predicates which allow *all NP* as its argument have sub-distributive entailments. In case of *gather*, C_* scopes only over DO so we get the reading that there is one event of gathering and the event of gathering has parts and for each of these parts it holds that an individual boy was the agent of DO. If we take DO to require, in this case, that a person goes to a place and waits there we correctly derive (106) as grammatical.

- (106) All the boys gathered.

I will simplify Brisson's system slightly. I am not going to assume DO as a new head in syntax. I will keep assuming that *v* which introduces the external thematic role, is present in activities and accomplishments. The only difference between these predicates and achievements and states is that the argument of Θ_1 of the latter predicates must make the predicates true but in the former case the argument of Θ_1 only has to make sure that the event took place. So, for example, the argument of Θ_1 of *be a group* must be a group while for the argument of Θ_1 of *gather* it suffices that he

went to some particular place and waited there, as did others. In other words, I am not going to introduce a new projection in syntax that is responsible for distributive sub-entailments but will assume that these sub-entailments are brought about by the entailments of Θ_1 in events (see also Kratzer 2003 for entailments of Θ_1 , which show that not every person that is the argument of Θ_1 makes the event that he is the agent of true).

Let us go back to achievements, where the insertion of C_* does play a role in blocking collective readings. The vP of (104c) *elect a president* is:

$$(107) \quad \lambda x. \lambda e. {}^{C_*}\Theta_1(x)(e) \wedge *ELECT(e) \wedge {}^{C_*}\Theta_2(A \text{ PRESIDENT})(e)$$

The subject *all NP* requires the insertion of C_* into the syntactic structure. There are two options: C_* can scope over v , in which case it pluralizes the thematic role (108a). Alternatively, it can scope at T' , after Existential Closure, which, as we know, normally derives distributive readings (108b).

$$(108) \quad \begin{array}{l} \text{a. } (\exists e)(*ELECT(e) \wedge {}^{C_*}\Theta_2(A \text{ PRESIDENT})(e) \wedge \langle \text{THE BOYS}, e \rangle \in [\{\langle \downarrow \\ a, e' \rangle : {}^{C_*}\Theta_1(a)(e) \wedge a \in Cov\}]) \\ \text{b. } (\forall x \leq \text{THE STUDENTS})(\exists e)(*ELECT(e) \wedge {}^{C_*}\Theta_2(A \text{ PRESIDENT})(e) \wedge \\ {}^{C_*}\Theta_1(x)(a) \wedge x \in Cov) \end{array}$$

Neither formula derives the collective reading if we add one extra assumption, following Brisson. The assumption is that C_* cannot be introduced in syntax if the argument of the predicate that C_* pluralizes is an atom. The argument could be an atom in two cases. Either it is an atomic individual or a single group. And in either case the insertion of C_* must be blocked. Thus, we can postulate the following requirement (from Brisson 2003, slightly rephrased to fit the terminology used here):

$$(109) \quad \begin{array}{l} \text{a. } \text{Economy-based condition on insertion of a } C_* \text{ operator in syntax:} \\ \text{A } C_* \text{ operator is licensed for a predicate P taking a plural argument Y} \\ \text{if Y has at least two and as many as n contextually relevant distinct} \\ \text{subparts } x_1 \dots x_n, \text{ and } P(x_1) \text{ or } P(x_2) \text{ or } \dots P(x_n) \text{ are live possibilities in} \\ \text{the discourse.} \end{array}$$

In case of (108a) this means that the boys, as a whole, are not in the set of $\{a : {}^{C_*}\Theta_1(e)(a)\}$. Rather, their subparts are. Thus, (108a) is not true if the boys, as a group, elected a president. It could only be true if there were multiple elections of one and the same president, and in each election a subpart of ‘the boys’ elected that president, which is a rather nonsensical reading. (108b) derives the reading ‘there were multiple elections and each subgroup of boys elected its own president in one of the elections’, which is a possible reading as we noticed above. Crucially, the collective reading in which the group of boys acted as one agent is not derived because (to repeat) (i) *all* requires the insertion of C_* , and (ii) the economy-based condition on insertion of C_* blocks it in case the group of the boys acted as one agent.¹³

¹³The economy condition on insertion of C_* is from Brisson (2003). Brisson (1998) assumes a different condition, which is also inspired by principles of economy, and which is spelled out as:

Let me move to the second example, which involves states. Again one might consider two formulas for the sentence (110a), (110b) and (110c). Again, the formulas cannot express that the students are one group because in that case, the theme is one entity, $\uparrow(\text{THE STUDENTS})$, and C_* cannot apply on the thematic role (110b) or on the whole predicate (110c). This leaves us only with one possible meaning, where subgroups of students form groups. However, this reading requires distribution over the nominal predicate, which is impossible in English for independent reasons (cf. (111)).

- (110) a. * All the students are a group.
 b. $(\exists e)(\text{*BE A GROUP}(e) \wedge C_*\Theta_1(\text{THE STUDENTS})(e))$
 c. $(\exists e)\langle \text{THE STUDENTS}, e \rangle \in [\{\langle \downarrow a, e' \rangle : \text{*BE A GROUP}(e) \wedge C_*\Theta_1(a)(e') \wedge a \in \text{Cov}\}]$
- (111) * All the students are a teacher.

3.3 Explaining the differences in distributive readings

3.3.1 Introduction

After introducing the semantics of events and pluralities we can finally come back to the explanation of why the acceptability of the distributive reading with covariation depends on the DP type. We want to explain why the scale in Table 3.6, discussed in Chapter 2, holds.¹⁴

Table 3.6: Acceptability of distributive readings with covariation

DQ-DPs	<	CQ-DPs	<	G-DPs
--------	---	--------	---	-------

An example of the distributive reading with covariation is given in (112). The distributive reading with covariation is the one in which the boys did not climb up the same tree. That is, at least two different trees were involved in climbing up.

- (112) The boys climbed up a tree.

- (1) Distributivity is permitted only when it is necessary.

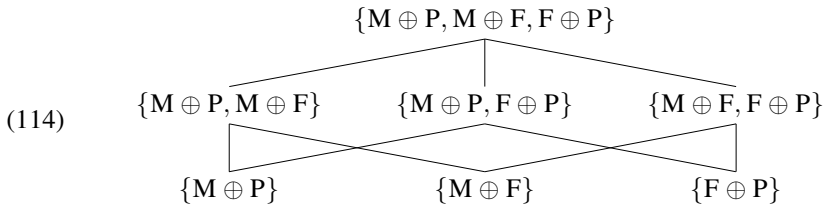
Where ‘being necessary’ equals ‘having an effect on interpretation’. This is similar to Reinhart’s and Fox’s assumption on quantifier movement (see Fox 2000, Reinhart 2006). Unfortunately, this much simpler condition does not work (not even in the cases for which it was designed). The reason is that *all* requires the presence of C_* . Thus, the insertion of C_* is necessary in these cases - without the insertion of C_* the derivation crashes. Or, to put in the terminology of Fox and Reinhart, where the meaning of derivations are compared and if they are equivalent, only the simpler derivation survives: one cannot compare the derivation with C_* to the one without C_* - because simply, there is no successful derivation without C_* . Thus, the more complicated condition of (109a) is necessary to block the collective reading of states and achievements with *all*.

¹⁴ ‘ $\alpha < \beta$ ’ should be read as α is preferred over β as the argument in whose scope a distributive reading with covariation occurs. DQ-DPs=DPs with universal distributive quantifiers, CQ-DPs=DPs with counting quantifiers, G-DPs=group-denoting DPs, see (19) in Chapter 2.

I will now define the distributive reading with covariation more precisely. First, we need two auxiliary notions: sets a quantifier lives on and a witness. These are given in (113) (based on the discussion in Introduction to Szabolcsi 1997b and Beghelli et al. 1997).

- (113) a. A generalized quantifier GQ lives on a set of entities A if, for any set of entities X ,
 $X \in GQ$ iff $(X \cap A) \in GQ$
 b. A set W is a witness of a generalized quantifier GQ iff $W \in GQ$ and $W \subseteq SL(GQ)$ where $SL(GQ)$ is the smallest set the GQ lives on

Let me give a few examples to show what (113) selects. Consider a model which consists of Morris, Philip, Frobisher and Hilary. In this model, the sets the generalized quantifier $\lambda P.P(\text{MORRIS})$ lives on are all sets that include Morris. The sets that the generalized quantifier $\lambda P.P(\text{MORRIS AND PHILIP})$ live on are the sets that include the plurality $\text{MORRIS} \oplus \text{PHILIP}$. The smallest set that the latter GQ lives on is the one-membered set $\{\text{MORRIS} \oplus \text{PHILIP}\}$, which is also the witness of the GQ. As another example, consider the generalized quantifier $\lambda P.(\exists x)(|AT(x)| = 2 \wedge *MAN(x) \wedge Px)$, which is the translation of *two men*. The sets the quantifier lives on are sets that include all pluralities each of which consists of two men. These are all the sets that have the following elements as their members: $M \oplus P$, $M \oplus F$ and $F \oplus P$. The smallest set the GQ lives on is the following: $\{M \oplus P, M \oplus F, F \oplus P\}$. Witnesses of the GQ are subsets of this set which are members of the GQ. That is, the GQ *two men* have the witnesses that form a join semilattice:



Finally, we will need minimal witnesses. These are sets which stop being witnesses if any element is subtracted. In the lattice above, the minimal witnesses are the bottom elements.

With the help of these auxiliary notions, the distributive reading with covariation is defined as follows:

- (115) The distributive reading with covariation where DP_1 distributes over DP_2 : for elements of x_1, \dots, x_n such that there is a minimal witness W_1 of $[[DP_1]]$ and $x_1 \oplus \dots \oplus x_n \in W_1$ the following holds: x_1, \dots, x_m (where $m \leq n$) are the first arguments of some Θ_i for events e_1, \dots, e_m and there are different witnesses $\{y_1\}, \dots, \{y_m\}$ of $[[DP_2]]$ and y_1, \dots, y_m are the first arguments of $C_*\Theta_j$ for events e_1, \dots, e_m , where $i \neq j$.

Consider (112). The minimal witness of $[[\text{the boys}]]$ is the set that has the plurality of all the boys. Witnesses of $[[\text{a tree}]]$ are sets of one or more trees. If the plurality of the boys can be split into parts where each part is the first argument of Θ_1 and for each part there is a different witness of $[[\text{a tree}]]$ then we deal with the distributive reading with covariation. This is true if more than one tree is involved in the boys' climbing.

Why should this reading be degraded? Recall that there are two possible representations that a sentence with two G-DPs can give rise to. These are repeated here (I am ignoring the possibility of the inverse scope reading here):

- (116) a. $DP_{\text{subj}}-DP_{\text{obj}}$
 b. $DP_{\text{subj}} \in \{\{x : \dots DP_{\text{obj}} \dots \wedge x \in Cov\}\}$

To get the distributive reading with covariation we need to assume (116b) and that the object that stays in-situ covaries with the subject. (116a) cannot give us the distributive reading with covariation simply because there the DP_{subj} does not take distributive scope over the object.

Why is it then that the distributive reading with covariation is marked when the argument that distributes is a G-DP? One possible answer is to simplify the number of interpretations that G-DPs can give rise to, in the following way:

- (117) a. $DP_{\text{subj}}-DP_{\text{obj}}$
 b. –

In (117) we assume that there is no (117b) as the second option. Since in this case G-DPs cannot scope and distribute over a predicate, the distributive reading with covariation is excluded. To capture that the distributive reading with covariation is also somewhat degraded with CQ-DPs we could assume that the web of interpretations of CQ-DPs is also limited to (117a).

However, this account cannot explain why the distributive reading with covariation is only dispreferred, not impossible. Second, it also does not capture the difference between CQ-DPs and G-DPs since they both would lack the possibility to distribute over a predicate. Another problem comes from the interpretation of *same*. *Same* can be interpreted sentence-internally. For instance, the sentence-internal reading of *same* in (118) is 'every professor read the same book that other professors did'. In this case, *same* is anaphoric to the books read by the other professors.

- (118) Every professor read the same book.

I am going to argue in Chapter 5 that a semantically plural DP must distribute over *same* otherwise its sentence-internal reading cannot be licensed. We can see this on the following examples.

- (119) a. # The students elected the same president.
 b. # We are the same band.
 c. # The same boy thought that every professor hated him.

In (119a) and (119b) no distribution takes place because of the collective predicates and the sentence-internal reading of *same* is not possible. In (119c) the sentence-internal reading is missing since the distributive quantifier cannot take matrix scope.

Crucially, the sentence-internal reading of *same* is fully acceptable with CQ-DPs and G-DPs. The following examples all allow the sentence-internal reading of *same*.

- (120) a. The men went to the same play tonight.
 b. Mike and Bob think that America was discovered by the same explorer.
 c. John and Mary found the same solution.

The first two examples are from Carlson (1987), the third example is from Moltmann (1992). We have also seen in Section 2.3 that the sentence-internal reading of *same* is possible with definite plurals in Dutch.

The problem is that if we assumed (117) none of these examples should have an available sentence-internal reading since definite plurals and coordinations should not be able to distribute over *same*.

Finally, we will later on see data from language acquisition that are problematic for the assumption that G-DPs cannot distribute. Children do not have a preference for distributive readings with covariation or non-distributive readings when the subject is a G-DP. If, however, G-DPs lacked the distributive interpretation in adults' language it is not clear why this should not hold for child language as well. If children started out with interpreting G-DPs as being compatible with both (116a) and (116b) it is not clear how they could later on exclude (116b) if only positive evidence shapes language acquisition.

For these reasons, I think that both (116a) and (116b) are in principle possible for G-DPs. What makes (116b) where DP_{obj} covaries dispreferred?

I am going to argue that it follows from two facts:

1. There are other readings apart from the distributive reading with covariation that G-DPs can give rise to.
2. There are other expressions apart from G-DPs that can give rise to the distributive reading with covariation.

Concerning the first point, we have seen that G-DPs are compatible with collective readings, as well as cumulative and branching readings. Concerning the second point, we have seen that CQ-DPs and DQ-DPs have more restricted options with respect to the readings they can give rise to.

To be able to derive the degraded status of the distributive reading with covariation from 1 and 2, we have to assume that the language grammar involves competition between the actual form and meaning and unrealized alternatives. That at least a limited competition takes place in grammar is nowadays commonly accepted (see, for example, Prince and Smolensky 1993; Chomsky 1995; Reinhart 2006 for differing viewpoints on this issue). From this viewpoint, the degraded status of the distributive reading in (121a) is simply due to the fact that there is a more specific form, (121b), that expresses precisely this reading. This can be taken as the first reason why the distributive reading is degraded.

- (121) a. The boys built a boat.
 b. Each boy built a boat.

There is another explanation of why the distributive reading in (121a) is degraded. In this second explanation, however, we also need to assume that some readings and some expressions are marked. Assume that distributive readings with covariation are more marked than the other, non-distributive readings. Compared to G-DPs, assume further that CQ-DPs are more marked expressions. In particular, the two expressions that have been tested (*all NP*, *both NP*) add an extra material, *all* and *both*, respectively, to G-DPs. Thus, for instance, *all the boys* adds *all* to the definite plural. Similarly, *the boys* can be extended by adding *both* in the form of a floating quantifier. It is known that if two expressions can express the same two meanings, where one meaning is more marked than the other, then the expression that is marked goes with the marked reading and the expression that is unmarked goes with the unmarked reading. This division is also known as the division of pragmatic labor, after Horn (1984). Thus, the degraded status of the distributive reading in (121a) can be just another case of the division of pragmatic labor.

I am going to argue that both explanations are right which not only accounts for the degraded status of the distributive reading in (121a) but also explains why the distributive reading is more degraded with G-DPs than with CQ-DPs. I will make this more precise later on, in the following two Sections, where I also motivate the two main assumptions, the marked status of CQ-DPs and the marked status of distributive readings.

Recently, linguists developed various formal systems that focus on competition in grammar. Here, I will consider two. First, formalizing the account in bidirectional Optimality Theory (Bi-OT, see Blutner 1998, 2000). Second, formalizing it in a game-theoretical approach. The section is organized as follows. In the next subsection I introduce a bidirectional account and show how it can capture the data in Table 3.6. However, as we will see, this explanation comes with a price. First, less importantly, we have to postulate one constraint on meanings that seem dubious. Second, and more importantly, we will have to assume that (121b) is more marked than (121a), and furthermore, DQ-DPs are more marked than CQ-DPs. This seems necessary to derive the fact that DQ-DPs combine with the distributive reading, unlike CQ-DPs and G-DPs. I do not know how one could motivate this marked status of DQ-DPs. The problem can be avoided once we move to a game-theoretic approach discussed in Parikh (2000). At the end, I shortly go through data on the acquisition of distributivity, which support the competition viewpoint presented here.

3.3.2 Bi-OT account

Introduction to Bi-OT

Optimality Theory (OT) was originally proposed as a phonological framework which operates with violable constraints (Prince and Smolensky, 1993). In OT, we assume an input (an underlying representation) and an output, which is a set of competing

candidates generated from the input by *Gen*. The winning candidate is selected by the application of violable constraints. The constraints are ordered by $>$ (a strict weak order) and if there are two constraints $C_1 > C_2$ then one violation of C_1 is worse than any number of violations of C_2 . The winning candidate is the least offending one given the constraints.

OT analyses usually take a production perspective in phonology and syntax. In OT semantics it is common to take a comprehension perspective, that is, the hearer perspective (Hendriks and de Hoop 2001; Blutner et al. 2006). Let me give one example of OT semantics and what it can derive.

Wolff (2003) discusses three experiments which test lexical and periphrastic causatives. He tested, among other things, the following scenario: there are three marbles in a row. One of them (the blue marble) is moving and hits the second (red) marble, which then starts moving and hits the third (green) marble. The tested sentences were as follows:

- (122) a. The blue marble moved the green marble.
 b. The blue marble made the green marble move.

When he asked the participants of the experiment which sentence best describes the scenario people mostly selected the second one. On the other hand, when the blue marble hit the green marble directly (that is, with no mediation of the red marble) the preference switched. While these were only preferences for interpretations they were clearly significant. The preferences were matched by people's perception of how many events took place. They claimed that in the first scenario two events took place and at the second scenario only one event took place. So, lexical causatives were selected mainly for direct causation which corresponded to people's intuition that only one event took place while periphrastic causatives were selected for the two event causation (the mediated causation). Similar results were also reported for example in Pinker (1989). Suppose we have the following constraint:

- (123) *REFERENT: Do not introduce new referents

This constraint is closely related to *ACCOMMODATE which has been introduced in dynamic semantics in van der Sandt (1992) and is quite commonly accepted in the OT literature, in various guises (see, for example, Zeevat 2000; Blutner 2000; Aloni 2005; Grønn 2008). It is also closely related to DOAP "Do not overlook anaphoric possibilities" of Williams (1997). Finally, it plays an important role in psycholinguistic literature where it is connected to the Principle of Parsimony introduced by Crain and Steedman (1985) and Altmann and Steedman (1988). The intuitive idea behind it is that in interpreting we are trying to keep models as small as possible, that is we do not introduce new entities into the interpretation models unless we are specifically told to do so. Assuming just this constraint captures the fact that lexical causatives do not express mediated causation because it violates the constraint more than direct causation (we would have to assume two events for the interpretation unlike in case of direct causation). However, we would predict the same pattern for periphrastic causatives. To avoid this prediction we add another constraint. We want a match between the number of verbs and the number of events, so we assume the following constraint:

(124) MATCH EVENTS IN PERIPHRASTIC CAUSATIVES:
periphrastic causatives=more than one event

The two constraints, when ordered MATCH > *REFERENT, give us the winning candidates (☞) summarized in the OT tableaux in (125). These results correspond to people’s intuition discussed above.

(125)

a.

<i>move</i>	MATCH	*REFERENT
a. ☞ direct causation		*
b. mediated causation		**

b.

<i>make X move</i>	MATCH	*REFERENT
a. direct causation	*	*
b. ☞ mediated causation		**

One might notice that the two constraints are of quite different nature. To know when *REFERENT is violated it suffices to check the output. To know when MATCH is violated we need to check input-output relations. To use the terminology of OT, the first constraint is a markedness constraint while the second one is a faithfulness constraint since it is violated when the output does not faithfully match the input.

Recently, more sophisticated models of OT were proposed which dispense with faithfulness constraints. I will focus here on Bi-directional OT of Blutner (2000) in which production and comprehension perspectives are evaluated at the same time (see Blutner and Zeevat 2003, especially Beaver and Lee 2003 for other frameworks within OT). We have two sets of constraints. One set consists of constraints on forms, the other set constrains meanings. The first set orders the candidates that differ in form by >, such that $\langle f', m \rangle > \langle f, m \rangle$ iff $f' > f$. Similarly for the ordering by the constraints on meanings. We can think of the first set of constraints as the one that is picking the best candidate from the production perspective. The second set gives us the best candidate from the comprehension perspective. In other words, the first set leads to blocking of other forms and the second set leads to the best interpretation. The optimal (winning) candidate is defined as follows.¹⁵

- (126) $\langle f, m \rangle$ is a weak optimal candidate iff
- a. $\langle f, m \rangle \in Gen$
 - b. there is no weak optimal candidate $\langle f', m \rangle$ such that $\langle f', m \rangle > \langle f, m \rangle$
 - c. there is no weak optimal candidate $\langle f, m' \rangle$ such that $\langle f, m' \rangle > \langle f, m \rangle$

The definition might seem circular, but in fact there is a simple algorithm to arrive at optimal candidates so the circularity is only apparent (for more discussion see Jäger 2002). Informally, we proceed as follows (see also Dekker and van Rooy 2000 for a discussion of this algorithm): we first consider all the candidates. We find the candidate

¹⁵This is a weak version of the bidirectional OT. There is also a strong version of Bi-OT but that is not useful for our goals. See Blutner (2000) for details.

$\langle f, m \rangle$ for which it holds that there are no alternative candidates that are better. ($\langle f, m \rangle$ is called a strongly-optimal pair in Blutner 2000). We move $\langle f, m \rangle$ to the set of the winning candidates and disregard it from the competition. We also disregard all the candidates that differ from $\langle f, m \rangle$ in one dimension (so we disregard all $\langle f', m \rangle$ and $\langle f, m' \rangle$). In the remaining set of candidates we find a new $\langle f'', m'' \rangle$ that is strongly-optimal and we proceed as before until the set of candidates to be considered is empty.

Let me show this on the example given above. We still assume the constraint *REFERENT but instead of MATCH we consider a constraint that operates only on forms. I call it SIMPLE EXPRESSION:

(127) SIMPLE EXPRESSION: Prefer lexical causatives over periphrastic ones

Intuitively, the simplicity is measured in terms of lexicalization or length and periphrastic causatives loose on both sides. The two constraints give us the results in (128). $\langle \text{lexical causative, direct} \rangle$ is the winning candidate since there is no alternative that is strictly better. The two candidates below it are disregarded since they differ from the winning candidate only in one dimension. This leaves us with the last candidate that ends up as another winner (since there are no other competitors left).

(128)

	SIMPEXP	*REFERENT
a. $\text{☞} \langle \text{lexical causative, direct} \rangle$		
b. $\langle \text{lexical causative, mediated} \rangle$		*
c. $\langle \text{periphrastic causative, direct} \rangle$	*	
d. $\text{☞} \langle \text{periphrastic causative, mediated} \rangle$	*	*

The pattern we get with bi-OT is the following: marked expressions go with a marked meaning and unmarked expressions go with an unmarked meaning. This is known as the division of pragmatic labor in Horn (1984) (see his paper for many more examples). I will also call it the Horn strategy (following van Rooy 2004).

In the next subsection I am going to show how the availability of distributive readings, summarized in Table 3.6 follows as just another instance of the division of pragmatic labor.

Explaining the generalization about distributive readings in Bi-OT

I am going to assume the following constraint:

(129) *COVARIATION: Avoid the distributive reading with covariation

*COVARIATION is connected to the constraint *REFERENT. For example, *the boys climbed up a tree* would introduce many trees into the interpretation model under the distributive interpretation, which is not true for other readings. Alternatively, distributive readings with covariation are marked because they create quantificational dependencies, which are taxing for memory. That distributive readings create quantificational dependencies has been argued for in dynamic semantics with pluralities

(van den Berg 1996, Krifka 1996, Nouwen 2003 and others). Wang et al. (2006) provide an empirical evidence that these quantificational dependency deteriorate quickly in discourse.¹⁶

We can also distinguish expressions based on markedness. *All the NP* is more marked than *the NP* since the former is longer/introduces more structure. Another possible viewpoint is that the alternative to *the NP* is a partitive structure (*all of the NP*) which would lose to the definite plural for the same reason. The situation is less clear with singular distributive quantifiers (*each NP*, *every NP*). They might be marked as compared to *the NP* since they represent a syntax-semantics mismatch - they are interpreted as a plurality but syntactically and morphologically they behave as singulars. Alternatively, the competition might be not between *each boy/the boys* but between *the NP* and the same DP with a floated quantifier (*the NP each*) or a partitive construction (*each of the NP*) where the latter is always longer than the former and thus higher on the scale of markedness. I will assume the latter option, so nothing beyond SIMPLE EXPRESSION needs to be assumed.

The assumptions so far are quite natural and do not diverge from the assumptions that are commonly made in Bi-OT literature. Unfortunately, they do not suffice. We have three expressions so if they all compete with each other we need to assume further differentiation. I add preciseness of interpretation as another factor. In particular, I am going to assume the constraint VAGUE INTERPRETATION:

- (130) VAGUE INTERPRETATION:
Terms are preferably interpreted in a vague way.

This constraint is taken from Krifka (2002) who uses it to derive the fact that, simplifying somewhat, shorter numeral expressions are preferably interpreted as approximate unlike longer numeral expressions (compare the intuitive difference between *one thousand kilometers* and *nine hundred sixty-five kilometers* where the former expression suggests a low precision level). For further motivation of this constraint see Krifka (2002). Finally, we need to differentiate between *each boy/the boys each* on the one hand and *all the boys* on the other hand. Since I am not aware of a constraint that would be independently needed and argued for in the OT semantics/Bi-OT literature, I will stipulate *DISTRIBUTIVE QUANTIFIER:

- (131) *DISTRIBUTIVE QUANTIFIER:
Do not use a distributive quantifier

¹⁶It would be interesting to connect *COVARIATION to Donka Farkas' work. Consider the following example:

- (1) Every boy read a paper.

In a series of papers (Farkas 1997, Farkas 2002, Brasoveanu and Farkas 2009), Farkas argues that we could derive the "wide scope reading" of the indefinite in cases like the one above by specifying what it covaries with. In this way, the indefinite can, as far as syntax is concerned, stay in the scope of the distributive quantifier, even in its wide scope reading. The wide scope reading is simply just a lack of covariation of the indefinite. This would naturally account for *COVARIATION as a prohibition on the mechanism that Farkas assumes induces covariation of arguments in its scope.




The specificity of this constraint suggests it is wrong. I am going to use the constraint only for illustrative purposes. My actual account later will dispense with it (as well as with VAGUE INTERPRETATION).

Finally, this is the ordering between the constraints:

- (132) *DISTRIBUTIVE QUANTIFIER > SIMPLE EXPRESSION
 *REFERENT > VAGUE INTERPRETATION

The resulting tableau with the winning candidates is given in (133).

(133)

	*DIST	SIMPEXP	*REF	VAGUE
a.  <i>each</i> , dist	*	*	*	
b. <i>each</i> , nondist-precise	*	*		*
c. <i>each</i> , nondist-vague	*	*		
d. <i>all</i> , dist		*	*	
e.  <i>all</i> , nondist-precise		*		*
f. <i>all</i> , nondist-vague		*		
g. <i>the</i> , dist			*	
h. <i>the</i> , nondist-precise				*
i.  <i>the</i> , nondist-vague				

The bottom element is the super-optimal pair. If we exclude the pairs that differ from *the*, nondist-vague only in one dimension, we get, as the next optimal pair *all*, nondist-precise. Finally, if we exclude the pairs that differ from *all*, nondist-precise only in one dimension, we get *each*, dist as the last optimal pair (since nothing else is left).

The winning candidates are the ones we want. Only DPs headed by distributive quantifiers lead to the distributive reading of a predicate. The DPs headed by *all* take the next marked meaning and definite plurals are the last ones. Furthermore, if we assume Stochastic Bi-OT we can straightforwardly derive why *all NP* and definite plurals are only dispreferred with distributive readings (not impossible) and why the former is more likely to be accepted with that reading than the latter.

In Stochastic Bi-OT every constraint is assigned a real number. This number does not only determine the ranking of the constraints but also the distance between the constraints. At each evaluation the placement of a constraint is modified by adding a normally distributed noise value. The ordering is established only after adding this noise value to the number assigned to the constraint. Thus, if we added constraints that are mirror images of *DISTQUANT and SIMPLE EXPRESSION (that is, *ALL NP and COMPLEX EXPRESSION) and assigned higher values to the first two we would derive the pattern given above. If, furthermore, the noise value on the last two constraints could promote them over the first two we would derive that *all NP* and definite plurals can also give rise to the distributive reading albeit this is less preferred than with

distributive quantifiers. Furthermore, the distributive reading is more likely to arise with *all NP* than with definite plurals. I will not go into this issue any further here.

Finally, even though I presented how Bi-OT works on definite plurals, other G-DPs are predicted to be disliked with distributive readings as well: coordination of proper names, numeral expression (where the alternatives are either partitive constructions or DPs with floated quantifiers), the pronoun *they* (where the alternatives are *they all* or *all of them* and *each of them*). And lastly, the analysis correctly predicts that DPs headed by *all* require a higher level of precision as compared to definite plurals, which is empirically supported in Brisson (1998), Brisson (2003) and Lasersohn (1999). In contrast to their accounts we do not need to postulate any extra semantic condition to derive this effect since it simply follows from the competition between the forms and is just one consequence of the division of pragmatic labor.

However, there are also many things left to be desired. First, *DISTQUANT is a suspicious-looking constraint. Basically, we stipulated it to get the results right. Secondly, it is questionable whether we want to have a general constraint like VAGUEINT. Is it really so that we prefer to interpret sentences in a vague manner? As Krifka (2007) says “[t]here are certainly situations in which the speaker wants to be interpreted in a precise way. For example, if someone offers to sell a car for one thousand euros, he would not be satisfied if the buyer offers him less than that, with the excuse that approximate interpretations are preferred.”

Finally, the way we used Bi-OT is dubious. We have ended up correctly predicting that distributive quantifiers give rise to distributive readings but this has been a consequence of treating them as marked expressions. In fact, if we just took their lexical semantics into account this conclusion would already follow. This points to a big flaw in our setup. The division of pragmatic labor is a resolution of forms *underspecified* for their meaning but we used some forms (distributive quantifiers) which are fully specified and unambiguous from the start. This flaw, as well as the other problems disappear when we switch to the game-theoretical analysis that Parikh (2000) proposed (see Benz et al. 2006 for an introduction to game theory).

3.3.3 Game-theoretical analysis

Introduction to game-theoretical analysis of Parikh

Parikh (2000) offers an account of why ambiguous or underspecified forms are assigned its most probable interpretation by considering markedness of expressions and probability distributions. The relevance of probabilities in communication has already been put forth in Shannon (1948). As Parikh shows, we can explain why underspecified forms are resolved towards most probable interpretations if we consider communication as a game with partial information and two players, the speaker and the hearer.

In a conversation, the speaker has a private information that she is trying to convey to the hearer. Consider a simple (albeit abstract) example built after Parikh (2000). In this example, what the speaker is trying to convey is one of the two meanings, *m* or *m'*. The hearer, of course, does not know which of *m* or *m'* holds but knows

	m	m'
S_1	f	f'
S_2	f	f
S_3	f''	f
S_4	f''	f'

Table 3.7: Speaker's strategies

	f	f'	f''
H_1	m	m'	m
H_2	m'	m'	m

Table 3.8: Hearer's strategies

which one is more likely. Let us assume that the probability of m taking place is 0.8 and the probability of m' is 0.2. The speaker can choose from three expressions, f , which is ambiguous between m and m' , f' which unambiguously means m' and f'' which unambiguously means m . To find out whether f means m the hearer has to consider these alternative expressions as well. f means m only if once we consider the alternatives, $\langle f, m \rangle$ pair is the optimal speaker-hearer strategy which resolves the communication game. Let us see how we decide that.

A strategy is an action in which we abstract away from concrete circumstances. A speaker's strategy, notated as S , is a function from meanings to forms, i.e., an element of $[\{m, m'\} \rightarrow \{f, f', f''\}]$. A hearer's strategy, H , is a function from forms to meanings, i.e., an element of $[\{f, f', f''\} \rightarrow \{m, m'\}]$. The strategies cannot contradict the lexical meaning of the forms. All in all, this gives us four strategies for the speaker and two strategies for the hearer. The strategies are displayed in Tables 3.7 and 3.8.

We are looking for a pair of speaker-hearer strategies which is a solution to the communication game. Such a solution should be a Nash equilibrium, which is defined as:

$$(134) \quad \langle S, H \rangle \text{ is a Nash equilibrium iff}$$

- a. $\neg \exists S' : \langle S', H \rangle > \langle S, H \rangle$
- b. $\neg \exists H' : \langle S, H' \rangle > \langle S, H \rangle$

Notice the parallel between the equilibrium and winners in Bi-OT. As with Bi-OT we now need to define $>$, that is, when one pair counts as better than another. Parikh defines this in terms of expected utilities. So $\langle S', H' \rangle > \langle S, H \rangle$ if $\langle S', H' \rangle$ has higher expected utilities than the latter. The expected utilities are defined in (135).

$$(135) \quad EU(\langle S, H \rangle) = \sum_m P(m) \times U(m, S, H)$$

$P(m)$ is the probability of m . In our example, $P(m) = 0.8$ and $P(m') = 0.2$. $U(m, S, H)$ is a utility function. Parikh assumes one and the same utility function for the hearer and the speaker because he considers communication as a game of co-ordination, that is, whatever is useful (has a high utility) for the speaker is also useful (has a high utility) for the hearer. The utility function gives us a positive value in case the hearer interprets the message as the speaker intended. That is, it is positive iff $H(S(m)) = m$. This should again make sense since the speaker and the hearer are co-operating, they want to understand each other. Furthermore, Parikh adds that the actual utility depends on the complexity of used expressions. For our example, let us assume that the complexity of f , $Compl(f) = 1$, and $Compl(f') = Compl(f'') = 2$. This

	H ₁	H ₂
S ₁	1	0
S ₂	1	0
S ₃	0.5	0.5
S ₄	0.5	0.5

Table 3.9: Utilities for m

	H ₁	H ₂
S ₁	0.5	0.5
S ₂	0	1
S ₃	0	1
S ₄	0.5	0.5

Table 3.10: Utilities for m'

	H ₁	H ₂
S ₁	0.9	0.1
S ₂	0.8	0.2
S ₃	0.4	0.6
S ₄	0.1	0.5

Table 3.11: Expected utilities

expresses that the latter expressions are more complex. The complexity corresponds to the constraints on forms that are assumed in OT and that we have discussed in the previous section. For example, f' and f'' are more complex because they are longer expressions than f .

We want the utility function to be the inverse function of complexity, which would capture the fact that it is more useful for the speaker to use less complex expressions. However, the utility function should also reflect whether the hearer understood the speaker. In case the hearer did not the utility should be zero irrespectively of whether the speaker used a complex or a simple expression since the goal of the communication was not reached. These considerations give us the following definition for U :

$$(136) \quad \begin{aligned} U(m, S, H) &= \frac{1}{\text{Complexity}(S(m))} && \text{if } H(S(m)) = m \\ &= 0 && \text{otherwise} \end{aligned}$$

Thus, the utility is the inverse of the complexity of the form in case the hearer correctly understands the speaker, otherwise it is 0.

At this point we can calculate expected utilities using the definition (135) as multiplication of utilities by probabilities for each meaning and each speaker-hearer strategy pair. The tables 3.9-3.11 summarize the utility and expected utility values.

The boxed values represent Nash equilibria. $\langle S_1, H_1 \rangle$ is the case where f expresses m and f' expresses m' . $\langle S_2, H_3 \rangle$ is the case where f expresses m' and f'' expresses m . The first case is the Horn strategy: the unmarked form f goes with the unmarked meaning m . The second case is the anti-Horn strategy: the unmarked form f goes with the marked meaning m' . Parikh suggests that the actually used strategy is the Nash equilibrium with the highest expected utility, which in this case gives us the pair $\langle S_1, H_1 \rangle$. This is what we want.

Explaining the generalization about distributive readings

In accounting for the generalization about distributive readings we could first consider a very simple solution: a sentence including a G-DP competes with the sentence that includes a DQ-DP. For example, consider the following two sentences:

- (137) a. The boys built a boat.
b. Each boy built a boat.

(137a) is ambiguous while (137b) is not. Thus, (137b) is more informative than (137a). In particular, (137b) can only express the distributive reading with covariation (different boys built different boats). I will label this reading m_d . (137a) can express m_d but it can also express the meaning where the boys cooperated on building one boat (the collective reading). I will label any non-distributive reading m_t . The subscript expresses that the meaning is parallel to the one that one might arrive at by adding the adverb *together*.

A cooperative speaker should prefer using (137b) over (137a) when possible. However, if she uses (137a) the hearer might conclude that there was a reason why the more informative form was not used. Since the hearer assumes that the speaker was cooperative he concludes that the speaker does not know whether (137b) is true. This can be further strengthened into the conclusion that the speaker knows that (137b) is not true (see Sauerland 2004 for this last step in the reasoning). Thus, (137a), while semantically ambiguous, ends up expressing only m_t .

This reasoning can also be captured in the cooperative game with partial information that we have introduced above. In this case, we have two forms, $f_{(137a)}$ and $f_{(137b)}$ and two meanings, m_d and m_t . It is not hard to see that if the two forms are equally costly the solution with the highest utility is to assign $f_{(137a)}$ to m_t and $f_{(137b)}$ to m_d regardless of the probability distribution. The reason is that this way of distribution gives us positive utility values for both meanings, while other options (for example, assigning $f_{(137a)}$ to m_d) would give us the positive utility only for one meaning, m_d .

Unfortunately, this reasoning runs into various problems. First, we only considered a very limited competition between forms. Assume now we added another competitor:

- (138) The boys built a boat together.

(138) expresses m_t . If we applied the same reasoning we used for (137a) and (137b) we would end up predicting that (137a) cannot mean m_t either. In fact, it should have no meaning at all.

The problem we just encountered are common for Gricean accounts and can be replicated with many instantiations of implicatures. The standard response is to switch to a neo-Gricean account (Horn, 1989) in which expressions that are in competition with each other are lexically specified as such. Thus, one could stipulate that *each* competes with *the* but neither of them competes with the adverbial *together*. However, I do not think that this solves anything, it just begs the same question. We would still need to answer why the first two expressions compete with each other. More principled

accounts of which elements count as competitors (like Matsumoto 1995) are of no use here.¹⁷

There is another problem which could not be solved even if we could explain why (138) is ignored in the competition. Consider the following pair of sentences:

- (139) a. All the boxes are connected to each other by a line.
b. The boxes are connected to each other by a line.

The first sentence has the interpretation in which *all the boxes* distribute over *a line*. This reading is dispreferred in the second sentence. This fact has been observed in Williams (1991), Moltmann (1992) and I will present a questionnaire in the next chapter that further supports it.

The problem is that to derive the low preference for the distributive reading with covariation in (139b) we have to assume that definite plurals compete not only with distributive quantifiers but also with CQ-DPs like *all NP*. However, if G-DPs compete with CQ-DPs the simple-minded viewpoint of competition cannot suffice because CQ-DPs are compatible with other readings apart from the distributive one. As we have seen in Section 3.2.6 they can give rise to the collective interpretation of a predicate, at least when the predicate expresses an activity or accomplishment. For this reason, the collective interpretation should be degraded as well for G-DPs, contrary to the facts. In fact, as we have seen in Chapter 2 that only the distributive reading with covariation is degraded. Collective and branching readings are fully accepted.¹⁸

We can avoid both problems if we assume that some expressions are more costly than others (as Parkih does) and we do similarly for meanings. I will start with the latter assumption.

In Parikh (2000) the markedness of meanings is read off only from probability distributions: if there are two meanings the marked one is the one with lower probability. However, using just probabilities can hardly account for all cases of the Horn strategy. Consider the flagship of the Horn strategy, lexical and periphrastic causatives. It is true that in some cases it looks like lexical causatives describe the stereotypical, hence more likely, causation. The classical example that is repeated in literature to illustrate this is the following pair:

- (140) a. Black Bart killed the sheriff
b. Black Bart caused the sheriff to die.

¹⁷Matsumoto (1995) argues that only expressions of the same monotonicity may compete with each other. But adding *together* does not change the monotonicity so this would not help us exclude (138). Recently, Katzir (2008) has argued for a structural account of what counts as alternatives in competitions in the grammar. An alternative to the structure *f* is another structure, *f'*, which has been derived from *f* by finite applications of deletions and contractions on the tree structure of *f* and substitutions in the lexicon. This should let (137a) compete with *Each boy built a boat*, which would derive the degraded status of *m_d* in (137a). Furthermore, (138) is excluded as a competitor since it expands, rather than contracts the structure. However, in the main text I am going to consider another issue that is problematic for Katzir's account.

¹⁸Cumulative readings are somewhat degraded even though the degradation is milder than with distributive readings with covariation, as one could see in Gil's questionnaires, discussed in Section 2.2.2. The degradation is likely due to the Strongest Meaning Hypothesis, as I argued in Section 3.2.5.

(140a) is used in more stereotypical cases, like shooting or knifing the sheriff, while (140b) is rather going to be used in less common cases. For example, (140b) could have been used when Black Bart had left the town and the sheriff, thrown into despair from losing his only opponent, started drinking and one day after drinking too much he was run over by a stage wagon.

However, there are other things than just stereotypicality which govern the choice of a causative. Experiments in Wolff (2003) show, among other things, that intentionality of the causer plays a role. For example, if a girl plays with a ball which accidentally slips out of her hands and breaks a vase then people prefer the periphrastic causative to describe the situation (*The girl caused the vase to break*). If the girl throws the ball and it breaks the vase, people prefer the lexical causative (*The girl broke the vase*). I do not see how intentionality could follow from likelihood. At least in the example given it seems to me that throwing a ball does not constitute a more likely way of destroying a vase than accidentally dropping a ball. Other tested examples in Wolff (2003) are equally problematic. On the other hand, the contrast corresponds to the number of events perceived. Wolff asked, as part of his experiment, how many events people thought occurred. In case of unintentional causation people were more likely to assume two events took place while they assumed one event took place in case of intentional causation. We can make sense of this if we assume that the two-event interpretation is more marked due to its violation of *REFERENT, which we have discussed in the OT account of causatives.

Modifying the utility function suffices to account for this behavior (see also van Rooy 2003). Utilities depend not only on complexities of expressions, but also on the cost of meanings. This gives us the following definition of $U(m, S, H)$:

$$(141) \quad U(m, S, H) = \begin{cases} \frac{1}{c(m) \times Compl(S(m))} & \text{if } H(S(m)) = m \\ 0 & \text{otherwise} \end{cases}$$

Consider the following example: f is a lexical causative and f' is a periphrastic causative. m is direct causation (only one event is involved) while m' is mediated causation (two events are involved). Because the latter violates the constraint *REFERENT more than the former, it is costly. For illustrative purposes, let us say that $Compl(f) = 1$ and $Compl(f') = 2$ and $c(m) = 1$ and $c(m') = 2$ even though any numbers not smaller than 1 would do as long as they satisfy the following two conditions: $Compl(f') > Compl(f)$ and $c(m') > c(m)$ (we argued above why both of these conditions should hold).

Now, we can correctly derive that lexical causatives express direct causation and periphrastic causatives express mediated causation even if we let the probability distribution be undecisive, that is $P(m) = P(m') = 0.5$.

The strategies that the speaker and the hearer can play differ from the previous example somewhat since we have only two meanings here both of which are underspecified with respect to m and m' . The strategies are summarized in Tables 3.12 and 3.13. The expected utilities are given in Table 3.14. Notice that the winning strategies correspond to the Horn-strategy, $\langle S_2, H_1 \rangle$, where the unmarked form goes with the unmarked meaning and the marked form goes with the marked meaning, and the

	m	m'
S ₁	f	f
S ₂	f	f'
S ₃	f'	f
S ₄	f'	f'

Table 3.12: Speaker's strategies

	f	f'
H ₁	m	m'
H ₂	m'	m

Table 3.13: Hearer's strategies

	H ₁	H ₂
S ₁	$\frac{4}{8}$	$\frac{2}{8}$
S ₂	$\frac{5}{8}$	0
S ₃	0	$\frac{3}{8}$
S ₄	$\frac{1}{8}$	$\frac{1}{8}$

Table 3.14: Expected utilities

anti-Horn strategy $\langle S_3, H_2 \rangle$, where the unmarked form goes with the marked meaning and the marked form goes with the unmarked meaning. This is the same result as our previous one. Besides, as before, we also derive that the Horn-strategy should be preferred if one considers the strategy with the highest payoff.

We can now derive that G-DPs and CQ-DPs preferably give rise to non-distributive readings (the branching reading, the collective readings) even if they are not more likely than distributive readings if we make the following assumptions: first, expressions of CQ-DPs are more costly than G-DPs. I argued for this in the previous section. In particular, the G-DPs *the NP*, *Num NP*, and coordinations can be expanded with *all* which shifts them to CQ-DPs, either if we add *all* as a quantifier, a floating quantifier or in a partitive construction. Thus, shifting G-DPs into CQ-DPs is only possible by adding extra material, and this is an additional cost. The second assumption is that the distributive reading with covariation is a costly interpretation. I offered arguments for this assumption in the previous section.

Let me start with a simplified situation, where only three forms and two meanings play a role. First, we consider the situation in which we have a DQ-DP and G-DP and the distributive and non-distributive readings. I will call this game $\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}$.

The non-distributive readings subsume a collective interpretation, a branching reading and a cumulative reading (even though the last is dispreferred from the start due to the SMH). I will not distinguish between them since this is irrelevant here. I will use the following notation: f_d is a sentence with a DQ-DP, f_t is the same sentence with a G-DP. Furthermore, m_d is the distributive reading with covariation, and m_t are the other readings that the G-DP can give rise to. Finally, I will assume that there is a specific form that expresses the non-distributive reading. Adding the adverb *together* can play such a role, as we have seen above in comparing (137a) and (137b) with (138), where the last sentence lacks the distributive interpretation (see Schwarzschild 1992, Lasersohn 1995 and Kratzer 2003 on this role of *together*). I will notate this

form as f_t . (137a), (137b) and (138) can be considered a particular exemplification of competitors in this game, where (137a) = f_g , (137b) = f_d , (138) = f_t . The subscripts use the following convention: g in f_g says that we deal with a sentence which includes a **G-DP**, (137a). d in f_d says that we deal with a sentence which includes a **DQ-DP**, (137b). t in f_t says that we deal with a sentence which includes a **G-DP** and the adverb **together**, (138). d in m_d says that we deal with the **distributive** interpretation. t in m_t says that we deal with non-distributive readings, which are the ones that occur if we add the adverb **together**, hence the subscript.

Given what we have said in Section 3.2 f_d is only compatible with m_d . On the other hand, f_g is ambiguous between m_d and m_t and f_t is compatible only with m_t . Furthermore, we need the following assumptions:

(142) Assumptions for $\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}$:

- a. $\text{Compl}(f_t)$ is greater than the complexity of f_g , that is, *the boys together* is more complex than *the boys*
- b. $\text{Compl}(f_d)$ is not greater than the complexity of f_t , that is, *each boy* is not more complex than *the boys together*
- c. $c(m_d)$ is greater than the cost of m_t

I argued for the third assumption before. The first and second assumption should not be controversial since f_t requires an extra adverb. Consider the following particular exemplification of this game: $f_g = f_d = 1$, $f_t = 2$, $m_t = 1$, $m_d = 2$. I will assume throughout that probabilities do not play a role, unless specified otherwise. In other words, any reading is equally likely as any other reading. For the case at hand, this means that $P(m_t) = P(m_d) = 0.5$. Obviously, the best combination of speaker-hearer strategy is the one where both of them agree on the meanings assigned to the forms (that is, where $H(S(m)) = m$ for every m). This is true in two cases: f_g expresses m_d (in which case f_t is used for m_t) or f_g expresses m_t (in which case f_d is used for m_d). The first case holds in the following speaker-hearer strategy pair:

- (143) a. Speaker: $m_d \rightarrow f_g, m_t \rightarrow f_t$
 b. Hearer: $f_g \rightarrow m_d, f_t \rightarrow m_t, f_d \rightarrow m_d$

The second case holds in the following speaker-hearer strategy:

- (144) a. Speaker: $m_d \rightarrow f_d, m_t \rightarrow f_g$
 b. Hearer: $f_d \rightarrow m_d, f_g \rightarrow m_t, f_t \rightarrow m_t$

In the first case, the expected utility is $\frac{1}{2}$ (see (145a)), in the second case the expected utility is higher, $\frac{3}{4}$ (145b). Thus, we correctly derive that to interpret f_g as m_d is dispreferred, even though f_g carries this meaning.

- (145) a. The expected utility of $\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}, f_g = m_d$
 $EU(S, H) = 0.5 \times \frac{1}{2 \times 1} + 0.5 \times \frac{1}{1 \times 2} = \frac{1}{2}$
 b. The expected utility of $\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}, f_g = m_t$
 $EU(S, H) = 0.5 \times \frac{1}{2 \times 1} + 0.5 \times \frac{1}{1 \times 1} = \frac{3}{4}$

We derive it for these particular numbers but it would follow for other numbers, as long as we keep the assumption given in (142). These have been substantiated throughout.¹⁹

Consider another game, where a G-DP competes with a CQ-DP. Again, we have two meanings to consider: m_d and m_t . But in this case two forms are ambiguous: f_g (corresponding to the sentence with a G-DP) can express m_d and m_t . But similarly, f_c (corresponding to the sentence with a CQ-DP) can express m_d and m_t . As before, we have f_t which unambiguously expresses m_t . Furthermore, we make the following assumptions:

(146) Assumptions for $\text{GAME}_{\text{CQ-DP} \times \text{G-DP}}$:

- a. $\text{Compl}(f_c)$ and $\text{Compl}(f_t)$ are greater than $\text{Compl}(f_g)$
- b. $\text{Compl}(f_c)$ is not greater than $\text{Compl}(f_t)$
- c. $c(m_d)$ is greater than $c(m_t)$

I argued for the first and third assumptions before. The second one is similar to the one we used in the previous game. Again, I do not consider it controversial since f_t must add an adverb to G-DPs while f_c adds a quantifier so I do not see a reason why the latter should be more complex than the former. These assumptions suffice to derive the correct results: f_c is interpreted as m_d while f_g is interpreted as m_t (see Appendix). In fact, this is just another case of the division of pragmatic labor that we have discussed in the case of causatives. Even though both f_c and f_g are underspecified, their meaning is resolved in such a way that f_c , the marked form, expresses the marked meaning, while f_g , the unmarked form, expresses the unmarked meaning.

To sum up, if we assume that DPs compete with each other, we derive that sentences with G-DPs, albeit ambiguous in semantics, lack the distributive reading with covariation. This follows in two ways. First, if G-DPs compete with DQ-DPs, the DQ-DPs block the G-DPs from expressing the distributive reading with covariation because that is the only reading that the DQ-DPs can express.²⁰ Second, if G-DPs

¹⁹To see that it follows for any numbers that satisfy the conditions given above consider the following (I am assuming that probabilities are equal and thus, ignore them in the calculation; $x = c(m_d)$, $y = \text{Compl}(f_d) = \text{Compl}(f_t)$, $\text{Compl}(f_g) = 1$):

$$\frac{\frac{1}{c(m_d)\text{Compl}(f_d)} + \frac{1}{\text{Compl}(f)}}{\frac{1}{xy} + 1} > \frac{\frac{1}{c(m_d)\text{Compl}(f)} + \frac{1}{\text{Compl}(f_t)}}{\frac{1}{x} + \frac{1}{y}}$$

Given the assumptions of the game, y must be higher than 1. Then the inequality has two possible solutions: one in which x is smaller than 0 and one in which x is greater than 1. Since x represents a costly meaning it must be higher than 1. Thus, the inequality holds for the case where we assume the complexity of DQ-DP is equal to the complexity of G-DP with the adverb *together*. Intuitively, it also holds for the cases where the complexity of DQ-DP is smaller than the complexity of G-DP with *together* since in that case the left side of the inequality becomes greater so the inequality must hold in these cases as well. I show this in Appendix. From now on other inequalities relevant in the game are shown only in Appendix.

²⁰That *each* expresses only the distributive reading with covariation was argued for in Sectin 3.2.6. This was based on examples like (1), where the strongly preferred interpretation is that more than one tree was climbed, that is, different boys climbed up different trees.

- (1) Each boy climbed up a tree.

There is another way to derive this fact, however. The preference for covariation might follow from pragmatic considerations. According to this viewpoint, *each NP* can express the distributive reading with covariation, as well as the distributive reading with no covariation. In the latter case, (1) would mean that

compete with CQ-DPs, the restriction of G-DPs to non-distributive readings is an instantiation of the Horn-strategy.

Finally, distributive readings should also be degraded with CQ-DPs in case we assume they compete with DQ-DPs. This is for the same reason that the distributive reading with covariation is degraded with DQ-DPs. CQ-DPs are ambiguous between distributive and non-distributive readings but since there is another form, DQ-DPs, that can already express the marked reading, sentences with CQ-DPs receive the non-distributive interpretation.

This brings me to the final question. What happens if all forms compete with each other? That is, all three competitors compete at the same time: G-DPs, CQ-DPs and DQ-DPs. For example, the competition takes place between these three forms:

- (147) a. f_g : The boys built a boat.
 b. f_c : All the boys built a boat.
 c. f_d : Each boy built a boat.

There are two other forms to be considered. These are the ones that disambiguate (147a) and (147b) towards non-distributive readings:

- (148) a. f_t : The boys built a boat together.
 b. f_{ct} : All the boys built a boat together.

We will again consider two meanings: m_d , the distributive reading with covariation, and m_t , non-distributive readings. I call this game $\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP} \times \text{G-DP}}$. It is accompanied by the following assumptions:

- (149) Assumptions for $\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP} \times \text{G-DP}}$:
- a. $\text{Compl}(f_g) < \text{Compl}(f_c) < \text{Compl}(f_{ct})$, that is, *the boys* is less complex than *all the boys* which is less complex than *all the boys together*
 - b. $\text{Compl}(f_g) < \text{Compl}(f_t)$, that is, *the boys* is less complex than *the boys together*
 - c. $\text{Compl}(f_d) \leq \text{Compl}(f_t)$, that is, *each boy* is not more complex than *the boys together*
 - d. $\text{Compl}(f_c) \leq \text{Compl}(f_t)$, that is, *all the boys* is not more complex than *the boys together*
 - e. $c(m_t) < c(m_d)$, that is, a distributive reading is more costly than non-distributive readings
 - f. $\text{Compl}(f_d) < \text{Compl}(f_c)$, that is, *each boy* is less complex than *all the boys*

each boy climbed up the same tree. However, the latter reading is dispreferred due to the competition. The distributive reading with covariation is marked. If we assumed that DQ-DPs are marked expressions, they would end up expressing only the marked meaning, while the unmarked one would be left for G-DPs. I will not pursue this analysis of (1) any further here and will keep the assumption that *each* is lexically specified as requiring the distributive reading with covariation.

Apart from the last condition, all the other conditions have been discussed before and I offered arguments why they should hold. However, I do think that the last condition is problematic. It is not clear to me why a DQ-DP should be less complex than CQ-DPs. This might not be obvious at this particular case, but it is problematic if we consider other examples. For instance, if we had a numeral DP like *three boys*, the competitors would be the CQ-DP *all three boys* or *three boys all* (with *all* as a floating quantifier) and the DQ-DP *three boys each* (with *each* as a floating quantifier). I do not see a reason why *all three boys* should be considered more complex than the version with a distributive quantifier.

For a moment, I will ignore this problem. We will see later on how we can avoid it. With these assumption, the winning strategy is the one where f_d is interpreted distributively and f_g has the non-distributive interpretation. This leads to the best strategy since every meaning is expressed and the complexity of the forms used is minimal (see Appendix). Thus, G-DPs are not used to express the distributive reading (I will discuss CQ-DPs below). Interestingly, we are also able to derive why CQ-DPs and G-DPs differ, that is, why in (147a) the preference for non-distributive readings is higher than with (147b), as we have seen in the experiments discussed in the previous chapter.

We have seen above that the Nash equilibrium with the highest utility corresponds to the Horn-strategy. It turns out that at least in some cases the Horn-strategy represents only a preference. For example, periphrastic causatives are preferred as interpreting mediated causation but they can also be understood as direct causation (in Wolff 2003, experiment 1, around 80% of the participants used periphrastic causatives when describing mediated causation, and around 40% of the participants used periphrastic causatives when describing direct causation). One way to understand this is that the strategies with lower utility are not discarded. They are only dispreferred. It turns out that in the game we are considering, the next strategies that are Nash equilibria (the Nash equilibria with lower expected utilities) include the ones where:

- (150) a. f_c interpreted as m_d , f_g interpreted as m_t
 b. f_g interpreted as m_d , f_c interpreted as m_t

Crucially, the first option has a higher utility value than the second one. This is because in the second strategy the unmarked form is combined with the marked meaning, and this is not true for the first one. Thus, while the best strategy is the one where neither of the two express the distributive reading with covariation (hence the dispreference for this reading for CQ-DPs as well as G-DPs) the strategy where CQ-DPs are interpreted as leading to the distributive reading with covariation, has a higher utility and thus should be preferred over the strategy where G-DPs are interpreted as leading to the distributive reading with covariation (see Appendix). This might be, I believe, behind one's preference to interpret CQ-DPs as leading to the distributive reading with covariation, rather than G-DPs.

There are still some unpleasant features in this game which we might want to get rid of. First, we had the problematic assumption that $Compl(f_d) < Compl(f_c)$. Second, and more seriously, the winning strategy is the one where f_d is interpreted as m_d and f_g is interpreted as m_g . Crucially, this does not tell us anything about how CQ-DPs are interpreted. The strategy is winning regardless whether the hearer interprets f_c

as m_d or f_c as m_g . The hearer's interpretation of f_c has no effect on the resulting value of expected utilities and thus, both interpretations are possible. This freedom leads to the undesired conclusion: we, in fact, do not derive how f_c is interpreted and therefore, we do not derive that its distributive interpretation should not be fully acceptable. This is a consequence of the fact that we considered just two meanings. We could avoid this problem if we split m_t into two sub-readings, depending on the level of precision. Consider (147a) and (147b). In the first case there could be some boys who did not participate in building the boat. In the second case all the boys without any exception built a boat. I notate the first case as $m_t[90 - 100\%]$ and the latter case as $m_t[100\%]$. The percentages express how many boys participated in the boat-building. Obviously, $m_t[100\%]$ is more marked than $m_t[90 - 100\%]$ since its probability is lower. Since f_c is the marked expression it will combine with $m_t[100\%]$, while f_g will express $m_t[90 - 100\%]$. Notice that we do not need to stipulate this, rather, it is going to follow from the game that we will consider. However, we need to assume that the probability distribution does not make $m_t[100\%]$ the most marked meaning, that is, m_d still is the most marked interpretation. If they are equally likely, any cost of m_d will suffice to make m_d the most marked interpretation. If one could show that there are good reasons to assume that $m_t[100\%]$ is less likely than m_d we might want to postulate the cost of m_d high enough to still make it the most marked reading. However, it might also turn out that in cases where the distributive reading with covariation is very likely it is so that CQ-DPs and G-DPs can be interpreted distributively. I discuss this option in the next subsection. For now on, let me assume that $m_t[100\%]$ is as likely as m_d . Using the same numbers for complexities and costs we used above, we get the Nash equilibrium with the highest expected utility to be the strategy where:

- (151) f_d interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_g interpreted as $m_t[90 - 100\%]$

As before, this gives us the correct result that (147a) and (147b) are dispreferred with the distributive interpretation. We derive this result for any game in which the following conditions are satisfied:

- (152) Assumptions for the modified $\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP} \times \text{G-DP}}$:

- a. $\text{Compl}(f_g) < \text{Compl}(f_c) < \text{Compl}(f_{ct})$, that is, *the boys* is less complex than *all the boys* which is less complex than *all the boys together*
- b. $\text{Compl}(f_g) < \text{Compl}(f_t)$, that is, *the boys* is less complex than *the boys together*
- c. $\text{Compl}(f_d) \leq \text{Compl}(f_t)$, that is, *each boy* is not more complex than *the boys together*
- d. $\text{Compl}(f_c) \leq \text{Compl}(f_t)$, that is, *all the boys* is not more complex than *the boys together*
- e. $c(m_t) < c(m_d)$, that is, a distributive reading is more costly than non-distributive readings
- f. *the boys together* does not express $m_t[100\%]$

All the conditions apart from the last one have been substantiated before. The last condition follows if we assume that *the boys together* competes with *all the boys together*, that is, there is a competition between f_{ct} and f_t . If this is so then f_{ct} ends up meaning $m_t[100\%]$ and f_t ends up meaning $m_t[90 - 100\%]$ as just another instantiation of the Horn-strategy.

Let me show that (151) constitutes the winning strategy with the assumptions given above. Obviously, in the case at hand, the best strategies are the ones where the speaker and the hearer agree on the meanings assigned to the forms and all meanings can be expressed. Let me start with the ones where f_d is interpreted as m_d . This gives us the following possibilities:

- (153) a. f_d interpreted as m_d , f_{ct} interpreted as $m_t[100\%]$, f_g interpreted as $m_t[90 - 100\%]$
 b. f_d interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_t interpreted as $m_t[90 - 100\%]$
 c. f_d interpreted as m_d , f_{ct} interpreted as $m_t[100\%]$, f_t interpreted as $m_t[90 - 100\%]$
 d. f_d interpreted as m_d , f_g interpreted as $m_t[100\%]$, f_c interpreted as $m_t[90 - 100\%]$
 e. f_d interpreted as m_d , f_g interpreted as $m_t[100\%]$, f_t interpreted as $m_t[90 - 100\%]$
 f. f_d interpreted as m_d , f_g interpreted as $m_t[100\%]$, f_{ct} interpreted as $m_t[90 - 100\%]$

(153a) and (153c) lose to (151) because of the assumption (152a). (153b) and (153c) lose to (151) because of the assumption (152b). (153d) is an anti-Horn strategy since f_g , the unmarked form, combines with the less probable meaning, and the same holds for (153e) and (153f). They are excluded for the same reason that anti-Horn strategies are excluded in $\text{GAME}_{\text{CQ-DP} \times \text{G-DP}}$. This requires the assumptions (152a,b,d).

The following list is the set of strategies in which f_c or f_g are interpreted as the distributive reading with covariation:

- (154) a. f_c interpreted as m_d , f_{ct} interpreted as $m_t[100\%]$, f_g interpreted as $m_t[90 - 100\%]$
 b. f_c interpreted as m_d , f_{ct} interpreted as $m_t[100\%]$, f_t interpreted as $m_t[90 - 100\%]$
 c. f_g interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_t interpreted as $m_t[90 - 100\%]$
 d. f_g interpreted as m_d , f_{ct} interpreted as $m_t[100\%]$, f_t interpreted as $m_t[90 - 100\%]$
 e. f_c interpreted as m_d , f_g interpreted as $m_t[100\%]$, f_t interpreted as $m_t[90 - 100\%]$
 f. f_c interpreted as m_d , f_g interpreted as $m_t[100\%]$, f_{ct} interpreted as $m_t[90 - 100\%]$

(154a) loses to (151) for the same reason that CQ-DPs are not interpreted in the game considered above, $\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP}}$ (see Appendix). To derive this, we need the assumptions (152a,b,c,e). (154b) loses to (151) because it loses to (154a) (due to the assumption (152b)). (154c) loses to (151) for the same reason that f_g is not interpreted distributively in the game considered above, $\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}$ (see Appendix). This requires the assumptions (152b), (152c) and (152e). Since (154c) loses to the winning strategy, so does (154d) because it differs from the previous one only by using the more costly form for $m_t[100\%]$. (154e) loses to (151) for the same reason that CQ-DPs are not interpreted distributively in $\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP}}$ (see Appendix) and the last strategy loses because the previous one does.

One might notice that we did not make use of the assumption (152f) so far. (152f) is present to exclude the strategy in which f_c is interpreted as m_d , f_t is interpreted as $m_t[100\%]$ and f_g is interpreted as m_t . If this strategy was allowed its utility would be equal to (151) in case the complexity of *all the boys* was equal to the complexity of *each boy* and the complexity of *the boys together*. However, there is a good reason to believe that *the boys together* cannot express $m_t[100\%]$ because this would represent an anti-Horn strategy. In particular, $m_t[100\%]$ is more marked (since less likely) than $m_t[90 - 100\%]$ and thus it should be expressed by the more marked expression *all the boys together*. Thus, we correctly derive that the winning strategy is the one where distributive quantifiers are interpreted as giving rise to the distributive reading with covariations and other expressions are left to express non-distributive readings.

As before, we also derive that the strategy in which f_c is interpreted distributively is not the winning one but is still preferred over the strategy where f_g is taken to give rise to the distributive reading since the latter case represents an anti-Horn strategy. In particular, consider the strategies with two different interpretations:

- (155) a. f_c interpreted as m_d , f_g interpreted as $m_t[100\%]$, f_t is interpreted as $m_t[90 - 100\%]$
 b. f_g interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_t is interpreted as $m_t[90 - 100\%]$

These differ in their interpretation of the first two meanings. Since m_d is the marked meaning the strategy with the first interpretation has a higher utility than the second one. If we assume that the strategy with the highest utility represents the most preferred interpretation and other strategies are also ordered by utility values we capture that the distributive reading is completely acceptable with DQ-DPs, somewhat dispreferred with CQ-DPs and even more dispreferred with G-DPs.

Finally, I want to point out that these somewhat complex games can be simplified a lot if we ignore that f_t and f_{ct} are part of the game. For instance, if we followed Katzir's approach which postulates structurally defined alternatives (see Katzir 2008), our game would consist only of three candidates: f_d , f_c and f_g . In this case, it suffices to have two assumptions: $\text{Compl}(f_c) > \text{Compl}(f_g)$ and m_d is the most marked meaning, that is, $c(m_d) > c(m_t)$ and m_d is not more likely than non-distributive readings. This derives that the winning strategy leads to the following interpretations (notice that this is identical to (151)):

- (156) f_d interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_g interpreted as $m_t[90 - 100\%]$

Notice that this also derives that following the winning strategy, (157a) will be preferred over (157b) since the latter has a lower utility than the former. This is for the same reason that in $\text{GAME}_{\text{CQ-DP} \times \text{G-DP}}$ CQ-DPs were preferred for interpreting the distributive reading with covariation.

