WAYS OF SCOPE TAKING

edited by

ANNA SZABOLCSI

Department of Linguistics,
University of California,
Los Angeles, California, U.S.A.

preprint of 1997

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

CONTENTS

CONTRIBUTORS

Filippo Beghelli
Department of Linguistics
UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1543
beghelli@babel.ling.upenn.edu

Dorit Ben-Shalom
Department of Linguistics
UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1543
dorit@cognet.ucla.edu

Jenny Doetjes
HIL/Department of Linguistics
University of Leiden
P.O. Box 9515
2300 RA Leiden
The Netherlands
doetjes@rullet.leidenuniv.nl

Donka Farkas
Linguistics Board
UCSC
Stevenson College
Santa Cruz, CA 95064
farkas@ling.ucsc.edu

Javier Gutiérrez Rexach
Department of Linguistics
UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1543
gutier@ucla.edu

Martin Honcoop
HIL/Department of Linguistics
University of Leiden
P.O. Box 9515
2300 RA Leiden
The Netherlands
honcoop@rullet.leidenuniv.nl

Edward Stabler
Department of Linguistics
UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1543
stabler@ucla.edu

Tim Stowell
Department of Linguistics
UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1543
stowell@ucla.edu

Anna Szabolcsi
Department of Linguistics
UCLA
405 Hilgard Avenue
Los Angeles, CA 90095-1543
annasz@humnet.ucla.edu

Frans Zwarts
Institute for Behavioral, Cognitive and Neurosciences
Rijksuniversiteit Groningen
P.O. Box 716
NL-9700 Groningen
The Netherlands
frans@let.rug.nl

PREFACE

The present volume is as much a book co-authored by all the contributors as it is an edited collection of their papers. Most of the contributors have been involved in regular discussions over the past years, often inspiring the questions, or some aspects of the proposals, in each other's papers or actually collaborating on co-authored papers.[1] For this reason, the contributions make related assumptions and explore highly related issues. The organization of the volume reflects this unity of aims and interests. It starts out with an overview of some of the shared formal background, and the chapters are arranged in a sequence that is intended to invite the reader to proceed from one directly to the next. Nevertheless, there has been no attempt to eliminate individual differences in either assumptions or choice of topic. All the chapters are entirely self-contained, so the reader will find it equally possible to read any of them in isolation.

Two members of the UCLA community do not appear in this volume but have been an important source of inspiration for this project: Ed Keenan and Feng-hsi Liu. Many of Keenan's works have drawn attention to the empirically diverse behavior of natural language determiners and developed theoretical tools for studying them. Liu's 1990 dissertation examined the abilities of a representative sample of noun phrases to participate in scopal dependencies and branching, coming up with provocative generalizations and pointing out their significance for then-standard theories in powerful terms. Three other linguists discussions with whom have been more important to several of us than routine acknowledgments might indicate are Carmen Dobrovie-Sorin, Barry Schein, and Frans Zwarts.

It was a privilege to have Sean Fulop as our copy-editor and type-setter, and Edward Garrett as our proof-reader and advisor concerning the preparation of the manuscript.

Significant parts of the research, as well as the preparation of a camera-ready manuscript, were supported by NSF grant #SBR 9222501 and by grants from the Academic Senate of UCLA, which we gratefully acknowledge.

[1] Names of co-authors are in all cases listed in alphabetical order.

INTRODUCTION

Syntactic and semantic theories of quantificational phenomena traditionally treat all noun phrases alike, thus predicting that noun phrases exhibit a uniform behavior. It is well-known that this is an idealization: in any given case, some noun phrases will support the desired reading more readily than others. Anyone who has lectured on quantifier scope ambiguities to a class of unbrainwashed undergraduates will recall the amount of preparation time that goes into coming up with two or three examples that the class will judge to be ambiguous in exactly the ways the theory under discussion predicts. The same experience with "good citizens" and "bad citizens" repeats itself in connection with branching, anaphora, distributive versus collective readings, extraction, event quantification, pair-list questions, and so on.

Is the assumption of uniformity a theoretically necessary idealization, then, or is it an overgeneralization based on a small body of initial data? There is no doubt that, to some extent, it is a necessary idealization. To what extent it is, though, depends on how systematic the patterns of deviation turn out to be, and how coherent and interesting theoretical accounts can be devised for those patterns.

The unique contribution of this volume consists in scrutinizing large bodies of data, both well-known and novel, from a theoretical perspective and arguing that the patterns emerging are systematic and significant enough to prompt rather fundamental revisions of the standard accounts.

In proposing alternatives, many of the papers follow a heuristic that may be summarized as follows: The range of quantifiers that participate in a given process is suggestive of exactly what that process consists in. Instead of devising omnivorous rules that apply to all quantifiers and then need to be constrained in various, sometimes ad hoc, ways, it is proposed that the grammar of quantification involves a variety of distinct, often semantically conditioned, processes. Each type of expression participates in those processes that suit its particular properties. The main specific claims are to be reviewed shortly.

There are important results in recent semantic and syntactic literature that point in a similar direction. On the semantics side, the empirical theories of discourse representation and plurals have pointed out significant respects in which different noun phrase types contribute differently to interpretation, and the mathematical theories of generalized quantifiers and partially ordered sets offer tools for making various distinctions that prove relevant. On the syntax side, the minimalist program postulates that movement is not input-blind and optional, but it is driven strictly by the specific properties of lexical items, as in a number of other lexicalist approaches.

The work reported here pulls many of these results together and applies their insights in a unified manner.

*

The issues addressed in the volume fall into two major categories, (i) THE SYNTAX/SEMANTICS INTERFACE and (ii) more or less pure SEMANTICS. Syntax is relevant also in the second category: in some cases a semantic account is offered for a phenomenon usually held to be syntactic, or vice versa.

Many papers in the volume make use of some simple tools of formal semantics. Often, their empirical predictions derive directly from formal semantic considerations. To make these arguments more accessible and, hopefully, pleasurable to the reader, the first chapter offers a fairly informal introduction to the pertinent background notions in lattice theory and generalized quantifiers.

*

The first set of papers pertains to the SYNTAX/SEMANTICS INTERFACE. They focus on data where the missing readings are, in and of themselves, as coherent as the available ones; the problem is that the grammar of English does not associate them with the given strings of words. The central issue in these papers is how different noun phrase types acquire their scope and, consequently, how they interact with each other and with negation.

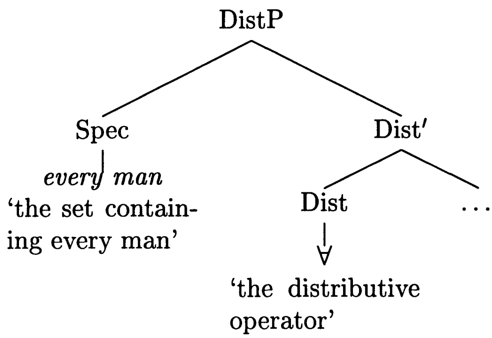
Traditionally, syntactic and semantic theories have assumed that all noun phrases are assigned scope by the same rule, that the scope assignment rule is optional, that it can "prefix" the quantifier to practically any syntactic domain, and that wide scope equals distributive wide scope. In a series of papers, Beghelli and Stowell have challenged these assumptions and developed a novel approach to Logical Form. The motivation for the proposed changes is empirical. As is shown by Kroch's, Ioup's, and Liu's work, as well as more recent research including Beghelli and Stowell's own, quantifier types differ in important respects. Consider a small sample of the contrasts. *More than three men* and *every man* differ in their readiness to take inverse scope: *More than three men read every book* easily admits the interpretation 'For every book, there are more than three men who read it,' but *Every man read more than three books* does not admit 'There are more than three books which every man read.' On the other hand, numerical indefinites and universals can both take inverse scope as far as existential import is concerned, but differ in supporting inverse distributive readings: compare *More than three men read every book* above, which allows the men to vary with the books, with *More than three men read two of the books,* which does not, although the two books can be picked independently of the men. Finally, *more than two men* and *every man* differ with respect to their interaction with negation: *More than two men didn't laugh* is a fine sentence, but *Every man didn't laugh,* with a non-contrastive intonation, is not; *I didn't read more than two books* is ambiguous, but *I didn't read every book* admits only a 'not every' interpretation.

The contents of the volume offer an account of these and similar contrasts in terms of a minimalist theory of Logical Form, whose distinctions can be correlated with those of Discourse Representation Theory and to some extent motivated by generalized quantifier theoretic considerations.

Working within the minimalist program of syntax, Beghelli and Stowell make the following basic assumptions. (a) Noun phrases acquire their scope as a by-product of moving into syntactic positions where they can check some scope-independent morphological and/or semantic feature, and (b) Distributivity is effected by a syntactically separate operator. These assumptions are useful in the following way.

Since noun phrases differ in morphological and semantic properties, (a) yields an account of the diversity of their behavior. To be more specific, Beghelli and Stowell claim that noun phrases fall into two larger categories. Members of the first have specific target landing sites distinct from the case positions; members of the second do not. The target landing sites include, along with the specifier positions of well-known categories like CP and NegP, those of a novel set of functional projections, RefP, DistP, and ShareP. Plurals like *(the) two men* move to the specifier of RefP or ShareP, and distributive universals like *every man* to DistP. Modified numerals like *more/less than three men* do not move beyond their case positions and thus scope in situ.

As regards (b), Beghelli and Stowell argue that both plurals and universals are associated with a set-denoting part and a phonetically null distributive operator. The distributive operator associated with plurals is the silent *each* known from the semantics literature; syntactically, it is shown to have the properties of floated *each*, an adverbial element that attaches to some heads, but not to others. On the other hand, the distributive operator associated with universals (and other noun phrases that pattern with them) is syntactically the head of DistP.



The fact that distributivity is factored out and represented in these particular syntactic ways allows the theory to account for a variety of subtle phenomena, including the separability of the existential import and the distributivity of noun phrases, the clause-boundedness of distributivity, the differential ability of noun phrases to induce referential variation when taking inverse scope, and the deviations from the basic patterns in interaction with negation and *wh*-phrases.

The outlines of the general theory and the interaction of universals with other quantifiers and negation are laid out in Beghelli and Stowell's **Distributivity and negation** in this volume. Beghelli's **The syntax of distributivity and pair-list readings** introduces the two types of distributivity in detail and applies the results to uncover and explain new data, along with the notorious syntactic asymmetries, in connection with *wh*/QP interactions. More on this paper below.

The one feature of Beghelli and Stowell's scope syntax that may appear strikingly baroque is the postulation of a multitude of new LF landing sites. **Strategies for scope taking** by Szabolcsi offers independent motivation for this feature. It is shown that the surface syntactic scope positions that have for long been postulated for noun phrases in Hungarian correspond to the positions Beghelli and Stowell posit for Logical Form in English.

Szabolcsi's paper further addresses the relation between this theory of scope and Discourse Representation Theory. On the basis of their commonalities as well as the syntactic advantages of Beghelli and Stowell's proposal, it is proposed that Beghelli and Stowell's way of constructing Logical Forms should, essentially, replace Kamp and Reyle's DRS construction algorithm. Concretely, movement into RefP or DistP corresponds to introducing discourse referents, while noun phrases that scope in their case positions are interpreted as performing a counting operation on predicates. Hungarian data play a crucial role in substantiating some of these claims. The upshot is that the independent structure that Beghelli and Stowell argue scope lives off of is, in semantic terms, a kind of discourse representation structure.

Some formal semantic groundwork for the above papers is laid by Beghelli, Ben-Shalom and Szabolcsi in **Variation, distributivity, and the illusion of branching.** They motivate breaking scope down into referential variation, distributivity and, in the case of non-upward entailing QPs, maximality. Then the same conceptual apparatus is applied in the study of branching readings, whose descriptive constraints have been observed by Liu. It is argued that (in the set of data considered) no specific branching quantifier needs to be, or indeed, should be, postulated in English. All the empirically attested branching readings are logically equivalent to some other reading that needs to be derived anyway: a scopally asymmetrical or a cumulative one.

In **Computing quantifier scope,** Stabler offers a different perspective on the issue of how semantic properties of noun phrases may affect their scopal syntactic abilities. Noting that the same properties manifest themselves in the inferential behavior of noun phrases, which can be represented syntactically, he proposes to reverse the order of explanation. He does not assume that the speaker has some grasp of the semantic value of the expression first and then decides where to put it in syntactic structure. Instead, the speaker uses the expression in a certain way, in the syntax according to the requirements specified in its features, and in inference. The proposal is implemented within a novel formalization of minimalist syntax, applied to Beghelli and Stowell's theory.

Whatever their take on the role of semantics, all the papers above assume that scope is a structural notion. Farkas, whose 1981 CLS paper contains some of the classical observations concerning the scope and distributivity of indefinites, proposes a non-structural approach. In **Evaluation indices and scope,** the relative scope of two expressions is a matter of possible dependencies of indices, seen as Kaplan-style coordinates of evaluation. In this way, Farkas's approach may be closer in spirit to Groenendijk and Stokhof's Dynamic Semantics than to Kamp and Reyle's DRT. This paper goes beyond the others in empirical coverage: it examines, in addition to noun phrases, the discourse scope of conditionals, modal and intensional expressions.

*

The above considerations pertain to the syntax/semantics interface. The second set of papers argues that scope assignment can go wrong in a DIRECTLY SEMANTIC way as well, namely, the intended meaning may be incoherent and, therefore, "unthinkable."

Such incoherence is the source of the ungrammaticality of *How much milk didn't you drink?,* in distinction to the well-formedness of *Which books didn't you read?,* argue Szabolcsi and Zwarts in **Weak islands and an algebraic semantics for scope taking.** The impossibility of *how*-extraction out of a negative island is assimilated to that of the combination of a numeral with a mass noun, as in *six airs.* In both cases, the explanation is that the interpretation of the construction requires us to perform an operation (complement formation in the first case, counting in the second) on a semantic structure that does not lend itself to that operation.

The paper explicates a denotational semantic limitation on scope interaction using some simple notions of lattice theory. The nature of the argument can be best illustrated by way of an example. Overt *wh*-extraction creates a syntactic configuration with an extraction domain D containing a gap α. Let D contain another scopal element β, which the filler of the gap is supposed to scope over.

[how much milk$_i$ [$_D$ did [$_\beta$ n't] you drink [$_\alpha$ $-_i$]]]
[which books$_i$ [$_D$ did [$_\beta$ n't] you read [$_\alpha$ $-_i$]]]

To calculate the denotation of the whole sentence, the denotation of D needs to be calculated. The question is whether this is possible, in view of what α and β are.

The kind of denotation D has is, to a large extent, determined by what kind of gap it contains. For instance, *did you read [gap of which books]* denotes a set of individuals. But *did you drink [gap of how much milk]* arguably does not; Szabolcsi and Zwarts argue it denotes an amount. Now, the general claim is that the narrow scope element β is interpreted by cashing out its contribution in terms of some operation(s) over the denotation of D minus β. For instance, *n't* in the examples above requires us to take the complement of that denotation. Sets of individuals form Boolean algebras, in which complement formation is defined, thus *didn't you read [gap of which books]* is perfectly coherent. Amounts, on the other hand, form join semi-lattices at best, in which complement formation is not defined. Hence, the denotation of *didn't you drink [gap of how much milk]* cannot be calculated. In general, this kind of conflict arises whenever the interpretation of β involves at least one Boolean operation not available in the structure that the denotation of D minus β belongs to.

The unacceptable extraction of amount and manner expressions out of negative islands, *wh*-islands, and factive islands is called a "weak island violation." Weak islands were traditionally thought to belong to the realm of pure syntax. More recently, it has been argued that they are due to the inability of the given *wh*-phrase to take scope over some other scopal element in the extraction domain. Szabolcsi and Zwarts concur with this view; the novel feature of the paper is the above reviewed algebraic semantic characterization of scope interaction, which explains why some expressions are unable to scope over certain others.

Naturally, the same considerations apply to covert scope assignment, in addition to the considerations discussed in the first set of papers.

The same semantic explanation extends, according to Honcoop and Doetjes, to the fact that event-related readings are sensitive to weak islands. **The semantics of event-related readings: a case for pair-quantification** proposes to treat the numeral in Krifka's famous *Four thousand ships passed through the lock* as quantifying over ⟨event, object⟩ pairs. Events are standardly thought to have a join semi-lattice structure without a bottom element, and ⟨event, object⟩ pairs inherit this from their event component. Thus *Four thousand ships didn't pass through the lock* has no event-related reading.

The pair-quantificational approach is argued to explain empirical constraints on event-related readings that go well beyond sensitivity to weak islands. Both the restriction and the scope of the pair-quantifier need to contain both an event and an object variable. Symmetric (weak) determiners (like *four thousand*) support event-related readings without further ado, because symmetry allows the copying of the verbal predicate that supplies the event variable into the determiner's restriction, by plain inference. Non-symmetric (strong) determiners support an event-related reading only when an event variable occurs in the restriction either due to "rebracketing" induced by focus (*Most ships passed through the lock YESTERDAY*) or because the noun is modified by an eventive relative clause (*Most ships that passed through the lock transported radio-active waste*). The specific treatment of the event variable is cast in terms of dynamic semantics, and parallelisms with donkey anaphora are explored.

Two papers in the volume are concerned with the phenomenon of pair-list readings. In addition to their interest as a further type of scope interaction, pair-list readings are directly relevant in connection with the scopal account of weak islands. There are two ways in which a scopal intervener β may turn out to be harmless. One, the interpretation of β may only involve operations that the relevant structure is closed under. Two, β may support an alternative wide scope reading, and thus "get out of the way." Such is the case with the intervening universal in *How much milk did every kid drink?*. This question is bad when *every kid* takes narrow scope, but good when it supports a pair-list reading.

Both papers on *wh*/QP interactions begin by showing that the actual distribution of pair-list readings is so different from what is assumed in the literature that it causes the standard syntactic and semantic accounts to lose much of their force.

Based on what QPs support a pair-list reading in what context, in **Quantifiers in pair-list readings** Szabolcsi shows that two quite different types need to be distinguished. Pair-list readings in matrix questions and in complements of *wonder*-type verbs are induced only by universals and can be assimilated to multiple interrogation. On the other hand, almost any QP induces a pair-list reading in complements of *find out*-type verbs; crucially, even non-increasing ones do. Compare *Where do fewer than five suspects live?* with *We only found* *out where fewer than five suspects live.* The standard analyses, according to which the quantifier in a pair-list reading contributes a set to restrict the domain of the question would work for all and only upward monotonic quantifiers in both contexts (too many for the first, too few for the second). It is argued that pair-list readings in *find out*-complements must be treated as quantificational. In the context of the present volume, this means that each QP supports a pair-list reading in the same fashion in which it takes scope in other, non-*wh* contexts. These observations in turn have some interesting consequences for weak islands.