- (157) a. f_c interpreted as m_d , f_g interpreted as $m_t[90 - 100\%]$
 b. f_g interpreted as m_d , f_c interpreted as $m_t[90 - 100\%]$

One wrinkle of this account might be that sentences with f_c can be interpreted in an imprecise way. The problem is not so severe, I believe, since this would be a dispreferred interpretation anyway. However, if we want to block it completely, we might assume that *all* has the semantic requirement for precision, as, for instance, Brisson (1998) and Brisson (2003) does. This would have only a slight consequence, namely, that (157b) would be substituted by:

- (158) f_g interpreted as m_d , f_c interpreted as $m_t[100\%]$

To conclude, we correctly derive that it is dispreferred to interpret sentences with CQ-DPs and G-DPs distributively. This result follows from the competition viewpoint, where (i) the distributive interpretation for CQ-DPs is dispreferred since there is another candidate that express precisely that, namely, DQ-DPs, and (ii) the distributive interpretation for G-DPs is dispreferred for the same reason as (i) and, furthermore, because CQ-DPs are marked expressions and thus are preferred to express the marked (distributive) reading over G-DPs. We also derive that CQ-DPs are more likely to be interpreted distributively than G-DPs if we assume that strategies which compete and lose are not discarded, only dispreferred.

Consequences of the analysis

Under this account, we predict that in languages where DQ-DPs and CQ-DPs are missing or not necessarily interpreted distributively, the distributive reading with covariation is not dispreferred with G-DPs. This is what might be going on in the acquisition of distributivity. First, I should point out that the acquisition of distributivity has not been tested to the extent that, for example, the interpretation of quantifiers or Condition B have been tested. I am aware of three experiments. Unfortunately, two of them, Miyamoto and Crain (1991) and Avrutin and Thornton (1994) do not tell us much about distributivity per se. Miyamoto and Crain (1991) studied children's knowledge of distributivity in sentence like (159).²¹

- (159) They lifted four cans.

²¹I could not get hold of the original paper. I base the description of the experiment on its summary in Thornton and Wexler (1999).

Short scenarios were acted out with toys, and children had to decide if the scenario matched a puppet's description by saying "Yes" or "No". For example, in the acted out scenario for (159) Big Bird and Ernie competed in lifting cans. Each of them lifted two cans. Afterwards the puppet said (159), which, according to Crain and Miyamoto, would be true under the collective interpretation. Children rejected the puppet's statement on average 70% of the time, where younger children rejected the statement more often and older children rejected it less often. Miyamoto and Crain (1991) take this to suggest that children prefer the distributive reading over the collective interpretation. I do not think that this conclusion is warranted. Contrary to what Miyamoto and Crain think, the statement (159) does not test the collective interpretation in the scenario. It tests the cumulative interpretation (i.e., there are some subevents in which Big Bird lifted cans, Ernie lifted cans, and in total, they lifted four cans). It is known that the cumulative interpretation is rejected to some extent also by adults (Gil, 1982), and in general, its acceptability is based on various issues like possible splits of events, as discussed in Section 3.2.5. The fact that children reject (159) then only tells us that they do not like the cumulative interpretation, for whatever reason, but does not tell us whether they prefer the collective interpretation or distributive interpretation. In fact, there was a second part of the experiment, which did test the collective interpretation. In this case, Big Bird and Ernie are trying to hold as many cans as possible. First, Ernie tries to hold all four cans, but they topple over. Big Bird comes to the rescue. He helps Ernie and together they lift all four cans. Now, the puppet's statement is *They are holding four cans*. In this case it is likely that we do not deal with the cumulative but with the collective reading of the sentence since Big Bird and Ernie act as a team. The tested children accepted this statement 89% of the time. Now, if both scenarios tested the collective interpretation then this dramatic change in responses would be surprising. But in the account here, it is not. The first part tests the cumulative reading. This is rejected by children. The second part tests the collective reading. This is accepted because the collective reading is the simplest one among the readings that pluralities can give rise to. But the experiment tells us nothing about children's ability to interpret distributive readings.

Another experiment was done by Avrutin and Thornton (1994), but it suffers from the same problem as the previous experiment. They tested children's preference between the distributive and, what they call, collective interpretation. But again, their collective interpretation is in fact the cumulative reading. They find out that when choosing between the two most children prefer the distributive reading. But if there are independent reasons why children reject the cumulative reading, and thus are left only with distributive interpretation, this result tells us nothing about the acceptability of the distributive interpretation.

The only experiment I know of which is relevant here is Brooks and Braine (1996), who tested children's preferences for distributive and collective interpretations. Unlike the previous experiments, this one did test for collective interpretations (I discussed their experiment in Section 2.2.3). They had pairs of pictures, one depicting the collective, another one depicting the distributive interpretation of a predicate, and they tested preferences for one interpretation over the other with three types of subjects, *each NP*,

all NP, *three NP*. As it turned out, while adults preferred the collective reading with the last two subjects and they accepted only the distributive reading with *each NP*, children accepted the collective interpretation with the first two subjects roughly as often as the distributive interpretation. Interestingly, unlike adults, they also accepted the collective interpretation with *each NP*. In other words, they did not differentiate between the type of subject (type of quantifier was non-significant). Table 3.15 shows the contrast in preferences between adults and 4-year olds.

Table 3.15: Percentage of answers in which subjects chose collective interpretations in Brooks and Braine (1996)

Group	<i>All NP</i>	<i>Each NP</i>	<i>Three NP</i>
4-year old	54%	37.8%	62%
adults	83.3%	0.8%	97.5%

If *each NP* is not specified for a distributive reading there is no reason it should block *all NP* from giving rise to this reading. Thus, in the account where the two compete to express a distributive reading, we not only explain why adults disprefer to interpret *all NP* this way but also why children show no preference. But why is it that children do not interpret *three NP* as giving rise to a distributive interpretation? Recall that we assumed that *all NP* is a marked expression. Should it not be then that this expression combines with the distributive reading and that leaves *three NP* with the collective reading? Blutner et al. (2006) discuss evidence that children do not consider cost on the speaker's perspective in comprehension tasks (see Blutner et al. 2006 and literature there). If this is so then the higher cost of *all NP* might be irrelevant. This does not steer kids away from comprehending *three NP* with a distributive interpretation, unlike adults.

Before concluding, I want to consider one more issue. In the games I considered, I ignored the effect of probability. I assumed that every interpretation is equally likely and the distributive reading is the marked one due to its violation of *COVARIATION. But do we really need to assume the cost of meaning? In other words, could this not follow just from the distribution of probabilities? I am quite convinced that we do need to have something extra outside of probability distributions. Consider, for example, *climbing up a tree*, *dancing one tango dance*, *baking several cakes*. It is hard to see why the distributive interpretation should be less likely than a non-distributive one.²² As the experiments discussed in Chapter 2 showed G-DPs still lead here to a non-distributive interpretation. Thus, we need to postulate the extra cost on meanings. In fact, even cases where any non-distributive reading seems very unlikely sound slightly odd when we combine them with G-DPs. This might suggest that we do not only need to assume that the distributive reading is costly but also that it is costly enough to override its low likelihood. Consider predicates like *break a finger*, *lose a tooth*, which can have a subject denoting one individual (*Perssy broke a finger*) but are degraded

²²Frazier et al. (1999) asked participants to grade likelihood of distributive and non-distributive interpretations in some of these predicates. The difference was non-significant.

with plural subjects:²³

- (160) a. ? The boys broke a finger.
b. ? Morris and Philip lost a tooth.

Linguists who argued that G-DPs can give rise to distributive readings used examples like the following ones:

- (161) a. The boys drank a beer.
b. The artichokes cost a dollar.

(161a) seems to contradict what I have just said. Here, we have a case where a non-distributive interpretation of the predicate is highly unlikely, since it does not make sense to interpret the predicate collectively and the branching reading is also odd because of the use of a verb of consumption. However, I suspect that in (161a) the distributive interpretation has something to do with the fact that *a beer* does not introduce a referent in the discourse. For example, it is odd to continue (161a) with ‘... and they liked them’, where *them* is interpreted as ‘the beers that they drank’. In this respect, the improved status of the distributive reading of (161a) as compared to, for instance, the sentence *the boys built a boat*, where the distributive reading is highly dispreferred, might not be that surprising. Similarly, *a dollar* is not a referential expression in (161b). Compare (161b) with the sentence *The artichokes were put in a small box*. In this sentence it seems to me that the distributive reading with covariation is much harder to get. If *COVARIATION is closely related to the constraint on adding referents in interpretation models, or complicating discourse models by building up quantification dependencies between referents, this effect is in fact expected. I will say a bit more about this in Chapter 5.

3.4 Alternative explanations

I am aware of four analyses of the data that I tried to account for in this chapter. I discuss them briefly here and then focus on one alternative explanation, which was suggested before (in Schwarzschild 1992).

Gil (1982) is the first to analyze the marginal status of the distributive reading (in sentences like *Two boys saw three girls*), as opposed to non-distributive readings. He

²³As Henriëtte de Swart points out to me, these sentences might improve if we substitute indefinites with definites. For instance, she finds the following example much better than (160a).

- (1) The boys broke their wrist.

Further testing might be necessary to see how indefinites differ from definites with respect to covariation. Before concluding anything about the distributive reading with covariation from (1), we also need to explain why the plural morphology on the pronoun *their* is possible. If the plurality was interpreted on *their* and we dealt with the distributive reading with covariation where the subject distributes over the rest of the clause, we should get an uninterpretable sentence, comparable to the infelicity of *The boy₁ broke their₁ finger* (the indices show the relevant interpretation). This might suggest that some other mechanism than the one which gives rise to the distributive reading with covariation is used in the interpretation of (1). I leave this issue for further research.

stipulates that particular readings are degraded with particular expressions. This model captures the data but there is no explanation why it should be the way it is and why, for example, the situation is not the other way round, with non-distributive readings marginal and distributive readings preferred.

Schwarzschild (1992) assumes that the marginal status of sentences like *John and Mary own a car* is due to the fact that indefinites take scope higher than the distributive operator.²⁴ This explains the basic data but offers no reason why it should be so. Besides, when this analysis is fully flashed out it runs into problems because it predicts scope relations that are not attested. I will show that in the next section (Section 3.4.1).

Moltmann (1992) suggests that C_* cannot be inserted with definite plurals. This account could be extended to other G-DPs. However, somewhat marginal status of distributive readings with *all NP* is unexplained, given that *all NP* requires the presence of C_* . Furthermore, I do not see why we would only deal with low preferences, not ungrammaticality if the insertion of C_* was simply impossible. Another problem is the acquisition of distributivity. Finally, in Chapter 5 I argue that the bound reading of *same* requires distributivity. The bound reading of *same* in (162) is the reading where Morris read the same book as Philip did (but there is no book in the discourse to which *same* is referring).

(162) Morris and Philip read the same book.

The bound reading of *same* can be licensed in the scope of G-DPs (coordinations of proper names, definite plurals). Thus, it seems that there is nothing wrong with the insertion of C_* in these cases, which makes Moltmann's assumption problematic.

Finally, Kaup et al. (2002) argue that distributive readings are degraded with plural expressions that are non-partitive. I am not completely sure what they mean by 'partitive' even though it seems from their discussion that the crucial property is that a quantifier/determiner can appear in a partitive construction headed by *of*. They compare *the*, which cannot head the partitive construction (**the of the boys*), and consequently, definite plurals are degraded with the distributive reading, with *most*, which can head the partitive construction (*most of the boys*) and which is fine with the distributive reading. I do not think that their descriptive generalization is correct. Numerals can head the partitive construction (*three of the boys*) but are degraded with the distributive reading (Gil, 1982). On the other hand, *every* cannot head the partitive construction but is perfectly acceptable with the distributive reading.

As far as I know, these are all analyses that have been proposed for the data I was concerned with here. In the rest of the section I focus on the one that has been proposed in Schwarzschild (1992).

3.4.1 Explanation in terms of scope preferences

In this section I want to contrast the presented analysis with an alternative, according to which distributive readings are degraded as a result of scope preferences. This

²⁴To be fair, Schwarzschild offers this only as a suggestion of accounting for the data. His paper focuses on other issues connected to distributivity.

alternative analysis is based on Schwarzschild (1992).

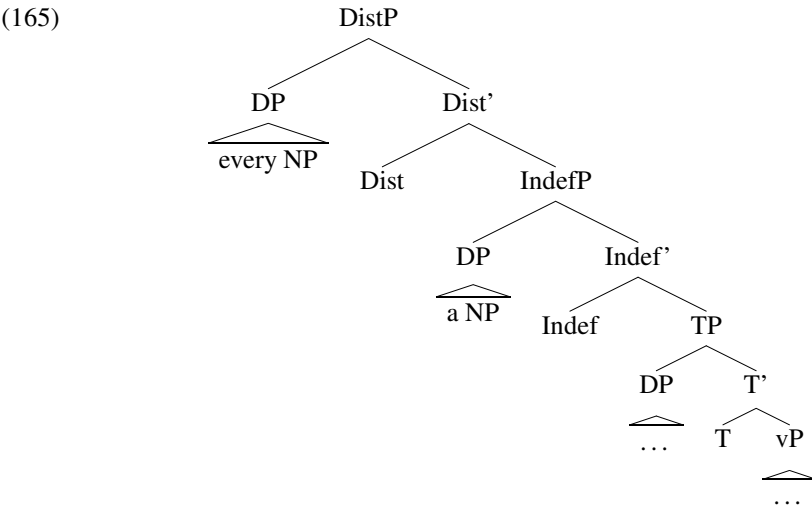
Under this account, the distributive reading is dispreferred in (163) because indefinites tend to take wide scope, higher than the subject's scope. Notice that this scope option should only be preference, and not an obligatory requirement, given the data in Chapter 2.

(163) The boys built a boat.

Clearly, such an account does not explain why the distributive reading in (164) is accepted by everyone. Therefore, we need to add that universal distributive quantifiers in the subject position scope even higher than indefinites. Finally, recall that DPs headed by *all* give rise to distributive readings slightly less preferably than DPs headed by *every*. There might be various ways to accommodate it but I will ignore this issue here.

(164) Every boy built a boat.

One way to capture this in the framework here would be to posit the functional hierarchy, $\text{Dist} > \text{Indef} > \text{T}$, as shown in (165). *Every NP* moves to the specifier of *Dist*, indefinites raise into the specifier of *Indef*, and definite plurals stay in the subject position. This is at least the preferred option, even though DPs might have other possibilities (like taking scope in their case positions). This is similar to Beghelli and Stowell's way of analyzing QR, even though the hierarchy here differs from the one that they propose (see Beghelli and Stowell 1997, Szabolcsi 1997a and other papers in Szabolcsi 1997b).



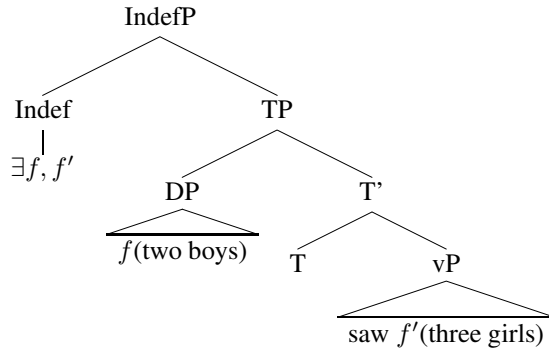
One relevant issue is what happens in case we have two indefinites in a clause, as in (166).

(166) Two boys saw three girls.

These data have been discussed in Section 2.2.2, where we saw that distributive readings are less preferred than non-distributive readings. Crucially, the inverse scope distributive reading is even less preferred than the surface scope distributive reading. But that is problematic for the given structure. As it is right now, we get *Obj* > *Subj* (the inverse scope reading, in case we assume the object indefinite rises to IndefP and the subject stays in TP) or *Subj* > *Obj* (the surface scope reading, in case the object stays in situ, or in case both indefinites rise and preserve the order of c-command). In any case, we are not able to get non-distributive readings. Of course, we could get non-distributive readings if we assumed that one indefinite takes scope over the other without distributing over it, but then the question would be why it cannot distribute – and this would push us back to the original explanation. If we want to pursue the alternative here we need to make sure that dispreferred distributive reading in (166) is also explained in terms of scope. In fact, this can be done. We need to make sure that indefinites target Indef but they distribute only lower in the structure. Thus, we need to split their scope into two parts: existential scope (what indefinites are scopally dependent on) and distributive scope (what scopally depends on distributive indefinites). Indef is the marker only of the first scope. This makes sure that we get non-distributive readings in (166) as the preferred reading.

Here is nowadays the most well-known way to deal with this fact. To analyze indefinites, we add choice functions, where choice functions are functions that choose a member from any non-empty predicate (Reinhart, 1997; Winter, 1997). IndefP is not a target of scope movement but a phrase in which choice functions are existentially closed. Furthermore, while QR cannot cross clause boundaries Indef can close any choice functions within its scope (deriving the fact that the scope of existentials is not restricted by syntactic islands, see Ruys 1992). This gives us the following structure for (166), interpreted in (167b). This derives the non-distributive reading, i.e., there are two boys and there are three girls, and each of the boys saw some of the girls and each of the girls was seen by some of the boys.

(167) a.



$$b. \quad (\exists f, f', e) \left(\begin{array}{l} C_{*SEE}(e) \wedge C_{*\Theta_1}(f(TWO\ BOYS))(e) \wedge \\ C_{*\Theta_2}(f(THREE\ GIRLS))(e) \end{array} \right)$$

So far so good. But now, the analysis leads us to expect that distributive readings are preferred less than non-distributive readings only when DPs in the scope of distributive

operators can be bound by choice functions. This is not correct. In the questionnaire that Øystein Nilsen and I conducted (see Section 2.3) we used objects with modified numeral DPs in some of the test items. One example is given here, with the modified numeral DP *exactly one unit of measurement*:

- (168) De steden hebben precies één lengte-eenheid.
 the towns have exactly one length-measure
 'The towns have exactly one length-unit.'

(168) was judged against the background that each town has its own unit of measurement, thus, requiring the distributive reading. Still, (168) was not accepted more than examples in which indefinites were used as objects. We cannot analyze this datum in the same way as indefinites because choice functions do not apply to modified numeral DPs (Reinhart, 1997). Consider the following pair. (169a) has a reading in which the indefinite *three relatives of mine* scopes outside of the conditional antecedent. This reading can be paraphrased as 'there are three relatives of mine such that if they die I will inherit a house'. In other words, it is not true that as long as any three relatives die I will inherit a house (narrow scope reading). Rather I will inherit a house only if some particular relatives of mine die, for example, my three uncles. Such reading seems to be missing in (169b). Since the wide scope in (169a) is commonly derived by postulating choice functions (see the discussion in Reinhart 1997) this is an argument against using choice functions in modified numeral DPs.

- (169) a. If three relatives of mine die I will inherit a house.
 b. If exactly three relatives of mine die I will inherit a house.

Thus, a choice functional analysis does not seem viable for modified numeral DPs. Now, in order to account for (168) we have, I believe, only one option, namely to assume that modified numeral DPs also scope above TP but below Indef. This gives us the hierarchy Dist>Indef>Mod>T (where Mod is the head hosting modified numerals).

As I mentioned above, we end up with a hierarchy of projections for quantifiers which is similar in spirit to the work of Beghelli and Stowell but different in details. I find one particular difference quite worrisome. Beghelli and Stowell argue that modified numeral DPs should always have a very narrow scope, not higher than their case position. This is for a good reason. In (170) *fewer than three books* must scope below the universal quantifier. If it could have wide scope the sentence would be true if every boy read for example ten books, as long as they all read the same two books. It is a very robust judgement that this reading is missing in (170) (see Beghelli and Stowell 1997).

- (170) Every boy read fewer than three books.

But in the account here it was crucial that modified numeral DPs scope higher than the subject. Now, there is only one option to keep this assumption and still account for the lack of inverse scope in (170). We could say that universal quantifiers take scope only in DistP. But this cannot work, either. We know that universal quantifiers can reconstruct as can be shown in the following example.

(171) Every player didn't score.

(171) is ambiguous. Crucially, we cannot derive the ambiguity by assuming that the negation moves around the universal quantifier because negation is normally taken not to vary its scope position, as witnessed by the fact that when it appears together with adverbs they all have a frozen scope determined by their surface position (see Sportiche 2005 and references therein). This suggests that it is rather the universal quantifier that can reconstruct below negation what enables the ambiguity in (171). Thus, universal quantifiers reconstruct. But where to?

Negation is rather interesting because it also takes scope below modified numeral DPs. (172) can be true if there was a football team in which, for instance, five players did not score and six players scored. This is the surface scope reading, in which negation takes scope below the subject.

(172) More than three players didn't score.

Now *every NP* can take scope below negation, and negation is at least below the highest position which modified numerals can take. This is ModP in our setup. Thus, the universal quantifier should be able to undergo reconstruction below ModP. In fact, it is commonly taken that reconstruction under negation involves reconstruction into thematic positions (Beghelli and Stowell, 1997). At any rate, since the universal quantifier can take scope lower than ModP this would also derive the inverse scope reading in (170), which is not what we want. Clearly, we could circumvent this situation if we did not postulate any scope positions for modified numerals, and let them scope only in their case positions, which is precisely what Beghelli and Stowell (1997) argue for. But in that case the object modified numeral DPs cannot scope above the subject and the fact that people disprefer the surface scope distributive reading in (168) cannot be accounted for using scope hierarchies.

There is another argument against the account of the data in terms of scope hierarchies. As we saw above modified numerals cannot take exceptional wide scope. Thus, if the preference for non-distributive readings was the result of the preference to have modified numerals take scope above the DP that is interpreted distributively, this preference should appear only when the distributive DP and the modified numeral are in the same clause. I think that this is not correct. Two examples here show why. The sentences (173a) and (174a) can only make sense in the contexts above when the matrix subject is interpreted distributively. However, they do sound slightly degraded in this reading, unlike (173b) and (174b) in which the matrix subject is headed by a universal distributive quantifier. I did not test it so thoroughly as the issues discussed in Chapter 2 but a few subjects that I asked confirmed this intuition. This is completely unexpected under the approach that tries to explain the preference for non-distributive readings by assuming that indefinites and modified numerals take higher scope than the distributive DP since in (173a) and (174a) the modified numeral DPs cannot take scope above the matrix subject.

(173) Mary has two dogs, a dachshund and a German shepherd, and she usually walks them separately. The other day, John was on a walk when he saw that Mary walked a dachshund in a park. Later on, Mary got home with the dachshund

and went out to walk her other dog, the German shepherd. At that moment Bill met her and saw her walk the German shepherd.

- a. The boys saw that Mary walked exactly one dog.
- b. Each boy saw that Mary walked exactly one dog.

(174) The husband of Susanne suspected that she cheated on him. He wanted to confirm that and asked two sleuths to check on his wife. Sherlock, the first sleuth, saw Susanne kiss a blonde, handsome guy on a street. Further than that, he did not see anyone else suspicious. Hercule, the second sleuth, saw Susanne kiss a guy on a street, too. But this guy was a different one than the guy that Sherlock saw. He was black-haired and pretty ugly. Further than that, Hercule did not see anyone else suspicious. Later on, both sleuths reported to the husband about the kissing they each saw.

- a. The two sleuths informed the husband that Susanne kissed exactly one guy on a street.
- b. Each sleuth informed the husband that Susanne kissed exactly one guy on a street.

I think that ultimately, scope preferences do play a role since there are indefinites (not discussed here) which obligatorily take the highest scope and thus are not distributed over. But scope preferences cannot be the sole explanation of the data, especially because this explanation runs into problems when the object is a modified numeral DP. This shows that some extra account is needed anyway. The account I offered in Section 3.3 is well-suited for that matter.

Another problem comes from the data concerning *same*. *Same* has a bound reading in the scope of a plural argument. This reading is also possible if the plural argument is a coordination of proper names or a definite plural. The bound reading of *same* in (175) can be paraphrased as ‘Morris read the same book as Philip did’ (but there is no book in the discourse to which *same* is referring).

(175) Morris and Philip read the same book.

On the other hand, the bound reading of *different* is degraded with coordinations of proper names or definite plurals. (176) normally gets only the discourse-anaphoric reading that there is some book and Morris and Philip read a book different from that one.

(176) Morris and Philip read a different book.

I will argue in Chapter 5 that in order to get the bound reading, *same* or *different* must be in the scope of a plural distributing argument. Under Schwarzschild’s analysis, to distinguish between the behavior of (175) and (176) one would have to stipulate that *different* has to scope out and *same* does not. I do not see any reason why this should be so. On the other hand, under the account I proposed the difference is due to the fact that *same* does not introduce covariation, unlike *different*.

3.5 Conclusion

In this Chapter I focused on two issues. First, I introduced the formal semantics' system of pluralities and events in which we could formalize the discussion about distributive and non-distributive readings. I introduced the system in Section 3.2. I will use the same system in the next chapter, when I turn to expressions of (non-)identity which get the bound reading when they are in the scope of a plural argument.

The second goal was to explain why various DPs differ in their ability to license distributive readings. I offered an analysis in Section 3.3. The analysis was based on the assumption that forms can compete with each other for their interpretations which narrows down their interpretation possibilities. Apart from accounting for the adult data, the analysis is potentially revealing in the explanation of the acquisition of distributivity. In Section 3.2.5 I compared my analysis to the analyses that were so far suggested in semantic and psycholinguistic literature and discussed shortcomings that these analyses face.

4.1 Introduction

After setting up the system for plurals, I will now move to the study of the expressions of (non-)identity. In this chapter, I focus on *the others* and the reciprocal *each other*. In the next chapter, I turn to the adjectives *different* and *same*. What these expressions have in common is their ability to be anaphoric to a previously introduced argument.

Anaphoric expressions pose a challenge to natural language semantics because they multiply the usage of lexical resources in interpretation. In normal cases, each lexical expression is interpreted exactly once. For example, (1) is interpreted as “Morris kissed Hilary”, and it cannot be interpreted as “Morris kissed Morris”, “Hilary kissed Hilary”, in which case one or more overt lexical expressions are not interpreted exactly once.

- (1) Morris kissed Hilary.

The fact that lexical expressions are interpreted exactly once follows implicitly from the way the interpretation of LF is set up (see Section 3.2), or, as in Type Logical Grammars, it is made explicit, by using “substructural logics” which cannot subtract or add information because the structural rules monotonicity and contraction, which are assumed in Classical Logic, are not valid (Linear Logic, Lambek calculus).

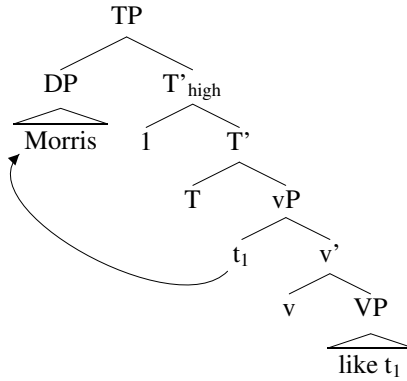
In this light, (2) comes as a surprise because here *Morris*, present only once in the clause, is in fact interpreted twice.

- (2) Morris likes himself.
‘Morris likes Morris.’

Obviously, the culprit of this double interpretation of *Morris* is the anaphoric expression. There are two strategies to deal with the multiplication of lexical resources caused by anaphora. In the first one, the multiplication is enabled in the syntax. In the second approach, the multiplication of resources is stipulated in the lexicon.

The first approach is commonly assumed in the generative grammar framework. One usual way is to translate anaphora as a variable which is bound by a higher lambda operator (see, for example, Heim and Kratzer 1998). (2) gets the structure in (3a). Since *himself* is interpreted as a variable bound by the lambda operator, T'_{high} is interpreted as in (3b), which, when combined with the subject, correctly derives the interpretation “Morris likes Morris”, see (3c). Furthermore, this approach does not overgenerate in cases like (1) simply because no overt lexical item there is interpreted as a variable.¹

(3) a.



b. $[[T'_{\text{high}}]] = \lambda x \lambda e. C_* \Theta_1(x)(e) \wedge * \text{LIKE}(e) \wedge C_* \Theta_2(x)(e)$

c. $[[\text{TP}]] = (\exists e)(C_* \Theta_1(\text{MORRIS})(e) \wedge * \text{LIKE}(e) \wedge C_* \Theta_2(\text{MORRIS})(e))$

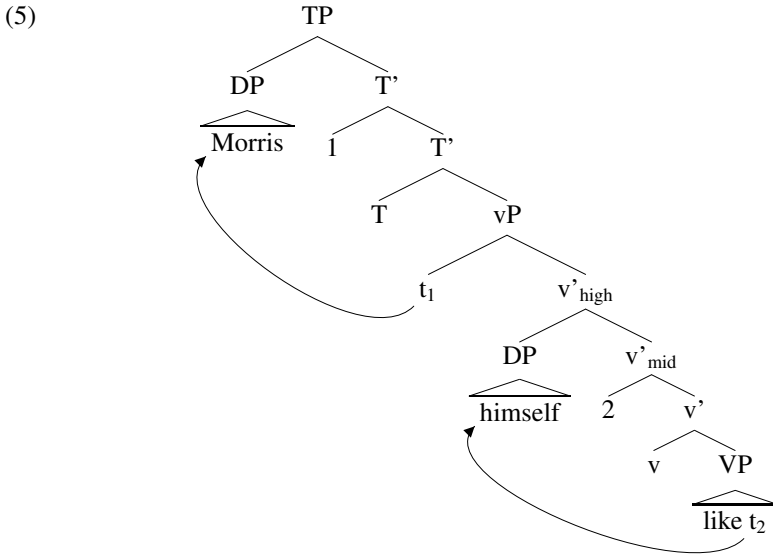
The second approach, in which the multiplication of lexical material is postulated in the lexicon, has appeared in Partee and Bach (1981). A more detailed analysis of this approach is due to Anna Szabolcsi (Szabolcsi, 1989, 1992). Szabolcsi focuses on reflexive anaphors, which she takes to be argument reducers: they take a relation and return a predicate, $\lambda R \lambda x. Rxx$. Since the system that we set up makes use of events, we need to modify the lexical semantics of *himself* to include an event argument. This is given in (4).

(4) $[[\text{himself}]] = \lambda R_{\langle e, \langle e, \langle v, t \rangle \rangle \rangle} \lambda x \lambda e. Rxxe$

Szabolcsi derives anaphora resolution in combinatory categorial grammar. In the system I introduced in Chapter 3, we can derive the meaning of (2) if we let *himself* undergo movement to scope over the relation of liking. This is shown in the tree in (5), where T'_{high} and TP receive the interpretation identical to T'_{high} and TP in (3a) as the reader can check. Unlike in the previous case, though, *Morris* is interpreted

¹ Another way still within the realm of the first approach is common in categorial grammars (Hepple, 1990; Jacobson, 1999; Jäger, 2001). Since their discussion would take me too far afield, I refer the interested reader to the actual papers, or to the summary in Jäger (2001) and Dowty (2007).

twice in the clause because it is an argument of *himself* and as such is applied twice to the relation. Other analyses which derive anaphora resolution by multiplication in the lexicon include Reinhart and Reuland's treatment of reflexives (Reinhart and Reuland, 1993), Moortgat's approach to anaphora (Moortgat, 1996), Reuland's recent work with Winter (Reuland and Winter, 2009), among others.



- (6)
- a. $[[v'_{mid}]] = \lambda x \lambda y \lambda e. {}^{C_*}\Theta_1(x)(e) \wedge *LIKE(e) \wedge {}^{C_*}\Theta_2(y)(e)$
 - b. $[[v'_{high}]] = \lambda x \lambda e. {}^{C_*}\Theta_1(x)(e) \wedge *LIKE(e) \wedge {}^{C_*}\Theta_2(x)(e)$
 - c. $[[TP]] = (\exists e)({}^{C_*}\Theta_1(MORRIS)(e) \wedge *LIKE(e) \wedge {}^{C_*}\Theta_2(MORRIS)(e))$

In the second approach the reflexive is a (polyadic) quantifier. Its argument is a relation and the reflexive requires the relation to apply twice to the same argument. Thus, it behaves as an argument reducer: it shifts relations to predicates. Since the multiplication of lexical resources is part of its meaning the only necessary operation for the interpretation of sentences with anaphora is a scope-taking mechanism (QR).

In the first approach, the reflexive is interpreted as a variable and does not need to undergo QR. However, in order to interpret the variable in the right way we need some extra mechanism that enables its binding. Standardly, Predicate Abstraction is taken to achieve this goal. Predicate Abstraction also always binds a trace in an argument position. Thus, since it binds two variables we end up multiplying lexical resources. However, in this case the multiplication is not only due to the lexical meaning of anaphora but also due to the mechanisms postulated in the syntax. QR, on its own, does not suffice. We also need a binding-in mechanism for variables introduced by

anaphora.^{2,3}

Both approaches to anaphora found their way into the study of reciprocal expressions like *each other* and its cognates in other languages. Reciprocals add another complication to the study of anaphora. Consider (7). It does not suffice to say that *each other* is interpreted as THE SOCCER PLAYERS (i.e., parallel to *himself*, but anaphoric to plural arguments) because in that case it would be left unclear who hated whom, which is, however, specified in (7). In particular, (7) is not true if each soccer player hated himself, which would be true if we only required that EACH OTHER=THE SOCCER PLAYERS.

- (7) The soccer players hated each other.

A successful analysis of reciprocals should make sure that the reciprocal is anaphoric to a plural argument *and* specify how members of the plural argument are related to each other. This has been done in two strategies. In the first one, reciprocals are analyzed as expressions doubly anaphoric: to the plural argument and to the members of the plural argument. This approach, exemplified by Heim et al. (1991a,b); Roberts (1991); Schwarzschild (1996); Sternefeld (1998); Beck (2001) among others, will be referred to as the *doubly-anaphoric approach* to reciprocals, or the DA-approach for short. The second approach takes reciprocals to be argument reducers, like reflexives. They differ from reflexives, though, because they furthermore specify what relations among the members of the plural argument are possible. This approach has been proposed in Dalrymple et al. (1998), Sabato and Winter (2005b) and others. Thus, we currently have two ways to deal with reciprocals. One might then wonder whether there are arguments that may force one to choose one theory over the other. Another issue is whether the R-approach and DA-approach can be broadened to other anaphoric expressions. One lexical item which could be studied in this respect is *the others* (or *the other NP* with a specified noun phrase). *The others* gives rise to the same reading as reciprocals when it is c-commanded by a plural antecedent which is interpreted distributively. I call this the bound reading of *the others*, and it is exemplified in (8). As one can see its reading is the same as the one with reciprocals, (7).⁴

- (8) Each soccer player hated the others/the other soccer players.

²There is another distinction made between multiplication in the lexicon and syntax which is irrelevant for the one that I make. Reinhart and Siloni (2005) argue that in some cases predicates are specified in the lexicon as having one of their thematic roles anaphoric to another one. In this case the anaphoric relation is expressed without filling in the anaphoric thematic role in the syntax. Reinhart et al.'s distinction is orthogonal to the one I pursue here and I will leave it aside.

³The two approaches can be distinguished in other frameworks than generative grammar. For example, in Jäger's Lambek Calculus with Limited Contraction (Jäger, 2001), which is the extension of Type Logical Grammar studied by Lambek, the *q*-operator, modelled after Moortgat (1996), is necessary if we assume the multiplication of resources in the lexicon, while \downarrow , a new implication introduced by Jäger, achieves binding "in the syntax".

⁴Apart from the bound reading, *the others* also has a discourse reading. In this case, no paraphrase with *each other* is possible.

- (1) The coach was angry with his soccer players. The goal keeper did not come to the training and the others were present, but completely drunk.

To facilitate the following discussion, I call *each other* and *the other(s)* **reciprocated anaphora**.

In this chapter, I argue that the R-approach and DA-approach should be maintained for the reciprocated anaphora. There are empirical arguments that can distinguish the two approaches and these show that the R-approach should be exclusively used for *each other*, while the DA-approach correctly captures the properties of *the others*. Furthermore, as we will see it is necessary that the R-approach combines with the multiplication of resources in the lexicon, while the DA-approach combines with the multiplication of resources in the syntax. This leads us to the conclusion that both ways of anaphoricity are present in language.

In the next section I briefly present the R- and DA-approaches to *each other*, and after that I shortly summarize their different predictions. In Sections 4.3-4.6 I show in detail that the predictions made by the two approaches force us to conclude that the R-approach works well for *each other*, while the DA-approach works better for *the others* but fails to account for *each other*.

4.2 Two research strategies on reciprocity

4.2.1 The R-approach to reciprocals

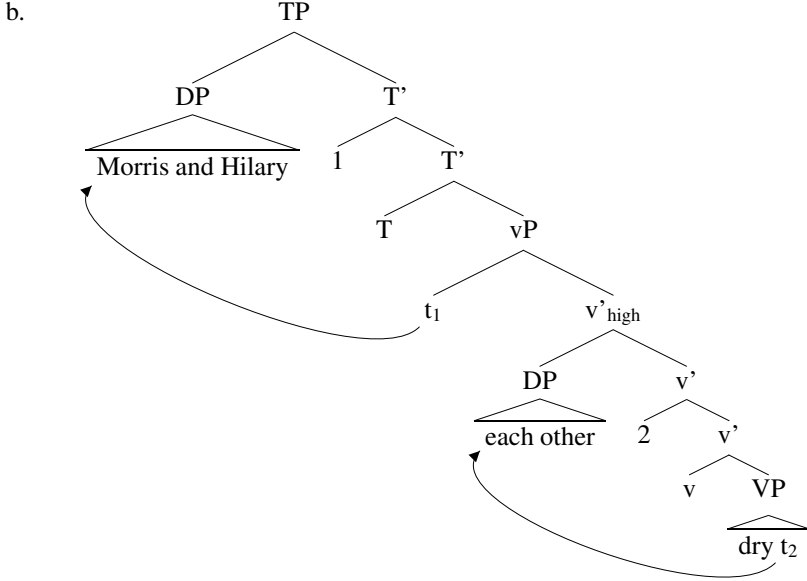
According to the R-approach, reciprocals multiply lexical resources in the lexicon, but they differ from reflexives. While reflexives state that the same argument applies to the relation twice, reciprocals state that the *parts* of the same argument apply to the relation. In Dalrymple et al. (1998) and Sabato and Winter (2005b) which parts of the plural argument apply to the relation is depends on the properties of the actual relation. I will come back to this point later. Let me start first with a simple meaning of *each other* in which every part of the plural argument is related to every other part, apart from the identical parts. The lexical semantics is given in (9). The type is identical to the reflexive as analyzed by Szabolcsi and others (4).

- (9) Interpretation of *each other*, the R-approach (to be modified)

$$[[\text{each other}]] = \lambda R \lambda x. \lambda e. (\forall y, z \leq x) (\exists e' \leq e) (y \neq z \wedge y, z \in AT \rightarrow R y z e')$$

The semantics requires that *each other* takes scope over a relation. This is done by letting the reciprocal QR and by abstracting over a predicate (one can see that the structure is the same as with reflexives in Szabolcsi's account):

- (10) a. Morris and Hilary dried each other.



The meaning of (11a) is in (11b). (11b) is true if Morris dried Hilary and Hilary dried Morris (but is not true if, for example, Morris dried himself and Hilary dried herself, which is what we want).

- (11) a. Morris and Hilary dried each other.
 b. $(\exists e)(\forall y, z \leq \text{MORRIS} \oplus \text{HILARY})(\exists e' \leq e)$
 $(y \neq z \wedge y, z \in AT \rightarrow *DRY(e') \wedge *\Theta_1(y)(e') \wedge *\Theta_2(z)(e'))$

4.2.2 The DA-approach to reciprocals

Another analysis of reciprocals assumes that they are doubly-anaphoric to their plural DP antecedent. Probably the best known approach that adheres to this viewpoint is Heim et al. (1991a), and Heim et al. (1991b) but the same position has been considered before (Fiengo and Lasnik, 1973) and supplemented with more details throughout the late nineties (Schwarzschild 1996, Sauerland 1998, Sternefeld 1998, Beck 2001). Here, I consider Beck's account of *each other*, for its simplicity and empirical coverage.

In Beck's account, *each other* is treated as a plural definite, with two free variables, viz. the *range argument* and the *contrast argument* (see (12)). Following my convention from the previous chapter, I will treat *each other* as a generalized quantifier. z and y in the formula are the range argument and the contrast argument, respectively, and σ is the maximality operator, i.e., an operator that picks the maximal element in the set, as was discussed before (see (43) in Section 3.2.4).

- (12) Interpretation of *each other*, the DA-approach
 $[[\text{each other}]] = \lambda P.P(\sigma(\lambda x.x \leq z_{\text{range}} \wedge \neg x \circ \downarrow y_{\text{contrast}}))$

(12) denotes the set of properties that hold of the maximal entity which is part of z_{range} (*the range argument*) and does not overlap with the entity $y_{contrast}$ (*the contrast argument*). Consider a situation in which z_{range} is a group of students and $y_{contrast}$ is one particular student, Angelica. In this case the meaning of *each other* corresponds to ‘all the students apart from Angelica’. The meaning of *each other* seems to be completely parallel to the meaning of *the others*, which also picks the maximal entity different from $y_{contrast}$ and among z_{range} ,⁵ as can be seen in the example (13) where *the others* picks two antecedents, ANGELICA and THE STUDENTS and is interpreted as ‘all the students apart from Angelica’. I say more on the comparison between *each other* and *the others* later.

- (13) The students are in general not very industrious. Angelica hands in all the assignments but **the others** usually forget.

The meaning of *each other* given in the DA-approach correctly works when the contrast argument is bound locally and the range argument is bound by the same argument as the contrast argument, see (15). Thus, we add the following restriction to (12):

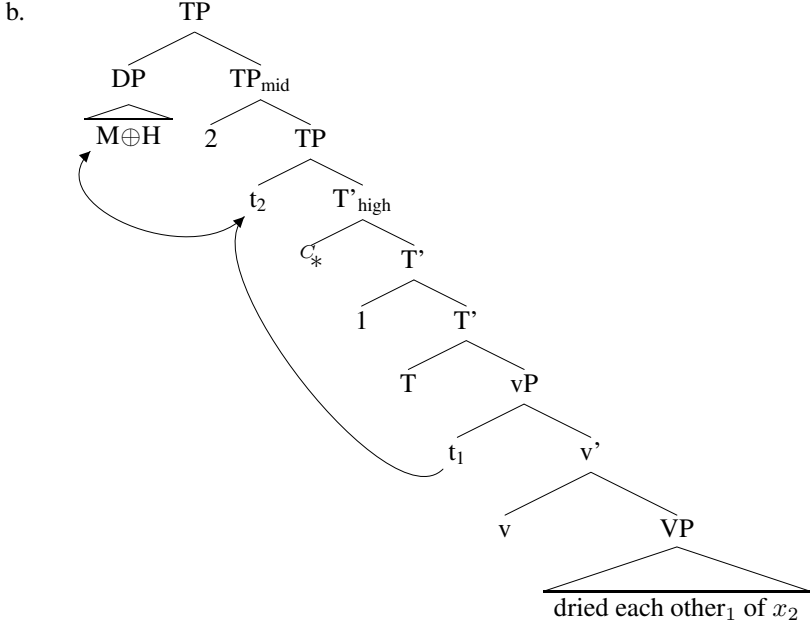
- (14) The contrast argument of *each other* must be bound in a local domain

For the purposes of this chapter, it suffices to assume that the local domain in (14) corresponds to a clause.

The contrast argument is bound by the lambda-operator which is in the scope of the * operator (1 in (15)). The range argument is bound by the highest lambda-operator (2 in (21)). The second binding leads to the interpretation of the range as MORRIS AND HILARY. The first binding leads to the interpretation of the contrast as ‘each of Morris and Hilary’. The resulting interpretation can be loosely paraphrased as ‘each of Morris and Hilary dried the other one among Morris and Hilary’.

- (15) a. Morris and Hilary dried each other.

⁵In fact, Nouwen (2003), who discusses the discourse reading of *the others* assumes a lexical meaning of *the others* that is virtually identical to (12).



The double binding of *each other* arises in the following way. First, the plural argument *Morris and Hilary* moves (by A-movement) to the subject position (Spec, TP). As discussed in Section 3.2 the moved phrase leaves a trace in its original position (t_1) and at the point where it takes scope a new node labeled as the numerical index (1) is introduced in the structure. This node is interpreted as lambda abstraction. The newly introduced lambda-operator binds every occurrence of the variable subscripted with the same index, which includes the contrast variable of *each other*. The predicate, which came to existence through the lambda abstraction, is the argument of $*$. Thus, T'_{high} is interpreted as (16a). The $[]$ in the formula creates a sub-lattice, i.e., works as a pluralizer and the whole formula can be simplified into the more readable (16b). For discussion on why this simplification is possible, see Section 3.2.5 and especially the definition of pluralization, (66) in Chapter 3. The contrast variable is bound at this point, the range variable remains free so far.

- (16) a. $[[T'_{high}]] = \lambda v.v \in [\{\downarrow y : (\exists e)(*dry(e) \wedge C_*\Theta_1(y)(e) \wedge C_*\Theta_2(\text{each other}_y \text{ of } z))(e)) \wedge y \in Cov\}]$
- b. $[[T'_{high}]] = \lambda v.(\forall y \in Cov)(\downarrow y \leq v \rightarrow (\exists e)(*dry(e) \wedge C_*\Theta_1(y)(e) \wedge C_*\Theta_2(\text{each other}_y \text{ of } z))(e)))$

Next, the plural argument should be the argument of the predicate. But to achieve binding of z_{range} , the plural argument moves one more time (QR). The trace left behind is bound by the higher lambda operator, as well as the range variable. Thus, TP_{mid} is interpreted as (17a). The interpretation is identical to (16b) apart from the fact that the range argument is bound. Finally, the predicate applies to the plural argument *Morris and Hilary*, giving (17b). (17c) is the same as (17b) but with the lexical semantics of *each other* filled in.

- (17) a. $[[\text{TP}_{\text{mid}}]] = \lambda z. (\forall y \in \text{Cov})(\downarrow y \leq v \rightarrow (\exists e)(\text{*dry}(e) \wedge C_{*}\Theta_1(y)(e) \wedge C_{*}\Theta_2(\text{each other}_y \text{ of } z)(e)))$
- b. $[[\text{TP}]] = (\forall y \in \text{Cov})(\downarrow y \leq M \oplus H \rightarrow (\exists e)(\text{*dry}(e) \wedge C_{*}\Theta_1(y)(e) \wedge C_{*}\Theta_2(\text{each other}_y \text{ of } M \oplus H)(e)))$
- c. $[[\text{TP}]] = (\forall y \in \text{Cov})(\downarrow y \leq M \oplus H \rightarrow (\exists e)(\text{*dry}(e) \wedge C_{*}\Theta_1(y)(e) \wedge C_{*}\Theta_2(\sigma(\lambda x.x \leq M \oplus H \wedge \neg x \circ \downarrow y_{\text{contrast}}))(e)))$

Assume that *Cov* consists of atomic individuals. Then, (17c) is identical to the interpretation that the R-approach gives (see (11b)). It is true if Morris dried Hilary and Hilary dried Morris, but false if, for example, each of them dried himself/herself. This is what we want.

4.2.3 Properties and predictions of the two strategies

The R-approach and DA-approach differ in two respects. First, in the R-approach reciprocals “are anaphoric” to their antecedent once, unlike in the DA-approach. Second, the R-approach postulates multiplication of resources in the lexicon, and the DA-approach achieves multiplication of resources by binding in the syntax. We can see this since the R-analysis of *each other* does not deal with any free variables that need to be bound by some external mechanism. In fact, QR suffices for the interpretation of the sentence. In the DA-analysis, some mechanism that can achieve binding of free variables is crucial for the correct interpretation.

One might wonder whether the two properties are related. For example, could the reciprocal be anaphoric just once but the multiplication of resources be achieved by lambda-binding? Or could the reciprocal be doubly anaphoric, as in DA-approaches, but the multiplication of resources be postulated in lexicon? In fact, neither option is possible.

First, if the reciprocal was anaphoric to the plural argument and the binding would be achieved by Predicate Abstraction, we would get (18) as its interpretation, i.e., a variable to be bound by some plural argument (recall that the underlined part expresses a presupposition). While something like this might suffice for the semantics of reflexives and pronouns it is not sufficient for the semantics of *each other*. The reason is that *each other* also expresses how atomic parts of the plural argument are related and for this reason it must operate on relations.

- (18) Incorrect interpretation of *each other*
 $[[\text{each other}]] = \textit{x is not atomic.x}$

Second, the DA-approach is incompatible with multiplication of resources in the lexicon. Recall the meaning of *each other* in the DA-approach:

- (19) Interpretation of *each other*, the DA-approach (repeated)
 $[[\text{each other}]] = \sigma(\lambda x.x \leq z_{\text{range}} \wedge \neg x \circ \downarrow y_{\text{contrast}})$

To combine this with the approach that achieves the multiplication of resources in the lexicon, the free variables would have to be bound by argument reducers. This

gives us either (20a) or (20b) depending on whether we bind the range or the contrast argument.

- (20) a. Incorrect interpretation of *each other*
 $[[\text{each other}]] = \lambda R \lambda z \lambda e. R(z)(\sigma(\lambda x. x \leq z_{\text{range}} \wedge \neg x \circ y_{\text{contrast}}))(e)$
 b. Incorrect interpretation of *each other*
 $[[\text{each other}]] = \lambda R \lambda y \lambda e. R(y)(\sigma(\lambda x. x \leq z_{\text{range}} \wedge \neg x \circ y_{\text{contrast}}))(e)$

Neither (20a) nor (20b) will work because in both of them one variable remains free which still calls for binding in the syntax.

The R-approach and DA-approach focus on the semantics of reciprocals. I want to broaden their perspective and see if they can also deal with *the others*, which, as we have seen, can give rise to the reading parallel to the interpretation of reciprocal sentences. For this reason, I will talk about the R-approach and DA-approach to reciprocated anaphora.

As we have seen, reciprocated anaphora connects the way we deal with multiplication of resources with the number of individual arguments the anaphor requires: either one, as in R-approaches, or two, as in DA-approaches. The different number of arguments the anaphor requires leads to various different predictions. I summarize five points here:

1. Distributivity

DA-approach: The reciprocated anaphor is doubly anaphoric to a plural argument. In one case, it is anaphoric to the plural entity that the argument refers to (the range argument), in the second case it is anaphoric to the entity that is part of the plural (the contrast argument). To make the antecedent of the contrast argument available, the antecedent of the reciprocated anaphor must scope and distribute over the clause.

R-approach The reciprocated anaphor takes the plural DP as its argument and its lexical meaning expresses that clausal relations hold for the parts of the plural argument. Thus, no distribution of the plural argument is necessary.

We will see that *the others* in the bound reading requires its antecedent to distribute. This supports the DA-approach to *the others* but it is problematic for the R-approach to *the others*. This point is discussed in detail in Section 4.3.

2. The range of readings

DA-approach: It is known that the relations in reciprocal sentences can be of various strengths. For example, while (21a) can be interpreted as ‘each of Morris, Hilary and Philip knows the other two (i.e., everyone knows everyone else)’, (21b) is accepted not only in a scenario in which each telephone pole is 500 feet from all the other telephone poles but also in weaker ones. For example, if the telephone poles are in a line and only each pair of neighbouring poles is at 500 feet distance, the sentence is judged as true. Thus, in this case it suffices that the reciprocal relation

holds only between members of each pair. Finally, (21c) is true if the chairs form one column, in which case the relation is not reciprocal at all, not even for the members of each pair.

- (21) a. Morris, Hilary and Philip know each other.
 b. The telephone poles are 500 feet from each other.
 c. The chairs are stacked on top of each other.

It has long been assumed that the strength of relations in reciprocal sentences is parallel to the strength of relations in sentences with transitive verbs and plural arguments (Fiengo and Lasnik, 1973; Langendoen, 1977; Williams, 1991). In Schwarzschild (1996); Sternefeld (1998); Sauerland (1998); Beck (2001), the DA-approach to the reciprocated anaphor is combined with the theory of plurals, which derives the range of readings that reciprocal sentences can give rise to. However, I am going to argue that these approaches cannot capture readings weaker than the strongest one (that is, (21b) and (21c) are unaccounted for).

R-approach: The range of readings shown in (21) is commonly derived in the R-approach through postulating an operation on the relation. This is either assumed solely for the interpretation of the reciprocated anaphor (like the operations on the relation in Dalrymple et al. 1998) or is postulated in another domain (Winter, 2001b), albeit not the semantics of transitive sentences with plural arguments. Thus, the parallelism between the strength of relations in reciprocal sentences and the strength of relations in sentences with plural arguments disappears. I will offer a novel account within the R-approach which allows the derivation of some weaker readings which cannot be derived in the DA-approach.

Since weaker readings can appear in reciprocal sentences, these data are problematic for the DA-approach to *each other* but are compatible with the R-analysis of *each other*. However, the bound *the others* lacks the weaker reading that the DA-analysis cannot derive, which supports the DA-approach to *the others*. This point is discussed in detail in Sections 4.3 and 4.4.

3. Co-reference in discourse

DA-approach: The DA-approach postulates the variables on the reciprocated anaphor which are bound during the derivation. Unless we stipulate extra conditions on binding the variables should be able to find antecedents introduced in a previous discourse.

R-approach: No binding is postulated for the interpretation of the reciprocated anaphor. Therefore, reference to entities in the discourse is not possible.

Since *each other* cannot refer to entities introduced in the previous discourse, unlike *the others*, this supports the DA-analysis of *the others* and the R-analysis of *each other*. This point is discussed in detail in Section 4.5.

4. Binding by more than one DP

DA-approach: The reciprocated anaphor is doubly anaphoric, therefore, each variable could be bound by a different argument.

R-approach: The reciprocated anaphor takes only one DP as its argument. Thus, the effect of binding by two different DPs, possible in the DA-approach, cannot be mimicked here.

Since *each other* cannot be anaphoric to two arguments, but *the others* can this supports the DA-approach to *the others* and the R-approach to *each other*. This point is discussed in detail in Section 4.5.

5. Long-distance reciprocity

DA-approach: Higginbotham (1981) discusses sentences like (22), which are ambiguous between the two readings indicated in (22a,b). The reading (22b), known also as long distance reciprocity, has been accounted for in Heim et al. (1991a,b); Dimitriadis (2000), which all belong to the DA-approach to reciprocals.

- (22) John and Mary think they like each other.
- a. John and Mary each think: “We like each other.”
 - b. John thinks he likes Mary and Mary thinks she likes John.

I will argue that the most successful account to this date, Dimitriadis (2000), overgenerates in many cases. This overgeneration, as we will see, correctly predicts the behavior of *the others*, which again supports the DA-approach to *the others*.

R-approach: Long-distance reciprocals are partly accounted for in Dalrymple et al. (1998). However, their approach cannot explain all the data that Dimitriadis (2000) discusses. I will offer a novel account within the R-approach that can deal with these data.

Long-distance reciprocity is problematic for the DA-approach to *each other* due to its overgeneration. As we will see, the same analysis does not overgenerate in case of the readings of *the others*, which again supports the DA-approach to *the others* and the R-approach to *each other*. This point is discussed in detail in Section 4.6.

I will elaborate on each of these points in the upcoming sections. As should be clear already now the conclusion is going to be that both the DA-analysis and the R-analysis are correct for reciprocated anaphora. However, they should be used for different lexical items. The DA-approach should be applied to *the others* and the R-approach to *each other*.

The strength of assuming two ways of anaphoricity lies in the predictions that the two approaches make. In particular, they lead us expect that the properties listed

above should bundle together. For example, once a reciprocated anaphor requires its antecedent to be distributive (property **Distributivity**) we also expect it to allow the discourse-anaphoric reading and disallow weaker readings, like the ones listed in (21b,c). These points are summarized in the Table 4.1. It is an empirical fact that

Table 4.1: Expectations of the two strategies

	DA-strategy	R-strategy
Distributivity	antecedent must distribute	antecedent does not distribute
The range of readings	weaker readings not possible	weaker readings possible
Co-reference in the discourse	possible	not possible
Binding by more than one DP	possible	not possible
Long-distance reciprocity	unlimited	limited

lexical items seem to bundle the properties together. Thus, for example, *the others* select all the properties that are listed under the DA-strategy heading, while *each other* selects the properties that are listed under the R-strategy heading. This supports the viewpoint that both strategies are relevant in dealing with anaphoricity in language.

The rest of the chapter is organized as follows. In the next section, I show in detail why the first two properties, **Distributivity** and **Range of readings**, support the DA-analysis of *the others*. The data from distributivity and the range of readings also support the R-approach to *each other* even though the actual implementation, I argue, should differ from the ones that have been proposed in the literature (Moltmann 1992; Dalrymple et al. 1998; Sabato and Winter 2005b). In Section 4.4 I offer a novel analysis of reciprocals cast within the R-approach. This analysis can explain the interaction of reciprocity and distributivity and the connection between the range of readings and the theory of plurality, both of which points remained unexplained under all the previous accounts, as far as I know. In Section 4.5 I show why the following two properties, **Co-reference in discourse** and **Binding by more than one DP** support the DA-approach for *the others* and the R-approach for reciprocals. This leaves the long-distance reciprocity as the last issue. I show in Section 4.6 that the DA-approach overgenerates in accounting for the long-distance reciprocity and offer a new analysis of the phenomenon, in which the reciprocal is analyzed in the R-approach.

4.3 Distributivity and the range of readings⁶

4.3.1 Distributivity

In the DA-approach, the reciprocated anaphor is doubly anaphoric to a plural argument. It is necessary that * is present in the structure, otherwise the contrast argument

⁶Parts of this section appeared in Dotlačil and Nilsen (2009).

would not be bound by the subparts of the plural argument, which is crucial for deriving the correct truth conditions of *each other*. To see that, consider what would happen if the $*$ operator was absent. Then, the LF structure would be as in (23a), and the range argument and contrast argument would have only one option for binding: the plural argument $M \oplus H$ itself. This leads to the interpretation (23b), in which *each other* would refer to the maximal entity among Morris and Hilary different from Morris and Hilary. Since this can never be satisfied the sentence would always cause presupposition failure.

- (23) a. $[_{TP} M \oplus H [1 [dry\ each\ other_1\ of\ 1]]]$
 b. $(\exists e)(*DRY(e) \wedge C_*\Theta_1(M \oplus H)(e) \wedge C_*\Theta_2(\sigma(\lambda x.x \leq M \oplus H \wedge \neg x \circ (M \oplus H)))(e))$

In the R-approach, there is no need for the $*$ operator to derive the correct meaning of reciprocal sentences because the reciprocated anaphor itself can, as part of its lexical meaning, express that the relation holds between parts of the plural argument.

The others

The $*$ operator requirement of the DA-approach correctly predicts how antecedents of *the others* differ in case of the bound reading. In Chapter 2, Section 2.3, I presented the questionnaire which tested, among other things, preference of *the others* for its antecedents (for more details see Section 2.3).

As we have seen the antecedents for the bound reading of *the others* were ordered on the scale shown in Table 4.2 ($\alpha < \beta$ should be read as α is preferred over β as the antecedent of the bound reading of *the others*, CQ-DPs=DPs with counting quantifiers, G-DPs=group-denoting DPs, see (19) in Chapter 2). To put it more concretely, the scale shows us that the bound reading of *the others* in (24a) is more acceptable than the bound reading of *the others* in (24b) which in turn is more acceptable than the bound reading in (24c).

- (24) a. Each boy hated the others.
 b. All the boys hated the others.
 c. The boys hated the others.

Table 4.2: Antecedents of *the others*

DQ-DPs	<	CQ-DPs	<	G-DPs
--------	---	--------	---	-------

This scale is identical to the scale known from distributive readings. I argued in Chapter 3 that the degradation of distributive readings is due to the fact that the distributive reading with covariation is marked and should not be expressed by unmarked expressions. The same reasoning can be transposed to *the others* since in its bound reading, it needs to be in scope of its antecedent and covary with it. This supports the DA-approach to *the others*. There are two more points I should clarify though.

First, in the discussion of the experiment in Section 2.3 one might notice that even though the antecedents of *the others* were preferred as shown in Table 4.2, *the others* and distributive readings did not show complete match because the bound reading of *the others* was generally less accepted than distributive readings (even with DQ-DPs; see especially Figure 2.1 in Chapter 2 which plots the means of people's judgements). The difference between distributive readings and the bound reading of *the others* in terms of absolute acceptability is not problematic for drawing the conclusion that both are arrived at by the same mechanism, i.e., * insertion. It only shows that the bound reading of *the others* is generally less preferred option for some speakers (who might prefer the discourse reading with *the others*).

The second point is the difference between definite plurals and coordination of DPs. We have seen that in case of the distributive reading, the definite plural was less preferred as the subject than the coordination of proper names. I argued at the discussion of the experiment that this has nothing to do with distributivity per se. It is well-known that proper names are preferred over definite descriptions in language use when they are anaphoric to previously introduced discourse referents, especially when the proper names are short (Ariel, 1990). In the questionnaire people got the description of the situation which included discourse referents and afterwards were asked whether they could say particular sentences. Thus, proper names used in the test sentences were likely understood as being anaphoric, and similarly for definite descriptions. Thus, the fact that people preferred to use coordinations of proper names is quite likely to be just an artifact of their general preference for names over descriptions in this case. This explanation, however, cannot be carried over to the bound reading of *the others*, where we have seen the situation is the opposite: definite plurals are preferred over coordination of DPs. I do not have an explanation of why this is so. Quite likely, it has to do with the fact that *the others* is anaphoric to the plural argument and the definite plural provides a description which might help picking up the whole group as the antecedent of *the others* more straightforwardly than coordination of proper names. No matter what explanation is correct, though, the difference between definite plurals and coordination of proper names is independent of the scale in Table 4.2. Thus, it does not threaten the argument that the scale in Table 4.2 and the same scale of DPs with predicates interpreted distributively have the same cause, namely, the dispreference for the insertion of the * operator and the preference for its low scope. It only tells us that there is another factor at play, which is irrelevant for the point I make here.⁷

⁷That the difference between definite plurals and coordinations is independent of distributivity can be supported by the following two sentences:

- (1) Morris, Philip and Perssy each like the others.
- (2) The professors each like the others (the other professors).

As Eric Reuland points out to me, he finds the second example better with the bound reading of *the others* than the first example. Crucially, in these examples we have an expression that forces distribution (floated *each*) so the difference cannot be due to unavailability of distributive readings.

Each other

Let us move to *each other*. If DQ-DPs are the most preferred antecedents, followed by CQ-DPs and G-DPs, then obviously this is again a point in support of the DA-approach. Obviously, this order of preference for *each other* is not correct since a DQ-DP is ungrammatical as the antecedent of *each other* (see (25a)), unlike CQ-DPs (see (25b)) or G-DPs (see (25c)):

- (25) a. * Each boy hated each other.
 b. All the boys hated each other.
 c. The boys hated each other.

Still, one might check if by using quantitative methods, any difference between CQ-DPs and G-DPs could be found. Suppose we do not find it. Does it mean that the DA-analysis of *each other* is ruled out?

In fact, it is not. Thus, while the presence of the above mentioned preference scale among antecedents is a support for the DA-approach to a reciprocated anaphor, its absence is a null result. The reason for this skewness is the following. We assume that the distributive reading with covariation in transitive sentences is the marked reading and there are other, unmarked readings in which arguments do not show covariation. Since G-DPs are unmarked expressions they combine with unmarked readings, which captures why people disprefer the distributive reading with covariation when a G-DP is the subject. However, since *each other* in DA-analyses must have the contrast variable bound, the distributive reading is the only possible reading in reciprocal sentences in (25). Since there is no other reading that G-DPs could make use of they have to combine with the distributive reading in this case. To put it differently, since this is the only reading available, it cannot be marked. Markedness is a relative, not an absolute notion. Concerning the R-analysis of *each other*, we also expect no degradation of reciprocal sentences with G-DPs because reciprocity is not a case of the distributive reading with covariation: the reciprocal is not a generalized quantifier that fills in an argument position and could vary with respect to its antecedent. In other words, the full acceptability of (25b) and (25c) tells us nothing about which analysis to reciprocals should be preferred.⁸

But there is still something left unexplained in the DA-approach. Why is (25a) ungrammatical? The interpretation of (25a) should be identical to (26), which *is* interpretable.

- (26) Each boy hated the others.

On the other hand, the R-approach has a ready-made answer: *each other* includes a distributive quantification over pluralities. Since the distributive quantifier also distributes, we would have two distributive quantification stacked atop each other. The interpretation we would get for (25a) is the following:

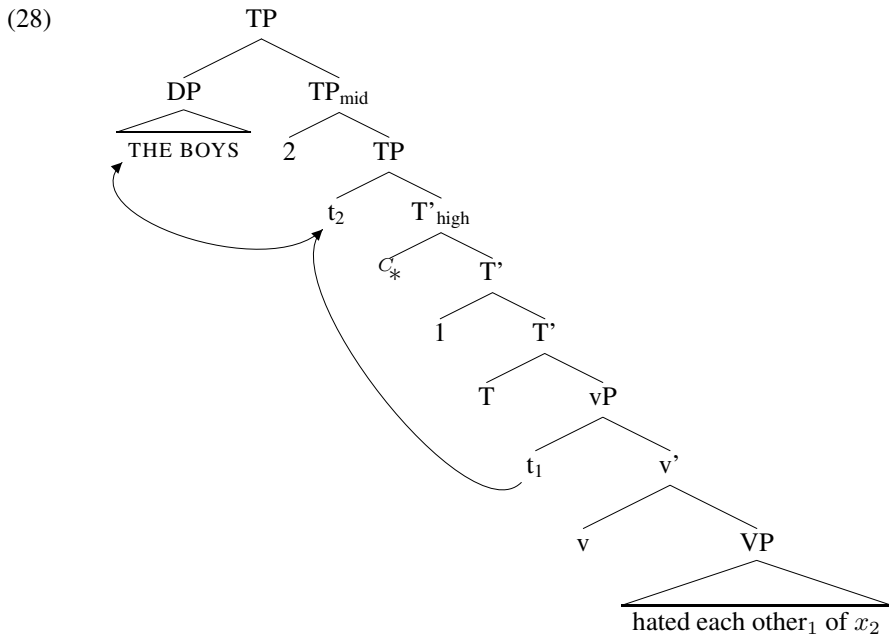
⁸This is contrary the analysis of Dotlačil and Nilsen (2009) where it was assumed that no difference in acceptability between (25b) and (25c) supports the R-approach. We claimed this because we had no analysis of why particular readings should be degraded and assumed that they always are degraded (a similar approach is taken in Moltmann 1992).

(27) $(\forall x \in \text{BOY})(x \in \text{AT} \rightarrow x) \text{ hated each other}$

Of course, this interpretation can never be true since *each other* is fed atoms, not a plurality.

One possible answer of the DA-approach as to why (25a) is ungrammatical is that the plurality BOYS is not available. Since this is necessary to interpret the range argument of *each other* we derive the ungrammaticality of (25a). However, this cannot be correct since in (26) we see that *the others* can pick up the plurality of BOYS as its range argument. But we could go one step further and assume that the range argument of *each other* must be *bound* (and this condition does not hold for the range argument of *the others*). Since binding requires c-command, we can derive that (25a) is ungrammatical: the plurality *boys* does not c-command *each other* at LF, at least in the assumption we make about the interpretation of the generalized quantifier *each boy*.

Assume we require binding of the range argument in the DA-approach to *each other*. How come (25b) and (25c) are possible? In this case, the range argument does c-command *each other*. This is so because the distribution is achieved by the insertion of * and is not inherent to the lexical meaning of the subject. For example, (25c) has the following LF structure. As one can see the plurality c-commands the reciprocal.



Here is one problem with this account. We expect (29), from Heim et al. (1991a) and Williams (1991), example 35, to be grammatical but it is not:

(29) * They each like each other.

The reason is that in this case, the subject *they* does c-command the reciprocal, under standard analyses of floating quantifiers. Of course, we could stipulate that binding of

they is impossible here as well. But why would it be? The problem is that *they each like each other* differs from *they like each other* mainly in the fact that where the latter would have a null operator, *, the former has the floating quantifier *each*. It is not clear to me what principle could exclude the former without excluding the latter, unless it is very specific, stating assumptions like ‘the floating quantifier *each* blocks binding’. Furthermore, this account could explain the ungrammaticality of (25a) only under specific assumptions about syntax, where the restrictor of *each* never c-commands the reciprocal (and this does not hold for the restrictor *all*, for example). This would make the account incompatible with approaches like Sportiche (2005).

On the other hand the ungrammaticality of (29) follows from the R-approach to reciprocals. The reason is the same as with (25a): the floating quantifier distributes down to atoms. Since *each other* also requires distribution we end up with conflicting requirements.⁹

Thus, to conclude, I consider the data like (29) and (25a) to be problematic for DA-approaches and supporting R-approaches to reciprocals.

There are more complicated data concerning distributivity which have been taken as problematic for both DA- and R-approaches to *each other*. Williams (1991) and Moltmann (1992) make the following observation: arguments in reciprocal sentences do not show covariation. (30), from Moltmann (1992), p. 426, shows this. For the interpretation that there were many presents, *Christmas presents* is preferred over *a Christmas present*. In other words, the reading in which *a present* varies for each child/each subpart of the children is marginal for this sentence.

- (30) The children gave each other Christmas presents/?? a Christmas present.
(each child giving a different present)

Williams (1991) notes that this requirement is parallel to non-reciprocal sentences. In (31), from Moltmann (1992), p. 426, the plural argument is preferred (in the indicated interpretation).

- (31) The men gave the women Christmas presents/?? a Christmas present.
(each man giving a different present)

⁹One might wonder why (25b) is grammatical in the R-approach given that we argued in Section 3.2.6 that *all* requires the insertion of *. There are two possible answers. First, the * operator can be present and scope over Θ_1 . We have used this option to derive collective readings compatible with *all*. Since * does not distribute over the whole clause, no clash with the distribution in *each other* arises. Alternatively, the requirement of *all* that * must be inserted in the structure can be satisfied by *each other* itself. Brisson (2003) discusses examples of *all NP* compatible with reciprocal states and achievements, like:

- (1) a. All the students recognized each other.
b. All the students look like each other.

We have claimed in Section 3.2.6, following Brisson, that *all* cannot combine with collective states and achievements since the insertion of * over Θ_1 does not lead to the collective interpretation. If the requirement of *all* for the presence of distribution in the clause could be also satisfied by *each other* the availability of reciprocal states and achievements would follow. This is also the approach to *all* for which Brisson argues in her 2003 paper.

The degraded status of the distributive reading in (31) should be familiar. It is parallel to the experimental data discussed in Chapter 2 and analyzed in Chapter 3.

(30) then shows that what holds for non-reciprocal sentences holds for reciprocal sentences as well: the interpretation in which a DP is scopally dependent on and covaries with another DP is dispreferred.

Williams (1991) and Moltmann (1992) claim that this is problematic for the approaches that require distribution in reciprocal sentences. The reason is that if reciprocals requires the distribution of the antecedent the antecedent should also be able to distribute over the indefinite. But similarly, one could argue that these data are problematic for R-approaches. For example, if we apply the semantics of *each other* that I used in the introduction, we get the interpretation in (32b) for (30) (with the LF structure given in (32a), where EO stands for *each other*).

- (32) a. [the children] EO $\lambda x \lambda y \lambda e. x$ gave y a present at e
 b. $(\exists e)(\forall x, y \leq \text{THE CHILDREN})(\exists e' \leq e)$
 $(x \neq y \wedge x, y \in AT \rightarrow x \text{ gave } y \text{ a present at } e')$

In (32b) the indefinite is in scope of the distributive quantifier and thus one might think that we get, incorrectly, the interpretation that each boy gave a different present. The same problem holds for other R-approaches, like Dalrymple et al. (1998) and Sabato and Winter (2005b). As far as I know the only analysis that can account for the lack of covariation in (30) is Moltmann (1992).

One approach to the data like (30) was to deny them. This is the approach of Dalrymple et al. (1998) and Beck (2001). Dalrymple et al. (1998) disagree with Moltmann's (and Williams') intuition. They claim that the relevant reading is available in reciprocal sentences, so the singular DP in (30) is acceptable under the indicated interpretation. A similar claim is made in Beck (2001).

The disagreement should be familiar from the discussion of the distributive reading with G-DPs, where some authors claimed that the distributive reading was dispreferred, marginal, or ungrammatical, while other authors found the reading possible. Because we are dealing with the same phenomenon here (the acceptability of the distributive interpretation) it is no surprise that the data are somewhat weak and not shared by all speakers. Thus, contrary to Dalrymple et al. (1998) and Beck (2001) I take the preference in (30) to be real and in need of accounting. In Section 4.3.3, I furthermore discuss a questionnaire which I designed together with Øystein Nilsen and whose results support Moltmann's and Williams' judgements regarding reciprocal sentences.

However, even if the data like (30) are valid I do not think we need to give up the idea that the antecedent of the reciprocal distributes over the clause. We have argued in the previous Chapter that covariation has a degraded status. Thus, one might argue that the antecedent does distribute over the whole clause in (30) and the indefinite is in its scope but it does not covary. This is for the same reason that indefinites do not covary in simple non-reciprocal sentences, like (31). I offered an explanation why this should be so in the previous Chapter.

There are other data discussed in literature which are more problematic for the viewpoint that the antecedent of a reciprocal distributes over the whole clause in re-

reciprocal sentences. Consider (33), repeated from above. Assume that there are only two children and each of them gave one present to the other child. The possible way of saying this is using the plural argument *presents*.

- (33) The children gave each other Christmas presents/?? a Christmas present.
(each child giving a different present)

Assume that the plural argument *presents* could take scope under *each other* or over *each other* (in case of the R-analysis of *each other*) or it takes scope under or over the * operator (in case of the DA-analysis of *each other*). Neither of these options derives the correct reading. Both of them would require that each child was giving more than one present to the other child. As far as I can see the reading that we are after, where one child was giving only one present but more than one presents were given in total, cannot be derived with the system we have right now. More concretely, we need a cumulative reading between *presents* and *the children*, but we also need *the children* to scope over *each other* and both requirements cannot be satisfied in the system we have. The same point is shown in the following example, from the Corpus of Contemporary American English (COCA):

- (34) Critics and defenders of the Catholic Church have been aligned against each other in two conflicting camps.

The reading in (69) is that critics aligned against defenders and defenders aligned against critics and in total there were two competing camps. Thus, the reciprocal does not distribute over *two conflicting camps*, otherwise, we would get the interpretation that critics aligned against defenders in two conflicting camps, and so did defenders. Cumulative readings in reciprocal sentences have been studied by Sternefeld (1998) and Sauerland (1998) and their account can deal with (33) and (34). However, their analysis requires that cumulative readings can be formed in other ways than by interpreting arguments in their thematic positions, and thus it is problematic in the system we are using. (I will discuss this way of deriving cumulative readings in Section 4.4.1.) Furthermore, even if we accepted their analysis, it only works for the DA-approach to reciprocals, which I am arguing against on independent grounds. For this reason, I will shy away from their analysis and assume that the relevant interpretations of (33) and (34) do lead us to the analysis in which reciprocals distribute only locally, excluding other arguments from their distribution.

To sum up, we have seen the data that are problematic for DA-approaches, as well as R-approaches. In Section 4.4 I offer a novel account of *each other* which is cast within the R-approach and which can deal with cumulative readings in reciprocal sentences but does not suffer from the problems that the DA-approach faces (for instance, the incompatibility of reciprocals with distributive quantifiers). Before doing so, I turn my attention to the range of readings which are possible in reciprocal sentences.

4.3.2 The range of readings

One of the biggest concerns in the literature on reciprocals is how to capture varying strength of reciprocal relations in reciprocal sentences (Fiengo and Lasnik, 1973; Lan-

gendoen, 1977; Sternefeld, 1998; Sauerland, 1998; Beck, 2001; Sabato and Winter, 2005b). The varying strength of reciprocal relations can be seen on examples (35a-c). While (35a) is interpreted as ‘each of Morris, Hilary and Philip know the other two (i.e., everyone knows everyone else)’, (35b) is accepted not only in a scenario in which each telephone pole is 500 feet from all the other telephone poles but also in weaker ones. For example, if the telephone poles are in a line and only each pair of neighbouring poles is at 500 feet distance, the sentence is judged as true. Thus, in this case it suffices that the reciprocal relation holds only between members of each pair. Finally, (35c) is true if the chairs form one column, in which case the relation is not even reciprocal for the members of each pair. These weaker readings cannot be arrived at by using (9) because in that meaning of the reciprocal every atomic individual is related to every other atomic individual. Following Dalrymple et al. (1998) I will call the reading exemplified by the first sentence Strong Reciprocity, or SR for short, the reading exemplified by the second sentence Intermediate Reciprocity, or IR for short, and the reading exemplified by the third sentence Intermediate Alternative Reciprocity, or IAR for short.

- (35) a. Morris, Hilary and Philip know each other. (example of SR)
- b. The telephone poles are 500 feet from each other. (example of IR)
- c. The chairs are stacked on top of each other. (example of IAR)

In the R-approach, the varying strength of the relation was captured by postulating operations on the clausal relation (Dalrymple et al., 1998; Sabato and Winter, 2005b). In the DA-approach the varying strength of the relation was captured by making use of the operators that are independently assumed in the theory of plurality (Schwarzschild, 1996; Sternefeld, 1998; Sauerland, 1998; Beck, 2001). I believe that it is correct to connect the readings of reciprocals to the theory of plurality, which can be supported from the interpretation of negated reciprocal sentences and close match between the range of readings in reciprocal and non-reciprocal sentences.

Negated reciprocal sentences

Consider the example (36). If a negated reciprocal sentence is true when its positive counterpart is not (36) should mean that not everybody among Morris, Hilary and Philip knows everybody else. But (36) is a stronger statement. It is only true if nobody among Morris, Hilary and Philip knows anybody else among these three.

- (36) Morris, Hilary and Philip do not know each other.

To my knowledge, this behavior of negated reciprocal sentences has been first observed by Krifka (1996), who notices that this makes reciprocal sentences parallel to sentences with plural arguments. For example, (37) requires that no boys are sleeping.

- (37) The boys are not sleeping.

As Fodor (1970) observes a definite plural noun phrase “does not [...] admit the possibility that the sentence might be true of some but not all things of the kind described”.

She suggests that definite plurals carry the presupposition that the ascribed predicate is true of all entities referred to by the definite plural or none of them (all-or-nothing presupposition in Fodor's words).

In a series of papers (see Löbner 2000 and references therein) Löbner argues that all-or-nothing presupposition is much more general phenomenon. It holds for any predication in natural language. He names this the Presupposition of Indivisibility in Löbner (2000) (in his previous work, he assumes Presupposition of Argument Homogeneity, which is less general).

(38) **Presupposition of Indivisibility:**

Whenever a predicate is applied to one of its arguments, it is true or false of the argument as a whole.

This presupposition plays the most important role when a predicate is distributive (summative in Löbner's terminology) since in that case we can clearly see that the predication is all-or-nothing phenomenon. Löbner guarantees the satisfaction of presupposition by building it into the meaning of the operator that derives summative (=distributive) predication. We can replicate this by adding the same presupposition in the * operator (for more discussion, see Schwarzschild 1992; Gajewski 2005). This gives us a modified version of the definition of pluralization. Let me start with the pluralization on the predicate (39).

$$(39) \quad C_*P = \lambda x : x \in [\{\downarrow a : Pa \wedge a \in Cov\}] \vee x \in [\{\downarrow a : \neg Pa \wedge a \in Cov\}]. \\ x \in [\{\downarrow a : Pa \wedge a \in Cov\}]$$

The presupposed part requires that either every entity (relativized to cover) of the plural x did P or did not do P . The sentence (40a) in its distributive reading is interpreted as none of the boys read a book. This follows because of the presuppositional part of the * operator, see (40b) (in the example, I omit covers and assume distribution to atomic individuals). While the non-presuppositional part only requires that there were some boys who did not read a book (because in that case THE BOYS is not in the pluralized set of entities that read a book), the presuppositional part adds the stronger requirement: THE BOYS is in the pluralized set of those entities that did not read a book.

(40) a. The boys did not read a book.

b. **Non-presuppositional content:** THE BOYS $\notin [\{\downarrow a : a \text{ read a book}\}]$

Presuppositional content: THE BOYS $\in [\{\downarrow a : a \text{ read a book}\}] \vee$
THE BOYS $\in [\{\downarrow a : \neg \text{read a book}(a)\}]$

The pluralization on relation with the presupposition added is given in (41).

$$(41) \quad C_*R = \lambda x \lambda y : \langle x, y \rangle \in [\{\langle \downarrow a, \downarrow b \rangle : R(a)(b) \wedge a, b \in Cov\}] \vee \\ \vee x \in [\{\downarrow a : y \in [\{\downarrow b : \neg R(a)(b) \wedge a, b \in Cov\}]\}] \\ \langle x, y \rangle \in [\{\langle \downarrow a, \downarrow b \rangle : R(a)(b) \wedge a, b \in Cov\}]$$

This is similar to the * operator combining with predicates. The presupposed part makes sure that the two plural arguments are homogeneous with respect to the relation to which they apply. Somewhat complicating factor, though, is the part of the

presupposition in which $R(a)(b)$ is negated. Instead of stating that the pair $\langle x, y \rangle$ is in the pluralized set of pairs for which $\neg R(a)(b)$ holds, it states that for every part of x and for every part of y $\neg R(a)(b)$ holds. In other words, this disjunct presupposes that nobody in x is related by R to anybody in y . This is parallel to the branching reading. The need for these stronger conditions becomes clear in the following example, repeated from above (see also Gajewski 2005, Chapter 5 for discussion).

- (42) a. The boys are not sleeping.
 b. **Non-presuppositional content:** $\neg(\exists e)(\text{*SLEEP}(e) \wedge {}^C\text{*}\Theta_1(\text{THE BOYS})(e))$
Presuppositional content: $\langle \text{THE BOYS}, e \rangle \in [\{\langle a, e' \rangle : \Theta_1(a)(e')\}] \vee$
 $\text{THE BOYS} \in [\{a : e \in [\{e' : \neg\Theta_1(a)(e')\}]]$

The non-presuppositional part states that there is no event in which the boys are Θ_1 of parts of e , e being the event of sleeping. This is true if most boys are sleeping, as long as one boy is not. The presuppositional part strengthens this reading. It further requires that in this event e no boy is related to any part of e by Θ_1 . This is only true if no boy is sleeping.

Finally, notice that the following sentence does not require that no boy is sleeping.

- (43) Every boy is not sleeping.

This is expected since in this case the predicate does not apply to the plurality of boys as its argument, rather, it applies to atoms. In general, all-or-nothing phenomenon is best visible with distributive predicates and G-DPs as arguments since these require the presence of the $*$ operator.¹⁰

Reciprocal sentences show the same presupposition of indivisibility as sentences with plural arguments. Another example is shown in (44). (44b) implies that none of the adults know any other adults. (44c) are odd answers. Compare this to the question *Do all the adults know each other?* where the negative response can mean that that not every adult knows every other adult. Since all-or-nothing presupposition shows up with G-DPs this is expected.

- (44) a. A: Do the adults know each other?
 b. B: No.
 c. B: # No, some of them do./# No, Morris knows Hilary and that is all.

(44) shows that the presupposition of indivisibility should somehow be part of the meaning of the reciprocal, which would make sense if the reciprocal required the $*$ operator, probably because the $*$ operator is built into its meaning. This is problematic for Dalrymple et al. (1998) and Sabato and Winter (2005b) who derive readings of

¹⁰The following sentence shows that *all NP* does not behave parallel to G-DPs (as already observed by Fodor 1970).

- (1) All the boys are not sleeping.

This might suggest that the distributive interpretation is not done by the $*$ operator, as I assumed in Chapter 3. Whichever way we bring about the distributive interpretation here it should lack the presupposition of indivisibility, similarly to the quantifier *every*. I will leave the modification of *all* that accounts for this open.

reciprocal sentences independently of the theory of plurality. I discuss their problems with these data in Section 4.4.3 in more detail.

The range of readings in positive reciprocal sentences

The second argument for relating readings in reciprocal sentences to the theory of plurality comes from their very close match. SR can be modelled as doubly-distributive reading or, as we will see, as a branching reading. IR can be modelled as a cumulative reading. To see the parallel between IR and the cumulative reading, consider a simple case of the cumulative reading in a sentence with two plural arguments: (45a) as interpreted in (45b). Assume that *Cov* provided by C_* on each thematic role is the set of atomic individuals, i.e., $\{M, P, H, D\}$. In that case, the sentence is true if, for example e consists of two subevents, e_1 and e_2 , such that Morris kissed Hilary in e_1 and Philip kissed Desiree in e_2 . Another option is that e consists of e_1 and e_2 where Morris kissed Desiree in e_1 and Philip kissed Hilary in e_2 . These are two cases of the cumulative reading.¹¹

- (45) a. Morris and Philip kissed Hilary and Desiree.
 b. $(\exists e)(\ast\text{KISS}(e) \wedge C_*\Theta_1(M \oplus P)(e) \wedge C_*\Theta_2(H \oplus D)(e))$

(45b) requires that each atom of the first plural DP (Morris and Philip) is the argument of Θ_1 and each atom of the second plural DP (Hilary and Desiree) is the argument of Θ_2 , without specifying anything more about the actual relations between the individuals. This relatively weak requirement seems correct to capture IR. We can see it on the example, repeated from above. (46) is true if (many) telephone poles are standing in a line, having the equal distance of 500 feet. Suppose we take *each other* to be the set of telephone poles and we assume that the relation holds between distinct individuals. Then, the scenario in which telephone poles are in a line just describes a cumulative reading: every telephone pole must be the argument of Θ_1 and every telephone pole must be the argument of Θ_2 which can be only satisfied if every telephone pole is 500 feet away from some other telephone pole.

- (46) Telephone poles are 500 feet from each other.

Consider now the following example: there is an audience in a lecture hall listening and looking at the speaker. The speaker is looking back at one person in the audience. In this scenario, (47) sounds very odd.

- (47) # The people in the lecture hall are staring at each other.

The odd status of (47) follows if we assume that IR is parallel to cumulative readings. In that case everyone of *the people* in (47) must be related by Θ_1 to staring, as well as by Θ_2 , i.e., everyone is staring (at someone) and being stared at (by someone). This is not satisfied in the context given above, where only the speaker is being stared at.

¹¹ Additionally, (45b) is true if in e_1 Morris kissed Hilary and Desiree and in e_2 Philip kissed Hilary and Desiree. This is the branching reading. It is irrelevant for the discussion now.

If we connect the readings of reciprocals to the readings in the theory of plurality, we expect that we cannot get weaker than the cumulative reading, i.e., IR. This is correct for most cases. Consider (48).

- (48) a. # The plates are stacked **underneath** each other.
 b. # The boys **preceded** each other into the room.
 c. # The men are each others **fathers**.
 d. # Mary and Desiree gave birth to each other.
 e. # John and Mary are **taller than** each other.

All these examples are excluded by the cumulative reading. So connecting the reciprocal readings to the theory of plurality correctly excludes these readings.

The cumulative reading cannot derive IAR. (49) is thus unaccounted for. Notice that in this case only the top chair is the argument of Θ_1 and only the bottom chair is the argument of Θ_2 and thus the cumulative reading is false.

- (49) The two chairs are on top of each other.

However, IAR is restricted to a small class of relations, in particular, IAR is restricted to spatial and temporal predicates. Furthermore, among these predicates it is only possible if they express figure-ground relation in such a way that the figure precedes the ground (*follow*, *be behind*) or the figure is vertically above the ground (*be on top of*). Given these limitations on IAR, Beck (2001) suggests that such readings should not be dealt with by a general theory of reciprocity, but rather thought of as a part of the semantic theory of the small set of spatial relations that actually give rise to it. I find this an entirely reasonable position.

DA-approaches to the range of readings with reciprocated anaphora

Most DA-approaches to reciprocals try to connect the variety of readings of reciprocal sentences to the theory of plurality. This could explain the data from negated reciprocal sentences as well as correctly capture the range of readings of reciprocals without stipulating new mechanisms, and thus they are preferable to the R-approaches to reciprocals, who fail in both respects. Unfortunately, DA-approaches are not particularly well-suited to derive weaker readings than SR. I will show this on the probably most well-developed system to this day, Beck (2001).

SR is, according to Beck, a case of a distributive reading. In the example of SR given above (35a) the $*$ operator applies to the predicate ‘knowing each other’. This gives us (50). To read the formula, recall that *each other_x* of y is an abbreviation for Beck’s semantic of *each other* with x the contrast argument and y the range argument, and could be paraphrased as ‘the other ones among y different from x ’. Thus the interpretation is ‘Morris knows Hilary and Philip, Hilary knows Morris and Philip, and Philip knows Morris and Hilary’ (granting *Cov* to consist of atomic individuals Morris, Hilary, Philip). This interpretation is correct for SR.

- (50) $[[(35a)]] = (\forall y \in Cov)(\downarrow y \leq M \oplus H \oplus P \rightarrow (\exists e)(*dry(e) \wedge C_*\Theta_1(y)(e) \wedge C_*\Theta_2(\text{each other}_y \text{ of } M \oplus H \oplus P)(e)))$

IR is assimilated to cumulative readings. This, as we have seen above, seems correct. Let us see how the parallelism can be captured in Beck's system.

In cumulative readings as discussed in Chapter 3, no $*$ operators are inserted in the structure. Plural DPs are arguments of (pluralized) thematic roles. For example, (51a) gets the cumulative reading if interpreted as in (51b).

- (51) a. Morris and Philip kissed Hilary and Desiree.
 b. $(\exists e)(\text{*KISS}(e) \wedge C_{*}\Theta_1(M \oplus P)(e) \wedge C_{*}\Theta_2(H \oplus D)(e))$

Unfortunately, this way of accounting for the cumulative reading is not usable for Beck (2001). The reason is that there is no $*$ operator which would distribute over the antecedent of the reciprocal and scope over *each other*, which, as we have seen above, is crucial for the interpretation of the reciprocal in DA-approach.

Beck derives IR using the $*$ operator higher in the structure. Unlike in case of distributive readings, the $*$ operator here takes a relation as its argument. (51a) gets the cumulative reading in structure (52a). This is interpreted in (52b). To simplify the discussion I ignore the event argument here.

- (52) a. $[M \oplus P [H \oplus D * \lambda x \lambda y. x \text{ kissed } y]]$
 b. $\langle M \oplus P, H \oplus D \rangle \in [\{ \langle x, y \rangle : x \text{ kissed } y \}]$

(52b) is truth-conditionally identical to (51b). One has to allow the $*$ operator to apply on relations derived in syntax by lambda abstractions. This would be an extra assumption in our system, just to derive IR. However, in Beck's analysis this is the only way to get cumulative readings to begin with and thus does not complicate the system just for the sake of deriving IR. Thus, I am going to assume it is possible.

Even admitting the $*$ operator to apply on the relation as in (52a) does not solve all the problems that Beck's approach faces when deriving IR. Let me show this on the example of IR repeated from above, (53).

- (53) Telephone poles are 500 feet from each other.

First, the relation to which $*$ applies cannot be $\lambda x \lambda y. x$ is 500 feet from y . The reason is that in that case *each other* is outside of the scope of $*$ and thus its contrast argument cannot be bound. Inversely, if *each other* stays in the scope of $*$, we would have no relation, but predication ($\lambda x. x$ is 500 feet from each other). The solution that Beck suggests is that only part of the meaning of *each other*, namely, the range argument, moves in front of the $*$ operator. This gives us the relation $\lambda x \lambda y. x$ is 500 feet from each other _{x} of y . Thus, we get the relation to which $*$ can apply, and the contrast argument is correctly bound (by λx). The full formula for (51a) is given in (54) (the range argument z is bound by the subject).

- (54) $[\text{telephone poles}]_z z * \lambda x \lambda y. x \text{ is 500 feet from } [\text{each other}_x \text{ of } y]$

Suppose there are four telephone poles ordered in a line: p_1 – p_4 . (54) is true if the two arguments, $(p_1 \oplus \dots \oplus p_4)_z$ and $z (= p_1 \oplus \dots \oplus p_4)$ can be split into parts for which the relation ' $\lambda x \lambda y. x$ is 500 feet from each other _{x} of y ' holds. This is true for the following pairs:

(55) Pairs that make the relation in (54) true:

- $\langle p_1, p_1 \oplus p_2 \rangle$
- $\langle p_2, p_2 \oplus p_3 \rangle$
- $\langle p_3, p_3 \oplus p_4 \rangle$
- $\langle p_4, p_3 \oplus p_4 \rangle$

If we pick these triples it is obviously true that every part of the plural individual ‘telephone poles’ is the argument of the relation, thus, the requirement of the $*$ operator applying to the relation is satisfied. For every triple the relation must be true, therefore, then each pole is 500 feet from another pole next to it in the line. These truth conditions are correct.

(54) can derive the correct truth conditions. However, it requires a dubious cover. Two types of contextually relevant subparts are necessary: first, single poles, and second, pairs of poles. While splitting a plurality into atomic individuals might be natural (Schwarzschild 1996 argues that it is one of the two default cases the other one being where we consider the group as a whole) splitting a plurality into pairs should require some context indications which I miss in the example above. On top of it, both ways of splitting the plurality need be present at the same time and I do not see how the sentence (53) enables that. Another problem is that we need the QR of the range argument. This QR seems a strange operation, for various reasons: First, it would require QR out of a definite island; secondly it would require QR of a variable; and finally, it would require QR of a proper sub-part of the semantic representation of a lexical expression. None of these are knock-down arguments for the postulation of the QR of the range argument but they make the operation very different from other cases of QR, as is quantifier scope, which does not require any of these special properties. Of course, we could assume that all these properties are allowed and that the range variable of *each other* in Beck’s analysis can move to derive IR. But I will try a different approach. I am going to pursue a more restrictive theory of the syntax-semantics interface, where such operations are disallowed.

I want to point out that even though I focus here on Beck’s account of IR other DA-approaches face problems with this reading as well. Heim et al. (1991a) and Heim et al. (1991b) do not consider weaker readings than SR so they are irrelevant for this discussion. Problems with Sternefeld (1998) are discussed in detail in Beck (2001), in fact, Beck’s account can be seen as improvement on Sternefeld’s approach. Sauerland (1998) does not discuss IR but based on his discussion on other readings, I presume he would deal with it by restricting the interpretation of *each other*. This is similar to the account as in Schwarzschild (1996). *Each other* would be interpreted in the last two approaches as follows:¹²

- (56) Interpretation of *each other* in Schwarzschild (1996) and Sauerland (1998), simplified

$$[[\text{each other}]] = \lambda P.P(\sigma(\lambda x.x \leq z_{range} \wedge \neg x \circ \downarrow y_{contrast} \wedge x = R(y)))$$

¹²I simplify here somewhat. Also I ignore differences in notation and background assumptions of Schwarzschild (1996) and Sauerland (1998) in order to make the notation to be as close to as Beck’s as possible, which, I hope, advances readability at this point.

The only difference from Beck's account is the addition of the last conjunct, $x = R(y)$, where R is some contextually given relation. If we let $R = a$ neighbouring telephone pole of in (53) we correctly derive IR. My problem with this account is that it does not say how we fill in R . Without that, the analysis is very weak. For example, *the people in the lecture hall are staring at each other* where everyone is staring at the speaker and the speaker is staring at one person in the audience becomes true, contrary to one's intuitions. This is because we could consider $R = \text{seeing}$ or *noticing*. Some further problems on choosing R are discussed in Beck (2001). There might be ways to restrict R in the correct way but these are still waiting to be developed. To conclude there is to this date no DA-account that, as far as I know, could capture IR by assimilating it to cumulative readings, which would avoid overgeneration faced by Schwarzschild (1996) and Sauerland (1998), and which would not suffer problems of Beck's. This means that if we want to follow the DA-approaches, we cannot derive readings weaker than SR. Obviously, this conclusion is wrong for *each other*.

But wait! We want to see if DA-analyses could be used for any reciprocated anaphora. We have seen that not deriving weaker readings makes it problematic for *each other*. Could it still work for *the others*? I believe that what is a weakness in case of *each other* becomes the strength of the analysis in case of *the others*. In the next section, I am going to discuss the results of the questionnaire that I developed with Øystein Nilsen and that tested the range of readings of *each other* and *the others*. There we will see that weaker readings in sentences with *the others* are much more rejected than the same readings in reciprocal sentences. But even before discussing the questionnaire we can show that it is preferable not to let *the others* be compatible with the cumulative reading. The argument comes from dependent plurals.

The term "dependent plurals" was coined by de Mey (1981), and it refers to bare plurals which are interpreted as singulars in scope of another plural DP. *Wheels* in (57), from Chomsky (1975) is a dependent plural since each unicycle only has one wheel.

- (57) Unicycles have wheels.

Chomsky (1975) analyzed dependent plurals through a syntactic rule which adds a plural morphological marker on nouns while leaving their semantics as singulars. While there are many problems with his original analysis (see Zweig 2008 for discussion), it sparked up research in which plural is taken to be number-neutral, and its requirement on non-singularity is only introduced when it is not dependent on other plurals (Roberts, 1990; Kamp and Reyle, 1993). There are good arguments that plurals are in fact number neutral. These come from sentences where plurals are inside downward entailing contexts. For example, in (58) the plural *children* is interpreted as meaning both a child and children.

- (58) a. A: Do you have children?
 b. B: Yes, I have one./# No, I have one.

However, Zweig (2008) argues that number-neutrality cannot account for dependent plurals. The reason is that dependent plurals *are* still interpreted as non-singular, albeit

this is not visible when one considers their number requirements in scope of the other plural. However, they must be interpreted as plural outside this scope. Consider (59a), from Zweig (2008). This sentence is true when each of the ten students lives in one New York borough (dependent plural reading). But the DP *NY boroughs* is interpreted as non-singular since the sentence is false if all ten students live in the same New York borough. (59b) is true in this case, which clearly shows that dependent plurals differ from singular DPs.

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

This behavior of dependent plurals can be captured if they are analyzed similarly to cumulative readings in event semantics.¹³ Thus, (59a) is analyzed as (60a) (before existential closure). Still, (60a) does not on its own derive the plurality requirement of dependent plurals. This is because $*$ just gives us a sublattice, a set consisting of both atomic and non-atomic entities. However, it is derived when we consider scalar conversational implicatures. The alternative of this sentence is (60b), in which the plural is substituted by a singular *NY borough*. Since (60b) is more informative than (60a) (every event that satisfies (60b) also satisfies (60a) but not vice versa) one can conclude that when a speaker utters (59a) she believes that (60b) is not true, and therefore, $*NY BOROUGH$ in (60a) must mean ‘more than one NY borough’.

- (60) a. $(\lambda e)(^C\Theta_1(\mathbb{E}(\text{TEN STUDENTS}))(e) \wedge *LIVE(e) \wedge \wedge^C\Theta_2(\mathbb{E}(\text{NY BOROUGHs}))(e)$
b. $(\lambda e)(^C\Theta_1(\mathbb{E}(\text{TEN STUDENTS}))(e) \wedge *LIVE(e) \wedge \wedge^C\Theta_2(\mathbb{E}(\text{NY BOROUGH}))(e)$

It has been noticed that a dependent plural reading of some plural DP is possible only if the other argument is non-distributive. For example, (61) lacks the dependent plural reading (every student lives in more than one NY borough).

- (61) Every student lives in New York boroughs.

Thus, it is known that the dependent plural reading depends on the properties of the other argument(s). But as far as I know there are no examples in the literature which would show that the dependent plural reading of some plural argument also depends on its own properties. In this respect, *the others* is particularly interesting. Suppose there are two boys. In that case (62) cannot get the bound reading not even for the speakers that allow the bound reading of *the others* with definite plurals. The bound reading reappears if *the boys* consist of three boys or more.

- (62) # The boys talked to the others/to the other boys.

¹³Even though they cannot be completely subsumed under cumulative readings, mainly because dependent plurals are licensed when the other plural argument is headed by *most* but cumulative readings are not licensed in this case. I will not go into details on how the difference is captured, see Zweig 2008, chapter 6 for details.

If the bound reading of *the others* cannot arise in cumulative readings, we have an explanation why (62) is odd: the dependent plurality in this case can only arise through the cumulative reading, and the cumulative reading is incompatible with the bound reading of *the others*.¹⁴ Notice that the lack of dependent plural reading shows up only in case of the bound reading. When interpreted as anaphoric to discourse referents, *the others* can be interpreted as a dependent plural. For example, in (63) the last sentence can be interpreted as ‘each of the boys read one book and they read two books in total.’

- (63) Students split into groups of three. Each group had to read five books. One group consisted of Angelica, Perssy and Adam. Angelica read three books. The boys read the other books.

Thus, *the others* should be analyzed by the DA-approach. This still leaves the situation with *each other* unresolved. On the one hand we should connect the range of readings to the theory of plurality to deal correctly with negated reciprocal sentences, as well as derive IR and SR without any additional mechanisms. But the DA-approach that attempts to connect the two runs into problems. In Section 4.4 I propose a new way of dealing with the range of readings which combines the theory of plurality with the R-approach to reciprocals. This, as we will see, avoids the problems that Beck’s approach faced.

Before going there, I turn to the questionnaire which tested whether the bound *the others* can give rise to weaker readings, unlike *each other*. Furthermore, it tested the intuition discussed in Section 4.3.1 that DPs appearing in reciprocal sentences are not interpreted as covarying with the antecedent of the reciprocal.

4.3.3 Questionnaire testing the range of readings and distributivity

This questionnaire tested whether the type of antecedent of *each other* influences acceptance of reciprocal sentences, in the same way as the antecedent of *the others* does. Second, the experiment tested which readings that are accepted with *each other* are accepted with *the others*. We focused on three readings, discussed above: SR, IR, IAR. The readings are exemplified here.

- (64) a. Morris, Hilary and Philip like each other. (SR)
 b. The telephone poles are 500 feet from each other. (IR)
 c. The boxes are stacked on top of each other. (IAR)

¹⁴This might not be the end of the story, though. Zweig (2008) claims that with quantified antecedents like the DPs headed by *both* or *most*, dependent plurality is derived differently than through the cumulative reading, basically by pluralizing a predicate by the * operator before the existential closure. This mechanism would also derive the dependent plural reading of *the others* but *the others* lacks the dependent plural reading, as the example here shows. It cannot be interpreted as ‘each of the boys talked to the other boy’.

- (1) Both boys talked to the others/the other boys.

Thus, if I am right about the analysis of *the others* this is an argument against Zweig’s way of dealing with dependent plural readings with the quantified antecedents.

Method

Øystein Nilsen and I designed a web-based questionnaire in which 40 undergraduate students of linguistics from University College of London participated. Before the experiment, each participant was instructed that she “will be presented with twenty eight situations consisting of a picture and a sentence describing it” and that she “will be asked whether the sentence is a true statement about the picture” and should bear in mind that “there is no “correct” or “incorrect” answer in any of the situations”. Each test item consisted of a picture under which the test sentence has been presented, as shown in Figure 4.1 (for the full list of test items, see Appendix). The participant had to click on the button to choose between the answer “true” or “false”, and could optionally fill in her commentary before moving to the next test item.

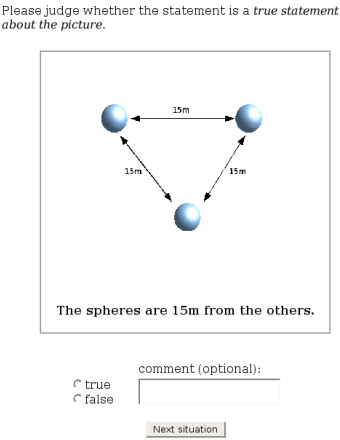


Figure 4.1: Example of a test item

The questionnaire had two lists, and each participant was sent to one of them. List 1 differed from List 2 in the type of the antecedent for *each other* and *the others*. List 1 consisted of test items in which the antecedent was a quantifier headed by *all* while List 2 consisted of test items in which the antecedent was a definite plural. In total, each picture was tested for 4 different sentences. For example, the picture in Figure 4.1 has been used as the background for testing truth value judgments in the following quadruple (where (65a) and (65b) appeared in List 1, while (65c) and (65d) appeared in List 2):

- (65) a. All the spheres are 15m from each other.
- b. All the spheres are 15m from the others.
- c. The spheres are 15m from each other.
- d. The spheres are 15m from the others.

Within each list, there were 14 test items, 3 control items and 11 fillers presented in randomized order. Of the 14 test items, 4 tested SR, 6 tested IR, and 4 tested IAR.

Table 4.3: *All NP* and *The NP*: Mean proportions of sentences judged as true in total

	antecedent = all NP		antecedent = the NP	sig.
all test items	.48	≈	.55	p=.16
each other	.55	<	.78	p=.001
the others	.36	≈	.33	p=.687

Results

Comparing the antecedents *all NP* and *the NP* Table 4.3 shows the mean proportion of sentences judged as true within the scenario. To test whether the mean proportion differed depending on the type of antecedent we applied independent t-test which compares two means of independent samples. When non-differentiated according to the type of reading and the presence of *each other* or bound *the others*, the sentences with *all NP* were less frequently judged as true than *the NP*, even though the result was not significant ($p = .16$). For the sentences in which only *each other* was used, the antecedent *all NP* significantly decreased the acceptability ($p = .001$, $t = -3.47$, $df = 35.56$). There was no significant difference between the two antecedents in case of bound *the others*.

When we further differentiate test items according to the type of reading, we get the following results (Table 4.4). First, *each other* is less accepted in SR with definite plurals than with quantified *all the NP*. The same holds for *the others*. However, this result is confounding because SR here consists of two subtypes, which are split into SR_1 and SR_2 . In SR_1 we have test items which could only be accepted under a distributive reading of indefinites in the clause. An example is given here in (66a), with the picture that formed the relevant context (Figure 4.2). As one can see the indefinite *a line* must be understood as covarying, otherwise the sentence is false.

- (66) a. The boxes are connected to each other by a line.
- b. All the boxes are connected to each other by a line.

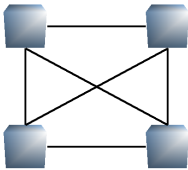


Figure 4.2: The picture accompanying (66a), (66b)

SR_2 involves test items in which no distribution over indefinites is necessary (see, for instance, (65a) above). In SR_1 both *each other* and *the others* showed effect of the antecedent: *all NP* was significantly better accepted in both cases. On the other hand, in SR_2 only the sentences with *the others* were judged as true in SR_2 more frequently

Table 4.4: *All NP* and *The NP*: Mean proportions of sentences judged as true for different readings

	antecedent = all NP		antecedent = the NP	sig.
<i>each other</i> in SR	.98	>	.72	p = .0007
<i>each other</i> in SR ₁	.95	>	.44	p = 0.0006
<i>each other</i> in SR ₂	1.0	=	1.0	–
<i>each other</i> in IR	.62	<	.88	p = .01
<i>each other</i> in IAR	.52	<	.88	p = .004
<i>the others</i> in SR	.84	>	.44	p = .0009
<i>the others</i> in SR ₁	.86	>	.33	p = .0005
<i>the others</i> in SR ₂	.81	>	.52	p = .05
<i>the others</i> in IR	.38	≈	.29	p = .4223
<i>the others</i> in IAR	.04	≈	.25	p = .06

Table 4.5: *each other* and *the others*: Mean proportions of sentences judged as true for different readings

	<i>each other</i>		<i>the others</i>	sig.
strong reciprocity	.98	≈	.84	p = .06
intermediate reciprocity	.62	>	.38	p = .006
intermediate alternative reciprocity	.52	>	.04	p = .0001

when the antecedent was *all NP* than when the antecedent was a definite plural ($p = 0.05$, $t = 2.02$, $df = 33.59$). *Each other* was fully accepted in SR₂ no matter which antecedent is chosen.

Furthermore, *all NP* is dispreferred as the antecedent for *each other* in Intermediate Reciprocity ($p = .01$, $t = -2.65$, $df = 35.63$) and Intermediate alternative reciprocity ($p = 0.004$, $t = -3.15$, $df = 31.03$). For *the others* the difference between the type of antecedent is non-significant in case of Intermediate Reciprocity and Intermediate Alternative Reciprocity.

Range of readings with *each other* and bound *the others* To compare which readings were accepted/rejected with *each other* and bound *the others*, we wanted to minimize the effect of the antecedent. Since we already knew that *the others* is less accepted when the antecedent is a definite plural than when the antecedent is a quantifier headed by *all* from the questionnaire discussed in Section 2.3 and this has also been confirmed in the first part of our Experiment 2, we decided to make use only of the data in which *all NP* is the antecedent. Table 4.5 shows the mean proportion of sentences judged as true for each type of reading. In Strong Reciprocity, the difference between *each other* and *the others* has been very close to significant ($p = .06$). In Intermediate Reciprocity, the difference was significant ($p = 0.006$, $t = 3.04$, $df = 21$). In Intermediate Alternative Reciprocity, the difference was highly significant ($p = 0.0001$, $t = 4.71$, $df = 21$).

Discussion

We have already established that *the others* in its bound reading requires that its antecedent is interpreted distributively, either thanks to the presence of the * operator or because it is a distributive quantifier. This explains why the antecedents for the bound *the others* were ordered on the scale identical to the scale of plural DPs that license distributive readings. To see whether *each other* requires the antecedent to be interpreted distributively, we can compare *all NP* and *the NP* as tested in this questionnaire 4.4). However, we must put aside the readings weaker than SR. The reason is that with IR and IAR there is a confounding factor, namely, some speakers do not accept IR or IAR when the antecedent of *each other* is *all NP*, presumably due to the preference of *all NP* to exhaustify the relation (see Brisson 1998). SR consisted of two types, SR₁ which could only be interpreted with a distributive reading over an indefinite, and SR₂ which included no indefinite. To start with the latter, we see that the reciprocal sentence is fully accepted independently of the type of antecedent. This makes sense in both DA- and R-accounts of reciprocals. The reason is that given the lexical semantics of *each other*, covariation is obligatory and thus it should be accepted with definite plurals even if it is the marked reading.

SR₁ depends on the type of subject. The acceptability was degraded if the subject was a definite plural and the difference was significant. This follows from our account, since SR₁ required the covarying interpretation of the indefinite, which is marked and therefore degraded with definite plurals, as argued in Chapter 3. Notice that the data here also argue for the viewpoint that the degradation of the reading is not due to the fact definite plurals compete with distributive quantifiers but it must be due to the fact that quantifiers headed by *all* are in competition with definite plurals (for more discussion, see Section 3.3 in the previous chapter).

The results from the availability of reading types support the DA-approach to *the others*. Table 4.5 shows that *the others* was significantly less accepted than *each other* with weaker readings. We can make sense of this if we restrict the DA-approach to *the others*. To account for IR Beck needs to assume that a QR of a variable takes place, something that runs into various problems which we discussed in Section 4.2. If we just accept that this QR is not possible we do predict low acceptance rate of IR for *the others*. Under Beck's account, we also expect that *the others* should not be accepted in IAR readings because, as Beck herself notes, these readings are not derivable in her account. *Each other* has much higher acceptability of both IR and IAR. In the next section we will develop the account of *each other* within R-approaches which can correctly derive the availability of weaker readings.

4.4 Novel R-approach to reciprocals

So far, we have seen that the DA-approach, as represented by Beck (2001) can very elegantly account for the properties of *the others*. If we want to use the R-approach for *each other* we need to make sure:

- that *each other* does not distribute over the whole clause

- that we derive IR and SR through the mechanisms used in the theory of plurality
- that we correctly derive the meaning in negated reciprocal sentences

Apart from Moltmann (1992) every R-approach I know of fails on these points, and Moltmann (1992), who correctly derives the first point (even though not fully satisfactorily, see Section 4.4.3) has problems to account for the other two.

I start with the first point. Consider the following example from Moltmann (1992), discussed in Section 4.3.1. Assume that there are only two children and each of them gave one present to the other child. We can use (67) to describe this situation.

- (67) The children gave each other Christmas presents/?? a Christmas present.
(each child giving a different present)

The problem is that in order to derive this reading we need a cumulative interpretation between *the children* and *Christmas presents*. We derive the cumulative interpretation by letting arguments be interpreted in situ. But if the theme was interpreted in situ in (67) we would incorrectly derive the following interpretation:

- (68) $(\exists e)(\forall x, y \leq \text{THE CHILDREN})(\exists e')(x \neq y \wedge x, y \in AT \rightarrow {}^{C_*}\Theta_1(x)(e') \wedge \text{GIVE}(e') \wedge {}^{C_*}\Theta_2(y)(e') \wedge {}^{C_*}\Theta_3(\mathbb{E}(\text{presents}))(e'))$

This is only true if one child gave more than one present to the other child. Consider another example, from the COCA, discussed above:

- (69) Critics and defenders of the Catholic Church have been aligned against each other in two conflicting camps.

The reading in (69) is that critics aligned against defenders and defenders aligned against critics and in total there were two competing camps. If we tried to interpret this sentence with our semantics of *each other* we could only get the interpretation that the critics were in two conflicting camps and so were the defenders, which is the incorrect interpretation here.

The problem is that in the R-approaches *each other* takes scope over a relation and distributes over it. If the relation includes another argument it cannot be interpreted cumulatively with respect to the antecedent of *each other*.

We could try to exclude the distributive quantification from the meaning of *each other* but I do not think that this is a viable option. The distributive quantification over the plural antecedent of the reciprocal is necessary if we want to express that distinct entities are related.¹⁵ Alternatively, we could let reciprocals distribute only over part of the clause. I think that this is what Williams (1991) informally suggests reciprocals should do so I will try to formalize his idea.

There are independent reasons that we might want a plural DP to distribute only over some but not all arguments in the clause. As far as I know, these data have been

¹⁵Faller (2007) assumes that the distinctness condition might follow independently of the meaning of reciprocal, namely from Condition B. This might work for the data she considers in her paper (reciprocity in Cuzco Quechua) and it might be a potential way to avoid distributive quantification in reciprocal sentences. However, English *each other* keeps the distinctness condition also in cases in which the reciprocal and antecedent are not co-arguments which makes resorting to Condition B problematic.

discovered by Schein (Schein, 1993). Accommodating them to the system will give us a handle to deal with reciprocals distributing only over part of the clause so I will start with discussing these data. After that I come back to *each other* and its meaning that is cast within the R-approach, and show how it derives the correct interpretation in negated reciprocal sentences and how it can capture the range of readings for which reciprocal sentences are so well-known.

4.4.1 Dependent subevents: Argument for a third type of distributivity

Right now, we use the * operator in two cases. In one case to pluralize thematic roles and events. This gives us a cumulative reading (and possibly a branching reading, as discussed in Section 3.2.5), i.e., the reading in which none of the DPs scopes over the others. The second case is when a DP undergoes QR and the * operator applies to the lambda-abstracted predicate. This derives a distributive reading, in which one argument scopes and distributes over the rest of the clause.

Now, following Schein (1993) and Landman (2000) I do not think that this exhausts all possibilities of the application of the * operator.

Consider the following sentence, from Roberts (1990).

- (70) Five insurance associates gave a \$ 25 donation to several charities.

Here, I am going to assume exactly-interpretation of each numeral argument, so the sentence could be paraphrased as ‘exactly five insurance associates gave exactly 25 dollars to several charities’.

It seems (*pace* intuition of Roberts and Kratzer) that (70) can be interpreted as follows: each of five insurance associates was donating money to charities in such a way that taken together, each charity received \$ 25 contribution. So, for example, there were three charities A, B and C. Two associates gave money to A and B, three associates gave money to B and C and in total, each of A, B and C had \$ 25 after the contributions.

This interpretation of (70) shows that some arguments can be interpreted cumulatively with respect to each other while another argument can be interpreted as scopally dependent on one of the arguments. In (70) *five associates* and *several charities* are interpreted cumulatively with respect to each other, so we get the interpretation that there are five associates and several charities and each of the associates donated to some of the charities and each of the charities got donation from some of the associates. Furthermore, \$ 25 is scopally dependent on the DP *several charities*, without being dependent on the first argument. If it was dependent on *five associates* we would get the interpretation that each of them donated \$ 25, but this is not the interpretation we are after. In fact, each of the associates can donate different sum, as long as each of the charities end up with 25 dollars, so, \$ 25 is dependent only on *several charities*.

Another example showing the same point is the following sentence, from Landman (2000).

- (71) Three boys gave six girls two flowers.

Again, I am going to assume exactly-interpretation of each numeral argument, so the sentence could be paraphrased as ‘exactly three boys gave exactly six girls exactly two flowers’.

Now, assume the cumulative interpretation for the first two arguments, so, three boys were giving flowers to girls, and six girls got flowers from boys. The question is, can the theme argument *two flowers* be scopally dependent? It seems that it can and in that case we get one of the interpretations paraphrased in (72).

- (72) a. Three boys gave six girls two flowers per girl.
b. Three boys gave six girls two flowers per boy.

Thus, flower-giving took place between three boys and six girls and in total, each girl received two flowers (72a). Alternatively, flower-giving took place between three boys and six girls and in total, each boy gave some girl two flowers (72b). The second case would be true if, for example, boy1 gave one flower to girl1 and one flower to girl2, boy2 gave one flower to girl3 and one flower to girl4 and boy3 gave one flower to girl5 and one flower to girl6. The first case would be true if, for example, each of girl1 and girl2 received one flower from boy1 and one flower from boy2, each of girl3 and girl4 received one flower from boy1 and one flower from boy3 and each of girl5 and girl6 received one flower from boy2 and one flower from boy3. It seems that these readings are possible. For more discussion and more examples, see Roberts (1990), Schein (1993), Landman (2000) and Kratzer (2003).

At this point, we cannot account for these readings. Let me focus on (71). Just to discuss the most obvious option we could assume that the DP *six girls* undergoes QR and is interpreted as distributing over the event. In that case, we would get the LF structure and interpretation in (73).

- (73) a. $[6 \text{ girls}] * \lambda x. 3 \text{ boys gave } x \text{ 2 flowers}$
b. $(\forall x \leq 6 \text{ GIRLS})(\exists e)(3 \text{ boys gave } x \text{ 2 flowers at } e)$

Obviously, in this case the DP *six girls* ends up distributing over the subject, thus we lose the cumulative interpretation between the two arguments. What we need is the following:

- cumulative interpretation between two arguments (*three boys* and *six girls* in (71))
- distribution of one of these DPs over the third argument (for example, *six girls* over *two flowers* in (71a))

So far none of the application of the * operator can achieve this. Pluralizing thematic roles is too local since in that case the * operator does not take scope over other DPs. On the other hand, using the * operator that derives the distributive reading does not help either since in that case the whole event is scopally dependent. Since the cumulative interpretation arises when the arguments are interpreted in thematic positions, we have to add a new distributive interpretation in which one argument distributes over another without undergoing QR to scope over the whole event. Furthermore, some

Table 4.6: Operators creating a dependent event type

MAKE _{Θ_{dep}}	$\lambda R_{\langle e, \langle v, t \rangle \rangle} \lambda y \lambda e' \lambda e. R(y)(e') \wedge e' \leq e$
ADD _{Θ_{dep}}	$\lambda R_{\langle e, \langle v, t \rangle \rangle} \lambda Q_{\langle v, \langle v, t \rangle \rangle} \lambda y \lambda e' \lambda e. R(y)(e') \wedge Q(e')(e)$
A-LIFT	$\lambda Q_{\langle e, \langle v, \langle v, t \rangle \rangle \rangle} \lambda T_{\langle \langle e, t \rangle, t \rangle} \lambda e' \lambda e. T(\lambda x. Q(x)(e')(e))$
* _{dep}	$\lambda Q_{\langle e, \langle v, \langle v, t \rangle \rangle \rangle} \lambda x \lambda e. \langle x, e \rangle \in [\{ \langle a, b \rangle : Q(a)(b)(e) \}]$

but not all arguments should be scopally dependent. What we need is what Landman (2000) calls dependent event types, where an event type is a lambda-abstracted event, or equivalently set of events.

- (74) **A dependent event type** is an event type which is scopally dependent on some, but not necessarily all of the arguments in an event type.

Even though I follow Landman's intuition in deriving dependent event types, my actual implementation differs from his. This is because his account requires various idiosyncratic properties of the syntax (for example, thematic roles must all be introduced by the same head) and it furthermore works only when the dependent argument (the argument over which another argument distributes) is a quantifier of type $\langle \langle e, t \rangle, t \rangle$. This works fine for the examples presented here but cannot be extended to reciprocal sentences in which *each other* is the dependent argument.

To derive the dependent event type we proceed as follows. First, we let the arguments that appear in dependent event types be related by a thematic role to a *subevent*. Second, we define a new * operator that distributes only over the subevents. The new operators that achieve this are defined in Table 4.6.

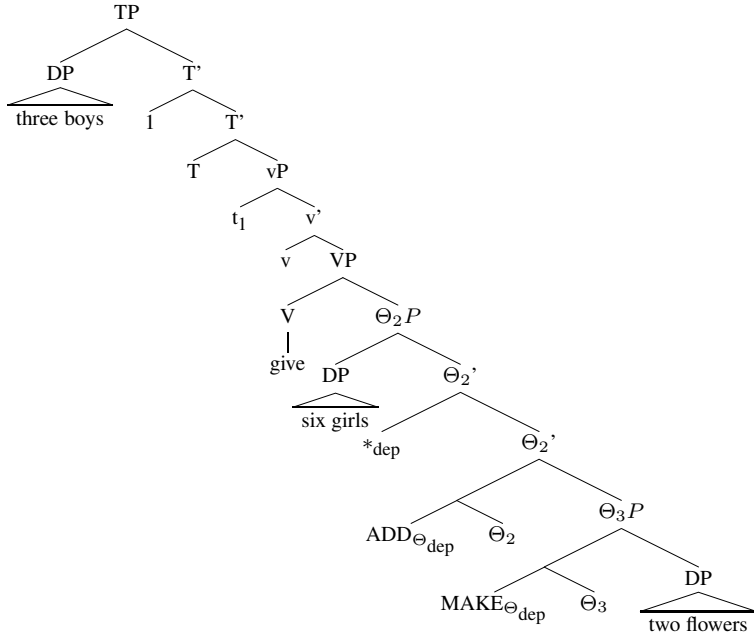
The first two operators change a thematic role in such a way that it now relates entities and subevents of e (why we need two operators will become clear in a minute). Furthermore, we now have a situation where thematic roles are not of type $\langle e, \langle v, t \rangle \rangle$ but of type $\langle e, \langle v, \langle v, t \rangle \rangle \rangle$. This is because they can be arguments of MAKE and ADD which adds the extra event argument. In this case, the old A-LIFT which allowed generalized quantifiers to apply to thematic roles, would not be usable due to type mismatch. We just need to adjust it to be applicable to thematic roles of type $\langle e, \langle v, \langle v, t \rangle \rangle \rangle$. The fourth operator is the * operator which differs from the old one only in that it distributes over subevents. Let me show how this works on the example from above, repeated here:

- (75) Three boys gave six girls two flowers.

I am interested in the interpretation (72a), that is, the one in which *three boys* and *six girls* are interpreted cumulatively and *two flowers* are scopally dependent on the girls.

The LF tree for this interpretation is given in (76).

(76)



Here is, how we get the meaning stepwise.

First, $\text{MAKE}_{\Theta_{\text{dep}}}$ combines with Θ_3 which gives us (77a). We now have a thematic role Θ_3 which relates an argument to e' , the subevent of e . This applies to the DP *two flowers* by function application with LIFT, as shown in (77b). In (77c) $\text{ADD}_{\Theta_{\text{dep}}}$ applies to Θ_2 and the result applies to Θ_3P , which is shown in (77d). This is taken as the argument of $*_{\text{dep}}$ (77e) and the result can combine by function application with LIFT with the object *six girls*. The rest proceeds as usual. I skip it here and give only the interpretation of the top node, TP.

- (77) a. $[[\text{MAKE}_{\Theta_{\text{dep}}}(\Theta_3)]] = (\text{FA})$
 $\lambda y \lambda e' \lambda e. C_{*} \Theta_3(y)(e') \wedge e' \leq e$
- b. $[[\Theta_3P]] = [[(77a)(\text{two flowers})]] = (\text{FAL})$
 $\lambda e' \lambda e. (\exists y)(|AT(y)| = 2 \wedge *_{\text{FLOWER}}(y) \wedge C_{*} \Theta_3(y)(e') \wedge e' \leq e)$
- c. $[[\text{ADD}_{\Theta_{\text{dep}}}(\Theta_2)]] = (\text{FA})$
 $\lambda Q \lambda z \lambda e' \lambda e. C_{*} \Theta_2(z)(e') \wedge Q(e')(e)$
- d. $[[\Theta_2']] = [[(77c)(77b)]] = (\text{FA})$
 $\lambda z \lambda e' \lambda e. (\exists y)(|AT(y)| = 2 \wedge *_{\text{FLOWER}}(y) \wedge C_{*} \Theta_2(z)(e') \wedge C_{*} \Theta_3(y)(e') \wedge e' \leq e)$
- e. $[[*_{\text{dep}}(77d)]] = (\text{FA})$
 $\lambda z \lambda e. \langle z, e \rangle \in \{ \langle a, e' \rangle : (\exists y)(|AT(y)| = 2 \wedge *_{\text{FLOWER}}(y) \wedge C_{*} \Theta_2(a)(e') \wedge C_{*} \Theta_3(y)(e') \wedge e' \leq e) \}$
- f. $[[\Theta_2P]] = [(77e)(\text{six girls})] = (\text{FAL})$
 $\lambda e. (\exists z)(|AT(z)| = 6 \wedge *_{\text{GIRL}}(z) \wedge$

$$\begin{aligned}
& \wedge \langle z, e \rangle \in [\{ \langle a, e' \rangle : (\exists y)(y = 2 \text{ flowers} \wedge C_*\Theta_2(a)(e') \wedge C_*\Theta_3(y)(e') \wedge e' \leq e) \}] \\
\text{g. } & [[\text{TP}]] = \\
& (\exists e)(\exists x, z)(|AT(x)| = 3 \wedge *BOY(x) \wedge C_*\Theta_1(x)(e) \wedge *GIVE(e) \wedge \\
& \wedge |AT(z)| = 6 \wedge *GIRL(z) \wedge \\
& \wedge \langle z, e \rangle \in [\{ \langle a, e' \rangle : (\exists y)(|AT(y)| = 2 \wedge *FLOWER(z) \wedge C_*\Theta_2(a)(e') \wedge C_*\Theta_3(y)(e') \wedge e' \leq e) \}])
\end{aligned}$$

(77g) is true in the following situation. There is an event e which has subevents $e_1 \dots e_6$. These are all events of giving. In every two subevents one boy is the argument of Θ_1 . The pair $\langle \text{SIX GIRLS}, e \rangle$ is a member of the i-join semilattice whose generators are members of the following set:

$$(78) \quad \{ \langle a, e' \rangle : (\exists y)(|AT(y)| = 2 \wedge *FLOWER(y) \wedge C_*\Theta_2(a)(e') \wedge C_*\Theta_3(y)(e') \wedge e' \leq e) \}$$

Since $\langle \text{SIX GIRLS}, e \rangle$ is a member of the i-join semilattice generated by this set it means that we can split the pair $\langle \text{SIX GIRLS}, e \rangle$ into parts and each part is in the set. Suppose we split them into the following pairs: $\langle \text{GIRL}_1, e_1 \rangle, \dots, \langle \text{GIRL}_6, e_6 \rangle$. In each one of these subevents Θ_3 holds of two flowers. Thus, we can visualize the reading as the following triples, where the first two positions represent boys and girls and the third position represent number of flowers given.

$$(79) \quad \begin{aligned} & \bullet \langle \text{BOY}_1, \text{GIRL}_1, 2 \rangle \\ & \bullet \langle \text{BOY}_1, \text{GIRL}_2, 2 \rangle \\ & \bullet \langle \text{BOY}_2, \text{GIRL}_3, 2 \rangle \\ & \bullet \langle \text{BOY}_2, \text{GIRL}_4, 2 \rangle \\ & \bullet \langle \text{BOY}_3, \text{GIRL}_5, 2 \rangle \\ & \bullet \langle \text{BOY}_3, \text{GIRL}_6, 2 \rangle \end{aligned}$$

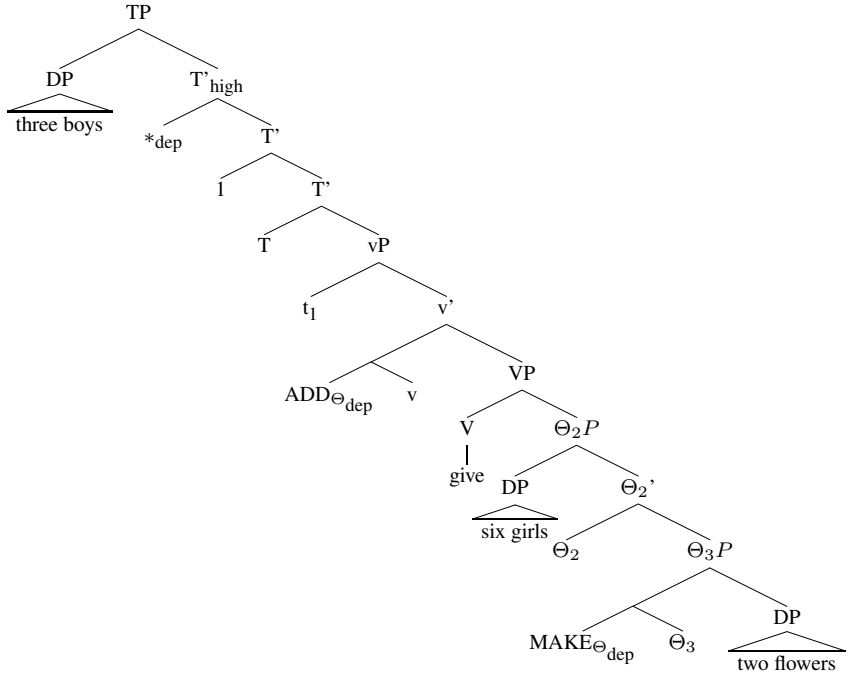
So, each boy gave flowers to some girl and each girl got some flowers. Each boy gave out four flowers in total, and each girl received two flowers in total. In other words, *three boys* and *six girls* are interpreted cumulatively but *2 flowers* is dependent on the latter argument. This is the reading we were after.

Intuitively we are getting this result in the following way: $\text{MAKE}_{\Theta_{\text{dep}}}$ and $\text{ADD}_{\Theta_{\text{dep}}}$ ensure that thematic roles to which apply relate its argument to some subevent e' . We need both operators because the first one creates a subevent but $\text{ADD}_{\Theta_{\text{dep}}}$ specifies that the thematic role to which it applies relates its argument to the same subevent. If we had only the first null operator, we could not ensure that both thematic roles are related to the same subevent.

Consider now the interpretation (72b), that is, the one in which *three boys* and *six girls* are interpreted cumulatively but *two flowers* are scopally dependent only on the boys.

The LF tree for this interpretation differs from the previous one only in having the top thematic role modified by $\text{ADD}_{\Theta_{\text{dep}}}$. It is given in (80).

(80)



The interpretation proceeds up to $\Theta_3 P$ in the same way as before. This is shown in (81a). (81a) combines with Θ_2 by event identification. Notice that (81b) is almost identical to (77d), which are both interpretations of Θ_2' , with one crucial difference: here, z is not related by Θ_2 to a part of event, but to e directly. This is the crucial step which will enable the subject to distribute over *two flowers* without distributing over *six girls*. On the other hand, Θ_1 relates x to the part of event. After that, the derivation should be straightforward. The interpretation of the top node with existential closure over event type is in (81f).