The syntax of distributivity and pair-list readings by Beghelli is an integral part of the theory of Logical Form that was reviewed in the first part of this introduction. As was mentioned, the theory distinguishes two types of distributivity: that induced by the Dist head associated with universals (called strong distributivity), and that induced by the covert counterpart of floated *each* associated with plurals (pseudo-distributivity). Among other things, the two types differ in what interactions they make possible between a subject and a complement on the one hand, and between two complements on the other. E.g. *Five of these students read every/two book(s)* 'for every book / *for each member of a set of two books, there is a possibly different set of five of these students who read it' and *John showed every book / five of these books to a student* 'for every book / ?for each of these five books, John showed it to a possibly different student.'

The paper lays out the general properties of the two types of distributivity and goes on to apply them in the study of pair-list readings. It is well-known that some pair-list readings exhibit robust syntactic asymmetries: *What did everyone read?* has a pair-list reading, but *Who read everything?* does not. These have been accounted for in the literature in terms of the Empty Category Principle and Weak Cross-over, for instance. Beghelli makes the surprising observation that a larger sample of data reveals that the patterns do not match either the ECP or WCO. Instead, the behavior of universals in *find out*-complements matches the pattern of strong distributivity; in matrix questions and in *wonder*-complements, it matches the pattern of pseudo-distributivity.

On the basis of such observations, Beghelli develops syntactic analyses that square well with the multiple interrogation versus quantification distinction established in the previous paper.

Several papers in the volume make use of the tools of the theory of generalized quantifiers in connection with standard noun phrases. In the literature, *wh*-phrases or questions do not fall under the scope of that theory. In **Questions and generalized quantifiers,** Gutiérrez Rexach argues that it is both possible and insightful to bring them into the fold. He interprets questions as functions that assign the value true or false to answer sets. E.g., in a world where John and Mary walk, *Who walks?* assigns true to a set if it is identical to the set of walkers, i.e. {j, m}. This yields the same interpretation of questions as Groenendijk and Stokhof's, but is formulated in a way that makes it possible to extend the apparatus of generalized quantifier theory to questions: notice that the *wh*-phrase relates two properties, the ones named by the question and by the answer set, as determiners do. The paper shows that the well-known properties of determiners carry over to *wh*-expressions. Finally, some cases of multiple interrogation, cumulative, and pair-list readings are shown to be irreducibly polyadic.

This concludes the summary of the main results in connection with scope at the syntax/semantics interface and in semantics.

1

BACKGROUND NOTIONS IN LATTICE THEORY AND GENERALIZED QUANTIFIERS*

Anna Szabolcsi

Department of Linguistics
UCLA

Many papers in this volume build on certain elementary notions of lattice theory and generalized quantifier theory; often, their empirical predictions derive directly from them. The goal of this chapter is to enable readers who have some background in formal semantics, but not in these particular areas, to appreciate the pertinent papers. But readers who are familiar with lattices and GQs may also find the discussion useful because, elementary as it is, it highlights certain aspects that other literature may not. On the other hand, precisely because this chapter is geared towards particular applications, it does not attempt to cover issues that a standard introduction would, when they do not seem directly relevant here.

The chapter consists of three parts. The first part familiarizes the reader with the relevant notions and their significance. The second is a set of problems. Some of them merely check the mastery of definitions, others touch on linguistic issues that are of theoretical relevance to the contents of this volume. The third part offers quite elaborate solutions. The gentle reader who is not in a problem solving mood is encouraged to read the problems and their solutions as if they were part of the main text.

*This chapter is based on my lecture notes for classes given at UCLA, the University of Budapest, and the 1993 LSA Linguistic Institute. I thank the participants of these courses for feedback. Work on the present version was partially supported by NSF grant #9222501.

A. Szabolcsi (ed.), Ways of Scope Taking, 1–27.

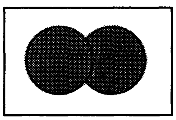
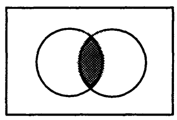
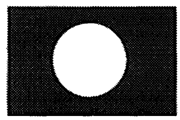
1 OPERATIONS IN PARTIALLY ORDERED SETS

1.1 Partially ordered sets: lattices, semi-lattices, Boolean algebras

Recall the basic set theoretical operations and their counterparts in the propositional calculus:

(1)

union: $A \cup B$	intersection: $A \cap B$	complement: $-A$
disjunction: $p \vee q$	conjunction: $p \wedge q$	negation: $\neg p$

What other operations are these related to? On what kind of entities can such operations be performed? What kind of structures do these entities form? These are the main questions we are going to ask.

The basic distinction to build on is between ordered and unordered sets. An unordered set is any set in the standard sense, e.g.,

(2) Unordered sets:

$$\begin{aligned} A &= \{\text{joe, ed, pat, sue}\} \\ B &= \{\emptyset, \{\text{joe}\}, \{\text{ed}\}, \{\text{pat}\}, \{\text{joe, ed}\}, \{\text{joe, pat}\}, \\ &\quad\ \{\text{ed, pat}\}, \{\text{joe, ed, pat}\}\} \\ C &= \{\text{joe, ed, pat, joe-and-ed, joe-and-pat, ed-and-pat,} \\ &\quad\ \text{joe-and-ed-and-pat}\} \end{aligned}$$

Sets become ordered if we explicitly assume some ordering relation on their members (whether or not there is a "natural ordering" that suggests itself anyway), e.g.,

(3) Ordered sets:[1]

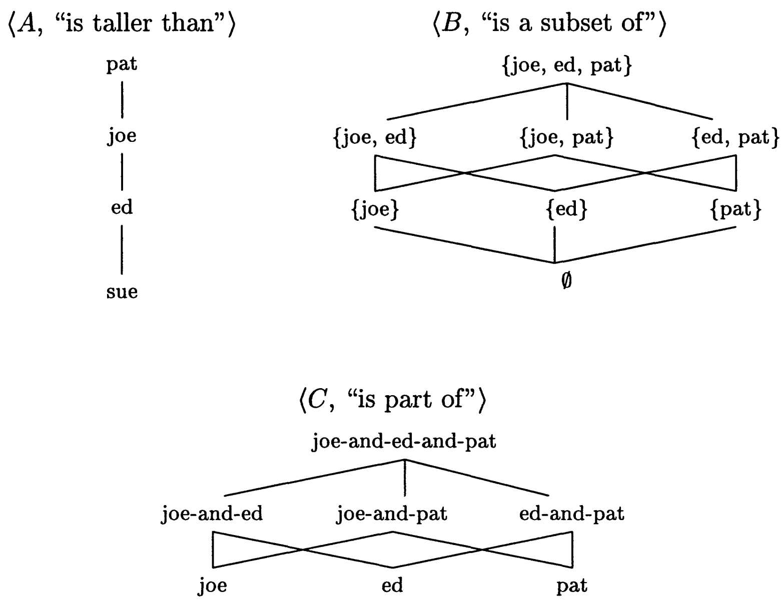
$\langle A$, "is taller than"$\rangle$	or	$\langle A$, "is likelier to cry than"$\rangle$
$\langle B$, "is a subset of"$\rangle$	or	$\langle B$, "has fewer elements than"$\rangle$
$\langle C$, "is part of"$\rangle$	or	$\langle C$, "is less happy than"$\rangle$

Clearly, different relations may order the same set differently. E.g., Joe may be taller than Ed (hence Joe $\geq_1$ Ed) but less likely to cry (Ed $\geq_2$ Joe). Or, {joe} is

[1]The non-atomic elements of C are called *collectives,* or *plural individuals,* or *sums.*

not a subset of {ed, pat} or vice versa (these two elements are not ordered with respect to each other by $\geq_3$) but has fewer elements ({ed, pat} $\geq_4$ {joe}). The two orderings may coincide, e.g., Joe is part of the collective Joe-and-Ed (Joe-and-Ed $\geq_5$ Joe) and may also be less happy on his own (Joe-and-Ed $\geq_6$ Joe). The ordering may be specified graphically, as in the Hasse-diagrams below. All lines can be read as upward arrows that point to the element ordered higher.

(4)



Two kinds of linguistic applications may be as follows. The elements of the set A are ordered with respect to an "extrinsic" property (in fact, these individuals cannot be ordered otherwise). Such an ordering may be invoked in the discussion of words like *even* (*Even Sue can reach this shelf* may be felicitous, because Sue herself is short relative to the others we are interested in). The elements of B and C can be ordered with respect to "intrinsic" properties such as "subset" and "part-of" as well as "extrinsic" ones. In this volume all linguistic applications happen to be of the "intrinsic" sort.

We now turn to more precise definitions. (Recall that R is reflexive iff $\forall x[Rxx]$, R is transitive iff $\forall xyz[(Rxy \ \& \ Ryz) \rightarrow Rxz]$, and R is anti-symmetrical iff $\forall xy[(Rxy \ \& \ Ryx) \rightarrow x = y]$.)

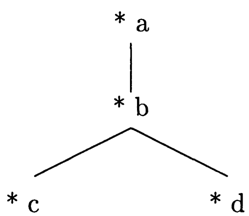
(5) A relation R is a *partial ordering* iff it is reflexive, transitive, and anti-symmetrical. A *partially ordered set* (partial order, or poset, for short) is any $\langle A, \leq \rangle$, where $\leq$ is a partial order.