- (81) a. $[[\Theta_3 P]] =$
 $\lambda e' \lambda e. (\exists y) (|AT(y)| = 2 \wedge *FLOWER(y) \wedge C_* \Theta_3(y)(e') \wedge e' \leq e)$
- b. $[[\Theta_2']] =_{(EI)}$
 $\lambda z \lambda e' \lambda e. (\exists y) (|AT(y)| = 2 \wedge *FLOWER(y) \wedge C_* \Theta_2(z)(e) \wedge C_* \Theta_3(y)(e') \wedge e' \leq e)$
- c. $[[\Theta_2 P]] =_{(FAL)}$
 $\lambda e' \lambda e. (\exists y, z) (|AT(z)| = 6 \wedge *GIRL(z) \wedge |AT(y)| = 2 \wedge *FLOWER(y) \wedge C_* \Theta_2(z)(e) \wedge C_* \Theta_3(y)(e') \wedge e' \leq e)$
- d. $[[v']] =_{(FA)}$
 $\lambda x \lambda e' \lambda e. (\exists y, z) (|AT(z)| = 6 \wedge *GIRL(z) \wedge |AT(y)| = 2 \wedge *FLOWER(y) \wedge *GIVE(e) \wedge C_* \Theta_1(x)(e') \wedge C_* \Theta_2(z)(e) \wedge C_* \Theta_3(y)(e') \wedge e' \leq e)$
- e. $[[T'_{high}]] =_{(FA)}$
 $\lambda x \lambda e. \langle x, e \rangle \in [\{ \langle a, e' \rangle : (\exists y, z) (|AT(z)| = 6 \wedge *GIRL(z) \wedge |AT(y)| =$

- $2 \wedge *FLOWER(y) \wedge *GIVE(e) \wedge {}^{C*}\Theta_1(x)(e') \wedge {}^{C*}\Theta_2(z)(e) \wedge {}^{C*}\Theta_3(y)(e') \wedge e' \leq e\} \}$
- f. $[[TP]] =$
 $(\exists e)(\exists x)(|AT(x)| = 3 \wedge *BOY(x) \wedge$
 $\wedge \langle x, e \rangle \in [\{ \langle a, e' \rangle : (\exists y, z)(|AT(z)| = 6 \wedge *GIRL(z) \wedge |AT(y)| = 2 \wedge$
 $*FLOWER(y) \wedge *GIVE(e) \wedge {}^{C*}\Theta_1(x)(e') \wedge {}^{C*}\Theta_2(z)(e) \wedge {}^{C*}\Theta_3(y)(e') \wedge$
 $e' \leq e) \}])$

(81f) is true iff three boys and an event are, as a pair, a member of the i-join semilattice whose generators are members of the following set:

$$(82) \quad \{ \langle a, e' \rangle : (\exists y, z)(|AT(z)| = 6 \wedge *GIRL(z) \wedge |AT(y)| = 2 \wedge *FLOWER(y) \wedge {}^{C*}\Theta_1(x)(e') \wedge {}^{C*}\Theta_2(z)(e) \wedge {}^{C*}\Theta_3(y)(e') \wedge e' \leq e) \}$$

We can simplify this. We are looking for a, e' that satisfies the conditions of the set. But for this, a few conjuncts are irrelevant, namely $|AT(z)| = 6$, $*GIRL(z)$, $\Theta_2(z)(e)$ and $*GIVE(e)$ since they include neither a nor e' . So we can rewrite the interpretation of the TP as follows:

$$(83) \quad (\exists e, x, z)(|AT(x)| = 6 \wedge *BOY(x) \wedge |AT(z)| = 6 \wedge *GIRL(z) \wedge *GIVE(e) \wedge {}^{C*}\Theta_2(z)(e) \wedge \wedge \langle x, e \rangle \in [\{ \langle a, e' \rangle : (\exists y)(|AT(y)| = 2 \wedge *FLOWER(y) \wedge {}^{C*}\Theta_1(x)(e') \wedge {}^{C*}\Theta_3(y)(e') \wedge e' \leq e) \}])$$

(83) is true if six girls are related by Θ_2 to e . Suppose that we have six subevents of e and in each of them one girl is related to e . Furthermore, the pair $\langle \text{THREE BOYS}, e \rangle$ is a member of the i-join semilattice generated from the following set:

$$(84) \quad \{ \langle a, e' \rangle : (\exists y)(|AT(y)| = 2 \wedge *FLOWER(y) \wedge {}^{C*}\Theta_1(x)(e') \wedge {}^{C*}\Theta_3(y)(e') \wedge e' \leq e) \}$$

For example, if boy₁ gave two flowers in $e_1 \oplus e_2$, boy₂ gave two flowers in $e_3 \oplus e_4$ and boy₃ gave two flowers in $e_5 \oplus e_6$, the pair $\langle \text{THREE BOYS}, e \rangle$ is a member of the generated semilattice. Thus, in that case, we can visualize the situation which satisfies (83) as follows, where the triples represent the giver, the one who received something and the number of flowers that were given.

$$(85) \quad \begin{aligned} &\bullet \langle \text{BOY}_1, \text{GIRL}_1, 1 \rangle \\ &\bullet \langle \text{BOY}_1, \text{GIRL}_2, 1 \rangle \\ &\bullet \langle \text{BOY}_2, \text{GIRL}_3, 1 \rangle \\ &\bullet \langle \text{BOY}_2, \text{GIRL}_4, 1 \rangle \\ &\bullet \langle \text{BOY}_3, \text{GIRL}_5, 1 \rangle \\ &\bullet \langle \text{BOY}_3, \text{GIRL}_6, 1 \rangle \end{aligned}$$

So, each boy gave flowers to some girl and each girl got some flowers. Furthermore, each boy gave out two flowers in total. This is the reading we were after.

There are two other accounts which try to explain why a DP can be interpreted distributively while other arguments stay in a cumulative interpretation. I shortly discuss them in the following sections.

Winter (2000), Kratzer (2003): Dependent definites

Winter (2000) assumes that in cases where two arguments are interpreted cumulatively and the third argument is scopally dependent, the cumulative interpretation is illusory and can be derived through a distributive interpretation if we treat definites as inherently relational, as assumed, for example, in Chierchia (1995). In that case *the girls* in (86a) could be paraphrased as ‘the girls of x ’, where x is bound by each boy, so we get interpretation that could be paraphrased in (86b). The interpretation is the one we were after: every boy gave two flowers to some of the girls and every girl got some flowers.

- (86) a. The boys _{i} gave the girls two flowers.
 b. Each boy gave his girls two flowers.

We get the right interpretation because of the following fact: the hallmark of the cumulative interpretation is that we can split two plural arguments in such a way that parts of one argument are related to parts of the other argument. Because *the girls of x* for any x is part of *the girls*, and each boy x is related only to *the girls of x* we correctly get that the parts of the plurality *the boys* are related to the parts of the plurality *the girls*. Furthermore, because *the boys* is interpreted distributively, they distribute over the argument *two flowers*, so we get the right interpretation.

However, I do not see how another interpretation, where *two flowers* is scopally dependent on *the girls* could be derived. We would have to assume that *the boys* is inherently relational and scopally dependent on the girls so we would end up with (87). But this is problematic since in this case, *the girls* would have to undergo QR over a DP that it binds and this should lead to Weak Crossover Violation.

- (87) For each girl, her boys gave her two flowers.

Another problem comes with the interpretation of plurals.¹⁶ Consider the following example (from Winter 2000, p. 56):

- (88) The circles are connected to the triangles by a dashed line.

Consider the following scenario: there are two circles and inside each of them there is one triangle and each circle is connected by a dashed line to the triangle that is inside it. Intuitively, the sentence is true in this case. However, the interpretation that Winter (2000) gets is ‘each circle is connected to its *triangles* by a dashed line’. Obviously, the problem is that the plurality of the dependent DP should still be interpreted and since *the triangles* are in Winter’s account dependent on *the circles* (which are interpreted distributively) there should be more than one triangle for each circle. Winter’s account would render the sentence false in this case, contrary to one’s intuition.

Winter (2000) suggests that this could be the effect of dependent plurals. Following Chomsky (1975) he assumes that in this case, the plurality on *the triangles* is only apparent, in reality it is just a syntactic plural number that is not interpreted semantically.

¹⁶This and the following problem were also discussed in Beck and Sauerland (2000).

But as I argued in Section 4.3.3, following Zweig (2008), this is an incorrect assumption. The plurality is still interpreted even with dependent plurals. The difference is that with dependent plurals the plurality requirement must be satisfied outside the scope of the other plural DP. Thus, in (88) the plural marker on *the triangles* still means that we must deal with a plural object here, as can be witnessed by the fact that the sentence is odd if two circles were connected to one triangle by a dashed line.

Now, the fact that the plurality requirement must still be satisfied, albeit not in scope of the other plural DP just follows from the cumulative analysis of these data. This is so because in a the cumulative interpretation plural DPs are not in scope with respect to each other. However, the same is a mystery in Winter's account: if the plural marker was simply not interpreted then the sentence should be fine in the scenario with one triangle but it is not. If the plural marker was interpreted then the sentence should be false also in the scenario where each circle is connected to one triangle which it is not. Thus, Winter's account either overgenerates or undergenerates, depending on one's assumption on interpreting plural markers. Either way, it does not account for the data.

Finally, this account cannot be extended to the cases where other DPs than definites enter cumulative interpretation and another argument is scopally dependent on them. Consider, for example, a sentence that we discussed before:

- (89) Three boys gave six girls two flowers.

I am interested in the interpretation where *three boys* and *six girls* are interpreted cumulatively and *two flowers* is scopally dependent on *three boys*, that is, each boy gave two flowers. To derive the reading, we could try to assume that *six girls* is relational, so it should really be interpreted as *six girls of x*, where *x* is bound by each boy among the three boys. But this does not lead to the right interpretation. It would give us the reading that each boy gave two flowers to (his) six girls, that is *three boys* and *six girls* would not be in the cumulative interpretation any longer.

The problem is that while *the girls of x* can be a proper part of *the girls*, which is crucial if we want to mimic the cumulative interpretation, *six girls of x* is not a proper part of *six girls*. Obviously, this is due to the presence of the numeral in the DP.

To be fair, Winter (2000) is well-aware of this fact and in fact, he argues that this might be what we want. He suggests that when we see what looks like a cumulative interpretation between two DPs with a third DP scopally dependent, the two DPs in the cumulative interpretation are preferably definite plurals and when we substitute them with conjunctions of proper names or numeral DPs, the cumulative interpretation disappears. This would follow from his account.

Now, this claim contradicts the data known from literature (Roberts, 1990; Schein, 1993; Landman, 2000; Beck and Sauerland, 2000; Kratzer, 2003), which I focused on so far. But let me ignore them for a moment and see some of the data with which Winter supports his claim. Consider the following context: in Figure 4.3, Mary and Sue are John's children, and Ann and Ruth are Bill's children (the picture and the example are from Winter (2000), p. 45). In this situation, (90a) seems true but the second example false, or at least rather odd. An objection in the latter case would be that John is not separated from Ann and Ruth, for example.

- (90) a. The fathers are separated from the children by a wall.
 b. John and Bill are separated from Mary, Sue, Ann, and Ruth by a wall.

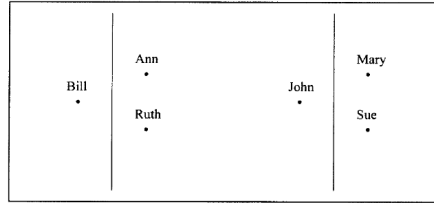


Figure 4.3: The picture accompanying (90),

This difference follows in Winter's account. To make the first sentence true in Figure 90 it suffices to let the subject distribute and interpret the object as inherently relational and bound by the subject. In the second sentence we cannot do the same because even if we treat the conjunction of proper names as relational, *Mary, Sue, Ann and Ruth of x* would not give us a proper part of *Mary, Sue, Ann and Ruth*.

Even though I find the contrast clear, I do not think that Winter's account of the difference between (90a) and (90b) is on the right track. Consider the following pair:

- (91) a. The fathers are separated from the children.
 b. John and Bill are separated from Mary, Sue, Ann, and Ruth.

It seems to me that in the scenario given above the same contrast holds, that is, (91a) is true and (91b) is false or odd. But in this case, there is no DP *a wall* that needs to be varied for each father so no dependent definites need to be assumed. If Winter wanted to extend his account of (90) to (91) he would have to say that a cumulative interpretation never exists. It is always derived as a case of distribution over dependent definites. But that would mean that no other DPs apart from definite plurals enter a cumulative interpretation, ever. This is too strong and goes against the data collected from naive subjects in questionnaires, attested examples etc, see, for example, Gil's questionnaire in Gil (1982). Winter suggests that there is another mechanism which can derive the cumulative interpretation in cases like (91) where there is no third DP that would be scopally dependent on one of the other arguments. But as long as we have this other mechanism, we cannot deal with the difference between ((91)a) and ((91)b).

My conclusion is this:

- Based on (91) we see that the likelihood of cumulative interpretation depends on the type of DP. Definite plurals are preferred over conjunctions of proper names (in this particular case; other factors, like the type of predicate might influence this preference). I do not know why this is.
- Whatever explains the difference in (91) carries over to (90). Here also, we see that the cumulative interpretation of the conjunctions of proper names is dispreferred.

- Treating (91) and (90) in parallel dissolves Winter's argument in favour of analyzing (90) exclusively as a case of dependent definites. The contrast simply stems from the fact that other types of DP than definite plurals are dispreferred in a cumulative interpretation, as we have established based on (91). This means that we can analyze (90) by the analysis I proposed (*the fathers* and *the children* in a cumulative interpretation, *a wall* scopally dependent on *the fathers*).

Cumulativity as a pluralized relation between DP arguments

There is long tradition to account for a cumulative interpretation by pluralizing the relation between two DP arguments. This way of accounting for a cumulative reading is especially common in eventless semantics (see, for example, Scha (1981); Schwarzschild (1996); Beck and Sauerland (2000); Ferreira (2007) and the literature therein). Consider the cumulative interpretation of (92a): Morris married Desiree and Philip married Hilary. Non-pluralized relations has only atoms in its extension. Thus, to get the cumulative interpretation we let the $*$ operator apply to the relation of marrying, which gives us (92c).

- (92) a. Morris and Philip married Desiree and Hilary.
 b. $[\text{Morris and Philip}] [\text{Desiree and Hilary}] * \lambda x \lambda y. x \text{ married } y$
 c. $\langle M \oplus P, D \oplus H \rangle \in [\{\langle a, b \rangle : a \text{ married } b\}]$

(92c) is true if we can split $\langle M \oplus P, D \oplus H \rangle$ into pairs of atoms and each pair is in the set $\{\langle a, b \rangle : a \text{ married } b\}$. This is true in the situation given above.

If we can form the relation to which the $*$ operator applies in the syntax nothing prevents us from letting the $*$ operator scope over other arguments. In the example below, we let *twice* be part of the relation which is pluralized. This can give us the interpretation Morris married Desiree twice and Philip married Hilary twice.

- (93) a. Morris and Philip married Desiree and Hilary twice.
 b. $[\text{Morris and Philip}] [\text{Desiree and Hilary}] * \lambda x \lambda y. x \text{ married } y \text{ twice}$
 c. $\langle M \oplus P, D \oplus H \rangle \in [\{\langle a, b \rangle : a \text{ married } b \text{ twice}\}]$

There is a problem with this account, which is discussed at great length in Landman (2000). I will present the problem here only briefly, for more discussion see the cited work.

The problem is that the DP must be scopally dependent on both arguments in this approach. Consider our previous example, repeated in (94a). To let *a \$25 donation* be scopally dependent, we have to have the LF structure in (94b), which is interpreted in (94c).

- (94) a. Five insurance associates gave a \$ 25 donation to several charities.
 b. $[5 \text{ associates}] [several \text{ charities}] * \lambda x \lambda y. x \text{ gave } y \$ 25.$
 c. $\langle 5 \text{ ASSOCIATES}, SEVERAL \text{ CHARITIES} \rangle \in [\{\langle a, b \rangle : a \text{ gave } b \$ 25\}]$

(94c) is true if we can split the pair $\langle 5 \text{ ASSOCIATES}, SEVERAL \text{ CHARITIES} \rangle$ into parts such that each of these parts is in the set:

(95) $\{\langle a, b \rangle : a \text{ gave } b \$25\}$

Consider the situation above, which is true for (94a): there are three charities and each one of them ends up with 25 dollars after donations. Furthermore, each of the five associates donated to at least one of these charities. For example, the following situation will make the sentence true:

- (96)
- $\langle \text{ASSOCIATE}_1, \text{CHARITY}_A, \$10 \rangle$
 - $\langle \text{ASSOCIATE}_2, \text{CHARITY}_A, \$15 \rangle$
 - $\langle \text{ASSOCIATE}_3, \text{CHARITY}_B, \$5 \rangle$
 - $\langle \text{ASSOCIATE}_4, \text{CHARITY}_B, \$10 \rangle$
 - $\langle \text{ASSOCIATE}_5, \text{CHARITY}_B, \$10 \rangle$
 - $\langle \text{ASSOCIATE}_3, \text{CHARITY}_C, \$25 \rangle$

The account of cumulativity we are considering here predicts that the sentence should be false in this situation. We are looking for associate-charity pairs, such that each pair is in the set (95) and when we sum up the pairs together, we get $\langle 5 \text{ ASSOCIATES, SEVERAL CHARITIES} \rangle$. This means that the relation ‘give 25 dollars to’ must hold for each pair, and this is not satisfied in the situation above, where in each associate-charity pair various sums were given. Instead of deriving this reading, this account derives only a very dubious-looking reading: it does not matter how much money each charity received and it does not matter how much money each associate gave as long as in each pair 25 dollars were given. So, for example, the following situation should make the sentence true:

- (97)
- $\langle \text{ASSOCIATE}_1, \text{CHARITY}_A, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_1, \text{CHARITY}_C, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_2, \text{CHARITY}_A, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_2, \text{CHARITY}_B, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_2, \text{CHARITY}_C, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_3, \text{CHARITY}_B, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_4, \text{CHARITY}_B, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_4, \text{CHARITY}_C, \$25 \rangle$
 - $\langle \text{ASSOCIATE}_5, \text{CHARITY}_C, \$25 \rangle$

So in this case, each of the five associate gave 25 dollars to some charity and each of the three charities received 25 dollars from some associate but only two associates gave 25 dollars (some of them gave 50 dollars, some of them gave 75 dollars) and none of the charities had 25 dollars in the end (some of them had 50 dollars, some of them had 75 dollars, some of them had 100 dollars). I do not think that there is such a reading. But even more worrisome is that the reading that is possible, namely the one in which each of several charities ended up with 25 dollars but the donations of associates differed, is not derived.

There is one way we could derive the relevant reading. It would fall out if we assumed event semantics with pluralized thematic roles, on top of which syntactically derived relations could be pluralized. In that case we would get the following:

$$(98) \quad \langle 5 \text{ ASSOICATES, SEVERAL CHARITIES} \rangle \in \\ \in [\{ \langle a, b \rangle : (\exists e, y) (\$25(y) \wedge *GIVE(e) \wedge * \Theta_1(a)(e) \wedge * \Theta_2(b)(e) \wedge * \Theta_3(y)(e)) \}]$$

This sentence is true in case each charity ends up with 25 dollars, regardless how the 5 associates donated the money. This is the relevant reading. Another way to derive the reading in eventless semantics is the following:

$$(99) \quad \langle 5 \text{ ASSOICATES, SEVERAL CHARITIES} \rangle \in \\ \in [\{ \langle a, b \rangle : *(\lambda x \lambda y \lambda z. gave(x)(y)(z))(a)(b)(\$25) \}]$$

Notice how (99) differs from (94c): the difference is that (99) includes pluralization of the ditransitive GIVE, which in turn is inside a pluralized relation. In (94c) we did not assume pluralization of GIVE. Thus, this difference is crucial for the correct derivation of the relevant reading. Champollion (to appear) offers a more detailed discussion on how other cases which Schein (1993) used as an argument for event semantics could be derived in eventless semantics if we let * operators apply to a relation which itself involves a pluralized ditransitive. I leave it open for future research whether this account would also be helpful in case of reciprocals, to which I turn now. In the next section, I will make use of dependent event types to capture the interpretation of *each other*.

4.4.2 Reciprocals in dependent event types

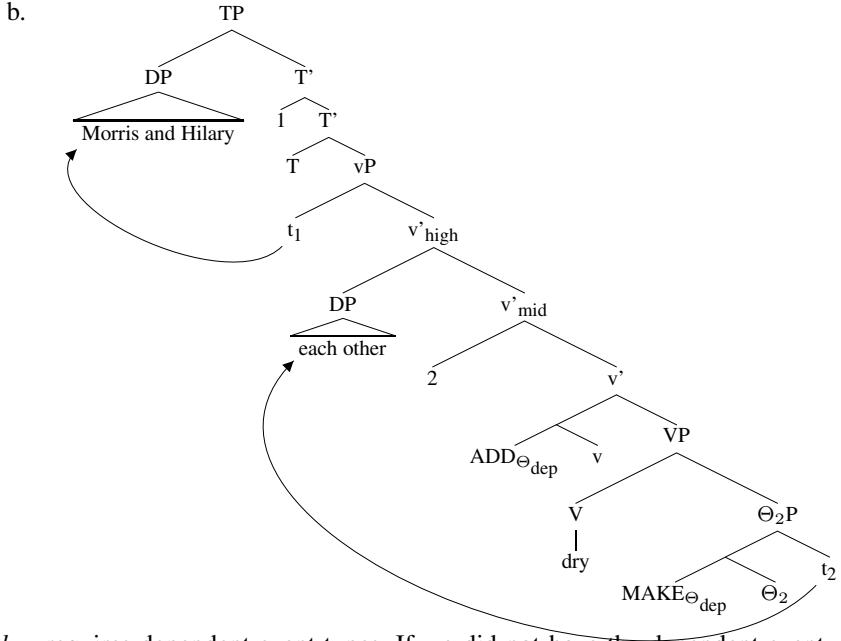
Each other is an anaphor set up in the R-approach that makes use of the dependent event types. Its meaning is given in (100).

$$(100) \quad \text{Interpretation of } each\ other, \text{ R-approach, the final version} \\ [[each\ other]] = \lambda Q_{\langle e, \langle e, \langle v, \langle v, t \rangle \rangle \rangle \rangle} \lambda x \lambda e. \langle x, x, e \rangle \in \\ \in [\{ \langle a, b, e' \rangle : Q(a)(b)(e')(e) \wedge \neg a \circ b \}]$$

Let me break this down. Reciprocals add two aspects to the meaning of a sentence. First, they express an anaphoric relation to a plural entity. Second, they specify how members of the plural entity are related – namely, each entity is related to some other entity than itself by the clausal relation. The first condition is expressed in the meaning by having x be twice part of the tuple that must be in the i-join semilattice generated by the set. The second condition is satisfied by $\neg a \circ b$. In the next subsection, I will show that this meaning of *each other* can be composed of parts that are needed independently and adds only one piece which we did not encounter so far, which is the condition $\neg a \circ b$.

Before going there, let me show the work of (100) on the following example, with the LF structure below.

$$(101) \quad a. \quad \text{Morris and Hilary dried each other.}$$



Each other requires dependent event types. If we did not have the dependent event type the argument of *each other*, v'_{mid} would be of the wrong type and the derivation would crash. Thus, I assume that the two thematic roles are lifted by the two operators to create the dependent event type. Later on, I will show that it is necessary that both the thematic role that applies to (the trace of) the reciprocal and the thematic role that applies to (the trace of) the reciprocal's antecedent must be lifted in this way. If they were not the resulting interpretation would be a contradiction. For the moment let us just ignore the issue and assume simply that both thematic roles are part of the dependent event type as it is in the LF structure here. The function to which *each other* applies is the following:

$$(102) \quad [[v'_{mid}]] = \lambda x \lambda y \lambda e' \lambda e. *DRY(e) \wedge {}^C * \Theta_1(x)(e') \wedge {}^C * \Theta_2(y)(e') \wedge e' \leq e$$

After *each other* applies to T'_{mid} we get (103a). After this function applies to the subject DP (and after the existential closure), we get the final interpretation, (103b).

$$(103) \quad \begin{aligned} \text{a.} \quad & \lambda x \lambda e. \langle x, x, e \rangle \in [\{ \langle a, b, e' \rangle : *DRY(e) \wedge {}^C * \Theta_1(a)(e') \wedge {}^C * \Theta_2(b)(e') \wedge e' \leq e \wedge \neg a \circ b \}] \\ \text{b.} \quad & (\exists e) (\langle M \oplus H, M \oplus H, e \rangle \in [\{ \langle a, b, e' \rangle : *DRY(e) \wedge {}^C * \Theta_1(a)(e') \wedge {}^C * \Theta_2(b)(e') \wedge e' \leq e \wedge \neg a \circ b \}]) \end{aligned}$$

We can rewrite (103b) into the more readable form which is below. The only change is that we put the condition $*DRY(e)$ outside of the set since this condition is irrelevant for deciding which elements are members of the set.

$$(104) \quad (\exists e) (*DRY(e) \wedge \langle M \oplus H, M \oplus H, e \rangle \in [\{ \langle a, b, e' \rangle : {}^C * \Theta_1(a)(e') \wedge {}^C * \Theta_2(b)(e') \wedge e' \leq e \wedge \neg a \circ b \}])$$

(104) is true if Morris and Hilary and an event e can be split into parts which are in the set:

$$(105) \quad \{\langle a, b, e' \rangle : C_{*\Theta_1}(a)(e') \wedge C_{*\Theta_2}(b)(e') \wedge e' \leq e \wedge \neg a \circ b\}$$

This is satisfied iff Morris dries Hilary and Hilary dries Morris. So the lexical semantics in (100) correctly accounts for the fact that *each other* is anaphoric on the level of pluralized relations but it requires distinctness on the atomic level.

Composition of reciprocity

The meaning of *each other* in (100) is in fact quite straightforward once we see what it is composed of. As I have said before, the reciprocal needs to express two things: anaphoricity and distinctness. The two conditions are meshed together in the meaning of the reciprocal but they can be teased apart. I separate the layers that together form the meaning of the reciprocal in Table 4.7. *REFL* expresses the anaphoric requirement. \overline{Id} expresses the distinctness condition. These two parts cannot combine because of type mismatch. Besides, even if they could the resulting function could not be satisfied in any situation since \overline{Id} adds non-overlap condition and *REFL* requires identity. Thus, to be able to combine the two, we have to make use of the $*$ operator that applies in dependent event types. This is identical to the $*_{\text{dep}}$ operator that we have used before, the only difference is that it pluralizes triples. These operators combine

Table 4.7: Parts of *each other*

$$\left| \begin{array}{ll} \text{REFL} & := \lambda Q_{\langle e, \langle e, \langle v, t \rangle \rangle \rangle} \lambda x \lambda e. Q x x e \\ \overline{Id} & := \lambda Q_{\langle e, \langle e, \langle v, \langle v, t \rangle \rangle \rangle} \lambda y \lambda z \lambda e' \lambda e. Q(y)(z)(e')(e) \wedge \neg y \circ z \\ *_{\text{dep}} & := \lambda Q \lambda y \lambda z \lambda e. \langle y, z, e \rangle \in [\{\langle a, b, e' \rangle : Q(a)(b)(e')(e)\}] \end{array} \right|$$

by means of function composition to give rise to the meaning of *each other*:

$$(106) \quad \begin{array}{l} \text{Composition of } \textit{each other} \\ [[\textit{each other}]] = \text{REFL} \circ *_{\text{dep}} \circ \overline{Id} = \\ \lambda Q_{\langle e, \langle e, \langle v, \langle v, t \rangle \rangle \rangle} \lambda x \lambda e. \langle x, x, e \rangle \in [\{\langle a, b, e' \rangle : Q(a)(b)(e')(e) \wedge \neg a \circ b\}] \end{array}$$

Of these parts of the meaning, *REFL* and $*_{\text{dep}}$ are needed independently, so it is only \overline{Id} that adds something new.

REFL is likely to be the meaning of reflexive-reciprocal constructions that are common cross-linguistically. As noted by Maslova (2007), reciprocals and reflexives are expressed by one and the same expression in many languages across the world (see also Chapter 1 and 3 and especially Volume II of Nedjalkov et al. 2007). Interestingly, as noted by Murray (2008), there are cases where one and the same expression seems

to do double duty as a reflexive and a reciprocal in one and the same sentence. She gives the following example from Cheyenne, where the verbal suffix *-achte* can serve both as a reflexive and a reciprocal.

- (107) Ka'èškóne-ho é-axeen-ahtse-o'o.
 child-PL.ANIM 3-scratch.ANIM-*ahte*-3PL.ANIM

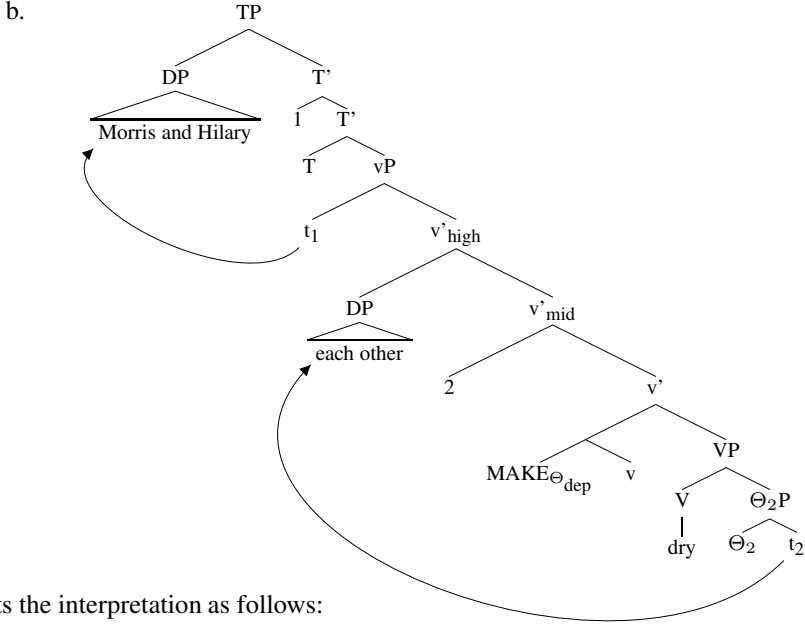
This sentence can mean that the children scratched themselves, or that they scratched each other. But it can also mean that some of them scratched themselves, and others scratched each other. As Murray points out, such mixed readings can also be found with reciprocal/reflexive expressions in French, Polish, and other languages. I can add Czech to her list. These mixed readings are particularly important, because they show that the reciprocal-reflexive relation, at least in those languages, should not be thought of as an ambiguity, or as homonymy. We can account for it if we specify the anaphoric expression as *REFL* since this leaves it open whether on the atomic level, one is related by the clausal relation to himself or to someone else.

$*_{\text{dep}}$ is a pluralizer that we have assumed elsewhere and is necessary to derive distribution in dependent event types, when combined with the two operators $\text{MAKE}_{\Theta_{\text{dep}}}$ and $\text{ADD}_{\Theta_{\text{dep}}}$.

One might still worry about the two null operators that create the dependent event type. As discussed in the previous section, we need them to derive distributive readings within cumulative readings: they are part of the derivation of dependent event types. The semantics of the reciprocal requires that the argument it applies to is a dependent event type (this is because of the presence of $*_{\text{dep}}$ in the meaning of the reciprocal). However, it does not specify how this dependent event type was created. For example, it does not specify which thematic roles should be lifted by the two silent operators. This might potentially lead to problems. If other possibilities led to non-existent readings we would have to block them somehow, which surely would complicate the picture.

Fortunately, it turns out that reciprocal sentences must include the dependent event type where both the thematic role that applies to (the trace of) the reciprocal and the thematic role that applies to (the trace of) the reciprocal's antecedent have been lifted, i.e., became the thematic roles that relate an argument to a subevent of *e*. Other options result in contradiction. Consider, for example, (108a) with the LF as follows. Notice that the head Θ_2 is not modified by $\text{MAKE}_{\Theta_{\text{dep}}}$.

- (108) a. Morris and Hilary dried each other.



This gets the interpretation as follows:

$$(109) \quad (\exists e)(\langle M \oplus H, M \oplus H, e \rangle \in [\{ \langle a, b, e' \rangle : *DRY(e) \wedge C_*\Theta_1(a)(e') \wedge C_*\Theta_2(b)(e) \wedge e' \leq e \wedge \neg a \circ b \}])$$

Compare (109) with (104). The only difference is that here Θ_2 is related to e while in (104) Θ_2 is related to e' . It is this difference which causes the formula here to end up being a contradiction. (109) says that Morris and Hilary can be split into parts such that each part dried the plurality Morris and Hilary, and furthermore, whoever did the drying must not overlap with the one that was dried. This cannot be satisfied since any part of Morris and Hilary does overlap with the plurality Morris and Hilary. Thus, we do not need to stipulate any mechanism that tells us where $MAKE_{\Theta_{dep}}$ or $ADD_{\Theta_{dep}}$ should appear in the structure – their position falls out independently.

To sum up, *each other* is built up from the pieces that we have seen and needed elsewhere and the only new part in its meaning (100) is \overline{Id} , which expresses the distinctness condition of the reciprocal. The distinctness condition is probably also not unique for reciprocals but plays a role also in the semantics of *together* (see Lasnik 1995, Kratzer 2003) and iterativity.

As the four volumes of Nedjalkov et al. (2007) show in detail, there are three main types of polysemy in reciprocal markers, apart from the polysemy of reciprocals and valency reducing markers, as are middles, passives etc. The first one is reflexive-reciprocal polysemy, which we discussed above and which could be derived if we assume that these markers express *REFL*. The other two are sociative-reciprocal polysemy and iterative/durative-reciprocal polysemy. The sociative is a term for the marker which expresses that the action was carried out jointly by a plurality of participants. An example of sociative-reciprocal polysemy can be found in Yakut:

- (110) a. Kiniler mælala-s-t-əlar.
 they soap-REC-PAST-3PL
 ‘They soaped each other.’
 b. Kiniler is tajaħ-ə mælala-s-t-əlar.
 they interior clothes-ACC soap-REC-PAST-3PL
 ‘They soaped the underwear together.’

The iterative/durative-reciprocal polysemy can be found in Chinese and Samoan, among other languages. It is likely that these polysemies arise due to the connection to distinctness condition. Sociatives require distinctness of an argument, which captures the fact that the argument must be expressed by a plurality in the presence of sociatives. Iteratives express distinctness of events. I will leave it open for future research how exactly the meaning composition of reciprocity should connect to these other domains where distinctness plays a role.

Cumulative and collective readings in reciprocal sentences

I would now like to move to more complicated examples, which involve an extra DP in a reciprocal clause. First, we have noticed that arguments can have a cumulative interpretation in reciprocal sentences. The following sentence, repeated from above, shows this point:

- (111) The children gave each other Christmas presents.
 (each child giving a different present)

The sentence is true if there are two children and each child gave the other child one present. This example is parallel to the cases from Section 4.4.1 where two arguments were in a cumulative relation and the third argument was dependent on one of them – the only difference is that in the cases in Section 4.4.1 the dependent argument was an indefinite and here it is *each other*. In case of reciprocal sentences, we derive the reading from the following LF:

- (112) the children EO_1 $ADD_{\Theta_{dep}}(\Theta_1)$ [give [t_1 $MAKE_{\Theta_{dep}}(\Theta_2)$ [Θ_3 Christmas present]]]

Thus, the (trace of the) reciprocal and the (trace of the) antecedent are related by thematic roles not to an event e that makes the whole proposition true, but to its subevent. Finally, the third argument which is scopally independent on the two is related to the event e . This gives us the following interpretation:

- (113) $(\exists e)(\ast GIVE(e) \wedge {}^C\ast\Theta_3(PRESENTS)(e) \wedge$
 $\wedge \langle THE\ CHILDREN, THE\ CHILDREN, e \rangle \in$
 $\in [\{ \langle a, b, e' \rangle : {}^C\ast\Theta_1(a)(e') \wedge {}^C\ast\Theta_2(b)(e') \wedge \neg a \circ b \}])$

Thus, there is an event of giving in which some presents were given and each child was a giver in some subevent e' and each child was a recipient in some subevent e' . This would be true, if we had two events, e_1 and e_2 . In e_1 the first child gives a present

to the second child. In e_2 the second child gives a present to the first child. This is what we want.

There is another issue which suggests that distributivity is limited in reciprocal sentences. This concerns collective predicates which can appear in reciprocal sentences. Consider (114a) and (114b). The first example is from Dimitriadis (2000) and the second example was found via Google search. As confirmed by three native speakers, both can be interpreted with the collective action of the subject. (114a) gets the interpretation ‘Bill and Peter carried the piano together across Bill’s lawn(s) and across Peter’s lawn(s)’. (114b) gets the interpretation ‘the sailors, as a group worked first on one sailor’s ship then on the next sailor’s ship etc.’ The last two examples are attested cases from the Corpus of Contemporary American English which show the same point.

- (114) a. Bill and Peter, together, carried the piano across each other’s lawns.
 b. The sailors have worked together on each other’s ships.
 c. Cooper and friends gather at each other’s homes to perform tunes and ballads. (COCA)
 d. ... instead of going out on the town, my girlfriends and I would meet at each other’s apartments to cook and catch up (COCA)

These examples lead Dimitriadis (2000) to complicate the distinctness condition of *each other*. Instead of assuming that two arguments must not overlap he assumes that one must not be part of the other or identical to the other and the part-of requirement is uni-directional. Dimitriadis’ condition is satisfied in (114a) since there the argument of Θ_1 (the plurality Bill and Peter) is not part of Bill and it is also not part of Peter, and similarly for the other examples.

Unfortunately, complicating the distinctness condition overgenerates. Consider the three examples in (115). As confirmed by the three native speakers who all accepted the collective predicates in (114), the collective predicates in (115) are ungrammatical.

- (115) a. * The boys are each other’s dream band.
 b. * The boys in our class outnumber each other’s families.
 c. # The students elected each other.

The first example cannot mean that the boys, as a group, are a dream band of each of the boys. The second sentence cannot mean that there are more boys in our class than there are family members in each boy’s family. According to three native speakers, the first two sentences cannot have any interpretation at all and are simply ungrammatical. The last one can be interpreted but in that case one shifts from the collective interpretation and understands *elect* as *voting for*, i.e., the sentence could be true if there was one election as long as most of the students voted for other students among them. Thus, in none of these sentences the predicate can be interpreted collectively.

The difference between (114) and (115) follows the difference between the two types of collective predicates studied by Winter (2001a) and Brisson (1998), among others. The first type, (114), consists of collective activities and accomplishments.

The second type, (115), consists of collective states and achievements. As we have seen in Chapter 3, only the first type is acceptable with *all NP* as the subject. We have analyzed this fact by assuming that *all NP* requires the presence of * which, in case of collective activities and accomplishments, scopes over a thematic role. However, the argument of Θ_1 does not need to bring about the event for which the predicate holds. It suffices that one acts in a way that supports the collective action. Thus, for instance, the argument of Θ_1 of *gather* does not need to do gathering himself, it suffices that he comes to the right place and stays there, along with other people.

Consider now (114c). This gets the following interpretation (where $C \oplus FR.$ stands for Cooper and his friends):

$$(116) \quad (\exists e)(\ast GATHER(e) \wedge \langle C \oplus FR., C \oplus FR., e \rangle \in \\ \in [\{ \langle a, b, e' \rangle : C_{\ast \Theta_1}(a)(e') \wedge C_{\ast AT}(\text{home of } b)(e') \wedge \neg a \circ b \}])$$

(116) is true if there was an event of gathering and in its subevents Cooper was the argument of Θ_1 in the house of one his friends, and similarly for other friends and the houses of their friends. This would be false if the argument of Θ_1 would have to be the person that brings about ‘gathering’. However, if we weaken the condition of what the argument of Θ_1 must satisfy, following Brisson’s work, and require only that the argument gets to some place and waits there (in this particular case), we correctly derive that the sentence is acceptable. There is one worrisome issue, though. In this viewpoint, the event qualifies as *gathering* if each person goes to a different place and waits there. Thus, for instance, *Morris, Philip and Perssy gathered* would be true if Morris was part of one gathering, Philip was part of another gathering and Perssy was part of yet another gathering. I am not sure that the sentence is true in this case. Thus, even though I believe it is correct to connect collectivity in reciprocal sentences to the distinction between two types of predicates and Brisson’s work, at the current state we have would only capture the acceptability of collective predicates in reciprocal sentences at cost of allowing rather weak interpretations of what agents of these collective predicates must satisfy. Thus, unlike in case of cumulative readings, I do not find the analysis of collective predicates in reciprocal sentences fully satisfactory. I have to leave further improvements to future research.

Deriving IR and SR

Recall that SR and IR in the title stand for Strong Reciprocity and Intermediate Reciprocity and are labels for two readings of reciprocal sentences which differ in strength of the clausal relation. Their examples are given in (117). The example of SR, (35a), is interpreted as ‘each of Morris, Hilary and Philip know the other two (i.e., everyone knows everyone else)’. On the other hand, (35b) is accepted not only in a scenario in which each telephone pole is 500 feet from all the other telephone poles but also in weaker ones. For example, if the telephone poles are in a line and only neighbouring poles are at 500 feet distance, the sentence is judged as true.

- (117) a. Morris, Hilary and Philip know each other. (example of SR)
 b. The telephone poles are 500 feet from each other. (example of IR)

Both readings follow from the semantics of plurality. IR is derived in the same way as cumulative readings. Unlike in Beck (2001), nothing extra needs to be assumed to combine the semantics of reciprocals with the cumulative reading. The relation to which *each other* applies in (117b) is (118). Notice that instead of Θ_2 the second argument is related to the event by a preposition modified by a measure phrase, *500 feet from*. I will ignore the composition of this prepositional phrase.

The function to which *each other* applies is the following:

$$(118) \quad \lambda x \lambda y \lambda e' \lambda e. * \text{STAND}(e) \wedge {}^C * \Theta_1(x)(e') \wedge {}^C * 500 \text{ FT FROM}(y)(e') \wedge e' \leq e$$

After *each other* applies to this relation and it further applies to the subject DP (and after the existential closure), the resulting interpretation is (119).

$$(119) \quad (\exists e) (* \text{STAND}(e) \wedge \langle \text{THE POLES, THE POLES}, e \rangle \in [\{ \langle a, b, e' \rangle : {}^C * \Theta_1(a)(e') \wedge {}^C * 500 \text{ FT FROM}(b)(e') \wedge e' \leq e \wedge \neg a \circ b \}])$$

This is true if the plurality of the telephone poles can be split into parts such that each part is 500 ft from some other part. This is true if we consider individual telephone poles where each one of them is 500 ft from some other telephone pole, namely from the neighbouring pole(s). Thus, the sentence is true if telephone poles are in a line and only neighbouring ones are at 500 feet distance. We derive that the most natural reading of (117b) is the one where we measure the distance between neighbouring poles due to our assumptions on Covers and events discussed in Section 3.2.5. In particular we want split of subevents to follow conditions that govern other domains of splitting plurals and vision situation in cognitive psychology. One of these principles that is relevant here is proximity which would lead us to consider poles next to each other rather than other pairs of poles. Notice that unlike Beck (2001) we derive the correct meaning without postulating QR of a variable. We also do not require Cover to consist of anything else than atomic individuals.

In general, reciprocal sentences are expected to be true if the antecedent of the reciprocal can be split into parts where each part relates to some non-identical part by the clausal relation, and each part is related to some non-identical part by the clausal relation. These are requirements identical to cumulative readings, thus we equate IR to cumulative readings which is what we want.

Since we cannot get weaker than the cumulative reading, we do not overgenerate in examples like (120), also repeated from above. The sentence is awkward if there is an audience in a lecture hall listening and looking at the speaker and the speaker is looking back at one person in the audience.

$$(120) \quad \# \text{ The people in the lecture hall are staring at each other.}$$

We correctly derive the oddness of the example, since to make the sentence true it must hold that everyone is looking at someone and everyone is being looked at by someone, which is not satisfied in (120). Have we changed the scenario and assumed, for example, that there were as many speakers up front as there were people in the audience, and everyone in the audience stared at a different speaker, and everyone among the speakers stared at a different person in the audience, the sentence becomes acceptable, exactly as predicted.

SR is derived in the same way as IR. Consider (121) repeated from above.

- (121) Morris, Philip and Hilary know each other. (example of SR)

As the interpretation, we get (122). This is identical to the way IR was derived.

- (122) $(\exists e)(\ast\text{KNOW}(e) \wedge \langle \text{M} \oplus \text{P} \oplus \text{H}, \text{M} \oplus \text{P} \oplus \text{H}, e \rangle \in [\{ \langle a, b, e' \rangle : {}^C\ast\Theta_1(a)(e') \wedge {}^C\ast\Theta_2(b)(e') \wedge e' \leq e \wedge \neg a \circ b \}])$

In Section 3.2.5 we have seen that in sentences with plural arguments the branching reading is preferred over a cumulative reading when possible and acceptable given our world knowledge. For example, in (123) the first sentence is preferably interpreted as each of the women saw each of the men (the branching reading) but the second one is acceptable in a weaker reading, i.e., each woman gave birth to some boy and each boy was given birth by some of the women (the cumulative reading).

- (123) a. Mary and Sue saw John, Bill and George.
b. Mary and Sue gave birth to John, Bill and George.

Whatever mechanism accounts for the difference in (123) and one's preference for branching readings is likely also to explain the difference between IR and SR and one's preference for SR when possible given our world knowledge. I suggested in Section 3.2.5 that we can derive the branching reading for (123a) if we assume that every subevent is exemplified. In that case, an event that exemplifies the proposition is the event where Mary saw John, Bill and George and so did Sue. I assumed that this event is chosen over the one which exemplifies a cumulative reading due to the Strongest Meaning Hypothesis. SR is parallel to branching readings (ignoring identities) so it should be chosen over IR when possible. In this way, the approach here is similar to the Strongest Meaning Hypothesis of Dalrymple et al. (1998) and the approach of Sabato and Winter (2005b). It also shares the weaknesses of these approaches. For example, it is not really clear what knowledge is relevant in deciding the strength of reading. For first attempt to clarify this issue that I am aware of, see Kerem et al. (2009), who argue that the semantic strength of reciprocal sentences does not follow from what is possible and what is not but rather, what relations are *typical*.

Finally, since we build the meaning of *each other* using the \ast operator from the theory of plurality we can also account for partitioned readings where covers play a role. I will only consider one example (for more examples, see Chapter six of Schwarzschild 1996 and Beck 2001). Consider the following sentence:

- (124) The prisoners on the two sides of the room can see each other.

The relevant reading is that the two groups of the prisoners can see each other. If there was an opaque barrier in the middle of the room the sentence would become false. We can get the reading if we consider covers in thematic roles. Suppose they consist of two groups of the prisoners. In that case we have to split a and b in the formula below, which is the interpretation of the sentence, in such a way that a is one group and b is the other group.

$$(125) \quad (\exists e)(\ast\text{SEE}(e) \wedge \langle \text{THE PRISONER, THE PRISONERS}, e \rangle \in [\{\langle a, b, e' \rangle : C_{\ast\Theta_1}(a)(e') \wedge C_{\ast\Theta_2}(b)(e') \wedge e' \leq e \wedge \neg a \circ b\}])$$

Notice that in the approach here, we have the lowest bound on the reciprocal readings. The least strong reading is IR, which is a cumulative reading with the extra condition added by the reciprocal that identical entities are not related to each other. In Section 3.2.5 I have argued that having the lowest bound on possible readings is a good thing in the domain of theory of plurality since we do not see that one can get a weaker reading than the cumulative one. The situation is somewhat different in reciprocal sentences. Here, we do have a weaker reading than IR. This is IAR, exemplified by (126).

(126) The chairs are on top of each other.

This reading, unlike the previous ones, cannot be derived. This makes the approach here different from Dalrymple et al. (1998) and Sabato and Winter (2005b) who can capture IAR. However, we have also seen before that the approaches which can derive IAR face the problem of overgeneration. All of the following sentences fail to give rise to IAR, even though IAR would make them true and is the strongest reading that is possible given our world knowledge. For Dalrymple et al. (1998) and Sabato and Winter (2005b) it is unclear how that could be, given the fact that (i) IAR is among the available readings of reciprocal sentences (ii) any reading can be used if it is the strongest one and compatible with our world knowledge.

- (127) a. # The plates are stacked **underneath** each other.
 b. # The boys **preceded** each other into the room.
 c. # The men are each others **fathers**.
 d. # John and Mary are **taller than** each other.

As we have discussed before, IAR is restricted only to spatial and temporal predicates. Furthermore, among these predicates it is only possible if they express figure-ground relation in such a way that the figure precedes the ground (*follow, be behind*) or the figure is vertically above the ground (*be on top of*). Beck (2001) suggests that such readings should not be dealt with by a general theory of reciprocity, but rather thought of as a part of the semantic theory of the small set of spatial relations that actually give rise to it. I already said before that I find this a reasonable option and. Just to operate with something more specific, we can proceed as follows. We could add an idiomatic meaning of *each other* in which the direction of the relation is irrelevant (as in Dalrymple et al. 1998) but the application of this meaning is restricted to the small set of spatial and temporal predicates.

$$(128) \quad [[\text{each other}_{\text{spatial/temporal}}]] = \lambda Q_{\langle e, \langle e, \langle v, \langle v, t \rangle \rangle \rangle \rangle} \lambda x \lambda e. \langle x, x, e \rangle \in [\{\langle a, b, e' \rangle : Q(a)(b)(e')(e) \vee Q(b)(a)(e')(e) \wedge \neg a \circ b\}]$$

in case the clausal predicate expresses spatial or temporal relations in which the figure precedes the ground (*follow, be behind*) or is vertically above the ground (*be on top of*)

This meaning of *each other* differs from the previous in allowing both $Q(a)(b)(e')(e)$ and $Q(b)(a)(e')(e)$, i.e., it leaves open which part is the figure and which part is the ground. This is just a stipulative way to account for the reading and does not follow from anything. In this respect, my approach is no better or worse than any previous approach. Every account so far has to either stipulate something akin to (128) to capture IAR, or has to stipulate a different rule which excludes IAR in all the cases where it is impossible (like the examples in (127)).

Negated reciprocal sentences

Since *each other* is built on the theory of plurality we correctly deal with negated reciprocal sentences. Recall that (129) repeated from above was problematic for the R-approaches because they derived a very weak reading, namely, that the sentence should be true if Morris does not know Hilary and otherwise everyone knows everyone else. But (129) requires that nobody knows anybody.

(129) Morris, Hilary and Philip do not know each other.

The meaning of (129) is correctly derived in our analysis of *each other* thanks to the presupposition on the * operator.

Before showing that, I want to say a few words on how we can deal with negation in event semantics. Negation adds complications to event semantics because negated expressions are not persistent: they can be true in some situation and turn out false if a larger situation is considered. We could assume that negation always scope above the event existential closure but this would be problematic if we considered event predicate modifiers (like the temporal modifier *for two hours*) which can scope over negation. For this reason I am going to assume the analysis of negation as in Krifka (1989).

We consider maximal events. Maximal events represent the join of all events in some specific reference point. This could be time, as assumed in Krifka (1989) but space is probably as relevant here. Consider as an example *nobody applauded*. This is true at some certain situation (in a lecture hall at 4pm) but might be false if we enlarge the situation beyond the relevant reference point, for example, to include the whole building or to include other time points, so we are considering only the maximal event with respect to some specified time/space coordinate. In the definition, which otherwise follows Krifka (1989), I will leave out how we get to this reference point.

(130) e is maximal ($MAX(e)$) iff
 $e = \sigma(\lambda e.e \text{ takes place at some contextually specified reference point})$

Negation is an event modifier which states that parts of the maximal event are not true:

(131) $[[\text{not}]] = \lambda P \lambda e. MAX(e) \wedge \neg(\exists e')(Pe' \wedge e' \leq e)$

Using this semantics of *not* we get the meaning of (129) as in (132). (132a) and (132b) are equivalent¹⁷ but the second formula is probably more transparent.

¹⁷Because $\neg(\exists e'')(e'' \leq e \wedge Pe'') \Leftrightarrow (\forall e'')\neg(e'' \leq e \wedge Pe'') \Leftrightarrow (\forall e'')(\neg e'' \leq e \vee \neg Pe'')$.

- (132) a. $(\exists e)(MAX(e) \wedge \neg(\exists e'')(e'' \leq e \wedge *KNOW(e'') \wedge \langle M \oplus P \oplus H, M \oplus P \oplus H, e'' \rangle \in [\{\langle a, b, e' \rangle : a \text{ knows } b \text{ at } e' \text{ and } \neg a \circ b\}])$
 b. $(\exists e)(MAX(e) \wedge (\forall e'')(\neg e'' \leq e \vee \neg *KNOW(e'') \vee \langle M \oplus P \oplus H, M \oplus P \oplus H, e'' \rangle \notin [\{\langle a, b, e' \rangle : a \text{ knows } b \text{ at } e' \text{ and } \neg a \circ b\}])$

(132b) holds for the maximal event whose subevents are either not events of knowing or they are not events in which everyone among Morris, Philip and Hilary knows someone among them and everyone among Morris, Philip and Hilary is known by someone among them. So, this is satisfied if, for example Morris knows Philip and Philip knows Hilary but Hilary does not know anyone. These truth conditions are too weak. However, they are strengthened if we consider the presupposition of $*_{dep}$. This is the Presupposition of Indivisibility. It requires that the function when applies to plural argument(s) is true or false of them as a whole. More concretely, we have the following presupposition:

- (133) $*_{dep}(Q)(y)(z)(e')(e)$
Presupposition: $\langle y, z, e \rangle \in [\{\langle a, b, e' \rangle : Q(a)(b)(e')(e)\}] \vee$
 $\vee y \in [\{a : z \in [\{b : e \in [\{e' : Q(a)(b)(e')(e)\}]\}]]$
 \Leftrightarrow
Presupposition: $\langle y, z, e \rangle \in [\{\langle a, b, e' \rangle : Q(a)(b)(e')(e)\}] \vee$
 $\vee (\forall a \leq y)(\forall b \leq z)(\forall e' \leq e)(Q(a)(b)(e')(e))$

For the example above, this would give us the following presupposition:

- (134) **Presupposition:** $\langle M \oplus P \oplus H, M \oplus P \oplus H, e' \rangle \in [\{\langle a, b, e' \rangle : a \text{ knows } b \text{ at } e' \text{ and } \neg a \circ b\}] \vee$
 $\vee (\forall a \leq y)(\forall b \leq z)(\forall e' \leq e)(a \text{ does not know } b \text{ at } e' \vee \neg a \circ b)$

The first disjunct contradicts the non-presuppositional content so we can ignore it. The second disjunct states that in the events e' either nobody among Morris, Hilary and Philip knows anybody among these three or the individuals overlap. Consider the situation in which everyone knows himself/herself and apart from that, Philip knows Hilary. This would lead to the presupposition failure because in one of the subevents, someone knows someone and the two are non-overlapping individuals (Philip and Hilary). This is what we want since negated reciprocal sentences should be true if nobody is related to anybody else (but it is irrelevant if anybody is related to himself/herself).

To conclude, I offered a novel analysis of *each other* that is built around the R-approaches to reciprocals. My account could deal with cumulative readings in reciprocal sentences, as well as it closely connects the range of readings in reciprocal sentences to mechanisms needed independently of reciprocal sentences. This, for instance, allows a simple account of negative reciprocal sentences. These issues, as we are going to see, are problematic for other R-approaches to *each other*.

4.4.3 Previous R-approaches to reciprocals

In this section I compare my account to previous accounts of reciprocals set up within the R-approach. I show that all these accounts fail to deal with negated reciprocal sentences and cumulative interpretations in reciprocal sentences.

Dalrymple et al. (1998)

I slightly modify the system of Dalrymple et al. (1998) to make it compatible with mereological approaches to pluralities. In the discussion of their system, I ignore events for the sake of simplicity. Dalrymple et al. (1998) offer the following basic interpretation of *each other*.

- (135) *Each other*, Dalrymple et al. (1998)
 $[[\text{each other}]] = \lambda R \lambda x. (R \cup Id \upharpoonright AT(x)) = AT(x) \times AT(x)$

$R \cup Id$ is the union of the relation with the identity relation. $((R \cup Id) \upharpoonright AT(x))$ restricts the relation unified with the identity relation to atomic parts of x and requires that $|AT(x)| \geq 2$:

- (136) a. $R \cup Id = \{\langle x, y \rangle : Rxy\} \cup \{\langle x, x \rangle\}$
 b. $R \upharpoonright AT(x) = R \cap AT(x) \times AT(x)$

Thus, (135) is true for the relations and plural entities in which every atomic entity of the plurality is related to every other atomic entity by R . Notice that this semantics of *each other* is within the R-approach (multiplication of resources in lexicon, i.e., in the meaning of *each other* because every variable is bound).

The range of readings in reciprocal sentences are derived as logical operations on relations. Dalrymple et al. (1998) consider three operations: the transitive closure of relations, ignoring the direction in which the relation holds (i.e., adding inverse pairs), and selecting the domain of the relation. Combining the operations can give us up to six readings. I focus first on the three readings that we have discussed before: SR, IR and IAR. The examples are repeated here.

- (137) a. Morris, Hilary and Philip know each other. (example of SR)
 b. The telephone poles are 500 feet from each other. (example of IR)
 c. The chairs are stacked on top of each other. (example of IAR)

(137a) follows if we make use of the semantics of *each other* in (135), that is, with no extra operations on relations. (137b) is derived by enabling the transitive closure of the relation. The transitive closure is given in (138).

- (138) $R^- = \{\langle x, y \rangle : (\exists z_1 \dots z_n)(R(x)(z_1) \dots R(z_n)(y))\}$

If we let the transitive closure apply on the relation that is the argument of the reciprocal, we get (139). This makes sure that the atomic parts of the plural argument, are related directly *or* indirectly, where the indirect relation means that the relation is arrived at through the transitive closure. Using (139) gives us the correct meaning for examples like (35b). Every part of the plural argument ‘the telephone poles’ is related to every other part because both direct and indirect relations count.

- (139) *Each other* used to derive IR
 $\lambda R \lambda x. (R \cup Id \upharpoonright AT(x))^- = AT(x) \times AT(x)$

The second operation I consider here is adding inverse pairs, i.e., ignoring the direction of the relation. This is achieved by the operation \vee :

$$(140) \quad R^{\vee} = \{\langle x, y \rangle : Rxy\} \cup \{\langle x, y \rangle : Ryx\}$$

If we combine the two operations (see (141)) we get the reciprocal which is true in (137c), i.e., can derive IAR. (141) works for (35c) because every chair is related to every other chair by the relation ‘being stacked on top’ if the relation can be indirect and the order of the two arguments is irrelevant.

$$(141) \quad \text{Each other used to derive IAR} \\ \lambda R \lambda x. (R^{\vee} \cup Id \upharpoonright AT(x))^{-} = AT(x) \times AT(x)$$

Finally, the third operation Dalrymple et al. (1998) consider is the one that gives us the domain of the relation. Combined with the previous operations, we have two possibilities: either we get the domain of the actual relation, or we can get the domain of the relation to which \vee applied (which gives us all the elements that are either in the domain or the range of the actual relation). Thus, in total Dalrymple et al. (1998) end up with five readings for reciprocal sentences.¹⁸ The three we have discussed above. Furthermore, they have One-Way Weak Reciprocity (OWR), in which case the domain of the relation equals the atoms in the plural antecedent of the reciprocal, and Inclusive Alternative Ordering (IAO), in which case the domain of the relation modified by \vee equals the atoms in the plural antecedent of the reciprocal.

However, in reality, we do not perceive that any reading is possible for every reciprocal sentence, i.e., this multiple ambiguity is missing in our intuitions. For example (137a) is true under SR (see above) but it is considered false under IAR, as can be seen from the fact that the following sounds as a contradiction: ‘Morris, Hilary and Philip know each other but Morris does not know Hilary.’ Dalrymple et al. (1998) develop a system in which the potential ambiguity of reciprocals disappears in actual sentences due to the Strongest Meaning Hypothesis (SMH). The SMH states that among the interpretations that reciprocals have, we always choose the strongest one which is compatible with our non-linguistic knowledge imposed on the relation. The readings of reciprocal sentences are ordered by entailment, shown in Figure 4.4, where the arrow represents the order of entailment. Given the SMH, if we have a relation where no restrictions seem to apply we choose SR (35a). In (137b) the strongest meaning is IR given our world-knowledge. SR is not possible because we live in three-dimensional space and we usually measure distances only on the horizontal plane in which only two or three poles could be separated by the equal distance. Finally, (137c) makes use of IAR because this weakest interpretation is the only one possible. IR would require that every chair is on top of another chair which would defy gravity.

There are good arguments that of these five readings, the two weakest ones, OWR and IAO, should be derived by mechanisms independent on reciprocity. Beck (2001) argues against One-way Weak Reciprocity as a reading to be derived by the interpretation of reciprocals and Sabato and Winter (2005a) argue that neither One-way Weak Reciprocity nor Inclusive Alternative Ordering should be derived by operations on reciprocals. They all show that the readings should rather follow from mechanisms used in the theory of plurality. I am not going to discuss these arguments here. Rather, I

¹⁸In fact, there is another reading derived in their system, Strong Alternative Reciprocity, SAR, but since it is not attested I ignore it here.

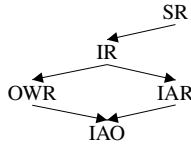


Figure 4.4: Entailment relations among DKKMP readings

want to point out that Dalrymple et al. (1998) might be forced to admit the same even if Beck's and Sabato et al.'s arguments do not hold. The reason is that their attested examples of OWR and IAO can also be true in IR and IAR, respectively and because of the Strongest Meaning Hypothesis, this means that they *have* to be true in IR and IAR (and false in OWR and IAO). Consider their example for OWR:

(142) The pirates were staring at each other.

Given that one can stare only at one person, we would derive IR for this reading: everyone is staring at someone and everyone is being stared at by someone. But Dalrymple et al. (1998) claim that the sentence can also be true in OWR: everyone is staring at someone. This is a problem for the SMH, which should in fact force the stronger reading. The only escape hatch I see is to say that OWR and IAO are derived by the mechanisms independent of reciprocity so the Strongest Meaning Hypothesis does not operate on them.

This leaves us with three readings for reciprocals. SR and IR are parallel to branching readings and cumulative readings in the sentences with plural DPs. It seems to me that we should rather let the theory of plurality capture SR and IR than stipulate mechanisms just to account for the two readings as Dalrymple et al. (1998) do. Finally, IAR cannot be accounted for in the theory of plurality (since there the weakest reading is the cumulative one). I assumed that in this case we deal with idiomatic reading of the reciprocal. I do not consider this the final solution. Dalrymple et al.'s account does not fully work either since in their case we expect that IAR should appear whenever it is the strongest reading possible. As we have seen above this is not the case (see the examples (127)).

Another problem for Dalrymple et al. comes from negative reciprocal sentences. (143) states that nobody among Morris, Hilary and Philip knows anybody else among these three.

(143) Morris, Hilary and Philip do not know each other.

Assume that reciprocals and negation can freely scope with respect to each other. If *each other* scopes over the negation Dalrymple et al. (1998) derive the correct truth conditions. This is given in (144). (144) is false if anyone among Morris, Hilary and Philip knows anyone else.

$$(144) \quad ((\lambda x \lambda y. \neg \text{KNOW}(x)(y)) \cup Id \upharpoonright AT(M \oplus P \oplus H)) = AT(M + P \oplus H) \times AT(M \oplus P \oplus H)$$

The problem comes when the negation scopes over the reciprocal. There are two possibilities what can happen in that case. First, the Strongest Meaning Hypothesis selects the strongest interpretation before the negation applies to the proposition. This would mean that the strongest reciprocal reading, SR, has been selected. If we negate that we get a proposition which is true as long as somebody among Morris, Hilary and Philip does not know somebody else. This is too weak. The other option does not work much better. In this case, the Strongest Meaning Hypothesis selects the strongest reading of the negated reciprocal sentence. Since the negation reverses entailment, the strongest reading is the one which uses IAO of the reciprocal. But this is not strong enough. IAO requires that every entity in the extension of the plural antecedent of *each other* is in the domain or range of the clausal relation. This means that IAO is false if among three individuals only one person is related by *R* to another one so this should make a negated reciprocal sentence true. But this is also too strong, witness the oddness of # Morris, Hilary and Philip do not know each other but Morris knows Philip.

One possible escape hatch for Dalrymple et al. (1998) is to show that *each other* must always scope over negation. (145) militates against this conclusion. (145a) seems false or odd if I think that one of Morris, Hilary and Philip met someone else, for example, if Morris met Philip. (145b) does not say that I do not care if some kissing is going to happen, as long as not every kid is kissed (or is kissing). Rather, my wish is that no kid is going to kiss any other kid, period. But in these cases negation is in the matrix clause while *each other* takes scope in the embedded clause.

- (145) a. I do not think that Morris, Hilary and Philip ever met each other.
 b. I do not want the kids to be kissing each other.

One option remaining to correctly derive the interpretation of (145) is to assume that negation takes scope in the embedded clause, through some covert operation, like neg-lowering. This has been proposed for attitude verbs like *think* which are neg-raising verbs. However, Horn (1989), Gajewski (2005) and others present very strong arguments against neg-lowering. Finally, even if one could show that the reciprocal must scope over negation we would still lack an explanation for the fact that strengthening of (143) from the negated SR to the meaning 'nobody knows anybody' is a matter of presupposition.

Another possibility how one could deal with negated reciprocal sentences in Dalrymple et al. (1998) is to make use of the presupposition of indivisibility. First, one could point out that we do assume that the relation of knowing satisfies the presupposition of indivisibility. Could this help? Not really. The reason is that *each other* checks whether *atoms* are in the extension of the relation and therefore, the presupposition of indivisibility is vacuously satisfied.

However, we could go one step further and require that the predicate created by *each other* must be true or false of the argument as a whole. This would mesh well with the general viewpoint of Löbner (2000), I believe, namely that every predication satisfies the presupposition of indivisibility. Thus, *know each other* would either have to be true for every part of Morris, Hilary and Philip or it would have to be false for all these parts. If we consider every part, the presupposition could never be satisfied

because it would clash with another presupposition of *each other*, namely, that only non-atomic entities are in the extension of the predicate. Assume that we restrict the presupposition of indivisibility only to those parts that have two members. In that case the presupposition would be satisfied if Morris knows Hilary and nobody else knows anybody else (because no pairs would make the predicate true, not even Morris and Hilary). But that is still not enough.

This is what would work for this particular example: we would keep the presupposition of indivisibility with all the assumptions but now we would not require that *know each other* must be true for every pair. Rather, we would require that *know EACH OTHER* must be true for every pair, where *EACH OTHER* is the reciprocal meaning that has been selected by the Strongest Meaning Hypothesis. Assuming that the Strongest Meaning Hypothesis selects the reciprocal reading after the negation applied to proposition, we would select IAO in (143). Now, in case anyone knows anybody else, *know EACH OTHER* would be true for this particular pair, which would lead to presupposition failure in negative sentences. Thus, we would correctly derive that (143) says that nobody knows anybody else.

Unfortunately, using the presupposition of indivisibility would cause havoc in any readings weaker than SR. Consider the example of IR from above:

(146) The telephone poles are 500 feet from each other.