The relations "larger than or equal to" and "subset of" are partial orderings. The relations "larger than" and "proper subset of" are strict orderings: they are not anti-symmetrical but asymmetrical.

How do we get to the desired operations from here, cf. (1)? They are definable in terms of partial ordering. The general lattice-theoretic names they come under are meet, join, and complement. Intersection is the realization of meet when applied to sets, and conjunction is meet when applied to propositions. Similarly, union is join for sets and disjunction is join for propositions; negation is complement for propositions.

(6) Let $\langle A, \leq \rangle$ be a poset. For any subset X of A,
a is a lower bound for X if for every element x of X, $a \leq x$.
The infimum of X, written $\bigwedge X$, is the *greatest lower bound* for X.
c is an upper bound for X if for every element x of X, $c \geq x$.
The supremum of X, written $\bigvee X$, is the *least upper bound* for X.

The lower bounds of the set X are elements of A (within X or outside X) which are smaller than or at best equal to all elements of X; the infimum is the greatest of these. Similarly for the least upper bound (supremum). E.g.,



The set of lower bounds for $\{a, b\}$ is $\{b, c, d\}$, of which b is the greatest.

(7) Let $a, b \in A$.

a. The *meet* of a and b, written $a \wedge b$, is the infimum of the 2-element set $\{a, b\}$.
Thus we have: $a \wedge b \leq a$ and $a \wedge b \leq b$.

b. The *join* of a and b, written $a \vee b$, is the supremum of the 2-element set $\{a, b\}$.
Thus we have: $a \vee b \geq a$ and $a \vee b \geq b$.

Meet is a special case of infimum: it is the infimum of some two-element set. Similarly for join and supremum.

Depending on what operations are available in a specific partially ordered set, we may have a Boolean algebra, a lattice, a meet or join semi-lattice, or none of these. "Available" means that the given poset is closed under that operation: whenever meet or join is applied to two elements of A, the result is also an element of A (the same for complement, which applies to one element). That is, these operations do not "lead out of" A.

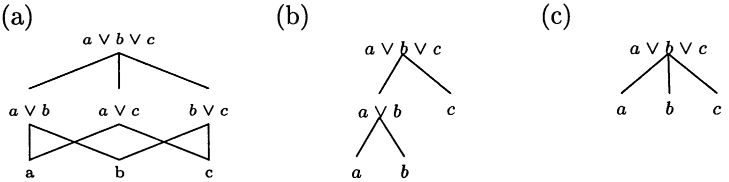
(8) A *lattice* defined in terms of partial ordering:
A lattice is a poset $\langle A, \leq \rangle$ which is closed under meet and join.
That is, for every $a, b \in A$, $\quad a \wedge b \in A \quad$ and $\quad a \vee b \in A$.
It follows that A is a lattice iff for any non-empty finite subset X of A, $\bigwedge X \in A \quad$ and $\quad \bigvee X \in A$.

E.g., both $\langle A$, "taller than"$\rangle$ and $\langle B$, "subset of"$\rangle$ are lattices. $\langle C$, "part of"$\rangle$ is not: it does not have meet.

Lattices (as well as semi-lattices and Boolean algebras) can be equivalently defined in algebraic terms. E.g. a lattice is an algebra $\langle A, \wedge, \vee \rangle$, where $\wedge$ and $\vee$ are two-place operations satisfying idempotency, commutativity, associativity, and absorption. This otherwise important fact does not concern us, so it will not be dwelt on further.

(9) A *join semi-lattice* is the "upper half" of a lattice:
a poset $\langle A, \leq \rangle$ where for every $a, b \in A$, $\quad a \vee b \in A$.

(10) (a) (b) (c)



All three structures in (10) are join semilattices. (10a) is said to be "free," which means that whenever two distinct pairs of elements can possibly have distinct joins, they do have distinct joins. E.g., $\{a, b\}$ and $\{a, c\}$ have distinct joins; $\{a \vee b, a \vee c\}$ and $\{a \vee b, b \vee c\}$ do not have distinct joins, but they could not possibly have, either.

(11) A *meet semi-lattice* is the "lower half" of a lattice:
a poset $\langle A, \leq \rangle$ where for every $a, b \in A$, $\quad a \wedge b \in A$.

Mathematically, meet semilattices and join semilattices are the same thing, only the relation is inverted. Linguistically, it may be interesting to note that while there are many applications for join semilattices, I do not know of applications of meet semilattices. For instance, observe that (4c) is the same as (10a). The *and* that occurs in the definition of collectives is a join, not a meet.

(12) A lattice is *bounded* if it has a bottom element 0 and a top element 1.
For any a, $a \wedge 0 = 0$ and $a \wedge 1 = a$

For instance, the lattice in (4a) is bounded but the lattice of natural numbers is not, since it has no top (greatest) element.

(13) A *Boolean algebra* is a poset $\langle A, \leq \rangle$ which is closed under meet, join and (unique!) complement, where
$a \in A$ is a *complement* of $b \in A$ iff $a \wedge b = 0$ and $a \vee b = 1$.

For any set S, its powerset is the domain of a Boolean algebra. $\langle B$, "subset of"$\rangle$ is an example: B is the powerset of {joe, ed, pat} .

You may now want to check Problems (58) and (59).

What properties entail what others? Can a structure turn out to be closed under more operations than we stipulated? Yes! For many applications this does not matter: all we are interested in is that a certain operation is available. But if we claim that some linguistic phenomenon is explained by the fact that a certain operation is unavailable, matters like the following need to be paid close attention to.

(14) A lattice is *complete* iff for any (not just finite) subset X of A,
$\bigwedge X \in A$ and $\bigvee X \in A$.

Some facts: Every complete lattice is bounded (= has both 1 and 0). Every finite lattice is complete and bounded. Infinite lattices need not be complete or bounded.

(15) A join semi-lattice A is *complete* iff for any subset B, the supremum of B is in A.

(16) A join semi-lattice A is *complete#* iff for any non-empty subset B, the supremum of B is in A.

E.g., (10a) is complete#; if we add a bottom element, it becomes complete.

Some facts: Every complete join semi-lattice is a lattice; it is even a complete lattice, hence bounded. Not every complete# join semi-lattice is a lattice. Similarly, not every finite join semi-lattice is a lattice. See Problem (60).

1.2 Quantifiers and negation in Boolean terms

Finally, let us highlight the connection between the three Boolean operations and quantifiers. It is well-known that universal quantification reduces to conjunction, and existential quantification to disjunction over the elements of a finite universe. If the universe of discourse E is $\{a, b, c\}$, i.e. it contains Andy, Belinda, and Carl, then *Everyone walks* is the same as *Andy walks, Belinda walks, and Carl walks*; and *Someone walks* is the same as *Either Andy or Belinda or Carl walks.* That is,

(17) $\exists x[fx] = fa \vee fb \vee fc$

(18) $\forall x[fx] = fa \wedge fb \wedge fc$

Similarly for numerical quantifiers, negative quantifiers, and negation:

(19) $\exists_2 x[fx] = (fa \wedge fb) \vee (fa \wedge fc) \vee (fb \wedge fc)$

(20) $\neg\exists x[fx] = \neg(fa \vee fb \vee fc)$

(21) $\neg fa = a \in (E - \{x : fx\})$

Consider now the case when another quantifier is to take scope over the above, as in *Someone/everyone/no one read three books* on its object wide scope reading, for instance. An intermediate step is to define the property of being read by someone/everyone/no one. In present terms this can be spelled out as follows:

(22) $\{y : \exists x[r(x, y)]\} =$
$\{y : r(a, y) \vee r(b, y) \vee r(c, y)\} =$
$\{y : r(a, y)\} \cup \{y : r(b, y)\} \cup \{y : r(c, y)\}$

(23) $\{y : \forall x[r(x, y)]\} =$
$\{y : r(a, y) \wedge r(b, y) \wedge r(c, y)\} =$
$\{y : r(a, y)\} \cap \{y : r(b, y)\} \cap \{y : r(c, y)\}$

(24) $\{y : \neg\exists x[r(x, y)]\} =$
$E - \{y : r(a, y) \vee r(b, y) \vee r(c, y)\} =$
$E - \{y : r(a, y)\} \cup \{y : r(b, y)\} \cup \{y : r(c, y)\}$

That is, the narrow scope quantifier is cashed out in terms of the operations that define it.

2 GENERALIZED QUANTIFIERS

2.1 The elements of a GQ

Montague introduced generalized quantifiers into his grammar of English in order to be able to assign a uniform denotation to all noun phrases, whether they refer to single individuals or not. Going beyond this, GQ theory provides the tools for studying various semantic properties of quantifiers.

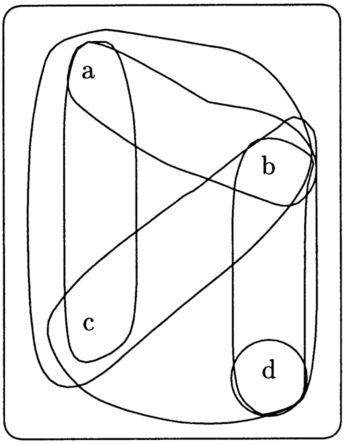
A generalized quantifier (henceforth, GQ) is not a syntactic object (an expression); it is a semantic object (something that expressions can denote). Specifically, a GQ is a set of properties, and noun phrases are claimed to denote such sets of properties. It is important to note that "property" is understood as nothing else than a set of individuals. E.g., if John, Bill, and Mary constitute the set of walkers, the property of walking is just {john, bill, mary}. In this sense, a GQ is a set of sets-of-individuals.

For instance, *every man* denotes the set of properties that every man has. The property of walking is in this set iff every man walks. Let us connect this to various terminologies and notations that are in use. In Montagovian terms, the denotation of *every man* is written as $\lambda P \forall x[\text{man}(x) \rightarrow P(x)]$. Here P is a variable of type $\langle e, t\rangle$: a variable over subsets of the universe. $\forall x[\ldots]$ is of type t. Hence the whole λ-expression is a function of type $\langle\langle e,t\rangle,t\rangle$. $\lambda P \forall x[\text{man}(x) \rightarrow P(x)]$ is the (characteristic function of the) set of properties every man has. Other ways of writing the same thing are $\lambda P[\textsc{man} \subseteq P]$ or $\{P : \textsc{man} \subseteq P\}$.

(At least) two men denotes the set of properties at least two men have, written as $\lambda P \exists x \exists y[x \neq y \ \&\ \text{man}(x) \ \&\ \text{man}(y) \ \&\ P(x) \ \&\ P(y)]$. Other ways of writing the same thing are: $\lambda P[|\textsc{man} \cap P| \geq 2]$ or $\{P : |\textsc{man} \cap P| \geq 2\}$.

Since GQs are sets (of sets of individuals), they have elements. For instance, *Every man walks* is true iff the set of walkers is an element of $[\![$*every man*$]\!]$, the GQ denoted by *every man*. When we are interested in (defining the conditions for) the truth of particular sentences, those sets that have "names" (that is, are denoted by the predicates in the sentence) are specifically interesting to us. However, when we are studying the GQs themselves, we are interested in all their elements and the structures they form. Hence no set is more interesting than the others. It is important to get into the habit of trading mnemonic names like *walk* for the corresponding sets and asking questions in the following form, "Is {john, bill, mary} an element of the quantifier denoted by *every man*?" (Yes, if the set of men is a subset of {john, bill, mary}.) For instance, consider a universe $E = \{a, b, c, d\}$ and some of its subsets (this example will be recycled in (41)):

(25) $\{a, b, c\}$ = man $\{d\}$ = dog $\{b, c, d\}$ = jump
$\{a, b, c, d\}$ = fat $\{a, b\}$ = run $\{b, d\}$ = laugh



On the other hand, the sets of all elements of a few quantifiers are as follows:

(26) $[\![$*at least two men*$]\!] = \{P : |\text{man} \cap P| \geq 2\}$
$= \{\{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}, \{a, c, d\}, \{a, b, d\}, \{b, c, d\}, \{a, b, c, d\}\}$

(27) $[\![$*every man*$]\!] = \{P : \text{man} \subseteq P\} = \{\{a, b, c\}, \{a, b, c, d\}\}$

(28) $[\![$*no man*$]\!] = \{P : \text{man} \cap P = \emptyset\} = \{\{d\}, \emptyset\}$

(29) $[\![$*andy and carl*$]\!] = \{P : P(a)\ \&\ P(c)\} = \{P : \{a, c\} \subseteq P\}$
$= \{\{a, c\}, \{a, b, c\}, \{a, c, d\}, \{a, b, c, d\}\}$

In the previous section it was noted that quantifiers are reducible to Boolean operations. In GQ-theoretic terms (18) may be rephrased as follows. We have a universe with three humans, a, b, and c. $[\![$*Everyone*$]\!]$, the set of properties everyone has, can be obtained by intersecting the sets of properties the individual humans have:

(30) $[\![$*everyone*$]\!] = \{P : P(a)\} \cap \{P : P(b)\} \cap \{P : P(c)\}$

And similarly for the other quantifiers.

2.2 Determiners (DETS)

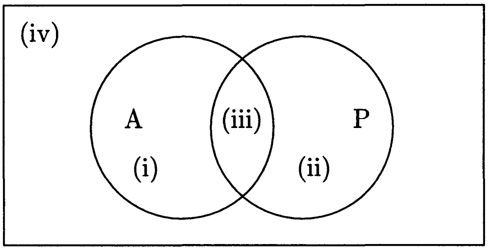
GQ theory does not concern itself only with GQs. It also deals with the denotations of determiners and with the denotations of noun phrases that are not exactly GQs (e.g. *himself*). Here we will not be directly concerned with determiners, but below is a small portion of necessary information.

In Montagovian terms, the denotation of *every* is written as $\lambda A \lambda P \forall x[A(x) \rightarrow P(x)]$. Here A is a variable of type $\langle e, t\rangle$, $\lambda P \forall x[\ldots]$ is of type $\langle\langle e, t\rangle, t\rangle$, hence the whole thing is of type $\langle\langle e, t\rangle, \langle\langle e, t\rangle, t\rangle\rangle$.

$\lambda A \lambda P \forall x[A(x) \rightarrow P(x)]$ is a function from properties to GQs or, equivalently, a relation between properties (A's and P's). Other ways of writing the same thing are, $\lambda A \lambda P[A \subseteq P]$ or $\{\langle A, P\rangle : A \subseteq P\}$.

Or, *two* denotes $\lambda A \lambda P \exists x \exists y[x \neq y \;\&\; A(x) \;\&\; A(y) \;\&\; P(x) \;\&\; P(y)]$. Other ways of writing the same thing: $\lambda A \lambda P[|A \cap P| \geq 2]$ or $\{\langle A, P\rangle : |A \cap P| \geq 2\}$.

Now consider the diagram below. It has four areas: (i) the individuals that have property A but not P, (ii) the individuals that have P but not A, (iii) the individuals that have both A and P, and (iv) the individuals that have neither A nor P.



Consider a sentence of the form $\text{DET}(A)(P)$. Do we need to check all four areas when we wish to determine whether it is true or false? It is an important empirical claim concerning natural language determiners (at least "simple" or "normal" ones) that they do not require the checking of all four areas. The following is a small but representative sample. The solidus in (c) indicates a fraction, and *n, m,* and *k* are natural numbers.

(31)	a.	At least two men walk.	$\lvert A \cap P\rvert \geq 2$
	b.	Every man walks.	$A \subseteq P$
	c.	Few men walk.	$\lvert A \cap P\rvert / \lvert A\rvert \leq n/m$ or $\lvert A \cap P\rvert \leq k$
	d.	No men walk.	$\lvert A \cap P\rvert = 0$

As the reader can easily check, none of these requires us to consider area (iv): their truth does not depend on non-walking non-men. We need not know anything beyond the properties explicitly mentioned: how big the surrounding universe is and what is going on in it are immaterial. A more surprising but equally intuitive fact is that none of these sentences requires us to check area (ii): their truth does not depend on walkers who are not men. On the other hand, (31b) and the first reading of (31c) require us to check (i): their truth is dependent on men who are not walkers. The irrelevance of areas (iv) and (ii) means that the two sets A and P do not play equal roles. The set A, the denotation of the noun that the determiner directly combines with, serves to restrict the universe to the largest part that can be possibly relevant: it serves as the determiner's restrictor. Natural language determiners are (overwhelmingly) restricted in this sense. Finally, observe that area (iii) is useful: (31a), (31d) and the second reading of (31c) require us to check only this.

Below are the definitions of the pertinent properties of determiners:

(32) DET has extension iff for any two universes E and E' where $A, B \subseteq E$ and $A, B \subseteq E'$, we have $D_E(A)(B) = D'_E(A)(B)$.

(33) DET is conservative iff $\text{DET}(A)(P) = \text{DET}(A)(A \cap P)$.

(34) DET is intersective iff $\text{DET}(A)(P) = \text{DET}(A \cap P)(P)$.

(35) DET is proportional iff $\text{DET}(A)(P)$ depends on $(A \cap P)/A$.

(36) DET is symmetrical iff $\text{DET}(A)(P) = \text{DET}(P)(A)$.

Two facts: A proportional DET cannot be symmetrical. If DET is conservative, symmetrical = intersective. See Problem (61).

Now back to GQs.

2.3 Live-on sets and witness sets

Conservativity is a property of determiners. Together with extension, it identifies DET's first argument as a restrictor set. A comparable notion for GQs is that of a live-on set.

(37) *Live-on*: A GQ lives on a set of individuals A if, for any set of individuals X,

$$X \in GQ \quad \text{iff} \quad (X \cap A) \in GQ.$$

(37) says that when a GQ lives on some set A, it makes no difference whether we check if a set X is an element of that GQ, or we check whether the intersection of X with A is an element of it; that is, we may safely restrict our attention to the smaller set $X \cap A$. What are a GQ's live-on sets? A linguistic way to check this is to instantiate the schema, as follows:

(38) More/fewer than two men run

	$\leftrightarrow$	More/fewer than two men are *men* who run
	$\leftrightarrow$	More/fewer than two men are *humans* who run
	$\leftrightarrow$	More/fewer than two men are *existents* who run
but:	$\not\leftrightarrow$	More/fewer than two men are *Frenchmen* who run

So $[\![$*more than two men*$]\!]$ and $[\![$*fewer than two men*$]\!]$ live on the set of men and its supersets. In general, the restrictor of the determiner is always a live-on set of the corresponding generalized quantifier. See Problem (62).