Because of the presupposition it would now be required that the predicate *be 500 feet from EACH OTHER* is true for every pair of telephone poles. But this is not true: non-neighbouring telephone poles are not in the extension of this predicate. Even though we consider the transitive closure of the relation here, it does not help because the transitive closure of the relation is restricted by the plurality to which the reciprocal applies. Since we are checking if the predicate holds for a pair of non-neighbouring telephone poles, we have to restrict the transitive closure of the relation by these two particular telephone poles. Thus, non-neighbouring telephone poles are not in the extension of *be 500 feet from EACH OTHER*. To put it more generally, the presupposition would then lead to the unfortunate result that every reciprocal sentence would have to end up as SR after all.

Thus, the only option which would work is to add a presupposition to the meaning of reciprocal. For example, we could change the meaning of *each other* as follows (shown for SR, similarly for weaker readings):

(147) *Each other*, Dalrymple et al. (1998), presupposition added

$$\frac{[[\text{each other}]] = \lambda R \lambda x. (R \cup Id \upharpoonright AT(x)) = AT(x) \times AT(x) \vee \vee (R \cup Id \upharpoonright AT(x)) \cap (\overline{AT(x) \times AT(x)}) = \emptyset.}{(R \cup Id \upharpoonright AT(x)) = AT(x) \times AT(x)}$$

Even though this does work, it begs the question why this particular presupposition should be part of the meaning of *each other*. On the other hand, this follows automatically if we pluralize the relation using the * operator.

Finally, consider the following sentence, which is true if there are two children and each of them gave one present to the other child.

- (148) The children gave each other Christmas presents.
(each child giving a different present)

Dalrymple et al. (1998) cannot derive this reading. If we allowed the plural argument *Christmas presents* to scope over the reciprocal or below, this would not help since we would only derive the reading that each boy gave the other boy Christmas presents. We could try other derivations. For example, we could try the following. We could assume cumulation on relations, as in Schwarzschild (1996), Beck and Sauerland (2000) and others (see Section 4.4.1). Then, we could say that *the children* and *Christmas presents* apply to the cumulative relation $\lambda x \lambda y. x$ gave EO y :

- (149) $\langle \text{THE CHILDREN, PRESENTS} \rangle \in [\{\langle x, y \rangle : x \text{ gave EO } y\}]$

But this also does not work. (149) is false if each child gave the other child one present. The reason is that we cannot consider proper parts of the plurality THE CHILDREN since the only parts are individual kids and atomic individuals are not in the extension of *each other*. So we can only consider the two children, and we are looking for y for which it would hold that the two children gave y to each other. Filling in the semantics of *each other* we get that we are looking for y that satisfies the following:

- (150) $(\{\langle x, z \rangle : x \text{ gave } z \ y\} \cup Id) \upharpoonright AT(\text{CHILDREN}) = AT(\text{CHILDREN}) \times AT(\text{CHILDREN})$

That is, we are looking for y for which it would hold that each kid gave y to every other kid. But there is no such y . Individual books do not satisfy this because it is not true that the children gave each other one and the same book. The two books that they gave each other do not satisfy this because it is not true that each of the kids gave the other these two books. In fact, I see no way how Dalrymple et al. (1998) could derive the crucial reading.

To conclude while Dalrymple et al. (1998) can account quite well for what Langendoen (1977) calls elementary reciprocal sentences (sentences of the form DP-Verb-Recip) they fail once we move beyond. Their semantics for negated reciprocal sentences is too weak. Finally, they cannot derive the interpretation in which cumulative readings appear inside reciprocal sentences, (148).

Sabato and Winter (2005b)

In recent work, Winter and Sabato suggest that we can derive the variety of readings without adhering to the ambiguity that Dalrymple et al. (1998) need to assume (see Winter 2001a, Sabato and Winter 2005a, Sabato and Winter 2005b). Winter (2001a) postulates a quantificational restriction Θ_V , which, for any expression V of type $\langle e, et \rangle$, provides all possible extensions of V compatible with our world knowledge. Now, it suffices to say that the actual extension of V restricted to the reciprocal antecedent is one of the the strongest relations in Θ , disregarding identities. The meaning of *each other* is given in (151), where $R_x^I = ((R \cup Id) \upharpoonright AT(x))$.

- (151) $[[\text{each other}]] = \lambda R \lambda x. \forall S \in \Theta_R (R_x^I \subseteq S_x^I \rightarrow R_x^I = S_x^I)$

(151) says that the extension of the relation to which *each other* applies is the maximal one when we restrict the relation to x , the antecedent of the reciprocal, ignore identities and take our world knowledge into account. Let me illustrate on two examples how this works. In (152a), *each other* applies to the relation *like* and the set *Morris, Hilary and Philip*. This means that the *like* relation, restricted to *Morris, Hilary and Philip*, and with the identities in it subtracted, should be one of the strongest relations that world knowledge about *liking* allows.

- (152) a. Morris, Hilary and Philip like each other. (SR)
 b. Morris, Hilary and Philip stood on top of each other. (IAR)

World knowledge places no constraints on how many persons one may like, or how many one may be liked by. So the strongest denotation we could have in this case is the full cartesian product of the set in question, minus identities. In other words, (152a) is true if each of Morris, Hilary and Philip likes each of the other ones. On the other hand, the *stand on top of* relation (SOT) in (152b) is restricted by world knowledge. It cannot denote the cartesian product of our set {Morris, Hilary, Philip}, for example. The strongest relations in Θ_{SOT} are strict linear orders.

Contrary to Sabato and Winter (2005b) I do not think that the extension of *relation* plays a role in searching for the strongest reading of reciprocal sentences. Consider the following example:¹⁹ there is an audience in a lecture hall listening and looking at the speaker. The speaker is looking back at one person in the audience. In this scenario, (153) sounds very odd.

- (153) # The people in the lecture hall are staring at each other.

This is problematic for Sabato and Winter (2005b). *Staring* is a function: one can only stare at one person/object at any moment. Thus, we would expect that the reciprocal sentence is satisfied if everyone among x is looking at someone among x . This is true in the scenario above yet the sentence is odd. On the other hand, if we modify the situation and assume that, for example, there were as many people in the audience as there were speakers and everyone in the audience was looking at a different speaker and every speaker was looking at a different listener, the sentence becomes fine. This is surprising for Sabato and Winter since in the two examples we do not vary the number of pairs that satisfy the relation and yet in one case (153) is odd and in the other case the sentence is fine. Obviously, it is not enough to count number of elements in the extension of a relation. What matters is which elements are related to each other, as is expected if we connect the readings in reciprocal sentences to the theory of plurality. Since the weakest reading possible is the cumulative one which requires that everyone was staring at someone and everyone was being stared at by someone, the difference in the two scenarios follows.

Apart from this problem, Sabato and Winter (2005b) suffer from the same problems as Dalrymple et al. (1998). First, negated reciprocal sentences ended up having too weak semantics. For example, if we negate (152a), we get the following interpretation:

¹⁹This example is due to Øystein Nilsen.

$$(154) \quad \neg \forall S \in \Theta_{\text{LIKE}} (\text{LIKE}_{\text{M} \oplus \text{P} \oplus \text{H}}^I \subseteq S_{\text{M} \oplus \text{P} \oplus \text{H}}^I \rightarrow \text{LIKE}_{\text{M} \oplus \text{P} \oplus \text{H}}^I = S_{\text{M} \oplus \text{P} \oplus \text{H}}^I)$$

This is true if the extension of the relation of liking is not the maximal one when restricted to Morris, Hilary and Philip, with identities ignored and given our world knowledge, which is satisfied if for example, everyone likes everyone else apart from Morris, who does not like Hilary. This is too weak since the sentence *Morris, Hilary and Philip do not like each other* is not true in this case.

Finally, Sabato and Winter (2005b) face exactly same problems as Dalrymple et al. (1998) with sentences like the following, discussed in the previous section:

- (155) The children gave each other Christmas presents.
(each child giving a different present)

The problem here is identical to the problem of Dalrymple et al. (1998). For more discussion, see the previous section.

Moltmann (1992)

Moltmann assumes a bipartite interpretation of reciprocal sentences. Consider the following example:

- (156) Morris and Hilary dried each other.

This example gets the following two parts as its interpretation. (157a) makes sure that *each other* is anaphoric to Morris and Hilary. (157b) says that within the event where Morris and Hilary dried Morris and Hilary (established by (157a)) it is true that every entity which is part of Morris and Hilary and in *Cov* dried a distinct entity and was dried by a distinct entity.

- (157) a. $(\exists e)(\text{*DRY}(e) \wedge C_{\text{*}\Theta_1}(\text{M} \oplus \text{H})(e) \wedge C_{\text{*}\Theta_2}(\text{M} \oplus \text{H})(e))$
b. $(\forall x \leq \text{M} \oplus \text{H})(\uparrow x \in \text{Cov} \rightarrow (\exists e' e'' x' x'')(x' \leq \text{M} \oplus \text{H} \wedge \neg x' \circ x \wedge x'' \leq \text{M} \oplus \text{H} \wedge \neg x'' \circ x \wedge e' \leq e \wedge e'' \leq e \wedge x \text{ dries } x'' \text{ at } e' \text{ and } x'' \text{ dries } x \text{ at } e''))$

We can get this reading if we assume the following interpretation of *each other*.

- (158) *Each other*, Moltmann (1992)
 $\lambda R \lambda x \lambda e. R x x e \wedge$
 $\wedge (\forall y \leq x)(\uparrow y \in \text{Cov} \rightarrow (\exists e' e'' y' y'')(y' \leq x \wedge \neg y' \circ y \wedge y'' \leq x \wedge \neg y'' \circ y \wedge e' \leq e \wedge e'' \leq e \wedge R y y' e' \wedge R y y'' e''))$

The first conjunct in *each other* expresses that the reciprocal is anaphoric. The second conjunct expresses the distinctness condition.

One reason Moltmann assumes the bipartite interpretation of *each other* is to explain why DPs in reciprocal sentences are not interpreted as covarying with the antecedent of *each other*, at least not preferably. Let me go through one example:

- (159) The children gave each other a Christmas present.

Moltmann wants to derive why (159) is preferably interpreted as ‘there was one present and the children gave this present to each other’. We have seen in Section 4.3.3 that her intuition is supported by the data in the questionnaire. In our account, this fact was captured because covarying reading is dispreferred. Moltmann assumes that inserting *, which would normally lead to a distributive interpretation, is dispreferred. This can derive why, for example, (160) has a dispreferred interpretation ‘each boy built a boat’. See Chapter 3 for more discussion.

(160) The boys built a boat.

However, blocking the insertion of * in the structure is normally irrelevant for reciprocal sentences since in standard accounts reciprocals distribute over the whole clause so the dispreferred interpretation of (159) should be fully acceptable no matter what we assume about the status of *.²⁰

How Moltmann’s analysis of *each other* deals with (159)? Assume that a *present* stays in situ so it is in scope of the reciprocal. In that case the relation $\lambda x \lambda y \lambda e. x$ gave y a present at e is an argument of the reciprocal. However, we do not get a distributive interpretation of a *present*. The interpretation of the sentence is in (161). This can only lead to the interpretation with only one present involves: it says that every kid gave a present to another kid and received a present from another kid (due to the second part of the meaning of *each other*) and in total one present was exchanged between all the kids (due to the first part of the meaning of *each other*).

(161) $(\exists e) * \text{GAVE A PRESENT}(\text{THE CHILDREN})(\text{THE CHILDREN})(e) \wedge$
 $(\forall x \leq \text{THE CHILDREN})(\uparrow x \in \text{Cov} \rightarrow (\exists e' e'' x' x'')(x' \leq \text{THE CHILDREN} \wedge$
 $\neg x' \circ x \wedge x'' \leq \text{THE CHILDREN} \wedge \neg x'' \circ x \wedge e' \leq e \wedge e'' \leq e \wedge x \text{ gave a}$
 $\text{present to } x'' \text{ at } e' \text{ and } x'' \text{ gave a present to } x \text{ at } e''))$

I am not sure that this way of dealing with (159) is on the right track. Here is a problem. Moltmann wants to derive the low preference for the distributive reading in (160), as well as why a similar reading is dispreferred in (159). In the first case, the low preference follows from the fact that people do not prefer to insert the * operator. However, this assumption is irrelevant for the second case. Here, the particular interpretation is missing because covarying presents with children in (159) simply derives a contradiction. Now, as Moltmann herself notes (Moltmann 1992, footnote 6) some people do accept distributive interpretations in (159) and (160). This corresponds to what we have seen in the questionnaires discussed in Chapter 2 and 3. While the readings are dispreferred, they are possible. This is a problem for Moltmann. She can account for the dispreferred, but possible distributive reading of (160) but she cannot do the same for (159). In the latter case, the reading is simply impossible and the correspondence between the two cases disappears.

²⁰This is not true for the analysis of *each other* that I proposed. If the insertion of * was dispreferred, which would derive that the distributive interpretation of (160) is marked and hard to get, this could still explain why the same interpretation is also marked in (159). This is so because in my analysis of *each other* the reciprocal does not scope over the whole clause, unlike in the account of Dalrymple et al. (1998) or Sabato and Winter (2005b) and others.

Another problem arises from negated reciprocal sentences. Consider the following example:

(162) Morris and Hilary did not dry each other.

(162) would get the following interpretation. This is in the case the negation scopes over the reciprocal. We have seen above that we need to allow for this option.

(163) $(\exists e)(MAX(e) \wedge \neg(\exists e') (e' \leq e \wedge *DRY(M \oplus H)(THE CHILDREN)(e') \wedge (\forall x \leq M \oplus H)(\uparrow x \in Cov \rightarrow (\exists e'' e''' x' x'')(x' \leq M \oplus H \wedge \neg x' \circ x \wedge x'' \leq M \oplus H \wedge \neg x'' \circ x \wedge e'' \leq e' \wedge e''' \leq e' \wedge x \text{ gave a present to } x'' \text{ at } e'' \text{ and } x'' \text{ gave a present to } x \text{ at } e'''))))$

We can rewrite (163) into the following formula:²¹

(164) $(\exists e)(MAX(e) \wedge (\forall e')(\neg e' \leq e \vee \neg *DRY(M \oplus H)(THE CHILDREN)(e') \vee \neg(\forall x \leq M \oplus H)(\uparrow x \in Cov \rightarrow (\exists e'' e''' x' x'')(x' \leq M \oplus H \wedge \neg x' \circ x \wedge x'' \leq M \oplus H \wedge \neg x'' \circ x \wedge e'' \leq e' \wedge e''' \leq e' \wedge x \text{ dried } x' \text{ at } e'' \text{ and } x'' \text{ dried } x \text{ at } e'''))))$

(164) is true if for every subevent e' of some maximal event e one of the two conditions are satisfied:

- either it is not true that Morris and Hilary dried Morris and Hilary in e'
- or it is not true that every atomic individual in Morris and Hilary dried someone else and was dried by someone else

(165) spells out the first condition in detail:

(165) $\neg(*DRY(e') \wedge C_*\Theta_1(M \oplus H)(e') \wedge C_*\Theta_2(M \oplus H)(e'))$

Recall that the $*$ operator comes with presupposition. Thus, the non-presuppositional part of (165) states that not everyone dried someone in e' (or not everyone was dried by someone in e') and it presupposes that noone dried anyone in e' . So far so good. The second condition is rewritten below, again spelled out in more detail:

(166) $\neg(\forall x \leq M \oplus H)(\uparrow x \in Cov \rightarrow (\exists e'' e''' x' x'')(x' \leq M \oplus H \wedge \neg x' \circ x \wedge x'' \leq M \oplus H \wedge \neg x'' \circ x \wedge e'' \leq e' \wedge e''' \leq e' \wedge (*DRY(e'') \wedge C_*\Theta_1(x)(e'') \wedge C_*\Theta_2(x')(e'')) \wedge (*DRY(e''') \wedge C_*\Theta_1(x'')(e''') \wedge C_*\Theta_2(x)(e'''))))$

(166) is true if not everyone dried someone else (and was dried by someone else). The presupposition of indivisibility on $*$ is irrelevant. This is because the presupposition only plays a role when thematic roles apply to plural entities but in this case x' and x'' are atoms.

The question is where the presupposition of (165) is interpreted. If it was interpreted outside of its disjunction we get the presupposition for (162) that nobody dried

²¹Because $\neg(\exists e')(e' \leq e \wedge Pe' \wedge Qe') \Leftrightarrow (\forall e')\neg(e' \leq e \wedge Pe' \wedge Qe') \Leftrightarrow (\forall e')(\neg e' \leq e \vee \neg Pe' \vee \neg Qe')$.

anybody. This is too strong: we would get presupposition failure if Morris dried himself, for example. If the presupposition was interpreted inside the disjunction we get the interpretation that either nobody dried anybody or not everyone dried somebody and was dried by somebody. This is too weak because now (162) should be true if Morris dried Hilary. Either way, the resulting interpretation is incorrect.

Moltmann also cannot derive a cumulative reading in reciprocal sentences for the same reason that Dalrymple et al. (1998) fail (see my discussion in the section where I criticized Dalrymple et al. (1998)).

4.4.4 Conclusion

I argued that a DA-approach, in particular, Beck's account of reciprocated anaphora, should be exclusively used to derive the interpretation of *the others*. This led me use an R-approach for the semantics of *each other*. I developed a new account, which could deal with cumulative readings in reciprocal sentences and negated reciprocal sentences, as well as the range of readings. I also suggested an analysis of collective predicates in reciprocal sentences, following Brisson's work. None of these issues can be accounted for in alternative R-approaches to *each other*, as far as I know.

Splitting the strategies in such a way that one should be solely used for *the others* while the other strategy should be solely used for *each other* helped us explain why

- sentences with *the others* lack weaker readings than SR
- sentences with *each other* can have weaker readings than SR
- *the others*, in its bound reading, can have distributive quantifiers as its antecedent and is dispreferred with non-distributive quantifiers
- *each other* can have non-distributive quantifiers as its antecedent, as long as they express a plurality, and it is incompatible with distributive quantifiers

4.5 Antecedents of *each other* and *the others*

In the DA-approach, the reciprocated anaphor is doubly anaphoric, by its range and its contrast arguments (see (167a)). This is not true of the R-approach where the reciprocated anaphor has only the plural argument as its antecedent (see (167b)).

- (167) a. Interpretation of *each other*, the DA-approach

$$[[\text{each other}]] = \lambda P.P(\sigma(\lambda x.x \leq z_{\text{range}} \wedge \neg x \circ \downarrow y_{\text{contrast}}))\sigma(\lambda x.x \leq z_{\text{range}} \wedge \neg x \circ y_{\text{contrast}})$$
- b. Interpretation of *each other*, R-approach

$$[[\text{each other}]] = \lambda Q_{\langle e, \langle e, \langle v, \langle v, t \rangle \rangle \rangle \rangle} \lambda x \lambda e. \langle x, x, e \rangle \in [\{ \langle a, b, e' \rangle : Q(a)(b)(e')(e) \wedge \neg a \circ b \}]$$

It is in principle possible, then, that the variables of the reciprocated anaphor in the DA-approach are bound by two different DPs. Consider (168). The first example can

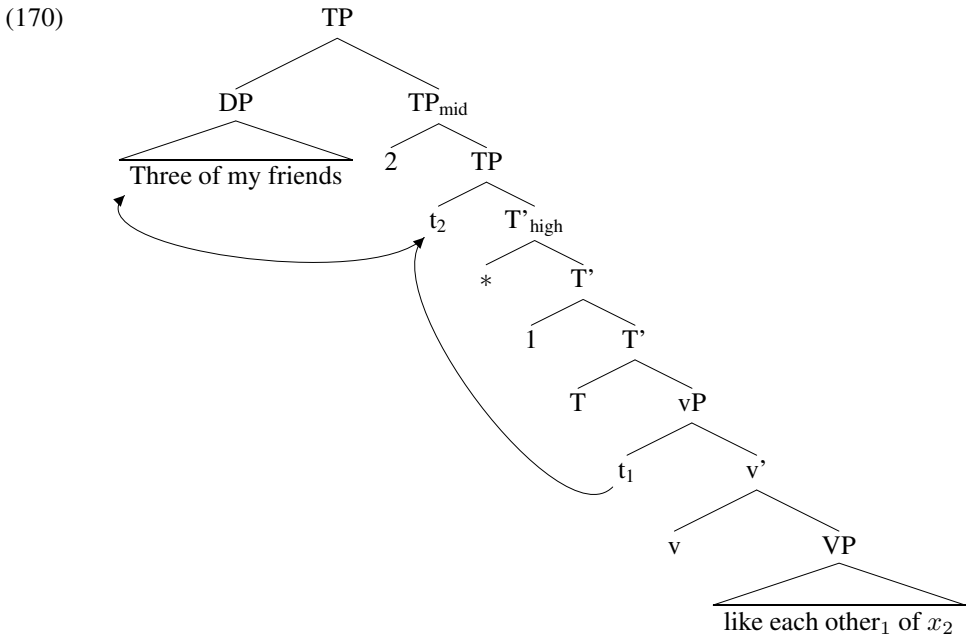
mean that each of my three friends likes all my other friends. Thus, *the others* is bound by two different DPs. The contrast argument is supplied by (each of my) three friends, while the range argument is supplied by all my friends. This reading is not possible in (168b) which can only be interpreted as ‘each of my three friends likes the other two friends’.

- (168) a. Concerning my friends, only three of them like the others.
 b. Concerning my friends, only three of them like each other.

Notice that the contrast cannot be simply explained by the fact that the preposed phrase *concerning my phrase* cannot bind. That it can bind is shown in the following example, where *it* is bound by *each cabinet*.

- (169) As concerning each cabinet, it is also provided with a switch fuse in order to isolate these charging points.

This difference between the behavior of *the others* and *each other* has been, to my knowledge, first noted by Mats Rooth (p.c. in Heim et al. (1991a)). Heim et al. (1991a), who assume the DA-approach to reciprocals, take these data to show that the range and contrast argument must be bound in specific way. In particular, the range argument of *each other* must be bound by the same plural argument that supplies the contrast argument for *each other*. (170) is a tree structure for (168). Since the contrast argument is bound by 1, which is the distributed part of *three of my friends* the range argument must be bound by 2 = three of my friends.



This account captures the data by adding a stipulation to binding that is unheard of in any other case. We know of restrictions that restrict binding to a local domain (Prin-

ciple A). There are also cases of binding which is restricted to particular antecedents. For example, it is cross-linguistically common that reflexives can be bound only by the subject. However, there are *no* cases, as far as I know, in which binding of one argument depends on binding of another argument. The DA-accounts of *each other* would represent a strange anomaly. Of course, we avoid this anomaly if we simply take *the others* to have the semantics (167a) (DA-approach) and *each other* to have the semantics (167b) (R-approach) because in the latter case there is simply one antecedent for the reciprocal, the plural argument.

A closely related observation is the following. The contrast and range arguments of *the others* do not need to be supplied locally, they can also be supplied in the discourse. This is shown in (171a) where the contrast argument of *the others* is ‘Angelica’ and the range argument is ‘the students’. The same is impossible for *each other* which needs to find its antecedent locally, within the clause.

- (171) a. The students are in general not very industrious. Angelica hands in all the assignments but **the others** usually forget.
- b. * The students are in general not very industrious. Angelica hands in all the assignments but **each other** usually forget.

This difference falls out if we assume that *the others* is derived in the DA-approach, and *each other* is derived in the R-approach. Standardly, discourse co-reference is achieved in the following way. We let discourse modulate the variable assignment. Pronouns that are anaphoric to previously introduced discourse referents contain free variables. Since free variables are interpreted by the variable assignment we can correctly achieve that they are interpreted as anaphoric to previously introduced discourse referents. The details on how this is done differ (see, for example, Discourse Representation Theory (Kamp, 1981; Kamp and Reyle, 1993), Compositional Discourse Representation Theory (Muskens, 1996), Dynamic Predicate Logic (Groenendijk and Stokhof, 1991)) but the gist is always the same. I will discuss one particular approach, Compositional Discourse Representation Theory with Pluralities in the next chapter. In any of these accounts what is crucial for discourse co-reference is the presence of free variables in anaphora. This straightforwardly distinguishes between (171a) and (171b). Since *each other* is analyzed in the R-approach, no free variables are available and thus, discourse-anaphoric reading must be missing. Since *the others* is analyzed in the DA-approach, the range and contrast argument can in principle be free and enable discourse-anaphoric reading.

4.5.1 Conclusion

We have seen that *the others* can be bound by two different arguments and it is compatible with discourse-anaphoric reading. On the other hand, *each other* cannot be bound by more than one DP and is incompatible with discourse-anaphoric reading. The properties follow if *the others* is analyzed in the DA-approach and *each other* is analyzed in the R-approach.

4.6 A null theory of long-distance reciprocity²²

4.6.1 Introduction

Higginbotham (1981) discusses sentences like (172), which are ambiguous between the two readings indicated in (172a,b). Henceforth, I will refer to the (172b) reading of (172) and similar sentences as *long distance reciprocity* or LDR, for short.²³

(172) John and Mary think they like each other.

- a. John and Mary each think: “We like each other.”
- b. John thinks he likes Mary and Mary thinks she likes John.

The best account of LDR to this date is, as far as I know, Dimitriadis (2000). However, it is crucial for this account that reciprocals are analyzed by the DA-approach. Dimitriadis (2000) himself builds the semantics of *each other* on Heim et al. (1991b) which is a DA-approach and a predecessor to Schwarzschild (1996); Sternefeld (1998); Beck (2001) and others. Dalrymple et al. (1998, Section 5) suggest how one could deal with LDR within the R-approach to reciprocals but their account has smaller empirical coverage than Dimitriadis (2000).²⁴ In this section, I propose a novel account of LDR cast within the R-approach to reciprocals, which accounts for the data problematic for Dalrymple et al. (1998). With Dimitriadis, it shares the crucial insight that reciprocals in LDR find its antecedent locally, i.e., within the embedded clause.

Traditionally, LDR has been accounted for by letting the reciprocal either take scope at the matrix clause or take the matrix subject as its antecedent. This was assumed in the analyses of LDR in Higginbotham (1981), Heim et al. (1991a), Heim et al. (1991b), Dalrymple et al. (1998). But there is a problem with these accounts. Binding of reciprocals is governed by local principles, Condition A, and reciprocals should not be able to look for antecedents outside of their clause, and quantifier scope normally cannot cross boundaries of a finite clause. To see the problem, take, as an example, Heim et al. (1991b). They assume the DA-approach, so the reciprocal is interpreted as in Beck’s account. To keep it simple I will just use a paraphrase here, *someone among z_{range} different from y_{contrast}* . The two variables are bound as discussed in the introduction to the DA-approach. In LDR the matrix clause is interpreted distributively. The lambda operator in scope of the matrix $*$ operator binds the pronoun *they* in the embedded clause. The pronoun *they* binds y_{contrast} of *each other*. The range argument of *each other* is bound by the matrix subject, i.e., the plurality John and Mary. This gives us (173a) interpreted in (173b). It should be clear that this is LDR.

- (173) a. $[\text{John and Mary}]_z * \lambda y. y \text{ thinks that } y \text{ likes someone among } z \text{ different from } y$
- b. $(\forall a \leq \text{John and Mary})(a \text{ is an atom} \rightarrow a \text{ thinks that } a \text{ likes someone among John and Mary different from } a)$

²²Parts of this section appeared in Dotlačil and Nilsen (to app.).

²³The term is intended as a purely descriptive label. I do not want to imply that LDR involves reciprocals with non-local antecedents, or that scope of the reciprocal is involved in the ambiguity.

²⁴I discuss both Dalrymple et al. (1998, Section 5) and Dimitriadis (2000) in more detail in Section 4.6.7.

LDR is captured but only at the cost of allowing non-local antecedents for reciprocals. This does not concern y_{contrast} , which has a local antecedent in this case (the embedded subject *they*) but it is problematic for z_{range} . In order to avoid over-generation, Heim et al. (1991a) constrain non-local binding of z_{range} . These add extra stipulations on binding and preferably, we should look for a way to avoid them. Furthermore, Dimitriadis (2000) shows that reciprocals always find its antecedents locally. Consider (174).

- (174) The lawyers who represent John and Mary think they will sue each other.
[Dimitriadis (2000), p. 58]

(174) has as one of its interpretations that ‘John’s lawyer thinks John will Sue Mary, and Mary’s lawyer thinks Mary will sue John’. Let us see how Heim et al. (1991b) could deal with (174). We could assume the embedded subject *they* is bound by the lambda operator in scope of the matrix $*$ and *they* in turn binds y_{contrast} . This gives us (175a) which means that the lawyers think that they (=the lawyers) will sue each other. This is a possible interpretation but not the one we want. Assume, instead, that pronouns are interpreted as functions. These could be functions from situations and descriptions to entities (Elbourne, 2006) but to simplify the matter let us assume that the function is from entities to entities. In particular, such a function can be applied to a lawyer and give, as its value, say, the entity that the lawyer represents (notated below as $f(a)$). Furthermore, assume that x_{contrast} can be interpreted in the same manner. If so, (174) is interpreted as (175b), which is also wrong. The reason is that x_{range} is still bound non-locally, and therefore, its value is *the lawyers*, which leads to the non-existent reading that John and Mary’s lawyers think John and Mary will sue the lawyers.

- (175) a. $\forall a \leq \text{the lawyers}(a \text{ is an atom} \rightarrow a \text{ thinks that } a \text{ will sue someone among the lawyers different from } a)$
b. $\forall a \leq \text{the lawyers}(a \text{ is an atom} \rightarrow a \text{ thinks that } f(a) \text{ will sue someone among the lawyers different from } f(a))$

Now, there might be ways to exclude the second reading. However, we still did not derive the reading we want. Clearly, we need the value of x_{range} to be *John and Mary*. Where could the reciprocal pick this antecedent? Not in the relative clause because the plural *John and Mary* does not c-command *each other*. The only other option is the functional pronoun *they*, whose range is *John and Mary*. But that means that *each other* finds antecedents for both x_{contrast} and x_{range} locally. Dimitriadis adds a mechanism to retrieve the range of functional pronouns and shows that this allows us to account for other cases of LDR, like (172), as well. Therefore, there is no need to analyze LDR as a case of non-local binding.

While Dimitriadis’ analysis is a step forward because it avoids non-local binding and can account for wide range of data, it still suffers from some problems. These are discussed in more detail in Section 4.6.7. Here, I only want to point out that it is not clear how his approach could be extended to R-analyses of *each other*. Consider the semantics of *each other* for which I argued before.

$$(176) \quad \lambda Q_{\langle e, \langle e, \langle v, \langle v, t \rangle \rangle \rangle \rangle} \lambda x \lambda e. \langle x, x, e \rangle \in [\{ \langle a, b, e' \rangle : Q(a)(b)(e')(e) \wedge \neg a \circ b \}]$$

If we assumed *each other* does not scope outside the embedded clause, pace Dimitriadis, we would get the following interpretation:

$$(177) \quad \forall a \leq_{\text{the lawyers}} \rightarrow a \text{ thinks that } f(a) \text{ EO}(\lambda x \lambda y \lambda e' \lambda e. x \text{ will sue } y \text{ in } e' \leq e)$$

(177) is true if one lawyer's client, let us say John, will sue each other, and the other lawyer's client, Mary, will also sue each other, which is nonsensical. The problem is that in Dimitriadis' account the embedded pronoun in the example must be interpreted as having a single value (the lawyer's client) but R-approaches require it to be a plurality since they bring in the distribution to its parts on their own.

In this section, I present a new analysis of LDR in which reciprocals find their antecedents locally, but unlike in Dimitriadis' account, this will be compatible with the R-approach to reciprocals. The account is built on work with Øystein Nilsen (Dotlačil and Nilsen, to app.). It differs from Dimitriadis' account in the following respect: while he assumed that LDR should be analyzed on par with distributive readings I believe that they should be analyzed on par with cumulative readings. This analysis has been briefly considered and rejected by Dimitriadis but I find his arguments against such an analysis are inconclusive. Moreover, treating LDR as a case of cumulation brings in advantages. It seems that LDR is restricted to the cases in which the reciprocal is in a complement clause of an attitude verb. For example, LDR is not possible when the reciprocal is embedded in a relative clause. (178) cannot mean the same as 'each of Ron, Tom and John read the book that they gave the others. It is only interpretable if Ron, Tom and John each read some book, tho one they had mutually exchanged as presents. This is a relative-clause-internal reading of the reciprocal.

$$(178) \quad \text{Ron, Tom, and John read the book that they gave each other.}$$

This is problematic for Dimitriadis (2000) and others who try to derive LDR by binding of a pronoun and/or reciprocal by a distributive operator because binding can span clause and island boundaries. For example, in (179) *they* can be interpreted as a bound pronoun. On the other hand, cumulative readings are more restricted, and usually are clause-bounded (Beck and Sauerland, 2000).

$$(179) \quad \text{Ron, Tom, and John read the book that they had bought.}$$

Since LDR shows up only with attitude verbs this suggests that the reciprocal is in fact interpreted locally, and we get the effect of long-distance reciprocity just because of some properties of attitude verbs. This is the analysis that we are going to pursue here. In a way this analysis of LDR is minimal since it does not postulate any special mechanism to allow for non-local antecedents, or special interpretations for pronouns, as is assumed in Dimitriadis (2000).

Before we move to the body of the analysis, one last issue should be clarified. One might not be convinced from what I have said so far that LDR exists as an independent reading. Recall (172). The LDR reading might be thought to be implied by the non-LDR reading (172a). Clearly, if John thinks that he and Mary like each other, he

must also think that he likes Mary. Could this be the reason that we find (172b) possible here? Higginbotham (1981) does not think so. One of the arguments against this conclusion is that we would then expect many more ‘readings’ of (172). For example:

- (180) a. “John thinks he likes Mary and Mary thinks they like each other.”
 b. “John thinks they like each other and Mary thinks John likes her.”

But intuitively, these readings are missing in our interpretation of (172). Another argument that something special is going on in (172) has been discussed in Heim et al. (1991a) and goes as follows. We know that there are reciprocal sentences that are semantically anomalous because they assign the subject contradicting requirements. (181) is one such an example. But when they are embedded in contexts which license LDR readings, the resulting sentence is non-contradictory (182b). But if there was only one reading we would expect the sentence to be understood only as contradiction (182a).

(181) John and Mary are taller than each other.

(182) John and Mary think they are taller than each other

- a. John and Mary each think: “We are taller than each other.”
 b. “John thinks he’s taller than Mary, and Mary thinks she’s taller than John.”

I conclude that LDR does not arise just because it is implied by other readings, and should be accounted for in a different manner. The rest of the paper discusses how.

This section is organized as follows. In the next section I introduce our background assumptions: semantics of attitude verbs. In section 4.6.3 I show how the semantics of attitude verbs combines with the semantics of pluralities to mimic readings that are normally analyzed by postulating bound pronouns. In section 4.6.4 I show how the same mechanism derives LDR.

4.6.2 Semantics of Attitude Verbs (Cresswell and von Stechow, 1982)

Perhaps the most obvious analysis of verbs like *believe* and the way they relate to the embedded clause was proposed by Hintikka (1962). According to this analysis, the embedded clause simply denotes the set of worlds where it is true, and a sentence like *John believes that p* means that all the worlds which John considers as candidates for the actual world satisfy *p*. The set of worlds John takes to be such candidates in world *w* is usually referred to as DOX_j^w , John’s doxastic alternatives in *w*. Thus, for *John believes that p* we get $DOX_j^w \subseteq p$. However, as has been widely discussed in the literature, this propositional account is not fine-grained enough to distinguish between the three modes of attitude, *de dicto*, *de re* and *de se*.

As Lewis (1979) showed, the differences between *de dicto* and *de se* can be captured if we switch from propositions (sets of worlds) to properties (sets of functions

from worlds to entities, or, equivalently, world-entity pairs). Let us use \mathcal{L}_a^w as the notation for the set of world-entity pairs $\langle w', b \rangle$ such that a keeps open the possibility that he is b living in w' . According to this analysis, a sentence like *John believes that P* means that $\mathcal{L}_{John}^w \subseteq P$. Cresswell and von Stechow (1982) show how Lewis' approach can deal with *de re* when combined with Kaplan's insights on acquaintance relations (Kaplan, 1969). In Cresswell and von Stechow (1982), a *de re* believes of b that P must satisfy three conditions:

- (183) a. There is a vivid acquaintance relation R (notation: $\mathcal{A}(R)$)
 b. R relates the agent of belief, a , and *res*, b
 c. $\mathcal{L}_a^w \subseteq \{ \langle w', x \rangle : x \text{ is related by } R \text{ to } y \text{ in } w' \text{ and } Py \text{ is true in } w' \}$

Let us see how this works on one example. Ralph sees Ortcutt walking down the street. Ortcutt is well dressed so Ralph forms a belief right away: Ortcutt must be a professor. We can report on this situation:

- (184) Ralph believes that Ortcutt is a professor.

This is a *de re* belief. In Cresswell and von Stechow's analysis (184) is true if the following holds:

- (185) $(\exists R) \left(\begin{array}{l} \mathcal{A}(R) \wedge R(\text{RALPH}, \text{ORTCUTT}) \wedge \\ \mathcal{L}_{Ralph}^w \subseteq \{ \langle w', x \rangle : professor(\iota b'. R(x, b')) \} \end{array} \right)$

(185) requires that there is some cognitively vivid acquaintance relation R which relates Ralph and Ortcutt. It is commonly assumed that direct cognitive relations (seeing, hearing etc.) satisfy this condition. Since in our case Ralph saw Ortcutt, we can take R to be ' x sees y walk down the street'. (185) says that Ralph believes that the person he is related to by R is a professor. Since this is Ortcutt in our case, (185) is true in the scenario described.

In Cresswell and von Stechow's analysis, attitude verbs are transitive and their internal argument is a structured proposition. A structured proposition is a pair $\langle \langle \alpha_1, \dots, \alpha_n \rangle, \eta \rangle$ s.t. $\eta(\alpha_1, \dots, \alpha_{n-1})$ is a proposition. To be able to read the formula bear in mind that $\alpha_1, \dots, \alpha_n$ are *res*, and η is the property that one believes to hold of these objects. To create a structured proposition one needs to assume a polymorphic functor that takes any number of arguments and a function and combines them into a structured proposition. Cresswell and von Stechow (1982) assume that *that* is the relevant polymorphic functor. Notice that in order to make this work *that* needs to be supplied with *res* separately from the predicate that applies to *res*. To achieve that one needs to assume movement into the position where the *res* and the property can be taken as arguments of the polymorphic *that*. *That* then returns a structured proposition.

- (186) $[[\text{that}]] = \lambda a_1 \dots \lambda a_n \lambda \eta. \langle \langle a_1, \dots, a_n \rangle, \eta \rangle$, if $\eta(a_1, \dots, a_n)$ is a proposition, undefined otherwise

The structured proposition is the internal argument of the attitude verb. In our event semantics, then, the sentence (184) comes out as follows:

$$(187) \quad (\exists e) \left(\begin{array}{l} *BELIEVE(e) \wedge C*\Theta_1(RALPH)(e) \wedge \\ C*\Theta_2(\langle ORTCUTT, \lambda x. PROFESSOR(x) \rangle)(e) \end{array} \right)$$

In (187) I introduced the internal argument of *believe* through thematic roles lacking labels. No matter what their labels are, what is relevant is that *believe* and other attitude verbs induce lexical requirements on the arguments of these thematic roles. This is similar to other verbs. For example, *kiss* requires that the agent touches the theme with her lips. In case of attitude verbs, the requirement is that Θ_2 must be a structured proposition, where the first member is a *res*, and the second member is a property. The *res* and the property must satisfy requirements we discussed above (183). These, as we have seen, are the right conditions for *de re* belief.

Now, notice that thematic roles is pluralized. Of course we could block the pluralization but since the $*$ operator is normally taken to apply freely, this would be an extra assumption. If we allow it, we expect to get cumulative readings with attitude verbs, similar to the ones that we discussed before (Section 3.2.4). This, as we are going to see, allow us to derive LDR.

4.6.3 Cumulation in *de re* Attitudes – Dependent Readings

I believe that LDR readings can be accounted for by applying cumulation between individuals and their beliefs. In fact, cumulation is not restricted to LDR readings. It can be detected in other sentences with attitude verbs embedding plural arguments. (188) has as one possible reading (188b). I am going to refer to (188b) as a *dependent reading*.

(188) John and Bill think they will win.

- a. John and Bill think: “We’ll win.”
- b. Each of John and Bill thinks: “I’ll win.”

The dependent reading of (188) falls out from what we assume so far about the semantics of attitude verbs, events and pluralities. Since this example is simpler than LDR readings I start with it.

(188) can be translated into the formula (189) (disregarding tense) which gives us a *de re* reading in which the pronoun is the *res*.

$$(189) \quad (\exists e)(*THINK(e) \wedge C*\Theta_1(JOHN \oplus BILL)(e) \wedge C*\Theta_2(\langle THEY, WIN \rangle)(e))$$

How is *they* interpreted? There are various options but to get the dependent reading, the simplest one suffices here. *They* is interpreted as $JOHN \oplus BILL$. Θ_1 , which relates John and Bill to e , is pluralized. That means that e could have subevents one of which has John satisfying Θ_1 and another one which has Bill satisfying Θ_1 . Similarly, Θ_2 is pluralized, so e could have subevents which have parts of the structured proposition as their Θ_2 . The most straightforward way to define subparts of structured propositions is a pointwise definition:

$$\langle \alpha, \beta \rangle \leq \langle \alpha', \beta' \rangle \text{ iff } \alpha \leq \alpha' \text{ and } \beta \leq \beta'$$

Then, the subparts of $\langle \text{THEY}, \text{WIN} \rangle$ are $\langle \text{JOHN}, \text{WIN} \rangle$ and $\langle \text{BILL}, \text{WIN} \rangle$ (provided *they* is interpreted as John and Bill, as said above). Is there any other way to split the structured proposition, in particular, could we find subparts of the property *win*? Probably not because an atomic property has no readily available subparts (this does not mean that we cannot define part of relation which would split a property, only that the subparts might not be available in this case, in which respect it is similar to atomic individuals). Even if we did have subparts of the property available it would not matter much. The reason is that the property plays a role only in one condition of *de re* beliefs (see (183c)). This condition states that the belief set of e (\mathcal{L}_a^w) is a subset of the set of $\langle w', x \rangle$, where x is related by R to someone who has the relevant property (*winning* here). Let us notate the set of $\langle w', x \rangle$ as Q . If we split the property *winning* into win_1 and win_2 and x is related by R to someone who win_1 in Q' , then Q' is a subset of Q . Now, let us say that a *de re* believes of b that win_1 . Then, a 's belief properties are a subset of Q' . But because *subset of* is a transitive relation, it holds that a 's belief properties are a subset of Q , therefore, a *de re* believes of b that *win*. Thus, splitting a property into subparts, even if we allow it, would add no new readings to the one we have, and the only relevant way of splitting the structured proposition is by looking at the subpart of the res.

Here is a situation in which (189) is true: there is a plural event, which consists of subevents e_1 and e_2 . e_1 is an event of believing, which has John as Θ_1 , and $\langle \text{JOHN}, \text{WIN} \rangle$ as Θ_2 . Thus, John is the res, and winning is the property. In other words, John thinks he will win. Furthermore, e_2 is an event of thinking, which has Bill as Θ_1 , Bill as the res, and winning as the property. In other words, Bill thinks he will win. This is the dependent reading of (188). Notice that it falls out just from the combination of standard semantics of attitude verbs with the semantics for pluralities that we assumed throughout. In fact, *any* semantics that can derive cumulation between co-arguments would derive this result.

Dimitriadis (2000) argues that one should not tie dependent readings to cumulation. According to him it should be treated as a distributive reading (cf. also Heim et al. 1991a,b). To get the dependent reading of (188) Dimitriadis assumes that $*$ is in the matrix clause taking scope over the embedded clause and *they* is bound accordingly. This gives us the following formula, which captures the dependent reading of (188).

$$(190) \quad (\exists e)(\forall y \leq \mathbf{J} \oplus \mathbf{B})(\exists e' \leq e) \left(\begin{array}{l} * \text{THINK}(e') \wedge \mathcal{C}_{*\Theta_1}(y)(e') \wedge \\ \mathcal{C}_{*\Theta_2}(\langle y, \text{WIN} \rangle)(e')(e') = \end{array} \right)$$

Dimitriadis' main reason to derive the dependent reading as a distributive reading is that this mechanism does not overgenerate. If one assumes cumulative readings for (188) as we do, Dimitriadis argues that one can't distinguish between the two following readings of (191):

- (191) John and Bill think they will win.
- a. Each of John and Bill thinks: "I will win."
 - b. John thinks: "Bill will win", and Bill thinks: "John will win."

While the first reading is attested (this is the dependent reading we just accounted for), the second "crossed" reading is missing, at least according to Heim et al. (1991a) and

Dimitriadis (2000). Let us see why we derive the crossed reading. (189) only requires that John and Bill are in the domain of the pluralized Θ_1 and the res of *de re* belief. But it says nothing about which of the two is, as the argument of Θ_1 related to which part of the res. Thus, *e* could be split into two subevents e_1 and e_2 , where e_1 has John as Θ_1 , Bill as the res (=John thinks Bill will win), and e_2 has Bill as Θ_1 , John as the res (=Bill thinks John will win).

However, I want to point out that there is a difference between (191a) and (191b) on our account, too. In the crossed reading, John is the agent of one subevent, and Bill is the res (and Bill is the agent of the other subevent, of which John is the res). On the other hand, in the dependent reading, John is both the agent and res of one subevent, and similarly for Bill. That means that the acquaintance relations between the res and agent are different in the two cases. In the dependent reading, we can have the relation of identity. In the crossed reading, another cognitive relation must be supplied that would connect John and Bill but would not connect John with himself. It has been noticed that the identity relation is the most prominent acquaintance relation and that this relation can be present out-of-the-blue (Maier, 2006). This can explain the preference for dependent readings. Alternatively, the dependent reading might be preferred over crossed readings because of the properties of the subevents in the two cases. Recall the discussion in Section 3.2.5. Splitting of subevents seems to follow Gestalt principles, which require, among other things, that we consider entities which are similar as belonging to one subevent rather than separating in different subevents.²⁵ Recall the example from Kratzer (2003). Suppose you and I each have a donkey and a cat, and these are all the animals we have. My donkey looks just like your donkey, and my cat looks just like your cat. In this situation, (192a) is true and (192b) is false. This is because in (192a) we have a plural event which can be split into two subevents. In one of the subevents we compare my cat to your cat, in the other subevent we compare my donkey to your donkey, and since they look alike the sentence is true. In (192b) we again need two subevents. But this time, in one of them we would have to have my donkey and your cat, and in the other one we would have to have my cat and your donkey. This latter split violates one of the Gestalt principles (principle of similarity).

- (192) a. My animals look like your animals.
 b. My animals look different from your animals.

Crossed readings violate the principle of similarity as well since we are putting together entities that are less similar (John as Θ_1 , Bill as res) even though we have an option to split the event in a way that does not violate the principle of similarity (John as Θ_1 , John as res). If this reasoning is correct we might expect that in some contexts, the crossed reading becomes possible. While such cases are not that easy to find, they do exist. The following example was found on Google:

- (193) Just about ten minutes ago I was talking to a Russian guy who was visiting his Uncle and cousin in Aktau. He is from a Russian city about 150 kilometers

²⁵This is a simplification because in reality Gestalt principles do not need to be satisfied absolutely when we parse a situation. Rather, we are striving for the optimal parse: the one which violates as few Gestalt principles as possible. For the discussion here the simplification suffices. For more details, see Section 3.2.5.

from the border with Kazakhstan. **We both thought that we were spies.** He was asking me if I had documents and where was I from. How did I get to Aktau. . .

<http://www.travelblog.org/Asia/Kazakhstan/blog-6884.html>

This example is from a weblog of an American who travels in Kazakhstan. The author is *not* a spy (neither does he think that), though he encounters local people who suspect him. Thus, the reading corresponding to (188a) ('I thought we were spies and he thought the same') does not make sense in this context. For the same reason the dependent reading corresponding to (188b) ('I thought I was a spy and he thought the same about himself') is excluded in this context. The only reading that is reasonable here is the 'crossed reading' ('I thought he was a spy and he thought the same about me'). Notice that in this case we have a vivid acquaintance relation, roughly, *meeting and talking to*. I think that this is the reason that the crossed reading emerges in this scenario. Dimitriadis (2000) offers the following minimal pair:

- (194) a. The voters who voted for Street and Weinberg thought that they would lose.
 b. The voters who voted against Street and Weinberg thought that they would win.

(194a) can be paraphrased as 'the voters of Street thought that Street would lose, and the voters of Weinberg thought that Weinberg would lose'. (194b) can be paraphrased as 'the voters voting against Street thought that Street would win, and the voters voting against Weinberg thought that Weinberg would win'. These readings contradict our world-knowledge about voters, namely, that they commonly believe that *their* candidates would win. Still, they are available while the readings conforming to our world-knowledge are absent. Dimitriadis (2000) notes that the embedded pronoun must bear some salient relation to the matrix subject, the argument of an attitude verb. He takes this to be a problem for a cumulative approach to dependent readings. But in fact, under our, cumulative, account we expect this state of affairs since the matrix subject must be related to the *res* (*they*) by some vivid acquaintance relation provided in the context, and *voting for* or *voting against* can do exactly that. All in all, we do make very similar predictions to Dimitriadis (2000). Apart from cases of identity, crossed readings only arise if *res* is related to the subject by some contextually prominent relation.²⁶

Even if we ignore such examples as (193), I think that the argument of Dimitriadis (2000) or Heim et al. (1991a,b) against cumulation is incomplete. In order to derive the dependent reading and block the derivation of the crossed reading it does not suffice to assume that this is a case of distributive reading. One also needs to show that other readings, and in particular, a cumulative reading, are impossible here. By what mechanism would it be blocked? Dimitriadis (2000) suggests that, for some reason,

²⁶In Dimitriadis (2000) the contextually salient relation must relate the subject and the embedded pronoun.

pronouns cannot enter cumulative readings. Consider the following pair, from Dimitriadis (2000). Both sentences can get a dependent reading: each father coached his son(s). But only (195a) can also get what looks like a crossed reading: each father coached sons of other fathers. This is clear if we try the continuation ‘For reasons of fairness, nobody coached his own son’. This is possible in (195a) but in (195b) it does sound like a contradiction.

- (195) a. The fathers coached the sons in Little League baseball.
 b. The fathers coached their sons in Little League baseball.

This could be an argument that pronouns never enter cumulative readings. But consider the following example (196). It seems that here, too, the pronoun blocks the crossed reading. It is hard to understand (196b) as ‘John’s kids kissed Bill and Bill’s kids kissed John’, and this is easier in (196a).

- (196) a. The kids kissed John and Bill.
 b. Their kids kissed John and Bill.

However, we cannot derive the dependent reading of (196b) (‘John’s kids kissed John and Bill’s kids kissed Bill’) as a distributive reading because that would lead to weak crossover violation. Thus, to derive the dependent reading in (196b) we are left only with one option, namely, this must be a cumulative reading. But since the crossed reading is still missing in this case that must mean that cumulative readings do distinguish between crossed and dependent readings with pronouns. The latter is strongly preferred. As I suggested above, this might have to do with the fact that in case of dependent readings the splitting of the plural event is more natural than in case of crossed readings. But if that is true then the missing crossed reading in cases like (188) tells us nothing. In particular, it does not argue against analyzing dependent readings as cumulative.

Dimitriadis (2000) presents another argument against cumulative readings. If we were right that the dependent reading is a case of cumulative readings then it should arise with other DPs than pronouns. Dimitriadis (2000) discusses the following contrast (p.c. to Anthony Kroch):

- (197) a. The voters who voted for Street and Weinberg thought that they would win the election.
 b. The voters who voted for Street and Weinberg thought that Street and Weinberg would win the election.

(197a) can get the dependent reading. On the other hand, (197b) only has the unrealistic reading that every voter thinks that both Street and Weinberg would win the election. In our account this is problematic since (197b) could have the following reading:

- (198) $(\exists e) \left(\begin{array}{l} *THINK(e) \wedge C_{*}\Theta_1(TH\text{E } VOTERS \text{ OF } S. \text{ AND } W.) (e) \wedge \\ C_{*}\Theta_2(\langle S. \text{ AND } W., WIN \rangle) (e) \end{array} \right)$

Since we should be able to split this into subevents where the voters of Street think that Street wins and the voters of Weinberg think that Weinberg wins (parallel to the

case in which *they* was the subject) we derive the dependent reading, which in reality seems to be missing.

I suspect that the reason for this overgeneration is not because we use cumulative readings to derive the dependent reading but because we did not constrain cumulative readings. It seems that cumulative readings are less preferred when the coordination of DPs are the arguments than when we use definite plurals or pronouns. Notice, for example, the following contrast. (199a) can be more easily interpreted with the cumulative reading ('the voters of Street supported Street and the voters of Weinberg supported Weinberg') than (199b).

- (199) a. The voters who voted for Street and Weinberg supported them.
 b. The voters who voted for Street and Weinberg supported Street and Weinberg.

That cumulative readings are harder to obtain with coordination of DPs has also been observed in Winter (2000) and Beck and Sauerland (2000). I do not have an explanation for this effect but it is likely that whatever the right explanation is it should transfer to (197).

Beck and Sauerland (2000) notice that adding the word *respectively* help bring out the cumulative reading in DP coordinations. They argue that constructions with *respectively* which relate two DP coordinations should in fact be treated as a hallmark of cumulative readings. If this is right then we should be able to add *respectively* to cases like (197b) to highlight the dependent reading which otherwise seems to be missing. In fact, such examples are attested. Consider the two examples in (200) from Google search.

- (200) a. ... emergentists [...] as well as non-naturalists [...] thought that emergent and non-natural properties, respectively, supervened on "physicalistically acceptable" properties.
 b. 36% and 25% thought that someone else and special health workers respectively should care for AIDS/HIV patients.

Both examples are interpreted according to the dependent reading. For example, (200a) says that emergentists thought that emergent properties supervened on "physicalistically acceptable" properties, and non-naturalists thought that non-natural properties do. This behavior is expected if the dependent reading can be dealt by assimilating them to cumulative readings.

4.6.4 Cumulation in *de re* Attitudes – LDR Readings

To recapitulate, the LDR reading of (201) is represented by the paraphrase (201b).

- (201) Burt and Clara think they like each other.
 a. Burt and Clara think: "We like each other."
 b. Burt and Clara each think: "I like the other."

The first option to derive (201b) that comes to mind is to try the same strategy as with dependent readings, where the subject pronoun (interpreted as Burt \oplus Clara) was the res:

$$(202) \quad (\exists e) \left(\begin{array}{l} * \text{THINK}(e) \wedge C_{*}\Theta_1(B \oplus C)(e) \wedge \\ C_{*}\Theta_2(\langle \text{THEY, LIKE EACH OTHER} \rangle)(e) \end{array} \right)$$

However, this will not work. We will get the following: there is a plural event e which consists of subevents e_1 and e_2 . e_1 is an event of thinking, which has Burt as Θ_1 , Burt is the res, and liking each other is the property. In other words, Burt thinks he likes each other. e_2 is an event of thinking, which has Clara as Θ_2 , Clara as the res, and liking each other as the property, i.e., Clara thinks she likes each other. This could be true if, for example, Burt and Clara suffer from schizophrenia. Of course, this is not really what we wanted to get.

Fortunately, this is not our only option. There is no reason why the res in (201) should be restricted to the subject pronoun. I suggest that what is needed is *de re* belief about events. This is close in spirit to Cresswell and von Stechow (1982) who, in fact, argue that many different kinds of objects, including properties can function as a res to explain why e.g. *Poirot believes that 59 is 59* does not entail *Poirot believes that 59 is a prime number* (for the details, see Cresswell and von Stechow, 1982, Section 3). Moreover, Abusch (1997) proposes an analysis in which one forms *de re* beliefs about times.

Recall that, on our approach a simple reciprocal sentence like (203a) is interpreted as (203b).

- (203) a. Burt and Clara like each other.
 b. $(\exists e)(* \text{LIKE}(e) \wedge \langle B \oplus C, B \oplus C, e \rangle \in [\{ \langle a, b, e' \rangle : C_{*}\Theta_1(a)(e') \wedge C_{*}\Theta_2(b)(e') \wedge \neg a \circ b \}])$

Here, then, is the main idea. In LDR, the res is some event in which Burt and Clara stand in some relation to each other. The property of the attitude is the embedded verb itself, which specifies what relation Burt and Clara have.

- (204) a. Property: $\lambda e. * \text{LIKE}(e)$
 b. Res: $\iota e. \langle B \oplus C, B \oplus C, e \rangle \in [\{ \langle a, b, e' \rangle : C_{*}\Theta_1(a)(e') \wedge C_{*}\Theta_2(b)(e') \wedge \neg a \circ b \}]$

To get (204) we have to separate the property and the res. First, the verb moves outside of the TP. The remnant TP is then the following:

$$(205) \quad \lambda e. \langle B \oplus C, B \oplus C, e \rangle \in [\{ \langle a, b, e' \rangle : C_{*}\Theta_1(a)(e') \wedge C_{*}\Theta_2(b)(e') \wedge \neg a \circ b \}]$$

The iota operator applies to (205), and this gives us one specific event. The application of the iota operator looks like a stipulation but is necessary as long as we want to allow the possibility that an event can be a res, which we might, independently of LDR. Finally, the event and the verb are combined into the structured proposition (206):

$$(206) \quad \left\langle \iota e. \langle B \oplus C, B \oplus C, e \rangle \in \left[\left\{ \begin{array}{l} \langle a, b, e' \rangle : C_{*}\Theta_1(a)(e') \wedge \\ C_{*}\Theta_2(b)(e') \wedge \neg a \circ b \end{array} \right\} \right], \lambda e. * \text{LIKE}(e) \right\rangle$$

(206) is taken as the argument $C_*\Theta_2$ of *think*:

$$(207) \quad (\exists e'') \left(*THINK(e'') \wedge C_*\Theta_1(B \oplus C)(e'') \wedge \left(C_*\Theta_2 \left(\left\langle \iota e. \langle B \oplus C, B \oplus C, e \rangle \in \left[\left\{ \langle a, b, e' \rangle : C_*\Theta_1(a)(e') \wedge C_*\Theta_2(b)(e') \wedge \neg a \circ b \right\} \right], \lambda e. *LIKE(e) \right) \right) (e'') \right)$$

In (207) we have a plural event of thinking, with Burt and Clara as pluralized Θ_1 , where the res of the thinking are (sub-)events which jointly form one event e in which Burt and Clara are somehow related to each other, and the property is *liking*. Now, (207) is true, for example, if e'' can be split into e'_1 and e'_2 , where in e'_1 , Burt is Θ_1 of thinking, the res is one part of event e , namely the one in which Burt is Θ_1 and Clara is Θ_2 , and this subevent of e is an event of liking. Furthermore, in e'_2 , Clara is Θ_1 of thinking, the res is one part of event e , namely the one in which Clara is Θ_1 and Burt is Θ_2 , and this subevent of e is an event of liking. This corresponds to LDR.

Apart from this reading, our setup gives us another reading: Burt thinks Clara likes him, and Clara thinks Burt likes her. This is again the crossed reading discussed in the previous section, and we can get it by simply splitting the res in a different way. In this case e'' would consist of two subevents, e'_1 and e'_2 , where in e'_1 , Burt is Θ_1 of thinking, the res is one part of event e , namely the one in which *Clara* is Θ_1 and *Burt* is Θ_2 , and this subevent of e is an event of liking, parallelly for e'_2 and Clara. Heim et al. (1991a) and Dimitriadis (2000) claimed that this reading does not exist but once again, the attested examples of the crossed reading exists:

- (208) a. We didn't work together but we met and he's shy, not unlike me, and so I think **we both thought we hated each other**, but were just too shy to really talk.
<http://movies.about.com/od/the producers/a/prodn121405.htm>
- b. I never really fell out of love with her, and I constantly thought about her. She tells me the same thing and that she always wanted to get back together. We would always talk to each other uncomfortably because she thought I hated her and I thought she hated me.(...) **During the time that we thought we hated each other** . . .
<http://www.loveshack.org/forums/t23746/>

In (208a), the context makes clear that the two people were in love with each other, and that they were unaware that the other person loved them back. Hence, we take it to be very unlikely that the highlighted text is intended to mean that *he thought he hated her and she thought she hated him*. It's also extremely unlikely that it should be read on a local construal of the reciprocal, i.e. they both thought: "We hate each other."

Similarly, in (208b), the context specifies that *she thought he hated her and he thought she hated him*. The time referred to by the highlighted text definitely refers to a time when this was true, but that would give us the crossed reading with LDR. In this case, a local construal (*the time that each of us thought: "We hate each other"*) would flatly contradict the preceding context.

These examples show that deriving crossed readings is rather than a problem a virtue of an analysis that connects LDR to cumulative readings.

4.6.5 Consequences of the analysis

Locality of LDR

The analysis of LDR requires no long-distance binding of *each other* so unlike Heim et al. (1991a), Heim et al. (1991b), Dalrymple et al. (1998) we need to add nothing to explain why LDR is possible even though *each other* must be bound within the clause where it appears. Also, unlike these approaches we have no problems to explain (209), repeated from above.

- (209) The lawyers who represent John and Mary think they will sue each other.
[Dimitriadis (2000), p. 58]

(209) has the LDR reading ‘The lawyer who represents John thinks John will sue Mary and the lawyer who represent Mary thinks Mary will sue John’. This can be derived in our analysis if we have *they* interpreted as John and Mary and *de re* belief of the lawyers where the res is the event of John and Mary being related to each other and the property is suing. Then, the proposition is true if there are two subevents e_1 and e_2 and in e_1 it holds that John’s lawyer thinks of the relation of John to Mary that the relation is suing, and in e_2 it holds that Mary’s lawyer thinks of the relation of Mary to John that the relation is suing.

Unlike Dimitriadis (2000) we do not overgenerate LDR to cases other than attitude verbs. Consider the following example:

- (210) Morris and Philip read the book that they had given each other.

This sentence does not have the LDR reading, that is, it cannot be paraphrased as ‘Morris read the book that he had given to Philip and Philip read the book that he had given to Morris’. We will see in the next section, where I discuss Dimitriadis’ approach in detail, why he derives this reading. Here, I only want to point out how we avoid it. For the simplicity, I will assume the approach to relative clauses as in Heim and Kratzer (1998): the relative clause *that they had given each other* is a predicate which intersects with the noun *book* and the resulting predicate is the argument of *the*. Given what we had assumed so far the relative clause is going to be:

- (211) $\lambda x. (\exists e) (*\text{GIVE}(e) \wedge C_*\Theta_3(x)(e) \wedge$
 $\langle \text{THEY, THEY}, e \rangle \in [\{ \langle a, b, e' \rangle : C_*\Theta_1(a)(e') \wedge C_*\Theta_2(b)(e') \wedge \neg a \circ b \}])$

In other words, the relative clause is the set of things that Morris and Philip gave to each other. If we intersect this with the noun *book* we get the set of the things that Morris and Philip gave to each other and which must be atomic entities and books. Finally, the definite article requires that a unique object satisfies this. So, the reading we derive for (210) is ‘Morris and Philip read a unique object which (i) is a book and (ii) they gave each other. This is the only possible reading of (210).

If we changed (210) into (212) where we substitute *the book* with *the books* we get the LDR reading back. Thus, (212) can mean ‘Morris read the book he had given to Philip and Philip read the book he had given Morris’.

- (212) Morris and Philip read the books that they had given each other.

This is predicted by the cumulative analysis of LDR. The relative clause remains the same as in (211). However, now we intersect it with the noun *books* so we get set of the things that Morris and Philip gave to each other and which must consist of at least two atoms and are books. So we have at least two objects in this set: one book that Morris gave to Philip and one book that Philip gave to Morris. The definite article picks the maximal set of books in the set. Suppose there are just two as assumed. The matrix clause would then express (213). Since we can cumulate between the two arguments of the clause we can get the reading that Morris read one book (for example, the one he had given to Philip) and Philip read the other book.

$$(213) \quad (\exists e)(\ast\text{READ}(e) \wedge {}^C\ast\Theta_1(\text{M} \oplus \text{P})(e) \wedge {}^C\ast\Theta_2(\text{THE TWO EXCHANGED BOOKS})(e))$$

Notice that unlike LDR, dependent readings are possible in a sentence parallel to (210). The following sentence can mean ‘each boy read the book that he had bought’.

$$(214) \quad \text{All the boys read the book that they had bought.}$$

The reason we get the dependent reading in (214) is due to the fact that the matrix subject can distribute over the sentence and the pronoun can be bound in which case it is interpreted distributively. In that case the relative clause is interpreted as ‘for each boy, the set of things which the boy had bought’. If we proceed further, it straightforwardly gives us the dependent reading. This strategy cannot work for cases like (210). If we let *they* be bound by the distributively interpreted matrix subject the interpretation crashes since *each other* cannot apply to a DP whose reference is an atomic individual. Thus, since we disassociate LDR and distributive readings we correctly block overgeneration in cases like (210), unlike approaches which assume LDR should be derived by distribution of the matrix subject and binding of an embedded pronoun (Heim et al. 1991b; Dimitriadis 2000). Furthermore, since we assume the DA-analysis of *the others*, we expect that ‘LDR’ should not be limited with *the others* in a way that it is limited with reciprocals. This is correct. For instance, (210) can be paraphrased with (215) in which *the others* substitute the reciprocal:

$$(215) \quad \text{The professors each read the book that they had given the others.}$$

I would like to conclude this section by pointing out one potential problem. Our account of LDR assimilates them to *de re* readings. Thus, we expect LDR to be impossible when *de re* reading is. This is not correct. For instance, (216) is grammatical and can be interpreted as ‘Morris wanted to kiss Hilary and Hilary wanted to kiss Morris’. But it is well-known that infinitives embedded under control predicates allow only *de se* reading (Chierchia 1989, Maier 2006).

$$(216) \quad \text{Morris and Hilary wanted to kiss each other.}$$

To account for (216) we need to assume that *each other* can find its antecedent non-locally, outside of its clause. We differ from previous accounts (Heim et al. 1991a, Heim et al. 1991b, Dalrymple et al. 1998) in assuming that *each other* has non-local antecedents only in infinitival clauses. It is commonly assumed that QR out of infinitival clauses is possible so this assumption might not be that difficult to swallow.

However, letting *each other* QR from infinitival clauses might still be too unrestrictive. The following example shows that *each other* cannot always have the matrix subject as its antecedent.