If we are interested in live-on sets as domains that we need not look beyond when checking the truth of a sentence, we do not need all of them: the smallest suffices and is thus the most efficient.[2]

In the above cases the restrictor set of DET is identical to the smallest set the GQ denoted by DET(A) lives on: A itself. So, do we need the notion of a smallest live-on set on top of the set with respect to which DET is conservative? The answer is Yes.

First, there are noun phrases that are not made up of a determiner and a noun, e.g., *John* and *John and Mary*. Here the question of what the determiner's restrictor is cannot arise. But the GQs that these noun phrases denote have run-of-the-mill smallest live-on sets: {john} and {john, mary}, respectively.

Second, the smallest live-on sets of some GQs are smaller than the restrictor sets of the corresponding determiners. Imagine a world in which the men are {john, bill, tim} and we are pointing at John and Bill:

(39) These two men run $\leftrightarrow$ These two men are *either John or Bill* and run

So, $[\![$*these two men*$]\!]$ lives on the set consisting of those two men who we are pointing at, which is smaller than the set of men. (This amounts to saying that *these two men* is interpreted as 'the two men I am pointing at'. Note though that while this interpretation is semantically justified, a syntactic analysis that mimics such a decomposition would not be.)

These discrepancies are understandable. Conservativity (with extension) may be regarded as a property of the syntax/semantics interface. It says that the syntactic unit that a determiner (or other two-place operator) directly combines with plays the semantic role of a restrictor, i.e. imposes a parallelism between syntax and interpretation. Live-on sets on the other hand are defined purely from the denotation of the noun phrase, without reference to its syntax and without requiring a direct syntactic correlate.

You may now want to tackle Problems (63)–(64).

With the notion of a smallest live-on set at hand, we may take another look at the elements of a GQ. What all the elements of a GQ are is characteristic of it; the individual elements themselves need not be. Take, for instance, the

[2] E. Keenan (p.c.) notes that the notion of a smallest live-on set is unproblematic as long as the universe is finite or at least our GQ does not crucially rely on infinity. But e.g. the intersection of the sets which $[\![$*all but finitely many stars*$]\!]$ lives on in an infinite universe is itself not a live-on set.

elements of $[\![$*two men*$]\!]$. They are those sets in the universe that contain two men—but note that they may as well contain tigers, stars, and forks.

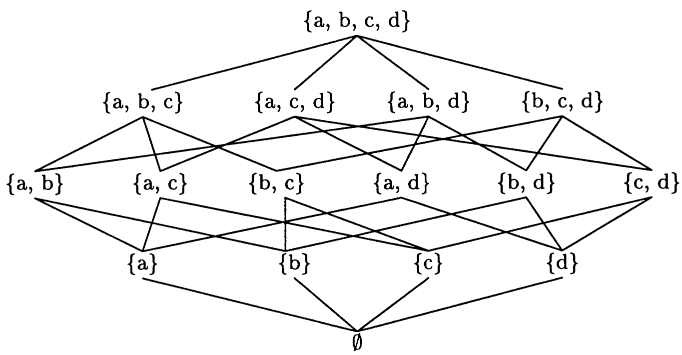
If we ask ourselves what sets the noun phrase "talks about", the elements of its GQ do not make a revealing choice. A natural alternative is to throw out the irrelevant beasts by restricting our attention to those elements that are also in the smallest live-on set:

(40) A set W is a *witness* of a GQ iff $W \in GQ$ and $W \subseteq SL(GQ)$, where $SL(GQ)$ is the smallest set the GQ lives on.

To compare elements and witnesses, we may consider a reincarnation of the four-element universe in (25). (41) is the Boolean algebra corresponding to its powerset. Its use is insightful, because it contains all subsets of the universe, not only those that have "mnemonic names"; and since it is partially ordered by the subset relation, it allows us to make inferences by simply going up or down in the diagram.

Recall that in our particular universe, a, b, c are men and d is a dog.

(41)



(42) a. $[\![$*more than one man*$]\!] = \{\{a,b\},\{a,c\},\{b,c\},\{a,b,c\},$ $\{a,c,d\},\{a,b,d\},\{b,c,d\},\{a,b,c,d\}\}$

b. the witnesses of $[\![$*more than one man*$]\!] =$ $\{\{a,b\},\{a,c\},\{b,c\},\{a,b,c\}\}$

(43) a. $[\![$*fewer than two men*$]\!] =$ $\{\{a\},\{b\},\{c\},\{d\},\{a,d\},\{b,d\},\{c,d\},\emptyset\}$

b. the witnesses of $[\![$*fewer than two men*$]\!] = \{\{a\},\{b\},\{c\},\emptyset\}$

See Problems (65) through (67).

Some GQs have a unique witness. This may be empty: such is the case with $[\![$*no man*$]\!]$: its only witness is the empty set. Or, the unique witness may be non-empty; in this case it coincides with the smallest set that the GQ lives on. GQs with a non-empty unique witness are called principal filters, and their unique witness A the generator set:

(44) A GQ is a *principal filter* iff there is a set of individuals A such that A is not necessarily empty and for any set of individuals X,
$X \in GQ$ iff $A \subseteq X$.

Every man, these$_{\text{deictic}}$ *two men, Andy, Andy and Carl,* etc. all denote principal filters. These always "talk about" the same sets, their generators. In terms of (41):

(45) a. $[\![$*andy and carl*$]\!]$ is a principal filter generated by $\{a, c\}$:
$[\![$*andy and carl*$]\!] = \{P : \{a, c\} \subseteq P\}$

b. the smallest live-on set of $[\![$*andy and carl*$]\!]$ = unique witness of $[\![$*andy and carl*$]\!]$ = generator set of $[\![$*andy and carl*$]\!]$ = $\{a, c\}$

You may now want to think about Problems (68) through (72).

2.4 Monotonicity properties and witnesses

An important property of functions is what monotonicity type they belong to. Suppose the domain of a function f is a partially ordered set with, say, $a \geq b$. If f is upward monotonic, it preserves this ordering in its value: $f(a) \geq f(b)$. If f is downward monotonic, it reverses the ordering: $f(b) \geq f(a)$. If f is non-monotonic, it obliterates the ordering. Since GQs are functions (characteristic functions of sets of properties), their monotonicity can be examined.

(46) GQ is monotone increasing (= upward mon.):
$(A \in GQ \ \&\ A \subseteq B) \Rightarrow B \in GQ$.

(47) GQ is monotone decreasing (= downward mon.):
$(A \in GQ \ \&\ B \subseteq A) \Rightarrow B \in GQ$.

(48) GQ is non-monotone: neither increasing nor decreasing.

Some examples: *John, at least two men, every man* denote increasing GQs. *No men, fewer than six men* denote decreasing GQs. *John and nobody else* and *exactly two men* denote non-monotonic GQs. Here is a linguistic way to show these:

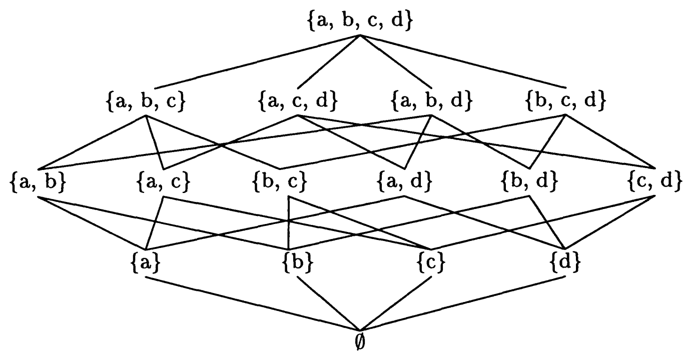
(49) (Every man runs & run $\subseteq$ run or sit) $\Rightarrow$ Every man runs or sits

(50) (Few men run or sit & run $\subseteq$ run or sit) $\Rightarrow$ Few men run

(51) (Exactly two men run & run $\subseteq$ run or sit) $\not\Rightarrow$
Exactly two men run or sit
(Exactly two men run or sit & run $\subseteq$ run or sit) $\not\Rightarrow$
Exactly two men run

A more general and also visualizable way to demonstrate monotonicity properties is to use a Boolean algebra as in (41), repeated here:

(41)



The algebra makes it easy to see, for instance, that if some set A is an element of $[\![$*at least two men*$]\!]$, i.e., has at least two men in it, then every set B that is larger than A (is above A in (41)) is also an element of this GQ; and conversely for, say, $[\![$*no man*$]\!]$:

(52) $[\![$*at least two men*$]\!]$ is monotone increasing:

for every A, B, ($A \in [\![$*at least two men*$]\!]$ & $B \supseteq A$) $\Rightarrow$
$B \in [\![$*at least two men*$]\!]$

(53) $[\![$*no man*$]\!]$ is monotone decreasing:

for every A, B, ($A \in [\![$*no man*$]\!]$ & $A \supseteq B$) $\Rightarrow$ $B \in [\![$*no man*$]\!]$

See Problem (73).

The best known linguistic application of monotonicity properties has to do with the licensing of negative polarity items. We will be making crucial use of another type of consequence of monotonicity differences.

Let a noun phrase contain a determiner that provides information concerning cardinality, e.g., *two, at least two, more than two, at most two, less than two, exactly two.*

(54) If the GQ denoted by a cardinality-indicating noun phrase is monotone increasing, then

$$\text{DET}(A)(P) = \exists X[X \subseteq A \ \& \ |X| = \text{det-many} \ \& \ X \subseteq P]$$

That is, *At least two men walk* can be equivalently stated as, 'There is a set of individuals whose elements are all men, whose cardinality is at least two, and whose elements all walk.' Or, using witnesses, 'There is a witness of ⟦*at least two men*⟧ and its elements all walk.'