(217) * Morris and Hilary recommended me to ignore each other.

Obviously, the difference between (217) and (216) is the presence of the object in the former case. Thus, the object should block movement of *each other*. Again, this might not be that surprising. The blocking effect of the object might be just a case of the relativized minimality (Rizzi 1990, Chomsky 1995). Alternatively, the difference between (216) and (217) might follow from the Specified Subject Condition (Chomsky 1973, Heim et al. 1991a). The Specified Subject Condition states that:

(218) No rule can involve X, Y in the structure $\dots X \dots [\alpha \dots Z \dots WYV \dots] \dots$ where Z is the specified subject of WYV in α .

In case of control predicates, it is assumed that the specified subject is the subject of the infinitival clause (PRO) in case X is not its controller. Thus, *each other* can move to the subject in (216) but it cannot move to the subject in (217) since in the latter case it crosses the domain with the specified subject. The Specified Subject Condition could also explain the following example:

(219) * I threatened Hilary and Desiree to kiss each other.

(219) does not allow *each other* to have the object as its antecedent since the object does not control PRO in the infinitival clause, and consequently, PRO is the specified subject.

Even though we do account for LDR in finite and infinitival clauses the fact that we achieve this by using two different accounts seems problematic to me. We could have only one strategy if we allowed cumulation in *de se* readings as well but that is impossible with the account we have right now where we need events to be the res. At this point, I have to leave unification of all LDR cases for future research.

When LDR readings are blocked

Recall that we derive LDR by assuming that the event which is the res of attitude predicates can be split into subevents. It is crucial for LDR that in case we have two participants each thematic role of each subevent has only one of them in its extension. However, there are reciprocal markers in various languages which do not allow split into such subevents. These are commonly verbal reciprocal markers, that is reciprocal markers that are not realized as full DPs like English *each other* (see Nedjalkov et al. 2007, Chapter 3). English null verbal reciprocal morphology is such a case. Consider (220). While the first example can be true if there were two consecutive kisses, that is, the two might be kissing each other on the cheek, this is not possible in (220b). The second example requires that only one kiss is involved (see also Carlson 1998).

- (220) a. Philip and Hilary kissed each other.
b. Philip and Hilary kissed.

I will not attempt to derive the difference between the two cases. The only thing that is crucial for me is that in (220b) there is no subevent where Philip is the sole agent and Hilary the sole theme of kissing. For derivations along these lines, see Dimitriadis (2005) and Rubinstein (to appear).

Given this, we expect that reciprocal sentences with these reciprocal markers should not lead to LDR. This is correct for English. (221) cannot give rise to LDR, that is, it cannot mean 'John thought that he had kissed Mary and Mary thought that she had kissed John' (Heim et al., 1991a; Carlson, 1998).

(221) John and Mary thought that they had kissed.

We expect this because to derive LDR we require that the event of the complement clause can be split into two subevents, such that in one of them John kissed Mary and in the other one Mary kissed John. But such splitting is incompatible with the property of the null reciprocal marker.

Another example of the same phenomenon comes from German. German has two markers to express reciprocity: *sich* and *einander*. The two can be semantically distinguished in (222). The difference is parallel to the difference in the English examples (220) (Kemmer, 1993).

- (222) a. Hans und Maria haben *einander* geküßt.
Hans und Maria have each-other kissed
b. Hans und Maria haben *sich* geküßt.
Hans und Maria have self kissed

As Sauerland (1998) observes, the two markers also behave differently with respect to LDR. While *einander* can give rise to LDR, this is not possible with *sich*. The contrast is shown in (223), where the first sentence is a contradiction, which would be avoided if LDR was a possible interpretation (as it is in (223b)). Again, this is parallel to English.

- (223) a. # Kai und Toni glauben, daß sie *sich* überragen.
Kai and Toni think that they self be-taller-than
b. Kai und Toni glauben, daß sie *einander* überragen.
Kai and Toni think that they each-other be-taller-than

Thus, English and German support the one-way implication of the analysis of LDR, namely, that if a reciprocal marker does not allow split into subevents with parts of the plural antecedent as arguments then this reciprocal marker cannot be used in LDR. It remains to be seen whether this can be extended to other languages and their reciprocal markers.

4.6.6 Conclusion

I have shown how a combination of a standard semantics for attitude verbs (Cresswell and von Stechow, 1982) with common assumptions on pluralities (Landman, 2000)

derives dependent readings in complements of attitude verbs and LDR. This analysis has several advantages over alternative approaches that derive LDR as non-local scope/binding. In particular, it obviates the need for ad hoc constraints on binding configurations for free variables (Heim et al., 1991b), and LDR-specific mechanisms for interpretation of functional pronouns (Dimitriadis, 2000). Furthermore, it is preferable to Dimitriadis (2000) because it restricts LDR to clausal complements of attitude verbs, which seems correct, while it can still derive the results that motivated his analysis.

4.6.7 Previous accounts of LDR

I will discuss two approaches: Dimitriadis (2000) and Dalrymple et al. (1998). I do not discuss Heim et al. (1991a) that has been criticized in Roberts (1991) and Williams (1991) and modified into Heim et al. (1991b). Heim et al. (1991b) is very similar to Dimitriadis (2000) even though the latter has a bigger empirical coverage. For this reason, I focus only on the latter. Most of the problems that hold for Dimitriadis' account also hold for Heim et al. (1991b) (in particular, Problems 1 and 2 discussed below).

Dimitriadis (2000)

The main aim of Dimitriadis (2000) is how to account for examples like (224), which can be paraphrased as 'the lawyer that represents John thinks John will sue Mary and the lawyer that represents Mary thinks she will sue John'.

- (224) The lawyers who represent John and Mary think they will sue each other.
[Dimitriadis (2000), p. 58]

Admittedly, examples like 224 are quite difficult to process to begin with and judgments are further complicated by the resemblance of truth conditions between LDR reading and non-LDR reading (the latter can be paraphrased as 'the lawyer that represents John thinks John and Mary will sue each other, and so does the lawyer that represents Mary'). The two readings can be more easily discerned in (225), where LDR reading is non-contradictory ('each coach thinks that the person he coaches will defeat the other(s)'), while non-LDR reading assigns contradictory belief to each coach. The sentence does have non-contradictory reading which suggests that LDR reading is real.

- (225) The coaches that trained them think they are faster than each other. [Dimitriadis (2000), p. 120]

Dimitriadis follows Engdahl (1986)'s functional analysis of pronouns developed for paycheck sentences. Simplifying somewhat *they* in 224 is translated as $f(u)$, where u is a variable bound by the distributive operator distributing over the matrix subject, *the lawyers who represent John and Mary*, and f is a contextually supplied reference function. In this particular case it would be a function that maps each lawyer to the

person that he/she represents. Thus the pronoun ends up denoting John for John's lawyer and Mary for Mary's lawyer.

Dimitriadis follows the DA-approaches to reciprocals, which means that in his semantics of *each other* there are contrast and range arguments which have to be bound. In the example (224) the contrast argument can be bound by the pronoun, i.e., by the value of $f(u)$. The range argument should then be whatever the range of f is. Dimitriadis makes this value accessible as follows. First, he restricts the function of the pronoun to the pronoun's antecedent (for other values it is undefined). The new restricted function is called r and defined as below (f in the definition is Engdahl's unrestricted reference function, discussed above):

$$(226) \quad r = \lambda x(\iota z)(x \leq \text{antecedent of pronoun} \wedge z = f(x))$$

Once having the restricted reference function we can retrieve the set of range elements, $\text{Range}(r)$ as in (227a). Using $\text{Range}(r)$ as the range argument of *each other* Dimitriadis (2000) offers new semantics for the reciprocal. I present it in (227b) in a slightly modified version.²⁷

$$(227) \quad \begin{array}{ll} \text{a.} & \text{Range}(r) = \lambda y.((\exists x)r(x) = y) \\ \text{b.} & [[\text{each other}]] = \sigma(\lambda z. * \text{Range}(r)(z) \wedge \neg z \circ y_{\text{contrast}}) \end{array}$$

The sentence (224) gets the following LF, which leads to the correct interpretation 'each lawyer thinks his client will sue the other client' (228b). Notice that Dimitriadis (2000) achieves this by assuming that *each other* receives both the contrast and range argument locally. The contrast argument is bound by the dependent variable x_2 , which is the value of r when it takes x_1 as its argument, and the range argument is Range applied to the restricted range function r .

$$(228) \quad \begin{array}{ll} \text{a.} & [\text{the lawyers that represent J and M}] * \lambda x.x \text{ thinks } [r(x)]_z \text{ likes } [\text{each other}_z \text{ of } r] \\ \text{b.} & (\forall x \leq \text{J} \oplus \text{M}'\text{S LAWYERS})(x \text{ thinks} \\ & \wedge [\text{LIKE}(\iota z.x \leq \text{J} \oplus \text{M}'\text{S LAWYERS} \wedge z = f(x))(\sigma(\lambda y.* \text{Range}(r)(y) \wedge \neg z \circ y))])) \end{array}$$

Problem 1. Dimitriadis (2000) assumes that the contrast argument of *each other* is passed up by the embedded subject *they* which, in turn, is bound by the distributed matrix subject. In the example above, the distribution was achieved by the insertion of the $*$ operator. If we used *other* means to induce distribution on the matrix subject, we should have the same results. For example, introducing the floated *each* in the matrix

²⁷Dimitriadis (2000) follows Heim et al. (1991a) and requires *each other* to scope over a predicate. In the version in 227b, I follow Beck (2001) and treat *each other* as a definite plural with unbound variables. This does not change any of the points but it enhances, I believe, the readability. Furthermore, I simplify the distinctness condition. I assume that the contrast argument and z must be non-overlapping. Dimitriadis assumes that the latter must not be part of the former. I discussed problems with this assumption in Section 4.4. $*\text{Range}(r)$ is the range argument (with the free variable r), to be bound by r of the embedded pronoun in cases like (224).

clause should be compatible with LDR. But clearly, there is a difference between the two. While the sentence without *each* can give rise to LDR, the one where *each* is present cannot, as seen in the contrast in (229) and (230), from (Heim et al., 1991a, ex. 69 and 71). While the first case can lead to a non-contradictory interpretation due to the possibility of LDR, the second sentence only has the contradictory reading.

(229) They think they are taller than each other.

a. Each of them thinks that he is taller than the other(s).

(230) They each think they are taller than each other.

a. # Each of them thinks that they are taller than each other.

To account for this contrast Dimitriadis (2000) would have to stipulate that the contrast argument must be bound only when distribution is achieved by the insertion of the * operator, but crucially not by the floated *each*. But this would simply restate the fact. There is no reason why this should be so.

Problem 2. Consider the example discussed before, (231).

(231) Morris and Philip read the books that they had given each other.

(231) presupposes the existence of a unique book that Morris and Philip exchanged. In particular, (231a) cannot give rise to LDR. It is unclear to me how Dimitriadis can account for the absence of the LDR construal in this case.

Recall that relative clauses are predicates that are intersected with the noun heading the relative clause. In Dimitriadis' account, the relative clause of (231) gets the following interpretation:

(232) $\lambda y.f(x)$ gave y to $\sigma(\lambda v.v \leq *Range(r)(v) \wedge \neg v \circ f(x)$

The pronoun is, as before, interpreted as $f(x)$. f is treated as the identity function. x is bound by the distributed matrix subject, i.e., each of Morris and Philip. $*Range(r)$ is the range of the pronoun, that is, Morris and Philip. These ways of binding are made explicit in (233), where the first example shows the LF structure and the second example is the interpretation of the LF structure. Notice that binding here is parallel to binding in LDR cases.

(233) a. $M \oplus P * \lambda x.x$ read the book that $[f(x)]_z$ gave to $[each\ other_z \text{ of } *Range(r)]$
 b. $(\forall x \leq M \oplus P)(x \text{ read } \sigma(\lambda y.BOOK(y) \wedge |AT(y)| = 1 \wedge f(x) \text{ gave } y \text{ to } \sigma(\lambda v.v \leq *Range(r)(v) \wedge \neg v \circ f(x)$

(233b) is true if for every x , where x is either Morris and Philip, x read the book that he had given to someone among Morris and Philip different from himself. This is the LDR interpretation, which was derived completely parallel to other cases of LDR but which should be blocked.

The only option I could think of is to argue that the matrix subject cannot bind the embedded pronoun in (231). But this assumption is problematic because normally, binding can span clause and island boundaries. Consider also (234), where the pronoun can be interpreted as bound by the distributed matrix subject.

- (234) All the boys read the book that they had bought.

Both Problems 1 and 2 arise for Dimitriadis because he assumes that LDR takes place when the subject in a higher clause distributes over the sentence and binds the pronoun, the antecedent of *each other*. The problems show that connecting LDR to distributive readings is unsatisfactory because other cases where we see that the subject in a higher clause can distribute and bind the pronoun cannot give rise to LDR. This behavior is expected if LDR is a case of cumulative readings. Given that cumulative readings are local, usually clause-bounded, we do not run into Problem 2. Problem 1 is also avoided given that the floated *each*, as a distributive quantifier, is incompatible with cumulative readings.

Problem 3. To account for LDR readings Dimitriadis (2000) makes use of the restricted reference function r and its range, $Range(r)$. It would be nice if each of those has been independently motivated. But as far as we know this is not the case. Dimitriadis discusses functional pronouns of Engdahl (1986) and suggests that the dependent pronoun *they* in LDR readings should be analyzed in the same way (Dimitriadis, 2000, Section 3.5). But in fact, it is crucial that these are *not* analyzed in the same way. Functional analysis of pronouns have been suggested to deal with donkey anaphora (235a) and paycheck sentences (235b).

- (235) a. Every woman who had a husband brought him to the party. ('him' = woman's husband)
 b. The man who put his paycheck in the bank was wiser than the man who fed it to his dog. ('it' = man's paycheck)

Him in (235a) could be analyzed using Dimitriadis' r in which case we would get the semantics as in (236). Thus, *him* would be the function that is restricted to the set of married women and for each married woman it gives her respective husband. But if this analysis was possible one would expect that donkey anaphora can antecede reciprocals. Thus, (237) should mean 'every woman who has a husband introduced him to the other women's husbands'. But of course, this sentence is ungrammatical. The same problem arises for paycheck sentences.

- (236) $[[him]] = \lambda x(\iota z)(x \leq \text{WOMAN WHO HAD A HUSBAND} \wedge z = x\text{'S HUSBAND})$

- (237) *Every woman who has a husband introduced him to each other.

As a way out of this problem, one might assume that the antecedent of the pronoun, which defines the domain of the pronoun, is not the set of married women. But it is unclear how this could be achieved since the antecedent of the pronoun is supplied as a free variable. To conclude, as it stands LDR is explained by mechanisms that are not employed for any other phenomena.

Dalrymple et al. (1998)

Dalrymple et al. (1998, Section 5) point out that one could treat *each other* in the R-approach and account for LDR by letting the reciprocal take matrix scope. The LF for (238a) would then come out as in (238b), where EO is *each other*. I discussed its interpretation in Section 4.4.3. Since I focus here on two-membered set, I will assume a simplified interpretation, that is, EO is true of a relation and its antecedent if everyone from the antecedent is related to everyone else from the antecedent by the particular relation.

- (238) a. John and Mary think they like each other.
 b. John and Mary EO $\lambda x \lambda y. [x \text{ thinks } x \text{ likes } y]$

(238b) states that John is related to Mary by the relation $\lambda x \lambda y. x \text{ thinks } x \text{ likes } y$ and Mary is related to John by the same relation. This is correct for the interpretation John thinks he likes Mary and Mary thinks she likes John.

Dalrymple et al. (1998) can account for the contrast between (229) and (230) (Problem 1 of Dimitriadis 2000). The floated *each* in the matrix should be impossible, since the argument of EO would be fed singularities. The second problem of Dimitriadis was that it cannot correctly capture the fact that LDR is restricted to attitude verbs and does not appear in other contexts. For example, (239) lacks LDR.

- (239) Morris and Philip read the book that they gave each other.

The account of Dalrymple et al. (1998) shares this problem. Since it assumes that EO can scope outside of clauses it needs to add some extra restrictions to block it in particular cases like (239).

Most crucially, however, the account of Dalrymple et al. (1998) cannot deal with examples like (240).

- (240) The lawyers who represent John and Mary think they will sue each other.

(240) would get LF as in (241) but this gets the incorrect interpretation ‘each lawyer thinks that his client will sue the lawyer of the other client’.

- (241) [the lawyers of John and Mary] EO[$\lambda x \lambda y. x \text{ thinks } f(x) \text{ will sue } y$].

Now, there might be ways to block this reading. But crucially, there is no way to derive the correct reading in Dalrymple et al. (1998). This just reconfirms the conclusion we have made before: *each other* finds its argument locally even in LDR.

4.7 Conclusion

I argued in this chapter that there are two strategies of resolving anaphoric reference. First, anaphoric expressions have free variables that are bound in syntax. Second, anaphoric expressions have all variables bound and undergo QR to find its antecedent. I claimed that the first strategy is used by *the others* while the second strategy is used

Table 4.8: Properties of *each other* and *the others*

	<i>the others</i>	<i>each other</i>
Distributivity	antecedent must distribute	antecedent does not distribute
The range of readings	weaker readings not possible	weaker readings possible
Binding by more than one DP	possible	not possible
Co-reference in the discourse	possible	not possible
Long-distance reciprocity	unlimited	limited

by reciprocals. This enabled us to explain why *the others* and *each other* have the properties summarized in Table 4.8. The strength of the analysis lies not in the fact that it can predict one or two properties but that we expect properties to bundle together and they do.

Consider an alternative explanation. We could, for example, assume the following: both *each other* and *the others* are analyzed in the DA-approach. This would give us the following semantics for the two expressions:

- (242) Tentative interpretation of *each other* and *the others*, the DA-approach
 $[[\text{each other}]] = \lambda P.P(\sigma(\lambda x.x \leq z_{\text{range}} \wedge \neg x \circ \downarrow y_{\text{contrast}}))$

We could furthermore assume that *each other* requires the contrast and range arguments to be bound in a local domain and this requirement does not hold for *the others*. Predictions of this account are summarized in Table 4.9.

Table 4.9: Predictions of the two strategies

	DA-strategy	DA-strategy with local binding
Distributivity	antecedent must distribute	antecedent must distribute
The range of readings	weaker readings not possible	weaker readings not possible
Binding by more than one DP	possible	possible
Co-reference in the discourse	possible	not possible
Long-distance reciprocity	unlimited	limited

By stipulating the locality condition we could capture the last two properties in which *each other* and *the others* differ but these would not be correlated with the first three properties. This is incorrect for *the others* and *each other*. It is worth pointing out that even though we mainly focused on English other languages show similar way of bundling properties like *the others* and *each other* do. For example, the following generalization seems to be true cross-linguistically: iff a reciprocated anaphor can have the bound reading with distributive antecedents it can also be co-referential to

discourse antecedents. Thus, the properties 1 and 4 in Table 4.8 are related in other languages than English. I show this in the appendix to this chapter.

We have now made a full circle to the beginning of this chapter. We started by discussing two strategies of anaphoricity, exemplified on reflexives. In the first one, the reflexive is a free variable and its binding is achieved by a mechanism in the syntax. I called this the multiplication of resources in the syntax. In the second one, the reflexive is a polyadic quantifier, has no free variables, and a QR suffices to derive its “bound” interpretation. I called this the multiplication of resources in the lexicon. We have wondered whether both strategies are necessary. Could we exclude one of them? Given our discussion on reciprocated anaphora the answer seems to be no. We need both of them. However, we need them for very different circumstances. Notice that the second strategy is necessary for *each other*, which must find its antecedent locally, while the first strategy is necessary for *the others*, which is not restricted in the way it finds its antecedents. If we want to exclude the possibility to derive Table 4.9 we might conclude the following: both strategies are possible but only the second one can be used for anaphora that must be bound within some local domain. This leads me to the following conjecture:

(243) **Conjecture on bound variables:**

There are no variables that must be obligatorily bound at LF

This conjecture explains why Table 4.9 does not exist. It explains, for instance, why discourse-anaphoric reading of reciprocated anaphora implicates that the anaphora will require a distributive quantifier as the antecedent for its bound reading or that it will lack weaker readings. The conjecture might in turn have consequences for the study of reflexives. Since ‘obligatorily bound variables’ are not existing species we need to resort to derive reflexives in the approach that postulates the multiplication of resources in the lexicon. I leave studying consequences of this step to future research.

In the next chapter I am going to turn to other expressions of (non-)identity, in particular *different* and *same*. As we will see the same strategy that we considered here, that is, splitting the way anaphoricity can be achieved, will be fruitful there as well.

4.8 Appendix – cross-linguistic data

Although we mainly focused on English in discussing the properties of reciprocated anaphora it is worth noting that other languages show similar way of bundling their properties together. Here, I focus on the following correlation: iff a reciprocated anaphor can have the bound reading with distributive antecedents it can also be co-referential to discourse antecedents. This, as we have seen, is true for *each other* and *the others* (properties 1 and 4 in Table 4.8). *The others* can appear in its discourse reading and it can give rise to the bound reading with plural distributive DPs (in fact, the plural DP *must* distribute over *the others*). That the latter is true follows from the acceptability of (244). *Each other* cannot appear in the discourse reading and it cannot give rise to the bound reading with distributive DPs.

(244) Each boy/The boys each talked to the others.

This generalization also holds for reciprocated anaphora in Italian ((245)-(246)), Czech ((247)-(248)), Dutch ((249)-(250)) and Greek ((251)-(252)). The data show that whenever a distributive quantifier cannot be the antecedent of a reciprocated anaphor, the discourse reading is missing. In case of Czech, (247b) and (247c) are not degraded but the reciprocal reading disappears and another reading, where *sebe* 'self' is interpreted as reflexive, becomes the only possible one.

- (245) a. I bambini hanno parlato l'uno all'altro.
the boys have talked one to-other
'The boys talked to each other.'
- b. ?? Ciascun bambino ha parlato l'uno all'altro.
each boy has talked one to-other
'Each boy talked to each other.'
- c. * Non ho parlato con tutti i miei studenti. Ho parlato con
Not have talked with all the my students. Have talked with
Angelica ma non l'uno con l'altro.
Angelica but not one with other
'I did not talk to all my students. I talked to Angelica but not to each other/*the others.'
- (246) a. Ciascun bambino ha parlato agli altri bambini
each boy has talked to-the other boys
'Each boy talked to the other boys.'
- b. Non ho parlato con tutti i miei studenti. Ho parlato con
Not have talked with all of my students. Have talked with
Angelica ma non con gli altri.
Angelica but not with the others
'I did not talk to all my students. I talked with Angelica but not with the others.'
- (247) a. Chlapci mluvili o sobě.
boys talked about self
'The boys talked about themselves/each other.'
- b. Každý chlapec mluvil o sobě.
each boy talked about self
'Each boy talked about himself/*each other.'
- c. # Nemluvil jsem se všemi svými studenty. Mluvil jsem s
not-talked aux with all self students. Talked aux with
Angelikou, ale ne se sebou
Angelica but not with self
'I did not talk with all my students. I talked with Angelica but not with myself/*the others.'

- (248) a. Každý chlapec mluvil o ostatních (chlapcích).
each boy talked about other (boys)
'Each boy talked about the other boys.'
- b. Nemluvil jsem se všemi svými studenty. Mluvil jsem s
not-talked aux with all self students. Talked aux with
Angelikou, ale ne s ostatními
Angelica but not with others
'I did not talk with all my students. I talked with Angelica but not with the others.'
- (249) a. De jongens helpen elkaar.
the boys help each-other
'The boys help each other.'
- b. *Iedere jongen helpt elkaar.
each boy help each-other
'Each boy helps each other.'
- c. *Ik heb niet met al mijn studenten gepraat. Ik heb met Angelica
I have not with all my students talked. I have with Angelica
gepraat maar niet met elkaar.
talked but not with each-other
'I did not talk with all my students. I talked with Angelica but not with each other/*the others.'
- (250) a. Iedere jongen helpt de andere jongens.
each boy help the other boys
the degraded status of (b) and (c) examples correlate. 'Each boy helps the other boys.'
- b. Ik heb niet met al mijn studenten gepraat. Ik heb met Angelica
I have not with all my students talked. I have with Angelica
gepraat maar niet met de anderen.
talked but not with the others
'I did not talk with all my students. I talked with Angelica but not with the others.'
- (251) a. I mathites milisan o enas ja ton alon
the students talked the one for the other
'The students talked with each other'
- b. *Kathe mathitis milise o enas ja ton alo
each student talked the one for the other
'Each student talked with each other.'
- c. *De milisa me olus tus mathites mu. Milisa me ton Jani, ala
not I-talked with all the students mine. I-talked with the John, but
de milisa o enas me ton alo.
not I-talked the one with the other

‘I did not talk with all my students. I talked with John but not with each other.’

- (252) a. Kathe mathitis milise ja tus alus (mathites)
 every student talked for the others (students)
 ‘Each student talked with the others (other students).’
- b. De milisa me olus tus mathites mu. Milisa me ton Jani, ala
 not I-talked with all the students mine. I-talked with the John, but
 de milisa me tus alus (mithites).
 not I-talked with the others (students)
 ‘I did not talk with all my students. I talked with John but not with the other students.’

CHAPTER 5

Different and same

5.1 Introduction

I argued in the previous chapter that there are two strategies of resolving anaphoric reference. In one of them an anaphoric expression has variables that are bound in the syntax. In the second strategy the anaphoric expression has all variables bound and undergoes QR to find its antecedent. I claimed that the first strategy is used by *the others* while the second strategy is used by reciprocals. In this chapter, I am going to follow the same dichotomy in anaphoricity to explain characteristics of *same* and *different*.

The chapter is organized as follows. In the next section, I summarize the properties of *same* and *different* that we should account for. In Section 5.3 I show how the two strategies on anaphoricity can explain differing behavior of *different* and *same* cross-linguistically. In that section we also offer an account of the sentence-internal reading of *different* that is built on using the strategy in which all variables are bound already in the lexical meaning. To account for the sentence-internal reading of *different* and *same* where binding happens outside of the lexicon, we will have to switch to dynamic semantics with the set of variable assignments. I do so in Section 5.4 and show how we deal with the sentence-internal reading of *same* and *different*, following Brasoveanu (2009). Section 5.5 is the conclusion.

5.2 *Different and same*

5.2.1 Introduction

Like *the others*, *different* and *same* can be anaphoric to a referent introduced in previous discourse but it can also be bound within a clause and express covariation (or lack of covariation) with some plural argument. The first use is represented in (1). (1a) shows the reading where *same* is anaphoric to *Deaf Sentence*. Obviously, *same* expresses that the book that Hilary started reading is identical to *Deaf Sentence*. In (1b) *different* is anaphoric to *Deaf Sentence* and expresses that the book I started reading today is distinct.

- (1) Yesterday, I finished reading Deaf Sentence.
 a. Today, Hilary started reading the same book.
 b. Today, I started reading a different book.

The second use is shown in (2).

- (2) a. Each boy read the same book.
 b. Each boy read a different book.

In this case, *same* and *different* are anaphoric to the referent that the object introduces, that is, the books that covary with boys. *Same* states that the covarying books are identical (so, in fact, no covariation takes place) and *different* states that they are distinct. I call the first reading, exemplified in (1), discourse-anaphoric reading. The reading exemplified in (2) is what Carlson (1987) calls the sentence-internal reading and I am going to follow this convention. Furthermore, I am going to use one term for *different* and *same*, DS-expressions.

5.2.2 *Different*

As already noticed in Carlson (1987), the sentence-internal reading of DS-expressions varies depending on type of the plural argument. For example, plural *different* gives rise to the sentence-internal reading with the DP headed by *all* while singular *different* is degraded. The situation reverses with distributive quantifiers. Notice that *same* can give rise to the sentence-internal reading in both cases.

- (3) a. All the men are from different towns/?? a different town.
 b. Each man is from a different town/?? different towns.
 c. All the men are/each man is from the same town.

Since Beck (2000) the focus in studies of *same* and *different* in semantics shifted towards their properties cross-linguistically (see also Toven and van Peteghem 2002; Matushansky 2008; Brasoveanu 2008b). As Beck (2000) shows, singular and plural *different* in English correspond to two lexical items in German: *verschieden* and *andere*. According to Beck, the first one can give rise to the sentence-internal reading when the plural DP is non-distributive, and in that case it must be part of a plural DP,

see (4a). The second one can be in singular and can give rise to the sentence-internal reading only with distributive quantifiers, see (4b).

- (4) a. Detmar und Kordula wohnen in verschiedenen Städten.
Detmar and Kordula live in different cities
- b. Jedes Mädchen hat ein anderes Buch gelesen.
every girl has a different book read

Brasoveanu (2008b) collects data from 11 languages and comes to the conclusion that there are three distinguishable strategies for a lexical item expressing difference: first, such a lexical item can give rise to the sentence-internal reading under distributive quantifiers and it can also give rise to the discourse-anaphoric reading. This holds, for example, for the German *andere* or the English singular *different*. Second, a lexical item can only give rise to the discourse-anaphoric reading. This holds, for example, for the English *other/another* or the French *autre*¹ or the Russian *drugoe*. Finally, a lexical item can have the sentence-internal reading with non-distributive quantifiers. In that case it modifies a plural DP, as the German *verschieden* or the English plural *different*.

The Dutch data discussed in Chapter 2 partly support Brasoveanu’s summary, but they add a few twists to it. The first addition concerns the Dutch *ander*. This can give rise to the sentence-internal reading with distributive quantifiers and is thus parallel to the German *andere* and English singular *different*. In fact, plural DPs that can license the sentence-internal reading of *ander* are ordered on the following scale ($\alpha < \beta$ ’ should be read as α is preferred over β as the subject of a predicate that includes *ander* DP in a sentence-internal reading. DQ-DPs=DPs with universal distributive quantifiers, CQ-DPs=DPs with counting quantifiers, G-DPs=group-denoting DPs, see (19) in Chapter 2):

Table 5.1: The sentence-internal reading of *ander*

DQ-DPs (not negative quantifiers)	<	CQ-DPs	<	G-DPs	<	neg. quantifiers
-----------------------------------	---	--------	---	-------	---	------------------

The scale in Table 5.1 should be familiar from Chapters 3 and 4. Ignoring negative quantifiers (‘no NP’) it is identical to the scale of acceptability of distributive readings and the sentence-internal reading of *the others*. I take it that these data point to the conclusion that the sentence-internal reading of *ander* should require distributivity (and

¹The situation with *autre* is more complicated. It can also give rise to the sentence-internal reading but only under distributive quantifiers whose members form a total order (antisymmetric, transitive, total). For example, the first example is degraded under the sentence-internal reading of *autre* but the second one can give rise to the relevant reading since temporal ordering satisfies the condition (for more discussion, see Tovenà and van Peteghem 2002).

- (1) ?? Chaque fille a lu un autre livre
each girl has read a *autre* book
- (2) Il sort chaque soir avec une autre copine
He goes-out every night with a different girl

something extra needs to be said about negative quantifiers). Following Brasoveanu (2009) I will show in Section 5.4 how we can capture the need of *ander* for distributivity. Probably, in this respect *ander* is similar to lexical items expressing difference in other languages, like the German *ander* or the English *different* or the Czech *jiný* since all these give rise to the sentence-internal reading preferably with distributive quantifiers and are degraded with CQ-DPs and even more with G-DPs. The degraded status of the sentence-internal reading with G-DPs has been noticed in literature, see, for example, Carlson 1987, Brasoveanu 2009, for English, Beck 2000 for German, Toven and van Peteghem 2002 for the French *autre*. The behavior of the Czech *jiný* is based on my own intuition of the difference between the following two examples, where the second one is somewhat degraded under the sentence-internal reading of ‘different’.

- (5) a. Každý žák přečetl jinou knihu.
 every boy read a different book
 b. ?Žáci přečetli jinou knihu.
 boys read a different book

I will call all expressions of *different* whose sentence-internal reading is degraded along the scale of Table 5.1 *different*₁.

The second expression tested in the questionnaire was *verschillend*. The acceptability of its sentence-internal readings is summarized in Table 5.2.²

Table 5.2: The sentence-internal reading of *verschillend*

CQ-DPs	≈	G-DPs	<	DQ-DPs
--------	---	-------	---	--------

This confirms Beck’s/Brasoveanu’s observation that it is common across languages to have one item expressing difference which has a sentence-internal reading only with non-distributive quantifiers. However, there is something what goes against their expectations: both Beck and Brasoveanu claimed that such lexical items must modify a plural noun. But we focused on *verschillend* modifying a noun that was singular (two examples are given in (6); for more discussion and contexts see Chapter 2). Contrary to what Beck and Brasoveanu predict the sentence-internal reading seemed possible: its acceptability with *all DP* or coordination was comparable to the sentence-internal reading of *ander* with distributive quantifiers. See Chapter 2 for exact numbers.

- (6) a. De steden in het noorden hebben een verschillende lengte-eenheid.
 the towns in the north have a verschillend length-unit
 b. Jan, Kees en Wim hebben een verschillend schilderij gekozen.
 Jan, Kees and Wim have a verschillend picture chosen

²I should caution the reader that the scale in Table 5.2 is a simplification: it does not reflect the fact that definite plurals were more degraded than coordinations. I claimed in Chapter 2 that the difference between definite plurals and coordinations is due to strategy preferences in referring and thus is irrelevant for the actual acceptability of the sentence-internal reading of *verschillend*. For this reason I omit this difference here. For the discussion why I claimed this see Section 2.4.

The existence of *different* for which the sentence-internal reading is fully accepted with CQ-DPs or G-DPs and can modify singular nouns is probably not a quirk of Dutch. citetovena+02 mention that the same holds for the French post-nominal *différent*:

- (7) Jean et Lea ont chanté la même chanson mais ils ont fait un
 John and Lea are sung the same song but they are drawn a
 dessin différent
 drawing different
 ‘John and Lea sang the same song but drew different drawings.’

Similarly, in Czech, a quick search in Czech Corpus reveals that *různý* shows the same properties: it has a sentence-internal reading with non-distributive DPs, yet it modifies singular nouns.

- (8) Pavel a Honza na tohle mají různý názor.
 Pavel and Honza on this have different opinion
 ‘Pavel and Honza have different opinions on this.’

All these expressions can also modify a plural noun to express the sentence-internal reading. I focus here on Czech (9a) and Dutch (9b) (see Tovena and van Peteghem (2002) for parallel French examples with *différent*).

- (9) a. Pavel a Honza mají různá auta.
 Pavel and Honza have different cars.
 b. Jan, Kees en Wim hebben verschillende schilderijen gekozen.
 Jan, Kees and Wim have verschillend pictures chosen

They can also give rise to another reading, not discussed before. This is one where no plural DP is present that could license the sentence-internal reading of *different* yet *different* is not discourse anaphoric. The English *different* can function in this role, see (10).

- (10) John went to different places on his shopping trip.

This reading is characterized as synonymous with *various* by Carlson (1987) and called a reciprocal reading by Beck (2000). I am going to call it *various* reading. Notice that this reading is possible with *různý* ‘different’ but only when it modifies a plural noun:

- (11) a. Honza šel do různých prodejen během nákupu.
 Honza went to different_{pl} shop_{pl} while shopping
 b. *Honza šel do různé prodejny během nákupu.
 Honza went to different_{sg} shop_{sg} while shopping
- (12) a. Deze stad heeft verschillende munt eenheden
 this town has different_{pl} coin unit_{pl}
 ‘This town has different units of currency.’

- b. * Deze stad heeft een verschillende munt eenheid
 this town has different_{sg} coin unit_{sg}
 ‘This town has a different unit of currency.’

I will call the type of expression exemplified by *verschillend* or the Czech *různý* *different*₂.

5.2.3 *Same*

Compared to *different*, *same* seems to be quite a dull experience. It can give rise to the sentence-internal reading with DQ-DPs, CQ-DPs and G-DPs:

- (13) a. The men are from the same town.
 b. All the men are from the same town.
 c. Each man is from the same town.

Unlike *ander* and *verschillend* the sentence-internal reading is also possible with negative quantifiers:

- (14) No boy read the same book.

Finally, it is also acceptable with discourse-anaphoric reading, as we have seen in the example (1).

Besides, as far as I know there is no cross-linguistic variation parallel to *different*₁ and *different*₂. That is, *same* does not split into more expressions where each one could have a sentence-internal reading only with the subset of plural DPs, as we have seen it with the cognates of *different* in languages other than English. Consider, for example, Dutch. The lexical expression for sameness, *dezelfde*, can have the sentence-internal reading with definite plurals, as established in the questionnaire from Chapter 2. It can also have the sentence-internal reading with distributive quantifiers (we did not test this but the data seem uncontroversial):

- (15) Ieder jonge heeft hetzelfde boek gekozen.
 every boy has the-same book chosen

This is similar to Czech:

- (16) a. Každý chlapec si vybral stejnou knihu.
 every boy refl chose same book
 b. Chlapci si vybrali stejnou knihu.
 boys refl chose same book

To be precise, it is common that there is more than one expression for sameness. One example could be *identical* and *same* in English. But these are not differentiated on the ground of what type of plural DP is necessary to license their sentence-internal reading.

5.2.4 Conclusion

We have discussed the properties of *same* and *different*. We found out the following:

- cross-linguistically it is common that there are two expressions of *different*, where *different*₁ has the sentence-internal reading preferably with distributive quantifiers and can have a discourse-anaphoric reading and *different*₂ has the sentence-internal reading with non-distributive quantifiers and lacks a discourse-anaphoric reading
- *different*₂ can always modify a plural noun but it can also modify a singular noun in some languages (Czech, Dutch, French)
- *same* is cross-linguistically expressed by one expression which can give rise to the sentence-internal reading with distributive, as well as non-distributive quantifiers and can give rise to the discourse-anaphoric reading

In the next section we will see how this behavior follows from the two strategies of anaphoricity that we discussed in the previous chapter.

5.3 Two strategies

5.3.1 Introduction

In this section, I discuss two strategies to analyze DS-expressions: the R-strategy and A-strategy. To recapitulate, the R-strategy builds in multiplication of resources into the interpretation of anaphors, and requires QR of anaphors. The A-strategy builds in multiplication of resource into syntax: it assumes a binding mechanism which can bind variables in anaphors. The first strategy has been used for *each other* and the second strategy has been used for *the others* in Chapter 4 where I also discuss in more detail the virtues of the two strategies.

It turns out that both strategies have already been proposed to account for DS-expressions. Barker (2007) assumes the R-strategy for *same* and *different*. Brasoveanu (2009) assumes the A-strategy for *different*₁ and one reading of *same*. In this Section I am going to argue that both approaches are correct for some DS-expressions. I show that Barker was right about the assumed strategy but wrong about lexical items that he applied it to: the R-strategy should be restricted to *different*₂. The A-strategy should be used for *different*₁ and *same*. This part is similar to Brasoveanu's account even though I differ from his approach in assuming that *same* is unambiguous. This is preferable since we have seen in the previous section that even in languages which differentiate lexically between *different*₁ and *different*₂ only one lexical item is used to express sameness.

5.3.2 R-strategy

Barker (2007) offers a fully compositional account of *same* and as he suggests, the same account should be extended to *different*. That a compositional account is pos-

sible might seem somewhat surprising in the light of Keenan's paper (Keenan, 1992) which includes a proof that there is no compositional analysis based solely on generalized quantifiers that can express the sentence-internal reading of *same* (as well as other expressions, like *different*, negative quantifiers in negative concord languages etc.) However, as Barker stresses, this does not mean that a compositional analysis is impossible, it just means that we cannot express the meaning of *same* purely in terms of generalized quantifiers. This is not that surprising since DS-expressions behave syntactically as adjectives after all. Making use of this fact, Barker argues that DS-expressions are best treated as quantificational adjectives. The denotation he assumes for *same* is somewhat closely related to Dowty's (Dowty, 1985) and I will start with that one since it is simpler and hopefully it will make Barker's own analysis more transparent.

Dowty (1985) presents an attempt to give a unified analysis to the sentence-internal and discourse-anaphoric reading of *same*. In Dowty's account *same* has the following denotation:³

- (17) Interpretation of *same* in Dowty (1985) *pace* Barker (2007)
 $[[\text{same}]] = \lambda Q_{\langle e, t \rangle} \lambda x. (\exists f_{\text{CH}})(\{x\} = f(Q) \wedge (\forall z' < z)(Rz'x))$

Same is of a type of an adjective $(\langle e, t \rangle, \langle e, t \rangle)$. However, it consists of two variables that need to be sentence-internal or contextually specified. The first one is z , which is called a comparison class in Dowty. The second one is the relation R . R relates each entity in the comparison class to x . What is x ? We know that $\{x\} = f(Q)$, where f is a choice function, albeit of a somewhat unusual type: $\langle \langle e, t \rangle, \langle e, t \rangle \rangle$. It is a function that takes a set of entities as its argument and gives us a set with one member as its value. The argument of f is the set denoted by the noun which *same* modifies.

Consider the following example, with the sentence-internal reading of *same*:

- (18) Morris and Philip read the same book.

Here, $c = \text{Morris and Philip}$, and R is reading. The interpretation we get is in (19): there is one book such that each of Morris and Philip read it.

- (19) $(\exists f)(\{x\} = f(\text{BOOK}) \wedge (\forall z' < \text{M} \oplus \text{P})(\text{READ}(z')(x)))$

Dowty extends his account to discourse-anaphoric readings. An example of this reading is in (20).

- (20) Morris and Philip reviewed *Deaf Sentence*. Hilary saw the same book in a shop.

We can get the correct interpretation if we let $z = \text{Morris and Philip}$ and $R = \text{reviewing}$.

As Barker discusses at length, leaving R and z to be contextually specified is quite problematic and overgenerates. For example, picking Morris and Philip as z and 'see in a shop' as R in (20) would lead to the reading 'Hilary saw the same book in a shop

³The denotation here is as presented in Barker (2007) which is truth-conditionally identical to Dowty's but differs in technical details. I prefer Barker's presentation since it will facilitate moving to his own analysis.

that Morris and Philip saw’. But there is no such reading. The same problem arises if we consider just the sentence-internal reading of *same*. In (21) it should be possible to pick Morris and Philip as *z* and ‘read’ as *R*. That would lead to the interpretation ‘Morris wanted to read the same book that he and Philip read and so did Philip’. But there is no such reading.

(21) Morris and Philip wanted to read the same book.

(21) shows that *R* cannot be chosen freely in the sentence-internal reading of *same*. Rather, it corresponds to the part of the clause that is left after *z*, the comparison class, is subtracted.

The denotation that Barker (2007) assumes for *same* is similar to (17) but with all variables bound. The semantics is given in (22):

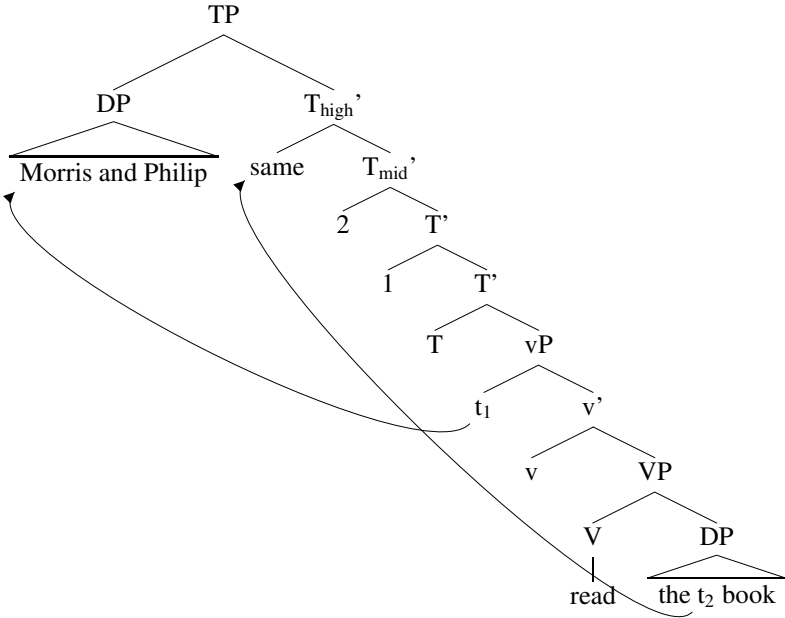
(22) Interpretation of *same*, after Barker (2007)

$$[[\text{same}]] = \lambda F_{\langle \langle e,t \rangle, \langle e,t \rangle \rangle, \langle e,t \rangle} \lambda z. |AT(z)| \geq 2 \wedge (\exists f)(\forall z' \leq z \wedge z' \in AT)(Fz'f)$$

In contrast to Dowty’s account, (22) is built up only for the sentence-internal reading. *F* corresponds to the part of the clause left after *z* and *same* are subtracted. *z* is the comparison class, and it is the plurality that is necessary to yield the sentence-internal reading of *same* (Morris and Philip in (21)).

How do we arrive at *F*? This is where it becomes crucial that *same* is considered a quantificational adjective. We assume that it undergoes QR from its surface position and scopes over the whole clause but under *z*. To visualize this, (23b) is the LF representation for (23a).

(23) a. Morris and Philip read the same book.
b.



Same undergoes QR and scopes over the predicate to which the subject would apply. This is a QR as we know it from previous chapters, the only difference is that since we deal with a quantificational adjective, t_2 , the trace left behind *same* is not of type e (entity). It must be of the type of adjective, i.e., $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$. Similarly, 2 is not a predicate abstraction, but abstraction over the type $\langle e, t \rangle, \langle e, t \rangle$ otherwise t_2 would not be bound. Thus, the interpretation of T'_{mid} is:

$$(24) \quad [[T'_{\text{mid}}]] = \lambda Q_{\langle e, t \rangle, \langle e, t \rangle} \lambda x. (\exists e) (C_{*}\Theta_1(x)(e) \wedge C_{*}\Theta_2(\sigma(\lambda y. Q(\text{BOOK})(y)))(e) \wedge *READ(e))$$

Notice that I assumed that the event closure applies below T'_{mid} . This makes the interpretation easier since we do not need to deal with event argument when discussing the semantics of *same*.

(24) is an argument of *same*. *Same* can apply to (24) since (24) is of the right type, $\langle\langle e, t \rangle, \langle e, t \rangle\rangle, \langle e, t \rangle$. The result of this application is given in (25):

$$(25) \quad [[T'_{\text{high}}]] = \lambda z. |AT(z)| \geq 2 \wedge (\exists f)(\forall z' \leq z \wedge z' \in AT)(\exists e)(C_{*}\Theta_1(z')(e) \wedge C_{*}\Theta_2(\sigma(\lambda y. f(\text{BOOK})(y)))(e) \wedge *READ(e))$$

(25), in words, says that there is a choice function and for each z' where z' is a proper part of z , it holds that z' read some entity that is a book and that is in one-membered set picked by the choice function from the set of books. Since the set has only one member and is a book, it must be so that each all z' in z read the same book. Supplying ‘Morris and Philip’ for z gives us the interpretation in (26). This correctly expresses that Morris read the same book as Philip did.

$$(26) \quad [[TP]] = |AT(\text{MORRIS} \oplus \text{PHILIP})| \geq 2 \wedge (\exists f)(\forall z' \leq \text{MORRIS} \oplus \text{PHILIP} \wedge z' \in AT)(\exists e)(C_{*}\Theta_1(z')(e) \wedge C_{*}\Theta_2(\sigma(\lambda y. f(\text{BOOK})(y)))(e) \wedge *READ(e))$$

Notice that the way *same* is set up is parallel to the R-strategy of reciprocals, and the only difference is that where reciprocals were relating two arguments, *same* relates an argument and a noun modifier (see the simplified interpretation of *each other* in (27) that brings out the similarity). The difference between the two should not be surprising since *each other* is occupying an argument position while *same* is an adjective.

$$(27) \quad \text{A simplified interpretation of } \textit{each other} \\ \lambda R_{\langle e, \langle e, t \rangle \rangle} \lambda x. (\forall y, z \leq x \wedge y, z \in AT)(y \neq z \rightarrow Ryz)$$

Let me pursue the parallel a bit more. In a reciprocal sentence like *The boys hated each other* we need to ensure that the arguments of Θ_1 and Θ differ. That is, we need to ensure that the relation $\lambda x \lambda y. (\exists e)(C_{*}\Theta_1(x)(e) \wedge C_{*}\Theta_2(y)(e) \wedge *READ(e))$ is true for distinct x and y . Furthermore, both x and y have to be values from the plurality *the boys*. We achieve this by letting *each other* scope over the whole relation and putting the distinctness condition into its lexical semantics. In *the boys read the same book* we need to ensure that the relation $\lambda Q \lambda x. (\exists e)(C_{*}\Theta_1(x)(e) \wedge C_{*}\Theta_2(\sigma(\lambda y. Q(\text{BOOK})(y)))(e) \wedge *READ(e))$ is true for one Q and various x . We achieve this again by letting *same* scope over the relation and putting the requirement on identical Q for various x in its lexical semantics. More concretely, the last is done by introducing the existential

quantification over choice functions which scopes above the universal quantification over proper parts of some plural argument. Because of the scope order we ensure that Q is the same for all x .

Thus, I consider Barker's approach as just another instantiation of the R-strategy. This is so since we achieve retrieving the value of the books that the other boys read in (28) by letting *same* scope over the relation and restricting it in particular ways.

- (28) The boys read the same book.

In the next section, I am going to argue that Barker is right in establishing this strategy but wrong in applying it to *same*. Contrary to Barker, I believe that the R-strategy is restricted to *different*₂.

R-strategy does not work for *same* and *different*₁ but it works for *verschillend*₂

One of the hallmarks of the R-approach to reciprocals is that the denotation of *each other* involves a distributive quantification. This is also true for Barker's *same*. Since the distributive quantification is part of the meaning of *same* we expect *same* to be incompatible in its sentence-internal reading with distributive quantifiers. Obviously, this prediction is wrong:

- (29) Each boy read the same book.

(29) would get the interpretation in (30). This is uninterpretable since *same* needs to distribute but atoms have no proper parts.

- (30) $(\forall x \in \text{BOY} \wedge x \in AT)(|AT(x)| \geq 2 \wedge$
 $\wedge (\exists f)(\forall x' \leq x \wedge x' \in AT)(\text{read}(x')(\sigma(\lambda y.f(\text{BOOK})(y))))$

Barker is aware of this problem and suggests a tentative solution. We could assume a different interpretation for distributive quantifiers. Rather than having a distributive quantification as part of their meaning, distributive quantifiers would denote a plurality and require that there is an independently established distribution to atoms over the plurality in the clause. In (29) the distributive quantification would be the one introduced by *same* so we derive the correct truth conditions.

I doubt that this solution could work. If *each* did not introduce a distributive quantification on its own, it could appear in sentences in which a distributive quantification is brought about independently of the quantifier *each*, for example, by introducing a floating quantifier, or binominal *each* or a reciprocal. These predictions are all wrong:

- (31) a. * Each boy each read a book.
 b. * Each boy read a book each.
 c. * Each boy talked to each other.

The situation, I believe, is parallel to the one we observed with reciprocated anaphora in Chapter 4. We have seen that reciprocation is expressed by two expressions, one of which (*the others*) prefers distributive quantifiers for its sentence-internal reading, while the other one is incompatible with distributive quantifiers. In that chapter we

have made use of the fact that the R-strategy needs to include distributive quantification as part of the lexical meaning of reciprocated anaphora to derive its incompatibility with distributive quantifiers. The problem for Barker is that he is using the same strategy (R-strategy) but now for an expression that *is* compatible with distributive quantifiers. This turns the strength of the R-approach to its weakness.

But there is no need to despair. Barker's semantics does not work for *same* but it might work for other DS-expressions. In fact, Barker suggests that the same strategy might account for the sentence-internal reading of *different*. The interpretation of *different* is given here:

- (32) Interpretation of *different*, Barker (2007)
 $[[\text{different}]] = \lambda F \lambda z. |AT(z)| \geq 2 \wedge (\forall f)(\forall z', z'' \leq z \wedge z', z'' \in AT)((Fz'f \wedge Fz''f) \rightarrow z' = z'')$

(32) cannot deal with *different*_I because *different*_I, as discussed in Section 5.2 requires distributive quantifiers for its sentence-internal reading while (32) predicts that the sentence-internal reading of *different* is incompatible with distributive quantifiers, for the same reason that *same* should be. Barker suggests that (32) might be correct for the plural *different*. For example, it might account for (33):

- (33) Morris and Philip read different books.

I do not think that this is correct. The interpretation we would get for (33) is that Morris read two or more books and Philip read two or more books that are different from what Morris read. However, (33) is also compatible with the reading that Morris read one book and Philip read another book and this reading cannot be captured by (32). The problem is that the plurality of the noun phrase *the different books* is interpreted in scope of a distributive quantification over Morris and Philip while it should be interpreted outside of it.

The problem points to the correct analysis of (33): (33) is a case of a cumulative reading. But cumulative quantification is incompatible with the denotation of *different* which Barker assumes so we have to change it. In fact, Beck (2000) already argued for the cumulative analysis of (33) and other cases so I will follow her here and assume that *different* is interpreted as:

- (34) Interpretation of *different* in (33), following Beck (2000)
 $[[\text{different}]] = \lambda P \lambda x. |AT(x)| \geq 2 \wedge (\forall y, z < x)(y, z \in AT \wedge y \neq z \rightarrow \text{distinct}(y)(z))$

Beck assumes that the distinctness condition in (34) is specified as one of the following:

- (35) Beck's definition of distinctness (to be modified):
 y is distinct from z iff (a) or (b)
 a. $y \neq z$
 b. y and z belong to kinds y' and z' and $y' \neq z'$

The (a) condition expresses a token difference. The second condition should express a type difference. However, this condition on distinctness will not work. Consider the sentence *My book is distinct from your book*. According to (35) this would be true if we shared one book. Suppose the book was *Deaf Sentence*. As it happens, this book belongs to many kinds that are different from each other. For example, it belongs to the kind BOOK and to the kind OBJECT. Thus, sharing one and the same book would still satisfy the condition (35b), contrary to one's intuition. Thus, instead of non-identity, we should require non-overlap over kinds (assuming that kinds are organized in lattice structures). This gives us (36):

- (36) The definition of distinctness, the final version
 y is distinct from z iff (a) or (b)
 a. $\neg y \circ z$
 b. y and z belong to kinds y' and z' and $\neg y' \circ z'$

Employing (34) for (33) would give us the following interpretation:

- (37) $(\exists e)(\exists x)(C_*\Theta_1(M \oplus P)(e) \wedge *READ(e) \wedge C_*\Theta_2(x)(e) \wedge |AT(x)| \geq 2 \wedge (\forall y, z < x)(y, z \in AT \wedge y \neq z \rightarrow \text{distinct}(y)(z)))$

The formula is true if Morris read one book and Philip read another book, which is what we want. Notice that the condition in the scope of the universal quantifier is always satisfied since $y \neq z$ then $\text{distinct}(y, z)$ must be true because of the condition (35a) of *distinct*. Thus, (33) does not seem to say anything else than *Morris and Philip read books* does. We might be avoiding the trivial contribution of *different* by requiring that the condition (35b) holds. There are clear cases where *different* does play a role. For example, in the following pair, from Krifka (1990), *different* blocks the interpretation in which ships are differentiated solely on stage-level and requires individual-level distinction.

- (38) a. 4,000 ships passed through the lock last year.
 b. 4,000 different ships passed through the lock last year.

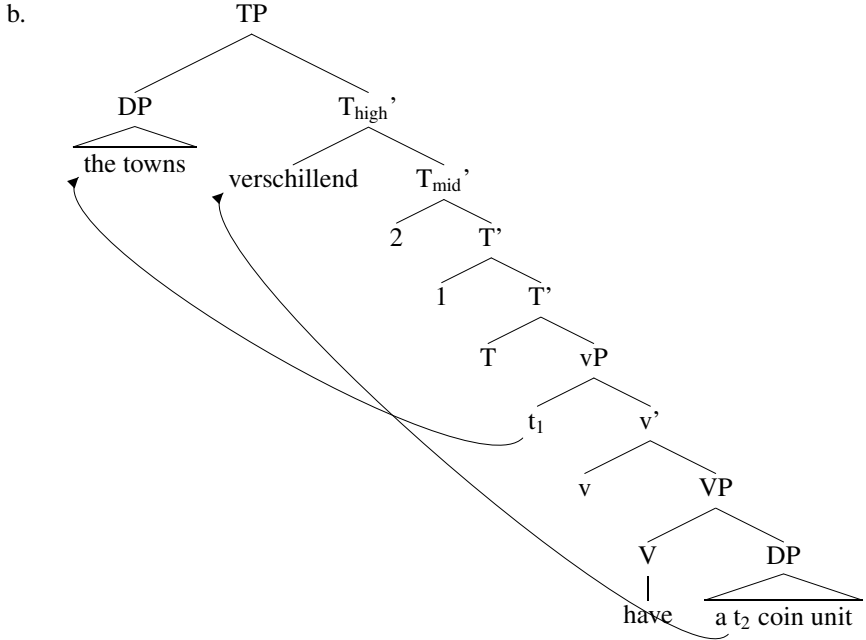
I will ignore this effect here. For more discussion of this, see Barker (2007).

However, Beck's analysis cannot be extended to *verschillend* and its cognates in other languages. The properties of *verschillend* were discussed in Section 5.2, where it was labeled as *different*₂. The problem is that *different* in Beck's analysis requires that the noun phrase it modifies refers to at least two objects. However, *different*₂ can modify singular noun phrases, see (6). This is a lexical item for which Barker's semantics (32) works perfectly. Before showing this, I modify Barker's semantics by combining it with the relation *distinct* as assumed in Beck's work:

- (39) Interpretation of *different*₂, the final version
 $[[\text{different}]] = \lambda F \lambda z. |AT(z)| \geq 2 \wedge (\forall f)(\forall z', z'' \leq z \wedge z', z'' \in AT)((Fz'f \wedge Fz''f) \rightarrow \neg \text{distinct}(z')(z''))$

Consider (40a), repeated from above, which gets the LF structure in (40b).

- (40) a. De steden hebben een verschillende munt eenheid.
 the towns have a verschillend coin unit



Using (39) as the semantics of *verschillend* and assuming the LF structure in (40b) we arrive at the reading (41).

- $$(41) \quad |AT(\text{THE TOWNS})| \geq 2 \wedge (\forall f)(\forall z', z'' \leq \text{THE TOWNS} \wedge z', z'' \in AT) \left(\begin{array}{l} (\exists e)(C_{*}\Theta_1(z')(e) \wedge C_{*}\Theta_2(A(f(\text{CURRENCY}))))(e) \wedge *HAVE(e) \wedge \\ (\exists e)(C_{*}\Theta_1(z'')(e) \wedge C_{*}\Theta_2(A(f(\text{CURRENCY}))))(e) \wedge *HAVE(e) \end{array} \right) \rightarrow \neg \text{distinct}(z')(z'')$$

(41) says that for every currency if two towns have this currency they cannot be distinct. This is true if each town has a different currency.⁴

In contrast to the Barker's analysis of *same* and the English *different*, this analysis does not run into problems with distributive quantifiers. Instead, it gains further support from them. The sentence-internal reading of *verschillend* is degraded with distributive quantifiers (see Table 5.2) which is what we expect. The reason is that *different*₂ includes a distributive quantification on its own. This means that the R-strategy plays a role in DS-expressions, albeit not for the ones that Barker (2007) considered.

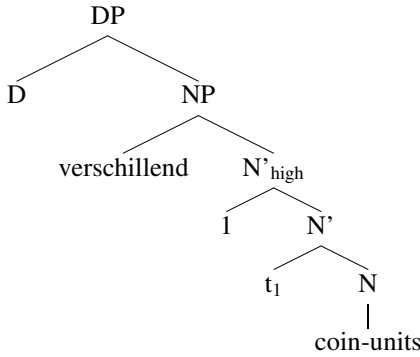
⁴(41) is also true if the towns have no currency at all. This reading can be excluded if we assume that *verschillend* in (40a) expresses (41) and presupposes that each town has a currency. There is some evidence that the latter might be accommodated in the discourse. For example, the nuclear stress does not fall on the noun but on the adjective *verschillend* which would make sense if the existence of currency in the context was given (see Schwarzschild 1999).

As noted in Section 5.2 *verschillend* and its cognates in other languages can also modify a plural noun phrase, in which case it leads to the *various* reading. Consider (42).

- (42) De steden hebben verschillende muunt eenheden.
the towns have different₂ coin unit

This sentence has the reading parallel to (40a), that is, each town has a different currency. Does it mean that *verschillend* is ambiguous between Barker's and Beck's analysis? This is unlikely, since the same ambiguity would have to appear in the French *différent* and the Czech *různý*. But as it turns out Barker's analysis can capture both uses of *verschillend*. We have seen the first one in (41). To interpret (42) correctly it suffices that *verschillend* takes a trivial scope (within the noun phrase which it modifies):

- (43) The tree structure of the DP *verschillende muunt eenheden* in (42)



The interpretation is given in (44). I present the interpretation of each subtree in the DP.

- (44) a. $[[N]] = \lambda y. \text{CURRENCY}(y)$
 b. $[[N']] = Q_1(\lambda y. \text{CURRENCY}(y))$
 c. $[[N'_{\text{high}}]] = \lambda Q. Q(\lambda y. \text{CURRENCY}(y))$
 d. $[[NP]] = \lambda x. (\exists x) |AT(x)| \geq 2 \wedge (\forall f)(\forall z', z'' \leq x \wedge z', z'' \in AT) ((f(\lambda y. \text{CURRENCY}(y))(z') \wedge f(\lambda y. \text{CURRENCY}(y))(z'')) \rightarrow \neg \text{distinct}(z')(z''))$
 e. $[[DP]] = \lambda P. (\exists x) |AT(x)| \geq 2 \wedge (\forall f)(\forall z', z'' \leq x \wedge z', z'' \in AT) ((f(\lambda y. \text{CURRENCY}(y))(z') \wedge f(\lambda y. \text{CURRENCY}(y))(z'')) \rightarrow \neg \text{distinct}(z')(z'')) \wedge Px$

When we plug in (44d) in the cumulative interpretation of the sentence (42) we derive the following truth conditions:

$$(45) \quad (\exists e, x)(C_{*\Theta_1}(\text{THE TOWNS})(e) \wedge *HAVE(e) \wedge C_{*\Theta_2}(x)(e) \wedge \\ |AT(x)| \geq 2 \wedge (\forall f)(\forall z', z'' \leq x \wedge z', z'' \in AT) \\ ((f(\lambda y.CURRENCY(y))(z') \wedge f(\lambda y.CURRENCY(y))(z'')) \rightarrow \neg \text{distinct}(z')(z''))))$$

This is true if each town has a different currency. It can also be true in case each town has more than one currency systems which differ from each other. This is the various reading discussed in Section 5.2 (see (12a)), which is possible with *verschillend* and *různý*.

To conclude, by using the R-strategy we could correctly derive that *different*₂ has the sentence-internal reading which is degraded with distributive quantifiers. Furthermore, we can also derive that it can give rise to the various reading. The latter reading is crucial for deriving what looks like the sentence-internal reading in (42) or the English (33). In fact, we can dispense with the semantics of *different* that Beck assumes since it can be derived by (39) when *different* scopes within the DP. This is summarized in Table 5.3.

Table 5.3: *Different* and *same* cross-linguistically

DS-expressions	Interpretation	Can it move outside of DP?
<i>verschillend, différent</i> (French), <i>různý</i>	(39)	yes
<i>verschieden, different</i> (English)	(39)	no
<i>ander, same</i>	??	-

Thus, the difference between the German *verschieden*, English *different*, on the one hand and the Dutch *verschillend* and the Czech *různý* boils down to the difference in scope. Since the latter can scope outside of its DP, at a clause level, we correctly derive its possibility to have the sentence-internal reading when being in singular. This is excluded for *verschieden*, as Beck (2000) notes. Finally, we still cannot account for *ander*, labeled throughout as *different*₁, and *same*. The R-strategy is unsuitable for these because they give rise to the sentence-internal reading with distributive quantifiers. Thus, we need to make use of the other strategy.

5.3.3 A-strategy

In the A-strategy we assume that some variables are free and its value is supplied by the assignment function. For *same* and *different*₁ this would mean that they have the following interpretation:

$$(46) \quad \text{Interpretation of } \textit{same}, \text{ A-approach} \\ [[\textit{same}]] = \lambda P \lambda x. Px \wedge x = y$$

$$(47) \quad \text{Interpretation of } \textit{different}_1, \text{ A-approach} \\ [[\textit{same}]] = \lambda P \lambda x. Px \wedge \neg x \circ y$$

y is free in these formulas. To get the sentence-internal reading right for (48) we need y to be the books that the other boys read.

- (48) Each boy read the same/a different book.

Brasoveanu (2009) recently showed how we might be able to retrieve ‘the books that the other boys read’ in (48) to get the correct sentence-internal reading of *different* and *same*. However, the analysis requires a shift to interpretation language with *sets* of assignments, where furthermore, assignments are in the object language so we can abstract and quantify over them. In Section 5.4, I switch to Compositional DRT with Pluralities to demonstrate how, following Brasoveanu, this enables us to deal correctly with *same* and *different*₁ in the A-strategy.

5.3.4 Consequences of the analysis

Suppose we can correctly derive the way *y* is retrieved in *same* and *different*₁, which would enable us to capture their sentence-internal reading in the A-strategy. Since we also can capture the interpretation of *different*₂ we can at that point answer why

- the sentence-internal reading of *different*₂ is degraded with distributive quantifiers and accepted with non-distributive quantifiers (see Table 5.2) while
- the sentence-internal reading of *different*₁ is accepted with distributive quantifiers and degraded with CQ-DPs and even more degraded with G-DPs (see Table 5.1) while
- the sentence-internal reading of *same* is accepted with both distributive and non-distributive quantifiers

Why the sentence-internal reading of *different*₂ is degraded with distributive quantifiers has been discussed in the previous section. The reason is that using the R-strategy forces us to let *different*₂ distribute over a plural argument and this is incompatible with another distribution induced by a distributive quantifier. We also derive that the sentence-internal reading of *different*₂ is not degraded with CQ-DPs or G-DPs. This follows for two reasons. As we have discussed in Chapter 3 based on the data from sentences with plural arguments the distributive reading with covariation is a marked option and should be not chosen for unmarked expressions like G-DPs. There are two possible reasons why the sentence-internal reading of *different*₂ is not degraded with G-DPs. First, this is because the reading is not really the distributive reading with covariation. We defined that reading as a case in which a witness of a generalized quantifier varies in scope of a distributive operator. This is not satisfied here for the simple reason that *different*₂ is not a generalized quantifier. Alternatively, even if we changed the definition of the distributive reading with covariation to subsume the sentence-internal reading of *different*₂ we would not derive that the reading should be degraded. This is because the sentence-internal reading is the only possible option and thus its costly status becomes irrelevant because simply there is no non-costly reading. In the terminology we used in Chapter 3 the interpretation that the hearer chooses must be one of the possible interpretations given the lexical meaning of expressions involved. Since the sentence-internal reading of *different*₂ is the only option the hearer has no other choice but to interpret it that way.