If, on the other hand, the GQ denoted by such a noun phrase is not increasing, then there is NO such equivalence. E.g., *Fewer than two men walk* is NOT equivalent to 'There is a set consisting of less than two men, and these men all walk'. Imagine the situation in which John walks, Bill walks, and Frank walks. The set {john} is surely one that has fewer than two elements, all of which are men and walk—but the existence of such a set does not make the sentence true here. The sentence does not allow us to ignore Bill and Fred, who also walk, but the proposed paraphrase allows us to ignore them. Or, *Exactly two men walk* is NOT equivalent to 'There is a set of individuals whose elements are all men, whose cardinality is exactly two, and whose elements all walk'. Imagine the same situation and pick the set to be {john, frank}, to see why not.

Note why this is so. The crucial property of upward monotonicity is that whatever is true in a small situation (say, one in which just two men walk) remains true when we embed that situation in a bigger one (in which three or more men walk). Neither downward monotonic nor non-monotonic quantifiers have this property, which means that to be safe, we must always look at the biggest possibly relevant situation.

The significance of these simple observations is that in the analysis of linguistic phenomena, one often wishes to associate existentially quantified sets with GQs. Great caution needs to be exercised in these cases. Either the phenomenon we are looking at is factually restricted to increasing GQs, or a maximality condition of some sort must be added to guarantee that no relevant individual gets ignored.

The relation between monotonicity and witnesses can be generally characterized as follows. Let W be a witness, and A the smallest live-on set, of GQ. Then,

(55) If GQ is monotone increasing, then for any X, $X \in GQ$ iff $\exists W[W \subseteq X]$.

E.g., *Two men run* is true iff there is a witness of ⟦*two men*⟧ whose members run.

(56) If GQ is monotone decreasing, then for any X, $X \in GQ$ iff $\exists W[(X \cap A) \subseteq W]$.

E.g., *Few men run* is true iff there is a witness of $[\![$*few men*$]\!]$ which contains all the men who run.

(57) If GQ is non-monotonic, then for any X, $X \in GQ$ iff $\exists W[(X \cap A) = W]$.

E.g., *Exactly two men run* is true iff there is a witness of $[\![$*exactly two men*$]\!]$ which equals all the men who run.

The observation in (54) is a special case of (55). On the other hand, the formulations in (56) and (57) ensure that we are looking at the maximal set: we are not "ignoring" anything. Just as (55) does not hold of decreasing GQs, (56) does not hold of increasing ones. MAN is a W, and the A, for $[\![$*at least two men*$]\!]$. Suppose that only one man runs. (RUN $\cap$ MAN) $\subseteq$ MAN does not entail that RUN $\in [\![$*at least two men*$]\!]$.

Since W is a subset of smallest live-on set A anyway, in (55) we might have used $\exists W[W \subseteq (X \cap A)]$, to bear out the pattern common to the three cases: there exists a witness W that contains, is contained by, or equals $X \cap A$.

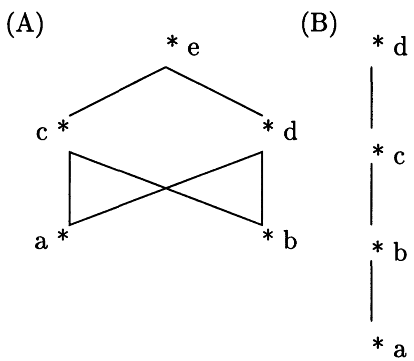
*

Finally, note that Section 1 of the next chapter (Beghelli et al. 1996) may be regarded as an extension of the present one: it is concerned with the use of witness sets in capturing some basic intuitions concerning scope.

3 PROBLEMS

Unmarked problems involve applying the definitions in the text. Those marked with an asterisk may require some creativity.

(58) Is (A) below a semi-lattice? Is (B) a Boolean algebra? Why?



(59) The structure in (10b) does not have elements labelled $b \vee c$ or $a \vee c$. How come it is still a join semi-lattice?

(60) Show that every *complete* semi-lattice A is a lattice. (Hints: Assume A is a meet semi-lattice. What is the $\bigwedge$ of the set of upper bounds of an arbitrary $X \subseteq A$? What is the $\bigvee$ of this set?)

(61) Show that for conservative DETs, symmetry = intersectivity.

(62) Show that $[\![$*more than one man*$]\!]$ and $[\![$*fewer than two men*$]\!]$ do not live on $\{a, b, d\}$.

(63) The textbook example of a "potential determiner" that is not conservative is *only*. Assume that *only men* is a noun phrase in *Only men run.* Demonstrate that *only* is not a conservative "determiner".

(64)* Formalize *Only men run* and *Only John and Bill run* using first order logic, and complete the following: If *only* was interpreted as ... , with restrictor ... and scope ... , it would turn out to be conservative. Are there linguistic arguments supporting this analysis?

(65) Is {John, Bill, Fido} (a) an element, (b) a witness of $[\![$*at least two men*$]\!]$?

(66) What are the witnesses of $[\![$*fewer than four men*$]\!]$ and of $[\![$*few men*$]\!]$ in the Boolean algebra (41)?

(67) The elements of a GQ are too big to be genuinely characteristic of it; the text suggests the use of witnesses. Couldn't we use minimal elements instead? The definition is this: X is a minimal element of GQ iff X is an element of GQ but ceases to be one if we take away even one individual from it. (Hint: What are the minimal elements of the GQs denoted by (i) *at least two men, more than one man, exactly two men,* (ii) *fewer than three men, at most one man, no man*? What are their witnesses?)

(68) (a) Is $[\![$*no man*$]\!]$ a principal filter? Why? (b) Is $[\![$*every man*$]\!]$ a principal filter? Why?

(69)* What set does *they* in (i) and (ii) refer to? Argue for your proposal with reference to whether they can be continued with *Perhaps there were others who did the same (i.e. both came in and were selling coke).* Formalize your proposal using notions introduced in 2.3.

(i) More than two people came in. They were selling coke.

(ii) At least two people came in. They were selling coke.

(70)* Is there a difference between the behavior of (i)–(ii) in (69) and that of (iii)? Sticking with the machinery of 2.3, come up with an interpretation for *two people* that makes the correct prediction without requiring a new rule, i.e. try to make (iii) a special case of (i)–(ii).

(iii) Two people came in. They were selling coke.

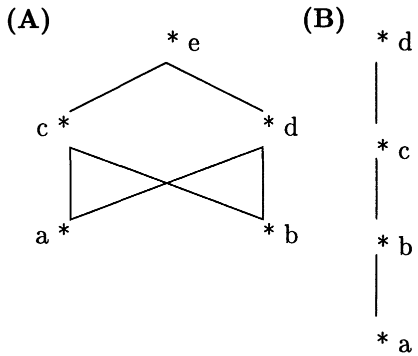
(71)* Compare the following sentences: (i) *A dog/every dog bit two women (you know, my neighbors)* and (ii) *A dog/every dog bit two or more/more than three women.* Do they have both a subject widest scope and an inverse, object wide scope reading? Set up a hypothesis that explains the data.

(72)* Examine now what readings *Every prof assigned more than two readings to three students* and *Every prof assigned three readings to more than two students* have. Do your findings make you change the hypothesis concerning inverse scope that was made in (71)?

(73) Is the smallest live-on set of (a) an increasing, (b) a decreasing, (c) a non-monotonic quantifier an element of that quantifier?

4 SOLUTIONS

(58) Is (A) below a semi-lattice? Is (B) a Boolean algebra? Why?



(a) No. The join is the least upper bound of a two-element set. $\{c, d, e\}$, the set of upper bounds of $\{a, b\}$, has no least element since c and d are equal. So there is no $a \vee b$ in A.

(b) No. B has no complements. Take b, for example. The complement of b ought to be another element of B (not a subset of B!). Now, $b \wedge a = a$ (and a is the bottom element 0) and $b \vee d = d$ (and d is the top element 1), but $a \neq d$. In fact, there is no element of B for which both equations would hold.

(59) The structure in (10b) does not have elements labelled $b \vee c$ or $a \vee c$. How come it is still a join semi-lattice?

Because $a \vee b \vee c$ is the least upper bound for $\{a, c\}$ and $\{b, c\}$: it is an upper bound, and there is no smaller upper bound in the structure. The fact that we could "imagine" a distinct $b \vee c$ does not matter: what matters is what elements the structure actually has.

(60) Show that every *complete* semi-lattice A is a lattice. (Hints: Assume A is a meet semi-lattice. What is the $\bigwedge$ of the set of upper bounds of an arbitrary $X \subseteq A$? What is the $\bigvee$ of this set?)

A being a complete meet semi-lattice means that not only every two-element subset, but any subset, of A has a greatest lower bound in A. What we need to show is that this guarantees that *every* subset also has a least upper bound. What is a least upper bound of $X \subseteq A$? It is the infimum of the set of upper bounds of X:

$$\bigvee X = \bigwedge (UB(X))$$

We can always define the set of upper bounds of X:

$$UB(X) = \{a \in A : \text{every } x \in X,\ a \geq x\}$$

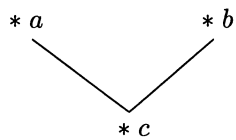
Since this is a subset of A, and A is complete, we know that its infimum is in A. So,

$$\bigwedge(UB(X)) \in A$$

Thus, if every subset has an infimum in A, then every subset has a supremum in A, too. The case of a join semi-lattice does not require new considerations: everything works the same, replacing $\bigwedge$ with $\bigvee$, upper by lower, etc.