We also derive why the sentence-internal reading of *different*₁ is best accepted with distributive quantifiers, somewhat degraded with CQ-DPs and more degraded with G-DPs. This is because the sentence-internal reading is a case of the distributive reading with covariation. In this case the generalized quantifier which involves the adjective *different* covaries so the resulting reading should be marked. Unlike in case of *different*₂, *different*₁ has an alternative option where no covariation takes place: the discourse-anaphoric reading. Given this option, the hearer should consider that interpretation somewhat preferred with CQ-DPs and even more so with G-DPs.

Finally, we expect the sentence-internal reading of *same* to be fully accepted with distributive quantifiers as well as CQ-DPs and G-DPs. This is because the sentence-internal reading induces no covariation and thus should not be dispreferred over the discourse-anaphoric reading. But this has another consequence. Since *same* induces no covariation it is accepted in every context where *different*₁ is accepted and it is also accepted in every context where *different*₂ is accepted. Suppose there was another expression, *same*₂ which would have Barker's interpretation of *same* (see (22)). Where would this expression be accepted? In fact, in the same context where *different*₂ is. But in contrast to *different*₂ there would never be a reason to use *same*₂ to signal the sentence-internal reading. This is because *same* does not induce covariation and thus the sentence-internal reading is not degraded with CQ-DPs and G-DPs, in contrast to *different*₁. This makes it less useful than *different*₂ which does appear in contexts where the sentence-internal reading of *different*₁ is degraded. This might explain why we cross-linguistically see two expressions for *different* but only one expression for *same*.

All this reasoning is based on the assumption that we can account for the sentence-internal reading using the A-strategy and the lexical semantics for *same* and *different*₁ which follows (46) and (47). The rest of the chapter offers such an analysis.

5.4 Compositional DRT with pluralities (PCDRT)

5.4.1 Introduction

To recapitulate we are in need of an account of the sentence-internal reading of *same* and *different*₁. We need to make sure that these expressions operate on the A-strategy which would enable us to capture that cross-linguistically, they can give rise to the sentence-internal, as well as discourse-anaphoric reading. Furthermore, we need to connect their sentence-internal reading with distributivity since the reading can be licensed by distributive quantifiers.

Brasoveanu demonstrates that these issues can be captured in Compositional DRT with pluralities, which is similar to other dynamic semantic systems with pluralities (Elworthy 1995; Krifka 1996; van den Berg 1996; Nouwen 2003, 2007) in the fact that unlike more traditional dynamic semantic systems (Discourse Representation Theory and its compositional variants, Compositional DRT or Discourse Predicate Logic, among many others) it operates with sets of assignments. I will make this point more precise below.

My aim is:

- to show that using the A-strategy for *same* and *different*₁ is feasible
- to explain why *same* and *different*₁ differ in the following respect: while the sentence-internal reading of *different*₁ is most accepted with distributive quantifiers and degraded with CQ-DPs and G-DPs (see Table 5.1), the sentence-internal reading of *same* is accepted with distributive quantifiers, as well as CQ-DPs and G-DPs (see Section 5.2.3)
- to explain why *same* is cross-linguistically realized as one lexical item and *different* is usually realized as two different items, called *different*₁ and *different*₂ throughout

The rest of the chapter is organized as follows. In Section 5.4.2 I introduce Compositional DRT with pluralities. In Section 5.5 I show how this system can deal with *same* and *different*. The chapter comes to an end in Section 5.6.

5.4.2 Compositional DRT with pluralities

In this section I am going to shortly introduce Compositional DRT with pluralities (abbreviated as PC DRT), as studied in Brasoveanu's work (Brasoveanu 2008a,c). To keep the system simple I will ignore the contribution of events throughout. Also, I am going to be rather brief on the discussion of the system. For more details, see the above mentioned work of Brasoveanu.

We assume three basic types. Two of them are our old friends: entities (including numbers) and truth values (type *e* and *t* respectively). Apart from these two we are also going to make use of information stacks (type *s*). Information stacks are a way of modelling partial variable assignment in which adding new information does not lead to the loss of the previous information. I use *i, j ...* for variables over information stacks.

Stacks

Information stacks are carried throughout a discourse and can be modified at various points in the discourse, and the modifications are carried on to the subsequent discourse. Stacks serve for retrieving anaphoric information. Pronouns can pick up an antecedent by choosing a position in a stack.

Information stacks can be seen as a special type of assignment function. Consider the following example of a stack:

- (49) Stack *i* =
 Morris Hilary Philip Perssy

i in this example is a stack with four positions. Treating the first position as '0' the second position as '1' etc. we can equivalently describe it as a function.⁵

⁵See Muskens (1996) for axioms we want this function to satisfy in order to behave intuitively as a total variable assignment and see Brasoveanu (2009) for axioms on the function that should describe stacks.

$$(50) \quad \{\langle 0, \text{MORRIS} \rangle, \langle 1, \text{HILARY} \rangle, \langle 2, \text{PHILIP} \rangle, \langle 3, \text{PERSY} \rangle\}$$

I will confuse a stack with the function that corresponds to it. Using this confusion we can describe a discourse referent as a function that retrieves the value from the stack position n as $u_n = \lambda i. i(n)$. Thus, a discourse referent is of type *se*. Consider a simple example. A discourse introduces a stack (49). Now we encounter a new sentence:

$$(51) \quad \text{He}_3 \text{ liked her}_1.$$

The pronouns are interpreted as discourse referents. He_3 is u_3 which in turn is, when applied to the stack i , Perssy. She_1 is Hilary. Thus, (51) will add to our discourse a new information, namely, that Perssy likes Hilary. Or, to be more precise, it will exclude those information stacks where u_3 and u_1 are not related in this way.

Recently, a lot of evidence has been gathered that a discourse should not handle just information stacks. Rather it should operate on *sets* of thereof (see Elworthy 1995; Krifka 1996; van den Berg 1996 for one of the first proposals along these lines). Consider the following example (from Wang et al. 2006 who also discuss questionnaires that confirmed the availability of the reading to be discussed below):

$$(52) \quad \text{Two students each wrote an article. They each sent it to a journal.}$$

Suppose we can get the following information stack after interpreting the first sentence, where s_1 and s_2 are students and s_1 wrote the article a_1 and s_2 wrote the article a_2 .

$$(53) \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_1 \oplus s_2 & a_1 \oplus a_2 \\ \hline \end{array}$$

I ignore the precise mechanism how we can get this stack (one obvious option is to use summation operation in DRT, for example). However, even if we can get this stack, it is of not much help for the second sentence in (52). Obviously, *they* is anaphoric to u_0 and *it* is anaphoric to u_1 . But there is more information in the second sentence than just that the two students sent the two articles. The sentence also expresses that each student sent the article that she wrote and this information disappears if we have only one information stack.

We can capture this interpretation if we consider a set of information stacks. In the example at hand the first sentence would then introduce two stacks:

$$(54) \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_1 & a_1 \\ \hline \end{array} \qquad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_2 & a_2 \\ \hline \end{array}$$

Rather than writing stacks separately, as we have done in (54) we will use the following notational convention: we will use multiple rows, where each row represents one stack. In this new convention, (54) looks as follows:

$$(55) \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_1 & a_1 \\ \hline s_2 & a_2 \\ \hline \end{array}$$

Using sets of information stacks give us a way to handle quantificational dependency of the example (52). In particular, each row, i.e., each stack keeps the information about dependency (we will see why when we consider the interpretation of distributive quantifiers in more detail). It suffices to require that *they* and *it* are anaphoric to u_0 and u_1 and that the relation ‘send to a journal’ holds between u_0 and u_1 in each information stack, which correctly gives us that each student sent the article she wrote.

This concludes my discussion on behavior of stacks. Before showing how they are used in PCDRT let me introduce a few definitions. Below, I will need to be able to expand stacks and sets of stacks with new discourse referents. I use the abbreviation $i[u_n]j$ to say that two stacks differ at most at the position n , defined as:

$$(56) \quad i[u_n]j = (\forall m < n)(u_m(j) = u_m(i)) \wedge (\forall m > n)(u_m(j) = u_{m-1}(i))$$

For sets of stacks we assume a pointwise definition:

$$(57) \quad I[u_n]J = (\forall j \in J)(\exists i \in I)(i[u_n]j) \wedge (\forall i \in I)(\exists j \in J)(i[u_n]j)$$

We will also need to be able to talk about all the entities that are at the same position in a set of stacks and to sum these entities up. This is given in the following two definitions (\oplus was defined in Chapter 3).

$$(58) \quad \begin{array}{ll} \text{a.} & u_n I = \{u_n(i) : i \in I\} \\ \text{b.} & \oplus u_n I = \oplus \{u_n(i) : i \in I\} \end{array}$$

There will be more operations on stacks, which I will introduce as we go along.

Interpretation fragment

Consider the following example:

$$(59) \quad \text{A man walks. He is tall.}$$

In DRT, we have the following discourse representation (DRS) for (59), where discourse referents newly introduced to discourse appear at top of the box and conditions appear below.

$$(60) \quad \begin{array}{|c|} \hline u_0 \\ \hline \text{man}\langle u_0 \rangle \\ \text{walk}\langle u_0 \rangle \\ \hline \end{array} \quad \text{tall}\langle u_0 \rangle$$

To save space, I will follow Muskens’ convention (Muskens, 1996) and use square brackets to notate DRSs. In this new notation, (59) is:

$$(61) \quad [u_0 | \text{man}\langle u_0 \rangle, \text{walk}\langle u_0 \rangle]; [|\text{tall}\langle u_0 \rangle]$$

We need to ensure that *he* in the second sentence is interpreted as the man that entered. To do so, we proceed as follows. We treat a DRS as a function of type $\langle \langle st \rangle \langle \langle st \rangle t \rangle \rangle$, that is, they take a set of stacks as input and output a set of stacks. Furthermore, we make use of ‘;’. ‘;’ operates as a dynamic conjunction. It makes discourse referents in the DRS to the left accessible for the DRS to the right.

To interpret (61) we can understand the boxes and conditions inside them simply as a neat way of abbreviations. These abbreviations are spelled out in (62):

$$\begin{array}{lll}
 [u_1, \dots, u_n | R_1, \dots, R_n] & \text{abbreviates} & [u_1]; \dots; [u_n]; [R_1]; \dots; [R_n] \\
 [u_n] & \text{abbreviates} & \lambda I \lambda J. I[u_n]J \\
 (62) \quad [R] & \text{abbreviates} & \lambda I \lambda J. I = J \wedge RJ \\
 [\dots]_1; [\dots]_2 & \text{abbreviates} & \lambda I \lambda J. (\exists K_{\langle st \rangle}) ([\dots]_1(I)(K) \wedge \\
 & & \wedge [\dots]_2(K)(J)) \\
 R\langle u_0, \dots, u_n \rangle & \text{abbreviates} & \lambda I. (\forall i \in I)(R(u_0(i), \dots, u_n(i)))
 \end{array}$$

Using these abbreviations, we can now rewrite (61) as (63):

$$\begin{array}{l}
 (63) \quad \lambda I \lambda J. (\exists K)(I[u_0]K \wedge (\forall k \in K)(\text{MAN}(u_0(k))) \wedge (\text{WALK}(u_0(k))) \\
 \wedge K = J \wedge (\forall j \in J)(\text{TALL}(u_0(j))))
 \end{array}$$

To see the interpretation of (63) we add the definition of truth:

$$\begin{array}{l}
 (64) \quad \text{A DRS } D (= \text{type } \langle \langle st \rangle \langle \langle st \rangle t \rangle \rangle) \text{ is true in a model } M \text{ under a set of infor-} \\
 \text{mation stacks } I \text{ iff there is some } J \text{ such that } D(I)(J)
 \end{array}$$

For (63) assume that $I = \emptyset$. Assuming semantics parallel to language \mathcal{L} discussed in Chapter 3 (with the extension of basic types to s), (63) is true in a model in which a man exists who walks and who is tall.

Since we know how a sentence should be interpreted it is not difficult to go into a sub-sentential level and derive the meaning in a compositional way. As a guideline, we can use the interpretation of LF we were using in Chapter 3 (but ignoring events) and the following convention: type t in static semantics corresponds to type $\langle \langle st \rangle \langle \langle st \rangle t \rangle \rangle$ (DRS) in PCDRT. Type e in static semantics corresponds to type $\langle se \rangle$, a discourse referent. When explicating lexical meaning I will follow this convention and use \mathbf{t} and \mathbf{e} as ‘meta-types’ abbreviating $\langle \langle st \rangle \langle \langle st \rangle t \rangle \rangle$ and $\langle se \rangle$ respectively. Table 5.4 summarizes translations of expressions used in (61), as well as some other most common expressions. When showing the translations I assume the convention that Muskens (1996) makes use of: if an expression introduces a discourse referent, I notate it by the superscript to that expression. Furthermore, I use brackets inside the translations, which are graphical notations of DRSs.

Notice that as in previous chapters, I treat numerals as a modifier. To lift them to a quantifier, I assume a null operator EX which is like an indefinite article but without number restriction.

To see how the interpretation is derived compositionally, consider now the first sentence in (59), repeated here:

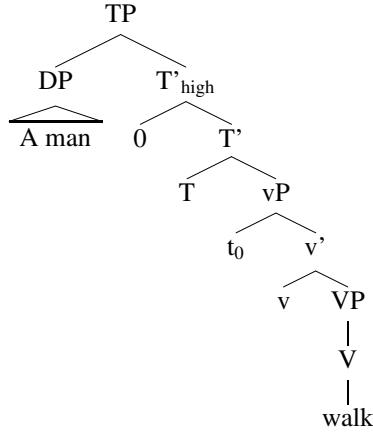
$$\begin{array}{l}
 (65) \quad \text{A man walks.}
 \end{array}$$

Table 5.4: Translations of some terminal LF nodes

Expression of LF	Translation
walk	$\lambda v_e. [[\text{walk}\langle v \rangle]]$
Philip ^{<i>u_n</i>}	$\lambda P_{\text{et}}. [u_n \text{Philip} = u_n]; P\langle u_n \rangle$
man	$\lambda v_e. [[\text{man}\langle v \rangle]]$
like	$\lambda Q_{\langle \langle \text{et} \rangle \text{t} \rangle} \lambda v_e. Q(\lambda v'_e. [[\text{like}\langle v, v' \rangle]])$
a ^{<i>u_n</i>}	$\lambda P_{\text{et}} \lambda Q_{\text{et}}. [u_n \text{singleton}\langle u_n \rangle]; P\langle u_n \rangle; Q\langle u_n \rangle$
the ^{<i>u_n</i>}	$\lambda P_{\text{et}} \lambda Q_{\text{et}}. [u_n \neq \emptyset]; P\langle u_n \rangle; Q\langle u_n \rangle$
three	$\lambda P_{\text{et}} \lambda v_e. [AT\langle v \rangle = 3]; P v$
EX	$\lambda P_{\text{et}} \lambda Q_{\text{et}}. [u_n]; P\langle u_n \rangle; Q\langle u_n \rangle$
t _n	$\lambda P_{\text{et}}. P\langle u_n \rangle$
he _n	$\lambda P_{\text{et}}. P\langle u_n \rangle$
singleton ^{<i>u_n</i>}	$\lambda I. AT(u_n I) = 1$
AT ^{<i>u_n</i>}	$\lambda I. AT(u_n I) = 3$

This sentence gets the following LF structure:

(66)



If we assume the rules of interpretation of non-terminal nodes that we assumed in Chapter 3 (with the modification that lambda-abstraction, applying here at T'_{high} , abstracts over type e), we can correctly derive the resulting meaning in a compositional way. The crucial steps are given in (67).

- (67) a. $[[vP]] = [[\text{walk}\langle u_0 \rangle]] =$
 $\lambda I \lambda J. I = J \wedge (\forall j \in J)(\text{WALK}(u_0(j)))$
- b. $[[T'_{\text{high}}]] = \lambda v_e. [[\text{walk}\langle v \rangle]] =$
 $\lambda v \lambda I \lambda J. I = J \wedge (\forall j \in J)(\text{WALK}(v(j)))$
- c. $[[DP]] = \lambda P. [u_0 | \text{man}\langle u_0 \rangle]; P\langle u_0 \rangle =$
 $\lambda P \lambda I \lambda J. I[u_0] J \wedge (\forall j \in J)(\text{MAN}(u_0(j))) \wedge (\forall j \in J)(P(u_0(j)))$

- d.
$$\begin{aligned} [[\text{TP}]] &= [u_0 | \text{man}(u_0)]; [[\text{walk}(u_0)]] = \\ &\lambda I \lambda J. (\exists K)(I[u_0]K \wedge (\forall k \in K)(\text{MAN}(u_0(j)))) \wedge \\ &\wedge (K = J \wedge (\forall j \in J)(\text{WALK}(u_0(j)))) = \\ &\lambda I \lambda J. (I[u_0]J \wedge (\forall j \in J)(\text{MAN}(u_0(j))) \wedge (\forall j \in J)(\text{WALK}(u_0(j)))) \end{aligned}$$

Assume that $I = \emptyset$. Then, (67d) is true in a model in which there is an individual that is a man and that walks. Furthermore, if we say that the application of ‘;’ is possible at the root level, we can dynamically conjoin (67d) with the next sentence to derive (61) which is true in a model in which there is a man that walks and that is tall. More examples with a more detailed discussion are given in Muskens (1996) and Brasoveanu (2008c).

5.5 Same and different₁ in PCDRT

5.5.1 Discourse-anaphoric reading in PCDRT

Same and *different*₁ can, in their discourse-anaphoric reading, get the meaning parallel to that of anaphors. Consider the following example, repeated from above.

- (68) Morris read Deaf Sentence.
 a. Hilary read the same book.
 b. Hillary read a different book.

In (68) *same* and *different* are anaphoric to the previously mentioned book, *Deaf Sentence*, and state that the book that Hilary read is identical and distinct, respectively. The interpretations of the two expressions are given in (69) and (70).

- (69) The interpretation of *same*

$$[[\text{same}_n]] = \lambda P_{\text{et}} \lambda x_e. \underline{Pu_n}; x = u_n; Px$$
- (70) The interpretation of *different*

$$[[\text{different}_n]] = \lambda P_{\text{et}} \lambda x_e. \underline{Pu_n}; \text{DISTINCT}(x, u_n); Px; Pu_n$$

The distinctness condition has been discussed in Section 5.3.2 and it is repeated here. It is similar to the distinctness condition we assumed for static semantics. The only difference is that where we assumed that x and y must not overlap, we now assume that the join of the discourse referents at the given position in all the stacks must not overlap. The difference simply stems from the fact that we operate with sets of stacks while before we made use of a single variable assignment.

- (71) The definition of distinctness
 y is distinct from z iff (a) or (b)
 a. $\neg \oplus yI \circ \oplus zI$
 b. $\oplus yI$ and $\oplus zI$ belong to kinds y' and z' and $\neg y' \circ z'$

In the formulas (69) and (70) the first condition Pu_n is underlined. This is because it is part of the presuppositional value of *same* and *different*. That the anaphoric element is presupposed to be in the extension of P has been argued in Brasoveanu (2009) and can be seen in the following examples:

- (72) a. Morris sang “Candy”.
 b. # Hilary did not read a different book./
 c. # Did Hilary read a different book?

In (72) the continuation of the first sentence is infelicitous since “Candy” is not a book. The awkward status of the continuation remains even if *different* is in the scope of negation or a question operator.

To see how *same* and *different* work consider (68). The first sentence gets the following interpretation:

- (73) $[u_1, u_2 | u_1 = \text{Morris}, u_2 = \text{Deaf Sentence}]$

(68a) is interpreted as (74a) and (68b) is interpreted as (74b).

- (74) a. $[u_1, u_2 | u_1 = \text{Morris}, u_2 = \text{Deaf Sentence}];$
 $[[u_3, u_4 | u_3 = \text{Hilary}, \underline{\text{book}\langle u_2 \rangle}, u_4 = u_2, \text{book}\langle u_4 \rangle]]$
 b. $[u_1, u_2 | u_1 = \text{Morris}, u_2 = \text{Deaf Sentence}];$
 $[u_3, u_4 | u_3 = \text{Hilary}, \underline{\text{book}\langle u_2 \rangle}, \text{DISTINCT}(u_4, u_2), \text{book}\langle u_4 \rangle]]$

(74a) is true in a model in which Morris read Deaf Sentence, Deaf Sentence is a book and Hilary read the same book. (74b) is true in a model in which Morris read Deaf Sentence, Deaf Sentence is a book and Hilary read a book different from the one that Morris read.

We will now move to the sentence-internal reading of *same* and *different*₁.

5.5.2 The sentence-internal reading of *same* and *different*₁ in PC-DRT

Introducing distributive quantification into PCDRT

Consider (75).

- (75) Two students each wrote an article.

The set of information stacks that should be the output of (75) could be pictured as follows, where u_0 is the discourse referent introduced by the subject *two students* and u_1 is the discourse referent introduced by the object *an article*.

- (76)
- | | |
|----------------|----------------|
| 0 | 1 |
| s ₁ | a ₁ |
| s ₂ | a ₂ |

An obvious obstacle in getting from (75) to (76) is that the object *an article* requires its discourse referent to be a singleton but u_1 consists of two elements. Obviously, this problem is at the heart of the fact that we deal with distributive quantification here: if we let the indefinite introduce its discourse referent within the scope of distributive quantification it is going to be true that the discourse referent introduced is one atom.

Thus, to get to (76) we need to isolate the stacks that have exactly one student in the 0 position. In the formulas below we make use of the following definition:

$$(77) \quad I_{u_n=x} = \{i : i \in I \wedge u_n(i) = x\}$$

That is, $I_{u_n=x}$ is the set of stacks that have the entity x at the n th position. In our example, we consider, for each student x , $I_{u_0=x} = \{i : i \in I \wedge u_0(i) = x\}$ and we interpret the rest of the clause in scope of distributive quantification only with respect to this subset of stacks. Finally, if this could be done we end up updating the whole set of stacks with (76).

So these are the steps: first, we have information stacks with the position 0 filled in which was introduced by the subject:

$$(78) \quad \begin{array}{|c|} \hline 0 \\ \hline s_1 \\ \hline s_2 \\ \hline \end{array}$$

Afterwards we consider the subset of the stacks where the position 0 consists of one student. That is, for s_1 we consider the singleton set of stacks:

$$(79) \quad \begin{array}{|c|} \hline 0 \\ \hline s_1 \\ \hline \end{array}$$

We proceed interpreting the rest of the clause which is in the scope of the distributive quantification. We update this subset of the original set of stacks with the discourse referent introduced by *an article*:

$$(80) \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_1 & a_1 \\ \hline \end{array}$$

And we proceed similarly for the subset of the stacks with s_2 in the position 0. This gives us the following singleton set of stacks:

$$(81) \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_2 & a_2 \\ \hline \end{array}$$

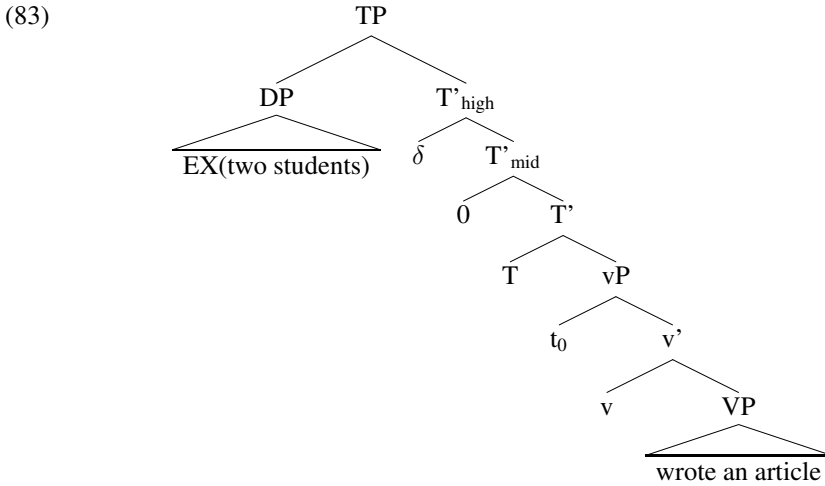
If this is satisfied we end up with (76).

We now need to define the distributive operator that allows considering subsets of stacks in its scope. Following other works, I call this distributive operator δ . It is interpreted as follows:

$$(82) \quad \text{The interpretation of the distributive operator, first version} \\ \delta(\phi_{\text{et}}) = \lambda v_e \lambda I \lambda J. vI = vJ \wedge (\forall x \in vI)(\phi(v)(I_{v=x})(J_{v=x}))$$

Let me break this down. δ is a distributive operator, parallel to the $*$ from the previous section. As $*$, it takes type et as its input and outputs type et . It consists of two conjuncts. The first one requires that the discourse referent which is the distributor is the same for the input and output stacks. Without this condition we could uncontrollably expand the discourse referent that serves as the distributor so this just makes sure that this discourse referent is not modified. For details why we need this condition, see Nouwen (2003). The second conjunct states that for each entity x in the discourse referent which is the distributor we interpret the predicate only with respect to the subset of the stacks where the discourse referent picks x .

Let me assume that the floating *each* is the overt manifestation of the distributive operator δ . Now, we get the following LF structure for (75).



This is interpreted in the following way:

- (84)
- a. $[[VP]] = \lambda v_e.[u_1 | \text{singleton}\langle u_1 \rangle, \text{article}\langle u_1 \rangle, \text{like}\langle v, u_1 \rangle] = \lambda v_e \lambda I \lambda J. I[u_1] J \wedge (\forall j \in J)(\text{ARTICLE}(u_1(j)) \wedge \text{LIKE}(v(j), u_1(j)))$
 - b. $[[vP]] = [u_1 | \text{singleton}\langle u_1 \rangle, \text{article}\langle u_1 \rangle, \text{like}\langle u_0, u_1 \rangle] = \lambda I \lambda J. I[u_1] J \wedge (\forall j \in J)(\text{ARTICLE}(u_1(j)) \wedge \text{LIKE}(u_0(j), u_1(j)))$
 - c. $[[T'_{mid}]] = \lambda v_e.[u_1 | \text{singleton}\langle u_1 \rangle, \text{article}\langle u_1 \rangle, \text{like}\langle v, u_1 \rangle] = \lambda v \lambda I \lambda J. I[u_1] J \wedge (\forall j \in J)(\text{ARTICLE}(u_1(j)) \wedge \text{LIKE}(v(j), u_1(j)))$
 - d. $[[T'_{high}]] = \lambda v_e. \delta(\lambda v'. [u_1 | \text{singleton}\langle u_1 \rangle, \text{article}\langle u_1 \rangle, \text{like}\langle v', u_1 \rangle]) = \lambda v \lambda I \lambda J. v I = v J \wedge (\forall x \in v I) (\lambda I \lambda J. I[u_1] J \wedge (\forall j \in J)(\text{ARTICLE}(u_1(j)) \wedge \text{LIKE}(v(j), u_1(j))))(I_{v=x})(J_{v=x}))$
 - e. $[[TP]] = [u_0 | \text{student}\langle u_0 \rangle, |AT(u_0)| = 2]; \delta(\lambda v [u_1 | \text{singleton}\langle u_1 \rangle, \text{article}\langle u_1 \rangle, \text{like}\langle v, u_1 \rangle])\langle u_0 \rangle = \lambda I \lambda J. (\exists K)(I[u_0] K \wedge |AT(u_0 K)| = 2 \wedge (\forall k \in K)(\text{STUDENT}(u_0(k))) \wedge u_0 K = u_0 J \wedge (\forall x \in u_0 K) (\lambda L \lambda M. L[u_1] M \wedge (\forall m \in M)(\text{SINGLETON}(u_1(m)) \wedge \text{ARTICLE}(u_1(m)) \wedge \text{LIKE}(u_0(m), u_1(m))))(K_{u_0=x})(J_{u_0=x}))$

(84e) is true in a model in which there are two students and each of them wrote an article. Notice how this is arrived at: first, we establish that there are two students by expanding the set of information stacks by the discourse referent u_0 which consists of two students. This is done in line three of (84e). After that, the lines 4-6 establish that for each value of u_0 , the set of information stacks that gives us this value can be expanded by a new set, J , which has a single article as u_1 and u_0 wrote u_1 .

Graphically, the set K which makes the third line of (84e) true gives us the following set of stacks:

$$(85) \quad \text{Stacks } K = \begin{array}{|c|} \hline 0 \\ \hline s_1 \\ \hline s_2 \\ \hline \end{array}$$

The resulting set of stacks is given in (86).

$$(86) \quad \text{Stacks } J = \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_1 & a_1 \\ \hline s_2 & a_2 \\ \hline \end{array}$$

Suppose we now follow the discourse with the following sentence (what anaphoric expressions are anaphoric to is given by indices):

$$(87) \quad \text{They}_0 \text{ each sent it}_1 \text{ to a journal.}$$

The interpretation of (87) is given in (88).

$$(88) \quad \begin{aligned} [[(86)]] = & \delta(\lambda v[u_1 | \text{singleton}\langle u_2 \rangle, \text{journal}\langle u_2 \rangle, \text{send}\langle v, u_1, u_2 \rangle])(u_0) = \\ & \lambda I \lambda J. u_0 I = u_0 J \wedge (\forall x \in u_0 I) \\ & (\lambda L \lambda M. L[u_2] M \wedge (\forall m \in M) (\text{SINGLETON}(u_2(m)) \wedge \text{JOURNAL}(u_2(m)) \wedge \\ & \text{SEND}(u_0(m), u_1(m), u_2(m)))) (I_{u_0=x}) (J_{u_0=x})) \end{aligned}$$

Combining (88) with the previous DRS, (84e), we correctly derive that the two formulas are true in a model in which there are two students and each of them wrote an article and sent it to a journal. This is the reading we were after. We could not account for it if DRSs did not operate on sets of assignments, as we discussed in Section where information stacks have been introduced.

There is a problem with this account, pointed out in Nouwen (2007). Like van den Berg (1996) and Krifka (1996) we cannot deal with the following discourse, under the reading that *them* in the second clause refers to both articles written by the two students.

$$(89) \quad \text{Two students each wrote an article. They each liked them.}$$

The problem is that δ temporarily ignores a lot of information introduced previously in the discourse because it keeps only the subset of stacks available to the material in its scope.

Brasoveanu (2009) suggests that we could avoid this problem if we subsume the resolution of the object pronoun in (89) under the case of bridging. In bridging, the listener fills in some referent in order to resolve pronoun reference. Bridging has been

brought to the attention of linguists in Clark (1977) who studied the principles that makes it possible (see also Asher and Lascarides 1998; Piwek and Krahmer 2000; Matsui 2000 and others). An example of bridging is given in (90) where ‘the ceiling’ can appear with a definite article since its existence can be bridged from the existence of a room.

(90) I looked into the room. The ceiling was very high.

Brasoveanu’s idea is that in the sub-discourse that is built in scope of δ in (88) we have one article present (that is, for each boy, the article that the boy wrote is present in the information stack). We might then assume that the plurality of articles is bridged from the original article, parallel to the way the existence of the ceiling is bridged from the existence of the room in (90).

I do not think that this could work. The problem is that bridging usually proceeds from an object to its parts, not the other way round (see Clark 1977 for the list of bridging). Furthermore, even if the opposite way of bridging was possible in (88) there would be no way to correctly restrict it. Thus, we would not only get ‘the articles that the students wrote’ but any other plurality that includes the article that a student wrote. It seems to me that if we want to avoid this overgeneration we will need to modify δ along the lines that Nouwen (2007) argues for. However, I am not going to do so here since this is irrelevant for the main point we want to make, that is, the sentence-internal reading of *same* and *different*₁.

The sentence-internal reading of *same* and *different*₁

Consider now (91).

(91) Two students each read the same/a different book.

The main insight of Brasoveanu (2009) is, I believe, that in order to account for the sentence-internal reading of *same* and *different*₁, we only need to ensure that these expressions can be anaphoric to the books that the other boys read while being in scope of δ . To achieve this, Brasoveanu modifies δ that we introduced in (82). Intuitively, what we want for (91) is to temporarily have the books that the other boys read in the stack so *same* and *different*₁ could be anaphoric to them. Thus, our distributive operator should, for each student x , concatenate the stacks in which u_n gives us the value x with the stacks in which u_n gives us a different value. We thus need the last operation on stacks, stack concatenation. We first define a stack concatenation of stacks, $+$, in (92).

$$(92) \quad i + i' = i \cup \{ \langle n, d \rangle : n > \text{Dom}(i) - 1 \wedge n \leq \text{Dom}(i) + \text{Dom}(i') - 1 \wedge i'(n - \text{Dom}(i)) = d \}$$

Intuitively, what (92) does is that it takes two stacks and fuses them into one in such a way that the second stack starts at the position where the first stack ends. Assume we have two stacks: $i = \{ \langle 0, \text{Hilary} \rangle, \langle 1, \text{Morris} \rangle \}$ and $i' = \{ \langle 0, \text{Philip} \rangle, \langle 1, \text{Perissy} \rangle \}$. Then, $i + i' = \{ \langle 0, \text{Hilary} \rangle, \langle 1, \text{Morris} \rangle, \langle 2, \text{Philip} \rangle, \langle 3, \text{Perissy} \rangle \}$.

The concatenation of sets of stacks, \bullet , is defined as follows:

$$(93) \quad I \bullet I' = \{i + i' : i \in I \wedge i' \in I'\}$$

Let me show on the example (91) how we make use of concatenation. We start with u_0 introduced by the subject. This gives us the following sets of stacks:

$$(94) \quad \begin{array}{|c|} \hline 0 \\ \hline s_1 \\ \hline s_2 \\ \hline \end{array}$$

We then consider the subset of these stacks for each student and we concatenate this subset with another subset, which is complementary to the first. Thus, in our example we consider the following subsets for s_1 and s_2 :

$$(95) \quad \begin{array}{l} \text{The subset of stacks in scope of the distributive operator for } x = s_1 \\ \begin{array}{|c|} \hline 0 \\ \hline s_1 \\ \hline \end{array} \quad \bullet \quad \begin{array}{|c|} \hline 0 \\ \hline s_2 \\ \hline \end{array} \\ \text{The subset of stacks in scope of the distributive operator for } x = s_2 \\ \begin{array}{|c|} \hline 0 \\ \hline s_2 \\ \hline \end{array} \quad \bullet \quad \begin{array}{|c|} \hline 0 \\ \hline s_1 \\ \hline \end{array} \end{array}$$

For each subset we proceed interpreting the whole predicate over which δ takes scope. After interpreting the object this gives us the following set of stacks:

$$(96) \quad \begin{array}{l} \text{The subset of stacks in scope of the distributive operator for } x = s_1 \\ \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_1 & b_1 \\ \hline \end{array} \quad \bullet \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_2 & b_2 \\ \hline \end{array} \\ \text{The subset of stacks in scope of the distributive operator for } x = s_2 \\ \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_2 & b_2 \\ \hline \end{array} \quad \bullet \quad \begin{array}{|c|c|} \hline 0 & 1 \\ \hline s_1 & b_1 \\ \hline \end{array} \end{array}$$

Given the definition of concatenation, we can also rewrite these stacks as follows:

$$(97) \quad \begin{array}{l} \text{The subset of stacks in scope of the distributive operator for } x = s_1 \\ \begin{array}{|c|c|c|c|} \hline 0 & 1 & 2 & 3 \\ \hline s_1 & b_1 & s_2 & b_2 \\ \hline \end{array} \\ \text{The subset of stacks in scope of the distributive operator for } x = s_1 \\ \begin{array}{|c|c|c|c|} \hline 0 & 1 & 2 & 3 \\ \hline s_2 & b_2 & s_1 & b_1 \\ \hline \end{array} \end{array}$$

Now, we only need to ensure that *different*₁ and *same* can be anaphoric to u_3 . Furthermore, *different*₁ expresses that the element it is anaphoric to is distinct, while *same* requires identity.

Thus, we modify δ . The only difference from the previous version is adding stack concatenation. This allows us to concatenate one subset of stacks with the complement set of the stacks.

$$(98) \quad \begin{array}{l} \text{The interpretation of the distributive operator, final version} \\ \delta(\phi_{\text{et}}) = \lambda v_e \lambda I \lambda J. vI = vJ \wedge (\forall x \in vI)(\phi(v)(I_{u_n=x} \bullet I_{\neg u_n \circ x})(J_{u_n=x} \bullet J_{\neg u_n \circ x})) \end{array}$$

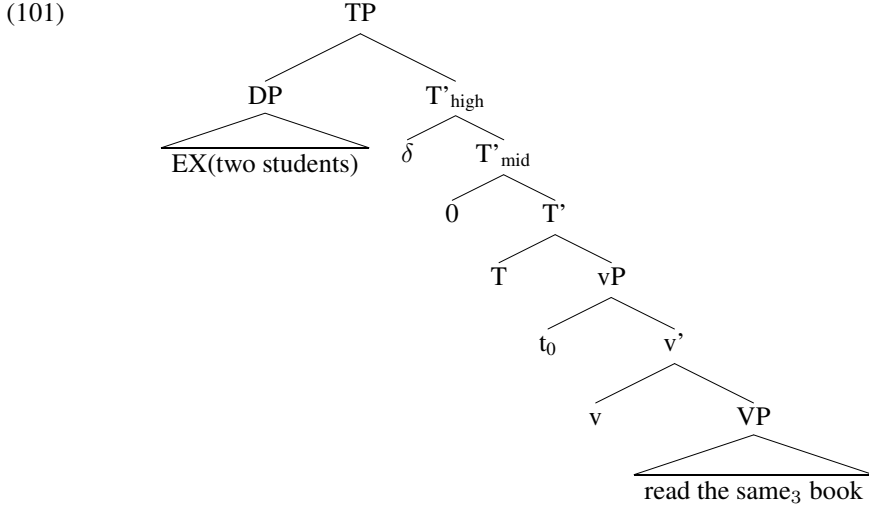
Thanks to the stack concatenation *same* and *different*_I have the possibility to be anaphoric to elements that are introduced in scope of δ and which represent ‘the book read by the other boy’ in the example (91).

Let me show how we derive the correct reading of (91) in more detail. We use the interpretation of *same* and *different* as introduced in (69) and (70) and repeated here:

(99) The interpretation of *same*
 $[[\text{same}_n]] = \lambda P_{\text{et}} \lambda x_e. \underline{Pu_n}; x = u_n; Px$

(100) The interpretation of *different*
 $[[\text{different}_n]] = \lambda P_{\text{et}} \lambda x_e. \underline{Pu_n}; \text{DISTINCT}(x, u_n); Px; Pu_n$

The LF structure for (91) with *same* in the adjectival position is:



This gets the following interpretation:

- (102) a. $[[\text{VP}]] = \lambda v_e. [u_1 | \text{article}\langle u_3 \rangle, u_1 = u_3, u_1 \neq \emptyset, \text{article}\langle u_1 \rangle, \text{like}\langle v, u_1 \rangle] = \lambda v_e \lambda I \lambda J. I[u_1] J \wedge (\forall j \in J) (\text{ARTICLE}(u_3(j)) \wedge u_1(j) = u_3(j) \wedge u_1(j) \neq \emptyset \wedge \text{ARTICLE}(u_1(j)) \wedge \text{LIKE}(v(j), u_1(j)))$
- b. $[[\text{vP}]] = [u_1 | \text{article}\langle u_3 \rangle, u_1 = u_3, u_1 \neq \emptyset, \text{article}\langle u_1 \rangle, \text{like}\langle u_0, u_1 \rangle] = \lambda I \lambda J. I[u_1] J \wedge (\forall j \in J) (\text{ARTICLE}(u_3(j)) \wedge u_1(j) = u_3(j) \wedge u_1(j) \neq \emptyset \wedge \text{ARTICLE}(u_1(j)) \wedge \text{LIKE}(u_0(j), u_1(j)))$
- c. $[[\text{T'midP}]] = \lambda v_e. [u_1 | \text{article}\langle u_3 \rangle, u_1 = u_3, u_1 \neq \emptyset, \text{article}\langle u_1 \rangle, \text{like}\langle v, u_1 \rangle] = \lambda v_e \lambda I \lambda J. I[u_1] J \wedge (\forall j \in J) (\text{ARTICLE}(u_3(j)) \wedge u_1(j) = u_3(j) \wedge u_1(j) \neq \emptyset \wedge \text{ARTICLE}(u_1(j)) \wedge \text{LIKE}(v(j), u_1(j)))$
- d. $[[\text{T'midP}]] = \lambda v_e. \delta([u_1 | \text{article}\langle u_3 \rangle, u_1 = u_3, u_1 \neq \emptyset, \text{article}\langle u_1 \rangle, \text{like}\langle v, u_1 \rangle]) = \lambda v_e \lambda I \lambda J. vI = vJ \wedge (\forall x \in vI) \lambda L \lambda M. L[u_1] M \wedge (\forall m \in M) (\text{ARTICLE}(u_3(m)) \wedge u_1(m) = u_3(m) \wedge u_1(m) \neq \emptyset \wedge \text{ARTICLE}(u_1(m)) \wedge \text{LIKE}(v(m), u_1(m))) (I_{v=x} \bullet I_{\neg v \circ x}) (J_{v=x} \bullet J_{\neg v \circ x}))$

- e.
$$\begin{aligned} [[\text{TP}]] &= [u_0 | \text{student}\langle u_0 \rangle, |AT(u_0)| = 2]; \\ \delta(\lambda v [u_1 | \text{article}\langle u_3 \rangle, u_1 = u_3, u_1 \neq \emptyset, \text{article}\langle u_1 \rangle, \text{like}\langle v, u_1 \rangle]) \langle u_0 \rangle &= \\ \lambda I \lambda J. (\exists K) (I[u_0] K \wedge |AT(u_0 K)| = 2 \wedge (\forall k \in K) (\text{STUDENT}(u_0(k)) \wedge \\ u_0 K &= u_0 J \wedge (\forall x \in u_0 K) \\ (\lambda L \lambda M. L[u_1] M \wedge (\forall m \in M) (\text{ARTICLE}(u_3(m)) \wedge u_1(m) &= u_3(m) \wedge \\ u_1(m) &\neq \emptyset \wedge \text{ARTICLE}(u_1(m)) \wedge \text{LIKE}(u_0(m), u_1(m)))) \\ (K_{u_0=x} \bullet K_{\neg u_0 \circ x}) (J_{u_0=x} \bullet J_{\neg u_0 \circ x})) \end{aligned}$$

A similar approach can be given to *different*₁. The only difference is that in (103) *different* expresses that the book a student read is distinct from the book that the other student read.

- (103) Two students each read a different book.

To conclude we see that PCDRT can account for the sentence-internal reading of *same* and *different*₁ in case the distributive operator scopes over these adjectives. Moreover, the interpretation is identical to the interpretation we use for the discourse-anaphoric reading. As Brasoveanu argues this is a good thing because we see a correlation between the sentence-internal reading of *different* and the availability of discourse-anaphoric reading: whenever the sentence-internal reading of some DS-expression is possible in scope of a distributive quantifier the discourse-anaphoric reading of that expression is also possible.

We can also derive why the sentence-internal reading of *different*₁ is degraded with non-distributive quantifiers, along the spine discussed in Table 5.1, while the sentence-internal reading of *same* is not. This follows from the assumptions made in Chapter 3 that the distributive reading with covariation is marked. In the context of this chapter, we can understand covariation as follows:

- (104) u_n covaries with u_m in a set of stacks I iff

- the quantifier introducing u_n is in scope of the distributive operator to which the quantifier introducing u_m applies and
- for some x and y that are distinct it holds that $u_n I_{u_m=x} \neq u_n I_{u_m=y}$

Given the discussion in Chapter 3 we assume that:

- (105) The reading in which some discourse referent covaries with another one is marked

I will not repeat the reasoning that helps us derive from this assumption that the distributive reading with covariation is most rejected when the distributor is a G-DP (a group-denoting DP) and less so when it is a CQ-DP (a DP with a counting quantifier) and it is completely acceptable when it is a distributive quantifier. For details, see Chapter 3. Notice only that assuming covariation as defined here, we also explain why some cases of what looks like the distributive reading with covariation seem to be fully acceptable, for example:

- (106) a. The artichokes cost a dollar.

- b. The ushers give you a hug.

Here, the covarying object is idiomatic or at any rate does not introduce a discourse referent, as one can see from the fact that we cannot refer to dollars or punches later on in these examples. For instance, *the ushers give you a hug* can only marginally be continued with *#They were very pleasant* (meaning, the hugs being pleasant).

The assumption derives why the sentence-internal reading of *different*₁ is degraded with CQ-DPs and G-DPs. This is simply because *different*₁ in its sentence-internal reading is a case of covariation. Since there is another reading where no covariation takes place, the discourse-anaphoric reading, the system expects this reading to be preferred. Straightforwardly, *same* is not degraded in its sentence-internal reading since unlike *different*₁ this reading does not lead to covariation. In this respect, we differ from Brasoveanu's analysis. Brasoveanu, following Moltmann (1992) and Beck (2000), assumes that *different* is ambiguous between the A-analysis and Beck's analysis (see (34)). I showed in Section 5.3.2 that Beck's account is not general enough and cannot explain the behavior of *different*₂ and substituted it with Barker's account. Concerning *same*, Brasoveanu assumes the same ambiguity: one interpretation follows the A-strategy, the second one is built on Beck's account. However, the second interpretation is solely assumed to explain why the internal reading of *same* is possible with CQ-DPs and G-DPs. But given our assumption on the degraded status of covariation, we can dispense with this ambiguity. We derive that the sentence-internal reading of *same* is possible whenever the sentence-internal reading of *different*₁ is possible (due to their parallel lexical semantics) and whenever the sentence-internal reading of *different*₂ is possible (since no degradation due to the presence of covariation is induced). This is a good thing because as far as I know, we do not see that *same* would be realized as two lexical items cross-linguistically, where one lexical item would appear in the contexts in which *different*₁ can appear and the second one would appear in the contexts in which *different*₂ appears.

5.5.3 Open problems

In this account we expect the sentence-internal reading of *same* to be acceptable whenever *different*₁ or *different*₂ is. I will conclude this section with discussion of data which go against this expectation and to which I can offer no account at this point. Most of the data, as far as I know, have not been noticed so far, unless stated otherwise.

First of all, Brasoveanu makes the following observation. While *same* can appear in the restrictor of a distributive quantifier and still give rise to the sentence-internal reading, *different*₁ cannot. He bases it on the following contrast:

- (107) a. (At the party) Every / Each boy with the same name was wearing a blue shirt / the same color shirt. (sentence-internal reading possible)
- b. (At the party) Every / Each boy with a different name was wearing a blue shirt. (sentence-internal reading excluded)

I do not see how we could account for this contrast, given the assumptions so far. Since *same* and *different*₁ get their sentence-internal reading in the same way we do not expect them to differ in these two cases. Notice that making use of the marked status of covariation is irrelevant here since here we deal with a distributive quantifier.

However, I am not sure that the observation Brasoveanu makes is fully correct. It seems to me that *different*₁ can get the sentence-internal reading but only under special conditions: it requires that covariation is signalled in the scope of a distributive quantifier as well. Consider the following example, found via Google search:

- (108) In Aquinas' view, since no one name can bespeak God's entire perfection, it must be that each name with a different meaning expresses a different part of that perfection.

The interpretation of this sentence is the sentence-internal one: if two names differ in meaning then they express two different parts of perfection. It seems to me that a parallel example with no covariation signalled in the scope of the quantifier lacks the sentence-internal reading. In other words, the following example is hard to read as 'if names differ in their meaning, then each name differing in this way expresses God's perfection'.

- (109) Each name with a different meaning expresses God's perfection.

Other examples showing the same point and found on Google are given below. In fact, once we know what we should look for it is not difficult to find more cases where *different* has the sentence-internal reading in the restrictor of a quantifier.

- (110) a. Each IgC molecule with a different specificity has a different amino acid sequence for the heavy and light chains in this part of molecule.
b. For a reliable evaluation of the wave front, knowledge of the exact position of each hole is necessary, because the information on the local tilt of the wave front is extracted from the spot pattern. Each beam with a different wave front casts a different shadow on the CCD, and by comparing the deviations, it is possible to reconstruct the wave front bit by bit.

It seems then that both *different* and *same* can have the sentence-internal reading in the restrictor of a quantifier but the sentence-internal reading of *different* is felicitous if we there is covariation taking place in the scope of the quantifier as well. At this point, I do not know why there is such requirement.

Another issue that I will not be able to resolve here concerns negative quantifiers. These license the sentence-internal reading for *same*, see (111):

- (111) In the class nobody read the same book.

This is not true for either *different*₁ or *different*₂, as we have found out in the questionnaire presented in Chapter 2. In fact, the sentence-internal reading of *different*₁ and *different*₂ was the worst with negative quantifiers. The same holds for English, see (112) which lacks the sentence-internal reading:

- (112) In the class nobody read a different book. (sentence-internal reading excluded)

The degraded status of the sentence-internal reading in (112) could be due to the fact that *nobody* is composed of a negation and a lower scoping indefinite. The main evidence for such decomposition comes from so called split-scope readings. These are cases where we find the negation scoping over some operator and the indefinite scoping below the operator. For example, (113) has two readings paraphrased below.

- (113) Mary need fire no nurses.
- a. There is no nurse that Mary needs to fire.
 - b. It is not the case that Mary needs to do some nurse-firing.

In the first reading, (113) says that there is no particular nurse which Mary needs to fire. In the second reading, (113) says that Mary is not forced to do any nurse firing. The second reading entails the first reading because it can be true that Mary needs to do some nurse-firing even though there is no nurse in particular that Mary needs to fire. The existence of the stronger second reading is crucial because it shows that the negation can be interpreted in the scope position above the modal verb while the indefinite scopes below the modal verb. Penka and von Stechow (2001), Penka and Zeijlstra (2005) (see also Zeijlstra 2004) take this as an argument that negative quantifiers are to be decomposed into a negation and an indefinite (even though this is not the only analysis of split-scope readings; see, for example, de Swart 2000 for an account of split-scope readings in which no decomposition needs to be assumed).

The relevant reading of (112) would then be degraded simply because no distributive quantification, which is necessary for the sentence-internal reading, occurs there. However, this reasoning should also make the sentence-internal reading of (111) degraded. Assuming ambiguity of *same*, as Brasoveanu does, does not help with this issue. In the ambiguity account, *same* can also be non-anaphoric and get its sentence-internal reading via a cumulative quantification, in the same way that *different₂* in Beck's account receives its sentence-internal reading. However, negative quantifiers generally are reluctant to appear in cumulative readings so this strategy does not help with (111). If we assumed that cumulative readings are possible with negative quantifiers to explain (111) we would run into problems with *different₂*. This does not allow for the sentence-internal reading with negative quantifiers, not even when it modifies a plural noun. Thus, as far as I can see, the difference between (112) and (111) is problematic for my account, as well as any other analysis of *same* and *different*.

5.6 Conclusion

I argued that DS-expressions represent another case of the application of the two strategies for anaphoricity, R-strategy and A-strategy. *Different₂* was analyzed using the R-strategy. This could explain the incompatibility of its sentence-internal reading with distributive quantifiers, as well as why it cannot give rise to the discourse-

anaphoric reading. Finally, it could also explain why *different*₂ can give rise to the *various* reading when modifying a plural noun.

The A-strategy was used for *same* and *different*₁. This could explain, following Brasoveanu, why both expressions can give rise to the sentence-internal reading with distributive quantifiers, as well as why they can give rise to the discourse-anaphoric reading. Finally, combining the account with the assumptions from Chapter 3 we could also explain why the internal reading of *different*₁ is degraded with non-distributive DPs and this is not true for *same*. As far as I know, it is not the case cross-linguistically that *same* would be expressed by two lexical items, where *same*₁ would have the sentence-internal reading available only in cases where *different*₁ does and *same*₂ would have the reading available where *different*₂ does. I speculated that this might be because *same* is not restricted in the way that *different*₁ is since the degraded status of the distributive reading with covariation is irrelevant here. This has the consequence that *same* can appear in any context where *same*₂, the hypothetical DS-expression parallel to *different*₂, could appear. This renders *same*₂ useless, and thus, we end up with only one lexical item interpreting sameness.

In the last chapter, I compared my analysis of *each other* and *the others* with the one where only one strategy of anaphoricity was possible, namely, the A-strategy, with an extra assumption that *each other* has to be bound locally, unlike *the others*. I argued that this account cannot explain the properties of *each other* and *the others*.

Using just the A-strategy in case of DS-expressions is similarly unsuccessful. We would predict the existence of *same* and *different*₁. We would not predict the existence of *different*₂ since it makes use of the R-strategy. If anything, the A-strategy in which variables are obligatorily bound would predict the non-existent lexical items which would be like *same* and *different*₁ but which would have to be bound locally. I call these hypothetical expressions *same* and *different*. One could, for example, say *Morris saw the asme person in the mirror*, which should mean that Morris saw himself in the mirror but could not say *Morris read Deaf Sentence. I read the same book* since in the latter case the variable of the DS-expression would not be bound by the c-commanding expression. I am not aware that such lexical items exist. This supports the viewpoint that there are no variables which have to be obligatorily bound. Thus, anaphora which seem to behave that way, as are reflexives, reciprocals, and *different*₂, in fact require the R-strategy, or in other words, they are not arguments with to-be-bound variables, rather they are polyadic quantifiers.

CHAPTER 6

Summary and conclusion

Anaphoricity is ubiquitous in natural language. It is common in linguistics to study its properties mainly on pronouns, reflexives, and reciprocals but anaphoricity is a much more abundant phenomenon than one would believe when reading linguistic papers. In this thesis, I focused on expressions of (non-)identity, that is, reciprocals, *others*, *different* and *same*. What these expressions have in common is their ability to license the sentence-internal reading if they appear in a clause with a semantically plural argument. I offered an analysis of this reading, which closely connected to and had consequences for the theory of plurality and semantics of anaphora.

We have seen that the expressions of non-identity cluster in pairs. One type can give rise to the sentence-internal reading if the plural argument distributes over the clause (*the others*, *different*₁). The other type can give rise to the sentence-internal reading if the plural argument does not distribute over the clause (*each other*, *different*₂). *Same* is insensitive to the type of the plural argument. These data were discussed in Chapter 2.

I offered an account of these data in Chapter 3 to Chapter 5. The analysis was built on two general assumptions: first, there are two different strategies to anaphoricity; second, the distributive reading with covariation is marked and plural arguments compete with each other when we consider their interpretation.

In Chapter 3 I showed that the distributive reading with covariation is degraded with various plural arguments. It is fully acceptable with distributive quantifiers, but somewhat degraded with quantifiers headed by *all* and even more degraded with definite plurals, numeral indefinites or coordinations of proper names. This is explained if one interprets sentences by considering alternatives where one plural argument can be substituted by another one, and interpretations are not equal: the distributive reading with covariation is a marked option. I offered a formalization of this idea, follow-

ing Parikh's account of resolving ambiguities in Parikh (2000). I claimed that we can not only derive why some expressions are fully acceptable with distributive quantifiers, while others are degraded, but we can also derive the scale of degradation if we assume that strategies which resolve ambiguities can be quantified. In particular, I followed Parikh in measuring expected utility values of interpretation strategies to capture the scale of degradation.

In Chapter 4 we analyzed *the others* and reciprocals. We made use of the analysis from Chapter 3 which could explain why the sentence-internal reading of *the others* depends on the type of plural argument in the same way that the distributive reading with covariation does, and it could also explain why the sentence-internal reading of *each other* differs in this respect. We also offered a new analysis of *each other* which could deal with collective predicates in reciprocal sentences, as well as cumulative quantification and negation. This, I believe, represents an improvement over previous analyses of reciprocals. We furthermore proposed a new account of so-called long-distance reciprocity.

In Chapter 5 we analyzed *different* and *same*. We showed that *different* should really be split into two items. The sentence-internal reading of the first one depends on the type of plural argument in the same way that the distributive reading with covariation does, which followed from the account of Chapter 3. The sentence-internal reading of the second one was similar to *each other* in its sensitivity to the type of plural arguments. We offered an account of this fact, as well as why *same* is not split into two lexical items and why its sentence-internal reading is insensitive to the type of plural argument, in contrast with all the other expressions of (non-)identity.

There are various issues I had to leave aside. To the ones I occasionally pointed out in the thesis, I would like to add two more. The first one is the question whether two strategies of anaphoricity show up outside of the realm of expressions of (non-)identity. Likely candidates for the distinction are reflexives and pronouns. While pronouns follow what we have called the A-strategy (and thus pattern with *the others* and *different*₁ and *same*), reflexives follow the second strategy, the R-strategy (and thus pattern with *each other* and *different*₂). This would explain the lack of the discourse-anaphoric reading of reflexives. Interestingly, the two differ in other respects. For example, reflexives must find their antecedent in a local domain, while pronouns cannot be bound locally (even though the matter is more complicated, see, for instance, Reinhart and Reuland 1993). We might expect something similar to be true for expressions of (non-)identity as well. But the match is only partial. *Each other* and *different*₂ behave like reflexives in looking for their antecedents locally. But *the others*, *same*, *different*₁ do not behave like pronouns: they can find their antecedents in a local domain.

Standardly, the anti-locality of pronouns is explained either by the fact that they compete with reflexives and are excluded in contexts where reflexives are accepted (Levinson 2000, Roelofsen 2008) or by postulating a separate condition which governs the distribution of pronouns, Condition B (Chomsky 1981, Reinhart and Reuland 1993 and others). Either way, it is not clear why, for instance, (114a) and (114b) are acceptable. Obviously, either the expressions of non-identity do not compete with their

alternatives or Condition B does not apply here. The question, of course, is why not. I do not want to resolve this issue here (for one possible solution, see, for instance, Reuland 2008). My point is more general. Studying other anaphoric expressions than pronouns and reflexives might bring in new issues to study of anaphoricity and shed a different light to old problems. I tried to show in this thesis that expressions of (non-)identity are, in fact, revealing, in our understanding of pluralities, strategies of anaphoric resolutions, and other issues.

- (114) a. All the boys hated the others/the other boys.
b. All the boys read a different book.

The second question I put aside in this thesis is cross-linguistic lexicalization of expressions of (non-)identity. It is interesting to note that *each other* is related diachronically to *others*. This is not true only for English but holds for other languages as well (see Nedjalkov et al. 2007). *Different* is connected to *others* in languages like Dutch, while, for instance, in Czech, one uses the same expression for *different* and *else*. It is likely that studying lexicalization patterns of expressions of (non-)identity might bring very interesting new generalizations to the fore. But nothing of that sort will happen in this thesis. Because this thesis is going to end very, very soon. In fact, now.

Bibliography

- Abusch, Dorit. 1997. Sequence of Tense and Temporal De Re. *Linguistics and Philosophy* 20:1–50.
- Aloni, Maria. 2005. A formal treatment of the pragmatics of questions and attitudes. *Linguistics and Philosophy* 28.
- Altmann, Gerry, and Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30:191–238.
- Anderson, Catherine. 2004. The structure and real-time comprehension of quantifier scope ambiguity. Doctoral Dissertation, Northwestern University, Evanston, Illinois.
- Andrews, Peter B. 1986. *An introduction to mathematical logic and type theory*. Orlando/London: Academic Press.
- Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Asher, Nicholas, and Alex Lascarides. 1998. Bridging. *Journal of Semantics* 15:83–113.
- Avrutin, Sergey, and Rosalind Thornton. 1994. Distributivity and binding in child grammar. *Linguistic Inquiry* 25:165–171.
- Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9:5–16.
- Barker, Chris. 2007. Parasitic scope. *Linguistics and Philosophy* 30:407–444.
- Beaver, David, and Hanjung Lee. 2003. Input-output mismatches in Optimality Theory. In *Pragmatics in optimality theory*, ed. Reinhard Blutner and Henk Zeevat. Palgrave Macmillan.
- Beck, Sigrid. 1999. Plural predication and partitioned discourses. In *Proceedings of the 12th Amsterdam Colloquium*, ed. Paul Dekker, 67–72. Amsterdam.
- Beck, Sigrid. 2000. The semantics of Different: Comparison operator and relational adjective. *Linguistics and Philosophy* 23:101–139.
- Beck, Sigrid. 2001. Reciprocals are definites. *Natural Language Semantics* 9:69–138.
- Beck, Sigrid, and Uli Sauerland. 2000. Cumulation is needed: A reply to Winter (2000). *Natural Language Semantics* 8:349–371.
- Beghelli, Filippo, Dorit Ben-Shalom, and Anna Szabolcsi. 1997. Variation, distribu-

- tivity and the illusion of branching. In *Ways of scope taking*, ed. Anna Szabolcsi, 29–69. Kluwer.
- Beghelli, Filippo, and Timothy Stowell. 1997. Distributivity and negation. In *Ways of scope taking*, ed. Anna Szabolcsi, 71–109. Dordrecht: Kluwer.
- Benz, Anton, Gerhard Jäger, and Robert van Rooy. 2006. An introduction to game theory for linguists. In *Game theory and pragmatics*, ed. Gerhard Jäger Anton Benz and Robert van Rooy, 1–82. Palgrave MacMillan.
- van den Berg, Martin. 1996. Some aspects of the internal structure of discourse: the dynamics of nominal anaphora. Doctoral Dissertation, ILLC, Universiteit van Amsterdam, Amsterdam.
- Blutner, Reinhard. 1998. Lexical pragmatics. *Journal of Semantics* 15:115–162.
- Blutner, Reinhard. 2000. Some aspects of optimality in natural language interpretation. *Journal of Semantics* 17:189–216.
- Blutner, Reinhard, Helen de Hoop, and Petra Hendriks. 2006. *Optimal communication*. Stanford: CSLI.
- Blutner, Reinhard, and Henk Zeevat. 2003. *Pragmatics in optimality theory*. Palgrave Macmillan.
- Boolos, George. 1984. To be is to be the value of a variable (or the values of some variables). *Journal of Philosophy* 81:430–450.
- Brasoveanu, Adrian. 2008a. Donkey pluralities: plural information states versus non-atomic individuals. *Linguistics and Philosophy* 31:129–209.
- Brasoveanu, Adrian. 2008b. Sentence-internal readings of *same/different* as quantifier-internal anaphora. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. Natasha Abner and Jason Bishop, 72–80.
- Brasoveanu, Adrian. 2008c. Structured nominal and modal reference. Doctoral Dissertation, Rutgers University.
- Brasoveanu, Adrian. 2009. Sentence-internal *different* as quantifier-internal anaphora. Ms., Santa Cruz.
- Brasoveanu, Adrian, and Donka Farkas. 2009. Scope and the grammar of choice. In *Proceedings of the tenth symposium of logic and language*. Budapest.
- Brisson, Christine. 1998. Distributivity, maximality, and floating quantifiers. Doctoral Dissertation, Rutgers University, New Brunswick.
- Brisson, Christine. 2003. Plurals, *all*, and the nonuniformity of collective predication. *Linguistics and Philosophy* 26:129–184.
- Brooks, Patricia J., and Martin D.S. Braine. 1996. What do children know about the universal quantifiers *all* and *each*? *Cognition* 60:235–268.
- Campbell, Donald T. 1958. Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science* 3:14–25.
- Carlson, Gregory. 1984. On the role of thematic roles in linguistic theory. *Linguistics* 22:259–279.
- Carlson, Gregory. 1987. Same and Different: some consequences for syntax and semantics. *Linguistics and Philosophy* 10:531–565.
- Carlson, Gregory. 1998. Thematic roles and the individuation of events. In *Events and grammar*, ed. Susan Deborah Rothstein, 35–51. Dordrecht: Kluwer.

- Casati, Roberto, and Achille Varzi. 1999. *Parts and places*. Cambridge, Massachusetts: The MIT Press.
- Cave, Kyle R., and Narcisse P. Bichot. 1999. Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review* 6:204–223.
- Champollion, Lucas. to appear. Cumulative readings of *every* do not provide evidence for events and thematic roles. In *Proceedings of the 17th Amsterdam Colloquium 2009*, ed. Maria Aloni and Katrin Schulz. Springer.
- Chierchia, Gennaro. 1989. Anaphora and attitudes *de se*. In *Semantics and contextual expressions*, ed. Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, 1–33. Dordrecht: Foris.
- Chierchia, Gennaro. 1995. *Dynamics of meaning: Anaphora, presupposition, and the theory of grammar*. Chicago: University of Chicago Press.
- Chomsky, Noam. 1973. Conditions on transformations. In *A festschrift for Morris Halle*, ed. Stephen R. Anderson and Paul Kiparsky, 232–285. New York: Holt, Rinehart and Winston.
- Chomsky, Noam. 1975. Questions of form and interpretation. In *Scope of American linguistics*, 159–196. Lisse: The Peter De Ridder Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Barriers*. Linguistic Inquiry Monograph 13. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz. Cambridge, Massachusetts: MIT Press.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads – a crosslinguistic perspective*. Oxford: Oxford University Press.
- Clark, Herbert H. 1977. Bridging. In *Thinking: Reading in cognitive science*, ed. P. N. Johnson-Laird and P. C. Watson, 411–420. Cambridge: Cambridge University Press.
- Crain, Stephen, and Mark Steedman. 1985. On not being led up the garden path: the use of context by the psychological syntax processor. In *Natural language parsing: Psychological, computational and theoretical perspectives*, ed. L. Karttunen David Dowty and Arnold Zwicky, 320–358. Cambridge: Cambridge University Press.
- Cresswell, Max, and Arnim von Stechow. 1982. *De Re* belief generalized. *Linguistics and Philosophy* 5:503–535.
- Dalrymple, Mary, Makoto Kanazawa, Yookyung Kim, Sam Mchombo, and Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21:159–210.
- Davidson, Donald. 1967. The logical form of action sequences. In *The logic of decision and action*, ed. Nicholas Rescher, 81–120. University of Pittsburgh Press.
- Dekker, Paul. 1993. Existential disclosure. *Linguistics and Philosophy* 16:561–587.
- Dekker, Paul, and Robert van Rooy. 2000. Bi-directional optimality theory: An appli-

- cation of game theory. *Journal of Semantics* 17:217–242.
- Dimitriadis, Alexis. 2000. Beyond identity: problems in pronominal and reciprocal anaphora. Doctoral Dissertation, University of Pennsylvania.
- Dimitriadis, Alexis. 2005. The event structure of irreducibly symmetric reciprocals. In *Event structures in linguistic form and interpretation*, ed. Johannes Dolling and Tatjana Heyde-Zybatow. Berlin: De Gruyter.
- Dotlačil, Jakub, and Øystein Nilsen. 2009. *The others* compared to *each other* – consequences for the theory of reciprocity. In *Proceedings of SALT XVIII, 2008*. Available at <http://hdl.handle.net/1813/13038>.
- Dotlačil, Jakub, and Øystein Nilsen. to app. A null theory of long-distance reciprocity. In *Proceedings of NELS 39, 2008*, ed. Suzi Lima, Kevin Mullin, and Brian Smith.
- Dowty, David. 1985. A unified indexical analysis of same and different: A response to stump and carlson. Ms., presented at University of Texas Workshop on Syntax and Semantics, Austin, Texas.
- Dowty, David. 1987. Collective predicates, distributive predicates, and all. In *Proceedings of the 3rd ESCOL*, ed. Fred Marshall. Columbus.
- Dowty, David. 2007. Compositionality as an empirical problem. In *Direct compositionality*, ed. Chris Barker and Pauline I. Jacobson, 23–101. Oxford: Oxford University Press.
- Elbourne, Paul. 2006. *Situations and individuals*. Cambridge, Massachusetts: MIT Press. Current Studies in Linguistics.
- Ellis, Willis D. 1938. *A source book of Gestalt psychology*. New York: The Humanities Press.
- Elworthy, David A. H. 1995. A theory of anaphoric information. *Linguistics and Philosophy* 18:297–332.
- Engdahl, Elisabet. 1986. *Constituent questions: The syntax and semantics of questions with special reference to Swedish*. Dordrecht: Reidel.
- Faller, Martina. 2007. The ingredients of reciprocity in cuzco quechua. *Journal of Semantics* 24:255–288.
- Farkas, Donka. 1997. Evaluation indices and scope. In *Ways of scope taking*, ed. Anna Szabolcsi, 183–215. Kluwer.
- Farkas, Donka. 2002. Varieties of indefinites. In *Proceedings of SALT 12*, ed. Brendan Jackson. Ithaca, NY: Cornell University.
- Farkas, Donka, and Henriëtte de Swart. 2007. Inclusive and exclusive plurals reconciled. In *Proceedings of the Sixteenth Amsterdam Colloquium*, ed. Paul Dekker Maria Aloni and Floris Roelofsen, 79–84. Amsterdam: University of Amsterdam.
- Featherson, Sam. 2007. Data in generative grammar: The stick and the carrot. *Theoretical linguistics* 33:269–318.
- Feldman, Jacob. 1997. Regularity-based perceptual grouping. *Computational Intelligence* 13:582–623.
- Feldman, Jacob. 1999. The role of objects in perceptual grouping. *Acta Psychologica* 102:137–163.

- Ferreira, Marcelo. 2007. Scope splitting and cumulativity. In *Proceedings of the workshop on quantifier modification, ESSLLI 2007*, ed. Rick Nouwen and Jakub Dotlačil. Dublin.
- Fiengo, Robert, and Howard Lasnik. 1973. The logical structure of reciprocal sentences in English. *Foundations of language* 9:447–468.
- Fodor, Janet Dean. 1970. The linguistics description of opaque contexts. Doctoral Dissertation, MIT, Cambridge, Massachusetts.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.
- Frazier, Lyn. 1978. On comprehending sentences: Syntactic parsing strategies. Doctoral Dissertation, University of Connecticut.
- Frazier, Lyn, Jeremy M. Pacht, and Keith Rayner. 1999. Taking on semantic commitments, II: collective versus distributive readings. *Cognition* 70:87–104.
- Frazier, Lyn, and Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14:178–210.
- Gajewski, Jon. 2005. Neg raising: Polarity and presupposition. Doctoral Dissertation, MIT.
- Gil, David. 1982. Quantifier scope, linguistic variation, and natural language semantics. *Linguistics and Philosophy* 5:421–472.
- Gillon, Brendan S. 1987. The readings of plural noun phrases in English. *Linguistics and Philosophy* 10:199–219.
- Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14:39–100.
- Grønn, Atle. 2008. Russian aspect as bidirectional optimization. In *Studies in Formal Slavic Linguistics*, ed. Franc Marušič and Rok Žaucer, 121–137. Frankfurt am Main: Peter Lang.
- Hamilton, David L., and Steven J. Sherman. 1996. Perceiving persons and groups. *Psychological Review* 103:336–355.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
- Heim, Irene, Howard Lasnik, and Robert May. 1991a. Reciprocity and plurality. *Linguistic Inquiry* 22:63–192.
- Heim, Irene, Howard Lasnik, and Robert May. 1991b. Reply: on ‘reciprocal scope’. *Linguistic Inquiry* 22:173–192.
- Hendriks, Petra, and Helen de Hoop. 2001. Optimality theoretic semantics. *Linguistics and Philosophy* 24:1–32.
- Hepple, Mark. 1990. The grammar and processing of order and dependency: A categorical approach. Doctoral Dissertation, University of Edinburgh, Edinburgh.
- Higginbotham, James. 1981. Reciprocal interpretation. *Journal of linguistic research* 1:97–117.
- Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16:547–931.

- Hintikka, Jaakko. 1962. *Knowledge and belief*. Ithaca, New York: Cornell University Press.
- Hintikka, Jaakko. 1974. Quantifiers vs. quantification theory. *Linguistic Inquiry* 5:153–177.
- Horn, Larry. 1989. *A natural history of negation*. University of Chicago Press. Reissued with new introduction (2001), David Hume Series, CSLI Publications, Stanford.
- Horn, Laurence. 1972. On the semantic properties of logical operators in English. Doctoral Dissertation, UCLA, Los Angeles.
- Horn, Laurence. 1984. Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In *Meaning, form and use in context: Linguistic applications*, ed. Deborah Schiffrin, 11–42. Georgetown: Georgetown University Press.
- Ioup, Georgette. 1975. Some universals for quantifier scope. In *Syntax and semantics* 4, ed. J. Kimball, 37–58. New York: Academic Press.
- Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22:117–184.
- Jäger, Gerhard. 2001. *Anaphora and Type-Logical Grammar*. Utrecht: UIL-OTS Working Papers.
- Jäger, Gerhard. 2002. Some notes on the formal properties of bidirectional optimality theory. *Journal of Logic, Language and Information* 11.
- Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal methods in the study of language*, ed. Jeroen Groenendijk, Theo Janssen, and Martin Stokhof, 277–322. Amsterdam: Mathematical Centre Tracts.
- Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic. Introduction to Model theoretic semantics of natural language, formal logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers.
- Kaplan, David. 1969. Quantifying in. *Synthese* 19:178–214.
- Katzir, Roni. 2008. Structural competition in grammar. Doctoral Dissertation, MIT, Cambridge, Massachusetts.
- Kaup, Barbara, Stephanie Kelter, and Christopher Habel. 2002. Representing referents of plural expressions and resolving plural anaphors. *Language and Cognitive Processes* 17:405–450.
- Keenan, Edward. 1992. Beyond the Frege boundary. *Linguistics and Philosophy* 15:199–221.
- Kemmer, Suzanne. 1993. *The middle voice*. Amsterdam/Philadelphia: John Benjamins.
- Kerem, Nir, Naama Friedmann, and Yoad Winter. 2009. Typicality effects and the logic of reciprocity. Talk at SALT 19, Ohio State University.
- Kratzer, Angelika. 2003. The event argument and the semantics of verbs. Ms. Four chapters available at <http://semanticsarchive.net>.
- Kratzer, Angelika. 2008. On the plurality of verbs. In *Event structures in linguistic form and interpretation*, ed. Johannes Dölling, Tatjana Heyde-Zybatow, and Martin Schäfer. Berlin/New York: Mouton de Gruyter.

- Krifka, Manfred. 1989. Nominal reference, temporal constitution, and quantification in event semantics. In *Semantics and contextual expressions*, ed. Renate Bartsch, Johan van Bentham, and Peter van Emde Boas. Dordrecht: Foris.
- Krifka, Manfred. 1990. Four thousand ships passed through the lock: object-induced measure functions on events. *Linguistics and Philosophy* 13:487–520.
- Krifka, Manfred. 1996. Pragmatic strengthening in plural predications and donkey sentences. In *Proceedings of SALT VI*, ed. J. Spence, 136–153. Cornell University.
- Krifka, Manfred. 2002. Be brief and vague! And how bidirectional optimality theory allows for verbosity and precision. In *Sounds and systems. Studies in structure and change. A Festschrift for Theo Vennemann*, ed. David Restle and Dietmar Zaefferer, 439–458. Berlin: Mouton de Gruyter.
- Krifka, Manfred. 2007. Approximate interpretation of number words: A case for strategic communication. In *Cognitive foundations of interpretation*, ed. Joost Zwarts Gerlof Bouma, Irene Krämer, 111–126. Amsterdam: KNAW.
- Landman, Fred. 1989a. Groups I. *Linguistics and Philosophy* 12:559–605.
- Landman, Fred. 1989b. Groups II. *Linguistics and Philosophy* 12:723–744.
- Landman, Fred. 1991. *Structures for semantics*. Dordrecht: Kluwer.
- Landman, Fred. 1995. Plurality. In *The Handbook of Contemporary Semantic Theory*, ed. Shalom Lappin. London: Blackwell.
- Landman, Fred. 2000. *Events and Plurality: The Jerusalem Lectures*. Dordrecht: Kluwer.
- Langendoen, Terence D. 1977. The logic of reciprocity. *Linguistic Inquiry* 9:177–197.
- Lasersohn, Peter. 1995. *Plurality, Conjunction and Events*. Dordrecht: Kluwer.
- Lasersohn, Peter. 1998. Generalized distributivity operators. *Linguistics and Philosophy* 21:83–93.
- Lasersohn, Peter. 1999. Pragmatic halos. *Language* 75:522–551.
- Levinson, Stephen. 2000. *Presumptive meanings. The theory of generalized conversational implicatures*. Cambridge, Massachusetts: MIT Press.
- Lewis, David. 1979. Attitudes *de dicto* and *de se*. *The Philosophical Review* 88:513–543.
- Link, Godehard. 1991. Plural. In *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, ed. Arnim von Stechow and Dieter Wunderlich, 418–440. Berlin: de Gruyter. English translation as Chapter 2 of Link (1998).
- Link, Godehard. 1998. *Algebraic Semantics in Language and Philosophy*. Stanford: CSLI Publications.
- Linnebo, Øystein. 2008. Plural quantification. In *Stanford encyclopedia of philosophy*, ed. Edward N. Zalta. The Metaphysics Research Lab, Stanford University.
- Löbner, Sebastian. 2000. Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences. *Linguistics and Philosophy* 23:213–308.
- Maier, Emar. 2006. Belief in context. Doctoral Dissertation, Radboud University Nijmegen.
- Maslova, Elena. 2007. Reflexive encoding of reciprocity: Cross-linguistic and

- language-internal variation. In *Reciprocals and reflexives: Cross-linguistic and theoretical explorations*, ed. Ekkehard König and Volker Gast. Berlin: Mouton de Gruyter.
- Matsui, Tomoko. 2000. *Bridging and relevance*. Amsterdam/Philadelphia: John Benjamins.
- Matsumoto, Yo. 1995. The conversational condition on horn scales. *Linguistics and Philosophy* 18:21–60.
- Matushansky, Ora. 2008. More of the same. Ms., presented at GLOW, Newcastle.
- de Mey, Sjaak. 1981. The dependant plural and the analysis of tense. In *Proceedings of NELS 11*. Amherst: GLSA: University of Massachusetts.
- Miyamoto, Y., and Stephen Crain. 1991. Children's interpretation of pronouns: Collective versus distributive. Paper presented at the 16th Annual Boston University Conference on Language Development, Boston.
- Moltmann, Friederike. 1992. Reciprocals and *sameldifferent*: Towards a semantic analysis. *Linguistics and Philosophy* 15:411–462.
- Moortgat, Michael. 1996. Generalized quantification and discontinuous type constructors. In *Discontinuous constituency*, ed. Wietske Sjitsma and Arthur von Horck. Berlin: De Gruyter.
- Murray, Sarah. 2008. Reflexivity and reciprocity with(out) underspecification. In *Proceedings of SuB 12*, ed. Atle Grønn. University of Oslo: ILOS.
- Muskens, Reinhard Anton. 1996. Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy* 19:143–186.
- Nedjalkov, Vladimir P., Emma Š. Geniušienė, and Zlatka Guentchéva. 2007. *Reciprocal constructions*. Amsterdam/Philadelphia: John Benjamins.
- Nouwen, Rick. 2003. Plural pronominal anaphora in context: dynamic aspects of quantification. Doctoral Dissertation, UIL-OTS, Utrecht University.
- Nouwen, Rick. 2007. On dependent pronouns and dynamic semantics. *Journal of Philosophical Logic* 36:123–154.
- Palmer, Stephen E. 1992. Common region: A new principle of perceptual grouping. *Cognitive Psychology* 24:436–447.
- Palmer, Stephen E. 1999. *Vision science: Photons to Phenomenology*. Cambridge, Massachusetts: MIT Press.
- Palmer, Stephen E., and Irvin Rock. 1994. Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review* 1:29–55.
- Parikh, Prashant. 2000. Communication, meaning, and interpretation. *Linguistics and Philosophy* 23.
- Parsons, Terence. 1990. *Events in the semantics of english: A study in subatomic semantics*. Cambridge, Massachusetts: MIT Press.
- Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in Discourse Representation Theory and the theory of generalized quantifiers*, ed. Jeroen Groenendijk, Dick de Jong, and Martin Stokhof, 115–143. Dordrecht: Foris Publications.
- Partee, Barbara, and Emmon Bach. 1981. Quantification, pronouns, and vp anaphora. In *Formal Methods in the Study of Language, Volume 1*, ed. Jeroen Groe-

- nendijk, Theo Janssen, and Martin Stokhof, 445–481.
- Penka, Doris, and Arnim von Stechow. 2001. Negativen indefinia unter modalverben. *Linguistische Berichte* 263–286.
- Penka, Doris, and Hedde Zeijlstra. 2005. Negative indefinites in Dutch and German. Ms.
- Pinker, Steven. 1989. *Learnability and cognition: the acquisition of argument structure*. Cambridge, Massachusetts: MIT Press.
- Piwek, Paul, and Emiel Krahmer. 2000. Presuppositions in context: Constructing bridges. In *Formal aspects of context*, ed. Pierre Bonzon, Marcos Cavalcanti, and Rolf Tomas, 85–106. Dordrecht: Kluwer.
- Prince, Alan, and Paul Smolensky. 1993. Optimality theory: Constraint interaction in generative grammar. Technical report, Rutgers University Center for Cognitive Science.
- Pritchett, Bradley L. 1992. *Grammatical competence and parsing performance*. Chicago, IL: The University of Chicago Press.
- Pylkkänen, Liina, and Brian McElree. 2006. The syntax-semantic interface: On-line composition of sentence meaning. In *Handbook of psycholinguistics*, 537–577. New York: Elsevier.
- Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335–397.
- Reinhart, Tanya. 2006. *Interface strategies: Optimal and costly computations*. Cambridge, Massachusetts: MIT Press.
- Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.
- Reinhart, Tanya, and Tal Siloni. 2005. The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry* 36:389–436.
- Reuland, Eric. 2008. Anaphoric dependencies: How are they encoded? Towards a derivation-based typology. In *Reciprocals and reflexives: Theoretical and typological explorations*, ed. Ekkehard König and Volker Gast, 503–561. Berlin: Mouton de Gruyter.
- Reuland, Eric, and Yoad Winter. 2009. Binding without identity: towards a unified semantics for bound and exempt anaphors. In *Proceedings of 7th Discourse Anaphora and Anaphor Resolution Colloquium*.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Massachusetts: MIT Press.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. L. Haegeman, 281–337. Dordrecht: Kluwer.
- Roberts, Craige. 1990. *Modal subordination, anaphora, and distributivity*. New York and London: Garland Publishing.
- Roberts, Craige. 1991. Distributivity and reciprocal distributivity. In *Proceedings of SALT I*, 209–229. Ithaca: Cornell University.
- Roelofsen, Floris. 2008. Anaphora resolved. Doctoral Dissertation, ILLC, Amsterdam.
- van Rooy, Robert. 2003. Conversational implicatures and communication theory. In *Current and new directions in discourse and dialogue*, ed. Jan van Kuppevelt and Ronnie W. Smith, 283–305. Dordrecht: Kluwer.

- van Rooy, Robert. 2004. Signalling games select horn strategies. *Linguistics and Philosophy* 27:493–527.
- Rubinstein, Aynat. to appear. Groups in the semantics of reciprocal verbs. In *Proceedings of NELS 38*, ed. M. Abdurrahman, A. Schardl, and M. Walkow. Amherst, MA: GLSA.
- Ruys, Eddie. 1992. The scope of indefinites. Doctoral Dissertation, Utrecht University, Utrecht.
- Sabato, Sivan, and Yoad Winter. 2005a. Against partitioned readings of reciprocals. In *Proceedings of the amsterdam colloquium 2005*.
- Sabato, Sivan, and Yoad Winter. 2005b. From semantic restrictions to reciprocal meanings. In *Proceedings of FG-MOL*.
- van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.
- Sauerland, Uli. 1998. Plurals, derived predicates, and reciprocals. In *The interpretative tract: MIT working papers in linguistics*. Cambridge, Massachusetts.
- Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27:367–391.
- Sauerland, Uli, and Paul Elbourne. 2005. Total reconstruction, PF-movement and derivational order. *Linguistic Inquiry* 33:283–319.
- Scha, Remko. 1981. Distributive, collective, and cumulative quantification. In *Formal methods in the study of language*, ed. T. Janssen and M. Stokhof, 483–512. Amsterdam: Mathematical Centre Tracts.
- Schein, Barry. 1993. *Plurals and events*. Cambridge, Massachusetts: MIT Press.
- Schlenker, Philippe. 2006. Scopal independence: A note on branching and wide scope readings of indefinites and disjunctions. *Journal of Semantics* 23:281–314.
- Scholl, Brian J. 2001. Objects and attention: the state of the art. *Cognition* 80:1–46.
- Schwarzschild, Roger. 1992. Plurals, presuppositions and the sources of distributivity. *Natural Language Semantics* 2:201–248.
- Schwarzschild, Roger. 1996. *Pluralities*. Dordrecht: Kluwer.
- Schwarzschild, Roger. 1999. GIVENness, avoid F and other constraints on the placement of focus. *Natural Language Semantics* 7:141–177.
- Shannon, Claude. 1948. A mathematical theory of communication. *Bell Systems Technical Journal* 27:379–432, 623–656.
- Sher, Gila. 1990. Ways of branching quantifiers. *Linguistics and Philosophy* 13.
- Sportiche, Dominique. 2005. Division of labor between merge and move: Strict locality of selection and apparent reconstruction paradoxes. Ms., Los Angeles.
- Sternefeld, Wolfgang. 1998. Reciprocity and cumulative predication. *Natural Language Semantics* 6:303–337.
- de Swart, Henriëtte. 2000. Scope ambiguities with negative quantifiers. In *Reference and anaphoric relations*, ed. Klaus von Heusinger and Urs Egli, 109–132. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Szabolcsi, Anna. 1989. Bound variables in syntax (are there any?). In *Semantics and contextual expressions*, ed. Renate Bartsch, Johan van Bentham, and Peter van Emde Boas. Dordrecht: Foris.

- Szabolcsi, Anna. 1992. Combinatory grammar and projection from the lexicon. In *Lexical matters*, ed. Anna Szabolcsi and Ivan Sag. Stanford: CSLI.
- Szabolcsi, Anna. 1997a. Strategies for scope taking. In *Ways of scope taking*, ed. Anna Szabolcsi, 109–154. Dordrecht: Kluwer.
- Szabolcsi, Anna. 1997b. *Ways of scope taking*. Dordrecht: Kluwer.
- Taub, Alison. 1989. Collective predicates, aktionsarten and all. In *Papers on Quantification*, ed. Emmon Bach, Angelika Kratzer, and Barbara Partee. University of Massachusetts at Amherst.
- Thornton, Rosalind, and Kenneth Wexler. 1999. *Principle B, VP ellipsis, and interpretation in child grammars*. Cambridge, Massachusetts: MIT Press.
- Tovena, Lucia, and Marleen van Peteghem. 2002. Facets of ‘different’ in french: différent and autre. In *Actes de cssp 2001: Empirical issues in formal syntax and semantics*, 63–79. Paris: Presses Universitaires de la Sorbonne.
- Tunstall, Susanne. 1998. The interpretation of quantifiers: Semantics and processing. Doctoral Dissertation, University of Massachusetts, Amherst.
- VanLehn, Kurt A. 1978. Determining the scope of english quantifiers. Technical Report, Cambridge, Massachusetts.
- Verkuyl, Henk J., and Jaap van der Does. 1996. The semantics of plural noun phrases. In *Quantifiers, logic and language*, ed. Jaap van der Does and Jan van Eyck, 337–374. Stanford: CSLI Publications.
- Wang, Linton, Eric McCready, and Nicholas Asher. 2006. Information dependency in quantificational subordination. In *Where semantics meets pragmatics*, ed. Ken Turner and Klaus von Heusinger, 267–307. Amsterdam: Elsevier.
- Wertheimer, Max. 1923. Untersuchungen zur Lehre von der Gestalt. *Psychologische Forschung* 4:301–350. Partial translation in Ellis (1938).
- Williams, Edwin. 1991. Reciprocal scope. *Linguistic Inquiry* 22:159–173.
- Williams, Edwin. 1997. Blocking and anaphora. *Linguistic Inquiry* 28:577–628.
- Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20:399–467.
- Winter, Yoad. 2000. Distributivity and dependency. *Natural Language Semantics* 8:27–69.
- Winter, Yoad. 2001a. *Flexibility Principles in Boolean Semantics*. Cambridge, Massachusetts: MIT Press.
- Winter, Yoad. 2001b. Plural predication and the strongest meaning hypothesis. *Journal of Semantics* 18:333–365.
- Wolff, Phillip. 2003. Direct causation in the linguistic coding and individuation of causal events. *Cognition* 88:1–48.
- Zeevat, Henk. 2000. The asymmetry of optimality theoretic syntax and semantics. *Journal of Semantics* 17:243–262.
- Zeijlstra, Hedde. 2004. Sentential negation and negative concord. Doctoral Dissertation, University of Amsterdam, Amsterdam.
- Zweig, Eytan. 2008. Dependent plurals and plural meaning. Doctoral Dissertation, New York University, New York.

Inequalities in Games discussed in Chapter 3

$\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}$

- (1) Assumptions for $\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}$:
- a. $\text{Compl}(f_t)$ is greater than the complexity of f_g , that is, *the boys together* is more complex than *the boys*
 - b. $\text{Compl}(f_d)$ is not greater than the complexity of f_t , that is, *each boy* is not more complex than *the boys together*
 - c. $c(m_d)$ is greater than the cost of m_t

We want to show that this game has the winning strategy where f_d is interpreted as m_d and f_g is interpreted as m_t . We have to show that this strategy has a higher utility than the strategy where f_g is interpreted as m_d and f_t is interpreted as m_t . Thus, we need to show the following:

$$\frac{1}{c(m_d)\text{Compl}(f_d)} + \frac{1}{c(m_t)\text{Compl}(f_g)} > \frac{1}{c(m_d)\text{Compl}(f_g)} + \frac{1}{c(m_t)\text{Compl}(f_t)}$$

We use the following notation: $x = c(m_d)$, $y = \text{Compl}(f_d)$, $z = \text{Compl}(f_t)$, $\text{Compl}(f_g) = c(m_t) = 1$. Thus, we rewrite the inequality as follows:

$$1/xy + 1 > 1/x + 1/z$$

Furthermore, given the assumptions above, the following must hold:

$$\begin{aligned}
x &> 1 && \text{(assumption (1c))} \\
z &> 1 && \text{(assumption (1a))} \\
y &\leq z && \text{(assumption (1b))}
\end{aligned}$$

The inequality is true for any numbers that satisfy these conditions. That is, under the assumptions given above, any chosen numbers for x, y and z will give the correct result.

GAME_{CQ-DP×G-DP}

- (2) Assumptions for GAME_{CQ-DP×G-DP}:
- $Compl(f_c)$ and $Compl(f_t)$ are greater than $Compl(f_g)$
 - $Compl(f_c)$ is not greater than $Compl(f_t)$
 - $c(m_d)$ is greater than $c(m_t)$

We want to show that this game has the winning strategy where f_c is interpreted as m_d and f_g is interpreted as m_t . We have to show that this strategy has a higher utility than (i) the strategy where f_g is interpreted as m_d and f_t is interpreted as m_t or (ii) the strategy where f_g is interpreted as m_d and f_c is interpreted as m_t . Thus, we need to show the following:

$$\frac{1}{c(m_d)Compl(f_c)} + \frac{1}{c(m_t)Compl(f_g)} > \frac{1}{c(m_d)Compl(f_g)} + \frac{1}{c(m_t)Compl(f_t)}$$

If we let $x = c(m_d), y = Compl(f_c), z = Compl(f_t), Compl(f_g) = c(m_t) = 1$, the game is identical to GAME_{DQ-DP×G-DP}, so we have that $x > 1, y > 1, z \geq y$. That is, under the assumptions given above, any chosen numbers for x, y and z will give the correct result. The second inequality we have to consider is the following:

$$\frac{1}{c(m_d)Compl(f_c)} + \frac{1}{c(m_t)Compl(f_g)} > \frac{1}{c(m_d)Compl(f_g)} + \frac{1}{c(m_t)Compl(f_c)}$$

If we let $x = c(m_d), y = Compl(f_c), Compl(f_g) = c(m_t) = 1$, the game can be rewritten as follows:

$$\frac{1}{xy} + 1 > \frac{1}{x} + \frac{1}{y}$$

We furthermore require that $x > 1$ and $y > 1$ (since these are costly meaning/expression). The inequality is true for any numbers that satisfy these conditions. Thus, under the assumptions given above we derive that the winning strategy is the one where f_c is interpreted as m_d and f_g is interpreted as m_t .

GAME_{DQ-DP×CQ-DP×G-DP} – the first version

(3) Assumptions for $\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP} \times \text{G-DP}}$:

- a. $\text{Compl}(f_g) < \text{Compl}(f_c) < \text{Compl}(f_{ct})$, that is, *the boys* is less complex than *all the boys* which is less complex than *all the boys together*
- b. $\text{Compl}(f_g) < \text{Compl}(f_t)$, that is, *the boys* is less complex than *the boys together*
- c. $\text{Compl}(f_d) \leq \text{Compl}(f_t)$, that is, *each boy* is not more complex than *the boys together*
- d. $\text{Compl}(f_c) \leq \text{Compl}(f_t)$, that is, *all the boys* is not more complex than *the boys together*
- e. $c(m_t) < c(m_d)$, that is, a distributive reading is more costly than non-distributive readings
- f. $\text{Compl}(f_d) < \text{Compl}(f_c)$, that is, *each boy* is less complex than *all the boys*

We want to show that the winning strategy is the following:

- (4) f_d interpreted as m_d , f_g interpreted as m_g

This strategy has a higher utility than the one where f_g expresses m_d and f_t expresses m_t . This has been shown in $\text{GAME}_{\text{DQ-DP} \times \text{G-DP}}$. It should be straightforward why any variant where either m_g or m_d is expressed by f_c has a lower utility than the winning strategy. It should also be straightforward why the strategy where f_{ct} or f_t expresses m_g is always worse. Thus, (4) is the winning strategy.

We further need to show that in the following two strategies in the game, the first one is preferred over the second one, that is, the utility of the first one is higher.

- (5) a. f_c interpreted as m_d , f_g interpreted as m_t
 b. f_g interpreted as m_d , f_c interpreted as m_t

Thus, we need to show the following:

$$\frac{1}{c(m_d)\text{Compl}(f_c)} + \frac{1}{c(m_t)\text{Compl}(f)} > \frac{1}{c(m_d)\text{Compl}(f_g)} + \frac{1}{c(m_t)\text{Compl}(f_t)}$$

If we let $x = c(m_d)$, $y = \text{Compl}(f_c)$, $\text{Compl}(f_g) = c(m_t) = 1$, the game is identical to $\text{GAME}_{\text{CQ-DP} \times \text{G-DP}}$, so it holds for any x, y greater than 1. That is, under the assumptions given above, any chosen numbers will give the correct result that (5a) has a higher utility than (5b).

$\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP} \times \text{G-DP}}$ – the second version

(6) Assumptions for the modified $\text{GAME}_{\text{DQ-DP} \times \text{CQ-DP} \times \text{G-DP}}$:

- a. $\text{Compl}(f_g) < \text{Compl}(f_c) < \text{Compl}(f_{ct})$, that is, *the boys* is less complex than *all the boys* which is less complex than *all the boys together*

- b. $Compl(f_g) < Compl(f_t)$, that is, *the boys* is less complex than *the boys together*
- c. $Compl(f_d) \leq Compl(f_t)$, that is, *each boy* is not more complex than *the boys together*
- d. $Compl(f_c) \leq Compl(f_t)$, that is, *all the boys* is not more complex than *the boys together*
- e. $c(m_t) < c(m_d)$, that is, a distributive reading is more costly than non-distributive readings
- f. *the boys together* does not express $m_t[100\%]$

Competition between (151) and (154a)

Consider the following two strategies. We want to show that the first one has a higher utility than the second one.

- (7) a. f_d interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_g interpreted as $m_t[90 - 100\%]$
- b. f_c interpreted as m_d , f_{ct} interpreted as $m_t[100\%]$, f_g is interpreted as $m_t[90 - 100\%]$

Let $x = c(m_d)$, $y = Compl(f_d)$, $z = Compl(f_c)$, $w = Compl(f_{ct})$, $Compl(f_g) = c(m_t) = 1$. To simplify the matter, I ignore probabilities and assume that the costs of $m_t[100\%]$ and $m_t[90 - 100\%]$ do not differ. The difference in probabilities is in fact irrelevant here. We want to show the following (in the formula, I abbreviate $Compl$ to C):

$$\begin{aligned}
 \frac{1}{c(m_d)C(f_d)} + \frac{1}{\frac{1}{xy} + \frac{1}{z} + 1} + \frac{1}{c(m_t)C(f_g)} &> \frac{1}{c(m_d)C(f_c)} + \frac{1}{\frac{c(m_t)C(f_{ct})}{xz} + \frac{1}{w} + 1} + \frac{1}{c(m_t)C(f_g)} \\
 &> \frac{1}{\frac{1}{xy} + \frac{1}{z}} &> \frac{1}{\frac{1}{xz} + \frac{1}{w}}
 \end{aligned}$$

We furthermore assume the following:

- $x > 1$ (assumption (6e))
- $z > 1$ (assumption (6a))
- $w > 1$ (assumption (6a))
- $w > y$ (assumption (6a+b+c))
- $w > z$ (assumption (6a))

The solution is any numbers for x, y, z, w that satisfy these assumptions and where y is 1 or greater than 1. The last condition must hold since f_d cannot be less costly than f_g (the minimal cost). Thus, (7a) is preferred over (7b).

Competition between (151) and (154c)

We want to show that the first strategy of (8) is preferred over the second strategy.

- (8) a. f_d interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_g interpreted as $m_t[90 - 100\%]$
 b. f_g interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_t is interpreted as $m_t[90 - 100\%]$

Let $x = c(m_d)$, $y = Compl(f_d)$, $v = Compl(f_c)$, $z = Compl(f_t)$, $Compl(f_g) = 1$. To simplify the matter, I ignore probabilities and assume that the cost of $m_t[100\%]$ and $m_t[90 - 100\%]$ does not differ. We want to show the following (in the formula, I abbreviate *Compl* to *C*):

$$\begin{aligned} \frac{1}{c(m_d)C(f_d)} + \frac{1}{c(m_t)C(f_c)} + \frac{1}{c(m_t)C(f_g)} &> \frac{1}{c(m_d)C(f_g)} + \frac{1}{c(m_t)C(f_c)} + \frac{1}{c(m_t)C(f_t)} \\ \frac{1}{1/xy + 1/v + 1} &> \frac{1}{1/x + 1/v + 1/z} \\ \frac{1}{1/xy + 1} &> \frac{1}{1/x + 1/z} \end{aligned}$$

This inequality is identical to the one in $\text{GAME}_{\text{D-DP} \times \text{G-DP}}$. Given the assumptions above, we derive that $x > 1$, $z > 1$, $z \geq y$. That is, under the assumptions given above, any chosen numbers for x , y and z will give the correct result.

Competition between (151) and (154e)

We want to show that the first strategy of (9) is preferred over the second strategy.

- (9) a. f_d interpreted as m_d , f_c interpreted as $m_t[100\%]$, f_g interpreted as $m_t[90 - 100\%]$
 b. f_c interpreted as m_d , f_g interpreted as $m_t[100\%]$, f_t is interpreted as $m_t[90 - 100\%]$

Let $x = c(m_d)$, $y = Compl(f_d)$, $z = Compl(f_c)$, $w = Compl(f_t)$, $Compl(f_g) = c(m_t) = 1$. To simplify the matter, I ignore probabilities and assume that the cost of $m_t[100\%]$ and $m_t[90 - 100\%]$ does not differ. We want to show the following (in the formula, I abbreviate *Compl* to *C*):

$$\begin{aligned} \frac{1}{c(m_d)C(f_d)} + \frac{1}{\frac{1}{xy} + \frac{1}{z} + 1} + \frac{1}{c(m_t)C(f_g)} &> \frac{1}{c(m_d)C(f_c)} + \frac{1}{\frac{1}{xz} + 1 + \frac{1}{w}} + \frac{1}{c(m_t)C(f_t)} \\ &> \frac{1}{\frac{1}{xy} + \frac{1}{z}} &> \frac{1}{\frac{1}{xz} + \frac{1}{w}} \end{aligned}$$

We furthermore assume the following:

- $x > 1$ (assumption (6e))
 $z > 1$ (assumption (6a))
 $w > 1$ (assumption (6b))
 $w \geq y$ (assumption (6c))
 $w \geq z$ (assumption (6d))

The solution is any number that satisfies these conditions as long as y is 1 or greater than 1. The last condition must hold since f_d cannot be less costly than f_g (the minimal cost).

Questionnaires

Questionnaire testing distributivity and antecedents of expressions of (non-)identity (Chapter 2)

Scenario 1

Jan, Wim en Kees houden met zijn drieën een paaltjesvoetbaltoernooi. Ze spelen drie rondes. Voor elke ronde geldt dat degene wiens paaltje het langste blijft staan heeft gewonnen, en dat de andere twee hebben verloren. Na de drie rondes was de einduitslag als volgt: Elk van de jongens had n ronde gewonnen.

Geen ander toernooi heeft plaatsgevonden en behalve Jan, Wim en Kees waren er geen andere spelers.

Translation: Jan, Wim, and Kees had a tournament in paaltjesvoetbal. They played three rounds. In each round, one of the players would win, and the other two would loose. After the three rounds, the results were the following: Each player had won exactly once. There were no other tournaments, and no other players had participated, except for Jan, Wim, and Kees.

Test items

- (10) Iedere speler heeft de anderen precies één keer verslagen.
 every/each player has the others exactly one time beaten
 'Every player beat the others exactly once.'
- (11) Elke speler heeft de anderen precies n keer verslagen.
 every/each player has the others exactly one time beaten
 'Every player beat the others exactly once.'
- (12) Alle spelers hebben de anderen precies één keer verslagen.
 all players have the others exactly one time beaten
 'All players beat the others exactly once.'

- (13) De spelers hebben de anderen precies één keer verslagen.
the players have the others exactly one time beaten
'The players beat the others exactly once.'
- (14) Jan, Wim en Kees hebben de anderen precies één keer verslagen.
Jan, Wim and Kees have the others exactly one time beaten
'Jan, Wim and Kees beat the others exactly once.'
- (15) Geen speler heeft de anderen meer dan één keer verslagen.
No player has the others more than one time beaten
'No player beat the others more than once.'
- (16) Iedere speler heeft twee rondes verloren.
every/each player has two rounds lost
'Every player lost two rounds.'
- (17) De spelers hebben twee rondes verloren.
the players have two rounds lost
'The players lost two rounds.'
- (18) Jan, Wim en Kees hebben twee rondes verloren.
Jan, Wim and Kees have two rounds lost
'Jan, Wim and Kees lost two rounds.'
- (19) Geen speler heeft meer dan één keer gewonnen.
no player has more than one time won
'No player won more than one time.'

Control items

- (20) Iedere speler heeft de anderen meer dan één keer verslagen.
every player has the others more than once beaten
'Every player beat the others more than once.'
- (21) Er zijn drie spelers in het paaltjesvoetbaltoernooi.
there are three players in the paaltjesvoetbaltoernooi
'There are three players in paaltjesvoetbaltoernooi.'

Scenario 2

In het land Ukhbar leven alle inwoners in één van de drie steden, die zijn vernoemd naar de burgemeesters: Appel, Sinaasappel of Peer. Dit zijn de enige steden in het land. In elke stad is er een kerk die de trots is van alle inwoners. De steden liggen in een driehoek. Elke stad ligt precies 50 km van de andere twee af.

Translation: In the country Ukhbar all inhabitants live in one of the three

towns which are named after their mayors: Appel, Sinaasappel of Peer. These are the only towns in the country. In each town there is a church which make all the town inhabitants proud. The towns lie in a triangle. Each town is exactly 50 km from the other two towns.

Test items

- (22) Iedere stad ligt precies 50 km van de andere.
every/each town lies exactly 50 km from the others
'Every town lies exactly 50 km from the others.'
- (23) Elke stad ligt precies 50 km van de andere.
every/each town lies exactly 50 km from the others
'Every town lies exactly 50 km from the others.'
- (24) Alle steden liggen precies 50 km van de andere.
all towns lie exactly 50 km from the others
'All towns lie exactly 50 km from the others.'
- (25) De steden liggen precies 50 km van de andere.
the towns lie exactly 50 km from the others
'The towns lie exactly 50 km from the others.'
- (26) Appel, Sinaasappel en Peer liggen precies 50 km van de andere.
Appel, Sinaasappel and Peer lie exactly 50 km from the others
'Appel, Sinaasappel and Peer lie exactly 50 km from the others.'
- (27) Geen stad ligt meer dan 50 km van de andere.
no town lies exactly 50 km from the others
'No town lies exactly 50 km from the others.'
- (28) Iedere stad heeft een kerk.
every/each town has a church
'Every town has a church.'
- (29) De steden hebben een kerk.
the towns have a church
'The towns have a church.'
- (30) Appel, Sinaasappel en Peer hebben een kerk.
Appel, Sinaasappel and Peer have a church
'Appel, Sinaasappel and Peer have a church.'
- (31) Geen stad heeft meer dan één kerk.
no town has more than one church
'No town has more than one church.'

Control items

- (32) Iedere stad ligt meer dan 100 km van de andere.
every/each town lies more than 100 km from the others
'Every town lies more than 100 km from the others.'
- (33) Er zijn drie steden in het land Ukhbar.
there are three towns in the country Ukhbar
'There are three towns in the country Ukhbar.'

Scenario 3

De matrozen Jip, Jaap en Joop kwamen terecht op een onbewoond eiland. Na een tijdje kregen ze ruzie en gingen ze ieder hun eigen weg. Jip ging in het scheepswrak wonen, Jaap nam zijn intrek in een grot, en Joop vond een verlaten hut. Ieder van hen dacht dat zijn nieuwe behuizing de slechtste was. Hierdoor werden ze jaloers op elkaar.

Translation: The sailors Jip, Jaap, and Joop were stranded on an uninhabited island. After a while, they started quarreling, and went their separate ways. Jip went to live in the shipwreck, Jaap moved into a cave, and Joop found himself an abandoned hut. Each of them thought his own dwelling was the worst one. Therefore they were jealous of each other.

Test items

- (34) Iedere matroos was jaloers op de anderen.
every/each sailor was jealous on the others
'Every sailor was jealous of the others.'
- (35) Elke matroos was jaloers op de anderen.
every/each sailor was jealous on the others
'Every sailor was jealous of the others.'
- (36) Alle matrozen waren jaloers op de anderen.
all sailors were jealous on the others
'All sailors were jealous of the others.'
- (37) De matrozen waren jaloers op de anderen.
the sailors were jealous on the others
'The sailors were jealous of the others.'
- (38) Jip, Jaap en Joop waren jaloers op de anderen.
Jip, Jaap and Joop were jealous on the others
'Jip, Jaap and Joop were jealous of the others.'
- (39) Geen matroos hield van de anderen.
no sailor loved from the the others
'No sailor loved the others.'

- (40) Iedere matroos vond een onderkomen.
every sailor found a dwelling
'Every sailor found a place to stay.'
- (41) De matrozen vonden een onderkomen.
the sailors found a dwelling
'The sailors found a place to stay.'
- (42) Jip, Jaap en Joop vonden een onderkomen.
Jip, Jaap and Joop found a dwelling
'Jip, Jaap and Joop found a place to stay.'
- (43) Geen matroos vond meer dan één onderkomen.
no sailor found more than one dwelling
'No sailor found more than one place to stay.'

Control items

- (44) Iedere matroos hield van de anderen.
every/each sailor loved from the others
'Every sailor loved the others.'
- (45) Er waren drie matrozen op het eiland.
there were three sailors on the island
'There were three sailors on the island.'

Scenario 4

Drie jongens, Jan, Wim en Kees, en drie meisjes, Janneke, Suzan en Iris, gaan met school naar een tentoonstelling van een bekende plaatselijke kunstenaar. Van hun docenten moeten ze na de tentoonstelling allemaal hun favoriete schilderij aankruisen op een formulier.

Jan, Wim en Kees willen er snel vanaf zijn. Ze kruisen met hun ogen dicht een schilderij aan op het formulier. Jan kruist schilderij A aan, Wim kruist schilderij B aan, en Kees kruist schilderij C aan.

In tegenstelling tot de jongens, zijn Janneke, Suzan en Iris veel voorzichtiger. Ze bespreken alle schilderijen en aan het einde besluiten ze dat ze allemaal hetzelfde schilderij het mooist vinden. Ze kruisen allemaal schilderij D aan.

Translation: Three boys, Jan, Wim and Kees, and three girls, Janneke, Suzan and Iris, go with school to an exhibition of a famous painter. The teacher assigns them the task to choose their favorite painting by putting marking it in their form.

Jan, Wim and Kees are quickly finished. They randomly mark a painting on the form with their eyes closed. Jan marks painting A, Wim marks painting B, Kees marks painting C.

Contrary to the boys, Janneke, Suzan and Iris proceed more carefully. They discuss all the paintings and in the end they decide they all like the same picture. They all mark painting D.

Test items

- (46) Iedere jongen heeft een ander schilderij gekozen.
every/each boy has a different picture chosen
'Every boy chose a different picture.'
- (47) Alle jongens hebben een ander schilderij gekozen.
all boys have a different picture chosen
'All boys chose a different picture.'
- (48) De jongens hebben een ander schilderij gekozen.
the boys have a different picture chosen
'The boys chose a different picture.'
- (49) Jan, Wim en Kees hebben een ander schilderij gekozen.
Jan, Wim and Kees have a different picture chosen
'Jan, Wim and Kees chose a different picture.'
- (50) Geen meisje heeft een ander schilderij gekozen.
no girl has a different picture chosen
'No girl chose a different picture.'
- (51) Iedere jongen heeft een verschillend schilderij gekozen.
every/each boy has a different picture chosen
'Every boy chose a different picture.'
- (52) Alle jongens hebben een verschillend schilderij gekozen.
all boys have a different picture chosen
'All boys chose a different picture.'
- (53) De jongens hebben een verschillend schilderij gekozen.
the boys have a different picture chosen
'The boys chose a different picture.'
- (54) Jan, Kees en Wim hebben een verschillend schilderij gekozen.
Jan, Wim and Kees have a different picture chosen
'Jan, Wim and Kees chose a different picture.'
- (55) Geen meisje heeft een verschillend schilderij gekozen.
no girl has a different picture chosen
'No girl chose a different picture.'
- (56) De meisjes hebben hetzelfde schilderij gekozen.
the girls have the-same picture chosen
'The girls chose the same picture.'

Control items

- (57) Iedere meisje heeft een ander schilderij gekozen.
 every/each girl has a different picture chosen
 'Every girl chose a different picture.'

Scenario 5

Het land Ukhbar is verdeeld in twee provincies: het noorden, en het zuiden. Het noorden heeft drie steden: Appel, Sinaasappel en Peer. Het zuiden heeft ook drie steden: Den, Berk en Palm. Er zijn geen andere steden of dorpen in het land Ukhbar.

Vroeger had elke stad in Ukhbar zijn eigen eenheid voor lengte. In het noorden konden ze het niet eens worden welke eenheid de beste was, dus zelfs vandaag de dag gebruikt elke stad nog zijn eigen lengte-eenheid. In het zuiden konden de steden het daarentegen wel eens worden over een lengte-eenheid. Alle steden daar gebruiken die eenheid, en de andere eenheden zijn afgeschaft.

Translation: The country Ukhbar is divided into two provinces: the north one, and the south one. The north province has three towns: Appel, Sinaasappel and Peer. The south province also has three towns: Den, Berk and Palm. There are no other towns in Ukhbar.

In the past, each town in Ukhbar had its own length unit. In the north, they could not agree which of these units is the best one. Therefore, each city kept its own length unit even nowadays. In the south province the cities agreed on one length unit. All the towns there use only one length unit, and the other ones are forgotten.

Test items

- (58) Iedere stad in het noorden heeft een andere lengte-eenheid.
 every/each town in the north has a different length-unit
 'Every town in the North has a different metrical system.'
- (59) Alle steden in het noorden hebben een andere lengte-eenheid.
 all towns in the north have a different length-unit
 'All towns in the North have a different metrical system.'
- (60) De steden in het noorden hebben een andere lengte-eenheid.
 the towns in the north have a different length-unit
 'The towns in the North have a different metrical system.'
- (61) Appel, Sinaasappel en Peer hebben een andere lengte-eenheid.
 Appel, Sinaasappel and Peer have a different length-unit
 'Appel, Sinaasappel and Peer have a different metrical system.'

- (62) Geen stad in het zuiden heeft een andere lengte-eenheid.
No town in the south has a different length-unit
'No town in the South has a different metrical system.'

- (63) Iedere stad in het noorden heeft een verschillende lengte-eenheid.
every/each town in the north has a different length-unit
'Every town in the North has a different metrical system.'

- (64) Alle steden in het noorden hebben een verschillende lengte-eenheid.
all towns in the north have a different length-unit
'All towns in the north have a different metrical system.'

- (65) De steden in het noorden hebben een verschillende lengte-eenheid.
the towns in the north have a different length-unit
'the towns in the north have a different metrical system.'

- (66) Appel, Sinaasappel en Peer hebben een verschillende lengte-eenheid.
Appel, Sinaasappel and Peer have a different length-unit
'Appel, Sinaasappel and Peer have a different metrical system.'

- (67) Geen stad in het zuiden heeft een verschillende lengte-eenheid.
No town in the south has a different length-unit
'No town in the South has a different metrical system.'

Control items

- (68) Iedere stad in het zuiden heeft een andere lengte-eenheid.
every/each town in the south has a different length-unit
'Every town in the south has a different metrical system.'

- (69) De steden in het noorden hebben dezelfde lengte-eenheid.
the towns in the north have the-same length-unit
'The towns in the north have the same length unit.'

Scenario 6

6 matrozen, drie Nederlandse en drie Scandinavische, overleefden een schipbreuk en dreven naar een eiland. Na een tijdje kregen de Nederlandse matrozen - Jip, Jaap, en Joop - ruzie. Jip nam zijn intrek in het scheepswrak, Jaap verhuisde naar een grot, Joop vond een verlaten hut. De Scandinavische matrozen - Kasper, Jesper en Jonathan - konden het goed met elkaar vinden en leefden harmonieus in n grote grot die ze hadden gevonden.

Translation: Six sailors, three Dutch ones and three Scandinavian ones, survived a shipwreck and were stranded on an island. After some time

the three Dutch sailors - Jip, Jaap and Joop - had a fight. Jip left in the shipwreck, Jaap moved to a cave, Joop found an empty hut. The Scandinavian sailors - Kasper, Jesper and Jonathan - could get along pretty well and lived happily together in a big cave that they found.

Test items

- (70) Iedere Nederlandse matroos vond een ander onderkomen.
every/each Dutch sailor found a different dwelling
'Every Dutch sailor found a different place to stay.'
- (71) Alle Nederlandse matrozen vonden een ander onderkomen.
all Dutch sailors found a different dwelling
'All Dutch sailors found a different place to stay.'
- (72) De Nederlandse matrozen vonden een ander onderkomen.
the Dutch sailors found a different dwelling
'The Dutch sailors found a different place to stay.'
- (73) Jip, Jaap en Joop vonden een ander onderkomen.
Jip, Jaap and Joop found a different dwelling
'Jip, Jaap and Joop found a different place to stay.'
- (74) Geen Scandinavische matroos vond een ander onderkomen.
no Scandinavian sailor found a different dwelling
'No Scandinavian sailor found a different place to stay.'
- (75) Iedere Nederlandse matroos vond een verschillend onderkomen.
every/each Dutch sailor found a different dwelling
'Every Dutch sailor found a different place to stay.'
- (76) Alle Nederlandse matrozen vonden een verschillend onderkomen.
all Dutch sailors found a different dwelling
'All Dutch sailors found a different place to stay.'
- (77) De Nederlandse matrozen vonden een verschillend onderkomen.
the Dutch sailors found a different dwelling
'The Dutch sailors found a different place to stay.'
- (78) Jip, Jaap en Joop vonden een verschillend onderkomen.
Jip, Jaap and Joop found a different dwelling
'Jip, Jaap and Joop found a different place to stay.'
- (79) Geen Scandinavische matroos vond een verschillend onderkomen.
no Scandinavian sailor found a different dwelling
'No Scandinavian sailor found a different place to stay.'

Control items

- (80) Iedere Scandinavische matroos vond een ander onderkomen.
every/each Scandinavian sailor found a different dwelling

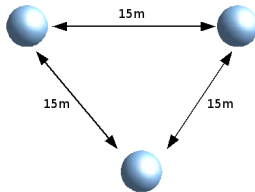
‘Every Scandinavian sailor found a different place to stay.’

- (81) De Nederlandse matrozen vonden hetzelfde onderkomen.
the Dutch sailors found the-same dwelling

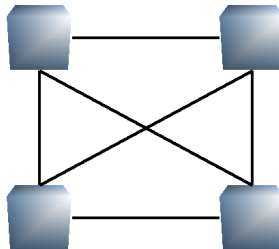
‘The Dutch sailors found the same place to stay.’

Questionnaire testing readings of *each other* and *the others* (Chapter 4)

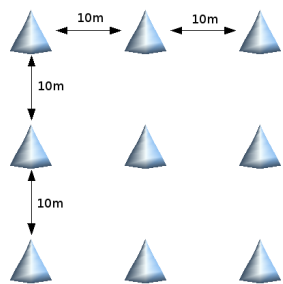
Test items (each list consisted of either test items with *all* or of test items with definite plurals)



- (82) a. All the spheres are 15m from each other.
 b. All the spheres are 15m from the others.
 c. The spheres are 15m from each other.
 d. The spheres are 15m from the others.



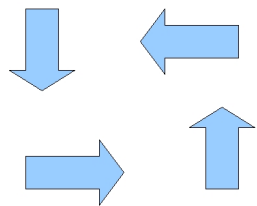
- (83) a. All the boxes are connected to each other by a line.
 b. All the boxes are connected to the others by a line.
 c. The boxes are connected to each other by a line.
 d. The boxes are connected to the others by a line.



- (84)
- a. All the pyramids are 10m from each other.
 - b. All the pyramids are 10m from the others.
 - c. The pyramids are 10m from each other.
 - d. The pyramids are 10m from the others.



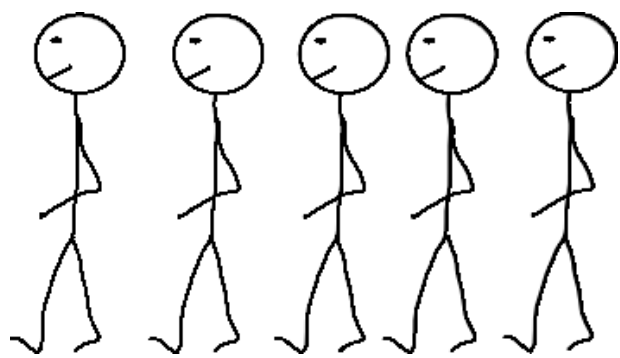
- (85)
- a. All the stars are touching each other.
 - b. All the stars are touching the others.
 - c. The stars are touching each other.
 - d. The stars are touching the others.



- (86)
- a. All the arrows are pointing at each other.
 - b. All the arrows are pointing at the others.
 - c. The arrows are pointing at each other.
 - d. The arrows are pointing at the others.

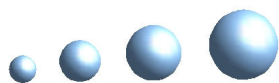


- (87)
- a. All the boxes are on top of each other.
 - b. All the boxes are on top of the others.
 - c. The boxes are on top of each other.
 - d. The boxes are on top of the others.



- (88) a. All the boys are following each other.
- b. All the boys are following the others.
- c. The boys are following each other.
- d. The boys are following the others.

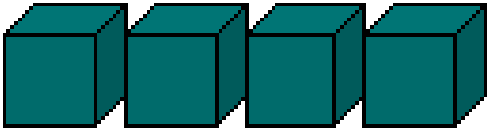
Control items



- (89) a. All the spheres are bigger than each other.
 b. The spheres are bigger than each other.



- (90) a. All the lines are parallel to each other.
 b. The lines are parallel to each other.

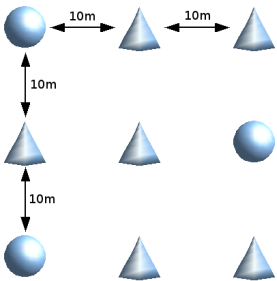


- (91) a. All the boxes are in front of each other.
-
-
- b. The boxes are in front of each other.

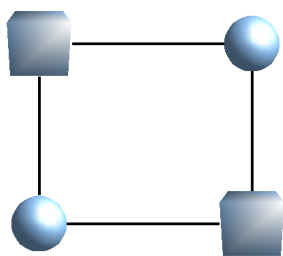
Filler items (selection)



- (92) a. All the circles are touching the stars.
b. The circles are touching the stars.



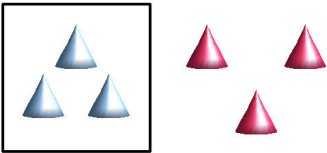
- (93) a. All the pyramids are 10m from the spheres.
b. The pyramids are 10m from the spheres.



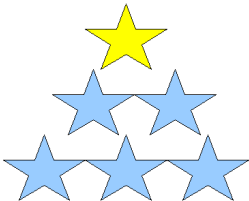
- (94) a. All the spheres are connected to the boxes.
- b. The spheres are connected to the boxes.



- (95) a. All the solid lines are parallel to the dashed lines.
- b. The solid lines are parallel to the dashed lines.



- (96) a. The blue pyramids are in a black square. All the other pyramids are not.
 b. The blue pyramids are in a black square. The other pyramids are not.



- (97) a. All the blue stars are underneath the yellow star.
 b. The blue stars are underneath the yellow star.

Samenvatting in het Nederlands

Dit proefschrift behandelt twee kwesties. De eerste is de interpretatie van zinnen met meervoudige argumenten. Bekijk het volgende voorbeeld:

- (1) Drie jongens bakten vijf cakes.

(1) kan op verschillende manieren genterpreteerd worden. Eén interpretatie is dat de jongens als groep samenwerkten en gezamenlijk vijf cakes bakten. Dit zou bijvoorbeeld het geval kunnen zijn als de jongens een team zijn dat in een restaurant werkt. Eén jongen maakte het deeg, de tweede het glazuur en de laatste bakte de cakes. Als team slagen ze er zo in om vijf cakes te maken. Een andere manier waarop (1) kan worden opgevat, is dat de jongens de taak om cakes te bakken verdeeld hebben: één jongen bakt één cake en de andere twee jongens bakken elk twee cakes. Ik noem de eerste interpretatie een collectieve lezing en de tweede een cumulatieve lezing. Ten slotte kan men (1) zo opvatten dat elk van de drie jongens vijf cakes heeft gebakken, zodat er in totaal vijftien cakes zijn gebakken door de jongens. Deze interpretatie noem ik de distributieve lezing met covariatie, omdat de referent waarnaar het object (de vijf cakes) verwijst (die het object (de vijf cakes) aanduidt), varieert met elke jongen. In de semantische literatuur wordt algemeen erkend dat al deze interpretaties mogelijk zijn. Maar de situatie is anders als we de volgende zin beschouwen:

- (2) Elk van de drie jongens bakte vijf cakes.

Bij (2) ontbreken de eerste twee lezingen die we bij (1) gezien hebben. Het ontbreken van bepaalde interpretaties bij (2) heeft tot een beter inzicht geleid in hoe we kwantoren zoals elk moeten analyseren, en ook tot een beter inzicht in de collectieve, cumulatieve en distributieve interpretatie. Dit proefschrift richt zich op de tegenovergestelde kwestie: de status van de verschillende interpretaties die zinnen als (1) zouden moeten hebben. In tegenstelling tot gangbare opvattingen in de semantiek, beargumenteer ik dat (1) niet al de hierboven beschreven interpretaties heeft, of dat de interpretaties in ieder geval niet allemaal even acceptabel zijn. Verschillende experimenten en vragenlijsten leiden tot deze conclusie. Die beschouw ik en bespreek ik in

detail in Hoofdstuk 2. Zoals ik aantoon, leiden ze tot de conclusie dat de distributieve lezing met covariatie in zinnen als (1) minder acceptabel is dan de collectieve lezing. Deze lezing is ook minder acceptabel dan de cumulatieve lezing. De analyse van dit feit wordt in Hoofdstuk 3 gepresenteerd. De analyse is gebaseerd op de aanname dat men bij het interpreteren van een zin met meervoudige argumenten, ook alternatieve manieren om hetzelfde te zeggen in overweging neemt, met andere uitdrukkingen die meervoudigheid uitdrukken. De redenering is als volgt. De distributieve interpretatie van (1) kan worden uitgedrukt door zin (2). (1) wordt daarom niet in de distributieve betekenis opgevat, omdat er een andere, eenduidige vorm is die deze betekenis kan uitdrukken. Maar er is nog een reden dat (1) niet de distributieve lezing met covariatie heeft: er is nu een vorm, (3), die overwogen wordt bij het interpreteren van (1):

- (3) Alle drie de jongens bakten vijf cakes.

(3) gebruikt een langere vorm, omdat het drie jongens vervangt door alle drie de jongens. Zoals vaak beargumenteerd wordt, is de langere vorm gemarkeerd. In dit geval zou dat betekenen dat (3) een gemarkeerde vorm gebruikt, terwijl (1) een ongemarkeerde vorm gebruikt. Het is lang geleden al opgemerkt dat er een interactie is tussen gemarkeerde vormen en gemarkeerde betekenissen (Horn, 1984). Om precies te zijn: we relateren gemarkeerde vormen aan gemarkeerde betekenissen, en ongemarkeerde vormen aan ongemarkeerde betekenissen. Ik beargumenteer dat de distributieve lezing met covariatie een gemarkeerde betekenis is, die dus gerelateerd wordt aan (3), terwijl de andere interpretaties worden gerelateerd aan (1). De twee verklaringen voor de observatie dat de distributieve lezing met covariatie, gedisprefereerd is, worden geformaliseerd in bi-directionele Optimaliteitstheorie en de benadering in Game Theory van Parikh (2000). De tweede kwestie die bestudeerd wordt in dit proefschrift is anaforiciteit in het Engels en het Nederlands. De meeste studies naar anaforiciteit richten zich op pronomina en reflexieven, maar dit proefschrift richt zich juist op de uitdrukkingen zelfde, ander, de andere(n) en elkaar. Wat deze uitdrukkingen gemeen hebben is dat ze anaforisch kunnen zijn ten opzichte van een referent in de deelzin als de deelzin een semantisch meervoudig argument bevat. Bekijk (4). Deze zin kan betekenen dat elke jongen hetzelfde boek las als de andere jongens. hetzelfde is anaforisch ten opzichte van de boeken die de andere jongens lasen. Het is duidelijk dat deze interpretatie alleen maar mogelijk is vanwege het meervoudige argument de jongens. Op vergelijkbare wijze kan ander in (4b) en verschillend in (4c) (althans voor sommige sprekers) anaforisch zijn ten opzichte van de boeken die de andere jongens lasen. De andere(n) en elkaar zijn anaforisch ten opzichte van de jongens in (4d) en (4e).

- (4) a. De jongens lasen hetzelfde boek.
 b. Elke jongen las een ander boek.
 c. % De jongens lasen een verschillend boek.
 d. Elke jongen praatte met de anderen.
 e. De jongens praatten met elkaar.

Geen van deze uitdrukkingen kan anaforsch zijn ten opzichte van een referent in de deelzin als de deelzin geen semantisch meervoudig argument bevat. Beschouw (5a) en (5e), waarin geen semantisch meervoudig argument voorkomt. Het gevolg is dat de zin ofwel ongrammaticaal is, zoals (5c) en (5e), ofwel we kunnen de zin alleen interpreteren als we aannemen dat er een voorafgaande discourse is waarin een element voorkomt ten opzichte waarvan de uitdrukkingen van non-identiteit anaforsch zijn ((5a),(5b) en (5d)).

- (5) a. Morris las hetzelfde boek.
- b. Morris las een ander boek.
- c. * Morris las een vereschillend boek.
- d. Morris praatte met de anderen.
- e. * Morris praatte met elkaar.

Het feit dat de uitdrukkingen van (non-)identiteit een meervoudig argument vereisen, brengt de studie naar anaforiciteit in verband met de studie naar de interpretatie van zinnen met meervoudige argumenten. Om precies te zijn, er is een verband tussen de uitdrukkingen van (non-)identiteit en het vermogen van meervoudige argumenten om de distributieve interpretatie met covariatie te licentieren, of de collectieve en cumulatieve interpretatie. Dit verband wordt in detail bestudeerd in de vragenlijst in Hoofdstuk 2. Ik toon aan dat verschillend en elkaar het soort meervoudig argument vereisen dat de collectieve en cumulatieve lezing licentieert. Aan de andere kant vereisen ander en de andere(n) het soort meervoudig argument dat de distributieve lezing met covariatie licentieert (maar niet de collectieve en cumulatieve lezing). Ten slotte is zelfde niet gevoelig voor het soort meervoudige argument. Ik beargumenteer dat we dit verschil tussen de uitdrukkingen van (non-identiteit) kunnen verklaren als we twee strategieën van anaforiciteit aannemen. Bij de eerste zijn anaforische uitdrukkingen quantificationale elementen die bereiksfafhankelijkheden creëren. Dit is identiek aan de manier waarop kwantoren genterpreteerd worden. De anaforische uitdrukkingen verschillen echter van kwantoren in die zin dat ze geen bereik hebben op het niveau van de propositie, maar bereik hebben over eigenschappen. Ik beargumenteer dat dit de juiste analyse is voor elkaar en verschillend. De analyse verklaart niet alleen dat deze uitdrukkingen het soort meervoudig argument vereisen dat de collectieve en cumulatieve lezing licentieert, maar ook verscheidene andere kwesties bijvoorbeeld waarom deze uitdrukkingen geen antecedent kunnen hebben in een voorafgaand discourse, en waarom elkaar het effect van zogenaamde reciprociteit op lange afstand te weeg kan brengen als het ingebed is onder een attitude-werkwoord, zoals bestudeerd in Heim et al. (1991a) en Dimitriadis (2000), onder andere. De semantiek van reciprociteit is een belangrijk onderwerp van Hoofdstuk 4. De semantiek van verschillend is een van de belangrijkste onderwerpen van Hoofdstuk 5. De tweede strategie voor anaforiciteit wordt gebruikt bij het naamwoord de andere(n) en de adjectieven ander en zelfde. Bij deze strategie wordt de anaforische relatie bewerkstelligd door een extern (syntactisch) mechanisme. Dit mechanisme kan predicaatabstractie en binding door een lambda-operator zijn, in de benadering waarin anaforen worden genterpre-

teerd als variabelen, zoals in Heim en Kratzer (1998). Ik toon aan dat als we dit mechanisme combineren met de analyse van distributieve lezingen in Hoofdstuk 3, we kunnen verklaren dat de andere(n) en ander het soort meervoudig argument vereisen dat de distributieve lezing met covariatie licentieert, en dat same/zelfde niet gevoelig is voor het soort meervoudig argument. Andere eigenschappen van deze uitdrukkingen volgen hier ook uit, bijvoorbeeld dat ze ook anaforsch kunnen zijn ten opzichte van een antecedent in een voorafgaand discourse. Dit is een belangrijk onderwerp in Hoofdstuk 4 (waarin het naamwoord elkaar wordt bestudeerd en vergeleken met elkaar) en Hoofdstuk 5 (waarin de adjectieven ander en zelfde worden bestudeerd en vergeleken met verschillend). Concluderend: dit proefschrift verbindt de studie naar anaforiciteit met de studie naar interpretaties van zinnen met meervoudige argumenten. Het laat zien dat de twee verschijnselen samenhangen, en hun samenhang vertelt ons iets over de analyse van distributieve lezingen en ook over de interpretatie van elkaar, ander, verschillend en zelfde. De samenhang tussen de twee verschijnselen geeft meer in het algemeen inzicht in anaforiciteit in natuurlijke taal.

Curriculum Vitae

Jakub Dotlačil was born on 14th May 1979 in Teplice v Čechách, Czechoslovakia. From 1997 to 2002 he studied Czech language and literature and Slovenian language and literature at Charles University in Prague, the Czech Republic. From 2002 to 2004 he studied Linguistics at the University of Tromsø in Norway, where he received his master's degree in 2004 with a thesis on the syntax of infinitival constructions in Czech. In 2005 he started his PhD project at Utrecht Institute of Linguistics OTS. He carried out his research there till 2009 with a short interruption from August 2008 to January 2009 when he was a visiting PhD student at Zentrum für Allgemeine Sprachwissenschaft in Berlin. This dissertation is the result of his PhD research.