This exercise highlights the fact that "greatest" (as in greatest lower bound) is defined using least upper bound, and "least" (as in least upper bound) is defined using greatest lower bound.

To give a concrete example of a meet semi-lattice that is not complete, consider



What subset lacks an infimum here? Well, the empty subset. Its lower bounds are a, b, and c since it is vacuously true of each of these that it is smaller than or equal to all the elements of the empty set. But this set of lower bounds $\{a, b, c\}$ has no greatest element, so the empty subset lacks an infimum. The missing top element would be the supremum of $\{a, b\}$, so its absence also prevents our structure from being a lattice, in accordance with the theorem just proved.

(61) Show that for conservative DETs, symmetry = intersectivity.

Symm: $D(A)(P) = D(P)(A)$
Int: $D(A)(P) = D(A \cap P)(P)$
Cons: $D(A)(P) = D(A)(A \cap P)$

Int $\Rightarrow$ symm:	$D(A)(P) =$	by cons
	$D(A)(A \cap P) =$	by int
	$D(A \cap (A \cap P))(A \cap P) =$	by def. of $\cap$
	$D(A \cap P)(A \cap P) =$	by def. of $\cap$
	$D(P \cap (P \cap A))(P \cap A) =$	by int "reversed"
	$D(P)(P \cap A) =$	by cons
	$D(P)(A)$	Symm!

Symm $\Rightarrow$ int:	$D(A)(P) =$	by symm
	$D(P)(A) =$	by cons
	$D(P)(P \cap A) =$	by symm
	$D(P \cap A)(P) =$	by def. of $\cap$
	$D(A \cap P)(P)$	Int!

(62) Show that $[\![$*more than one man*$]\!]$ **and** $[\![$*fewer than two men*$]\!]$ **do not live on** $\{a, b, d\}$**.**

$\{a, b, d\}$ has one of the men, c missing. Now, $\{a, c\}$ is an element of $[\![$*more than one man*$]\!]$, but $\{a, b, d\} \cap \{a, c\} = \{a\}$ is not. And conversely, $\{a, b, d\} \cap \{a, c\} = \{a\}$ is an element of $[\![$*fewer than two men*$]\!]$, although $\{a, c\}$ is not.

(63) The textbook example of a "potential determiner" that is not conservative is *only.* **Assume that** *only men* **is a noun phrase in** *Only men run.* **Demonstrate that** *only* **is not a conservative "determiner."**

Only men (if interpreted as a semantic constituent) is a GQ that does not live on the set of men at all, to wit:

Only men run $\not\Leftrightarrow$ Only men are *men* who run
Only men run $\Leftrightarrow$ Only men are *existents* who run

That *only* is not conservative is not very problematic: we can argue that it is simply not a determiner but a noun phrase modifier.

(64)* Formalize *Only men run* **and** *Only John and Bill run* **using first order logic, and complete the following: If** *only* **was interpreted as ... , with restrictor ... and scope ... , it would turn out to be conservative. Are there linguistic arguments supporting this analysis?**

$\forall x[\text{run}(x) \rightarrow \text{man}(x)]$
$\forall x[\text{run}(x) \rightarrow (x = \text{john} \vee x = \text{bill})]$

If *only* was interpreted as a universal quantifier, with the VP as its restrictor and the subject as its scope, it would turn out to be conservative:

Every runner is a man $\Leftrightarrow$ Every runner is a runner who is a man
Every runner is either John or Bill $\Leftrightarrow$ Every runner is a runner who is either John or Bill

(NB: We are not arguing that *only* is a determiner in syntax; we are arguing that semantically it is a conservative operator.)

This analysis is quite plausible. First, spelling out the contribution of *only* is obviously necessary anyway, and there is no reason why *only* should not be interpreted as 'all.' Second, *only* is a well-known focusing operator, that is, the phrase it combines with is the focus and the rest of the sentence is the focus frame. Since the focus–focus frame partition is assumed to be reflected in the syntax of Logical Form, and since many operators (e.g. adverbs of quantification) are assumed to have the focus frame as their restrictor, the same is natural in connection with *only.* So we are positing the following analogy:

Only MEN run	$\text{every}(_{\text{focus frame}}\{x : \text{run}x\})(_{\text{focus}}\{x : \text{man}x\})$
John always cites MEN	$\text{every}(_{\text{focus frame}}\{x : \text{John cites}x\})(_{\text{focus}}x : \text{man}x\})$

The only objection might be that *Only men run* requires the existence of runners, and the formula $\forall x[\text{run}(x) \rightarrow \text{man}(x)]$ does not. But this does not need to be specified in the meaning of *only*: it is generally assumed that sentences with focus presuppose that the property denoted by the focus frame is not empty. This analysis suggests that conservativity (or generally, domain restriction) may be far more pervasive than generally thought. It may be characteristic of all two-place operators, not only of determiners. This is natural if conservativity (domain restriction) indeed characterizes the syntax/semantics interface. This hypothesis suggests that the syntactic analyses of potential counterexamples should be checked and possibly recast.

(65) Is {John, Bill, Fido} (a) an element, (b) a witness of ⟦*at least two men*⟧?

(a) Yes, because the intersection of {John, Bill, Fido} with MAN has at least two members.

(b) No, because a witness of ⟦*at least two men*⟧ is an element of it that contains only men, and here we have a dog, too.

(66) What are the witnesses of ⟦*fewer than four men*⟧ and of ⟦*few men*⟧ in the Boolean algebra (41)?

(a) No element of this algebra has more than three men in it, so in this respect all qualify. But we need to discard those that contain d, the dog.

(b) *Few men* may mean either of two things. (i) 'fewer than a set number k'—if we set k as, say, 7, then the witnesses will be those sets that contain six men or less and no dog. This is independent of how many men we have in fact. (ii) 'few of the men'—we may stipulate that, say, 30% or less of the men counts as few of them; since we have 3 men, this will come down to 'at most one man'. So the witnesses are those sets that contain at most one man and no dog.

(67) The elements of a GQ are too big to be genuinely characteristic of it; the text suggests the use of witnesses. Couldn't we use minimal elements instead? The definition is this: X is a minimal element of GQ iff X is an element of GQ but ceases to be one if we take away even one individual from it. (Hint: What are the minimal elements of the GQs denoted by (i) ***at least two men, more than one man, exactly two men,*** **(ii)** ***fewer than three men, at most one man, no man*****? What are their witnesses?)**

All the GQs in (i) have minimal elements consisting of exactly two men, and all the GQs in (ii) have the empty set as their unique minimal element. This indicates that the notion of a minimal element does not only eliminate irrelevant individuals but also eradicates "size" distinctions and therefore collapses noun phrases it should not. On the other hand, while witnesses eliminate individuals outside the restrictor, they retain "size" distinctions.

(68) (a) Is $[\![no\ man]\!]$ a principal filter? Why? (b) Is $[\![every\ man]\!]$ a principal filter? Why?

For GQ to be a principal filter, there must be a non-empty set A such that (i) if some X is an element of GQ, A is a subset of X, and (ii) if A is a subset of some X, X is an element of GQ.

(a) What sets come to mind in connection with $[\![no\ man]\!]$? Say, $\emptyset$ (its unique witness) and MAN (its smallest live-on set). $\emptyset$ satisfies (i), because it is a subset of any X, but not (ii), for the same reason (say, $\emptyset \subseteq$ WALK does not entail that no man walks). In addition, the generator set should be non-empty. So try MAN. MAN clearly does not satisfy (ii): MAN $\subseteq$ HUMAN does not entail that no man is a human. Indeed, the fact that the unique witness and the smallest live-on set differ already indicates that $[\![no\ man]\!]$ is not a principal filter.

(b) The smallest set $[\![every\ man]\!]$ lives on is MAN. The definition of *every* says that for any set X, $X \in [\![every\ man]\!]$ iff $X \in \{P : \text{MAN} \subseteq P\}$, and the latter is equivalent to MAN $\subseteq X$. MAN is also the GQ's unique witness. There may be models in which MAN is empty. The definition in the text is "modalized" in order to allow for this: it requires the set the GQ is preoccupied with to be not always empty. This practically allows us to ignore models without men. Alternatively, we might say that *every man* denotes a principal filter in those models where there are men. In any case, we take *every man* to be an uncontroversial principal filter.

(69)* What set does ***they*** **in (i) and (ii) refer to? Argue for your proposal with reference to whether they can be continued with** *Perhaps there were others who did the same (i.e. both came in and were* *selling coke).* **Formalize your proposal using notions introduced in 2.3.**

(i) More than two people came in. They were selling coke.

(ii) At least two people came in. They were selling coke.

They refers back to all the people who came in. This is confirmed by the fact that the continuation "perhaps there were others ... " is no good: if all are referred to, there cannot be others. Noun phrases like *more than two people* are said to support only "maximal reference anaphora." A formalization may be: $SL(GQ) \cap VP$, where $SL(GQ)$ is the smallest set the subject of the first sentence lives on (here: MAN), and VP is the predicate of the first sentence (here: CAME IN).

(70)* Is there a difference between the behavior of (i)–(ii) in (69) and that of (iii)? Sticking with the machinery of 2.3, come up with an interpretation for *two people* that makes the correct prediction without requiring a new rule, i.e. try to make (iii) a special case of (i)–(ii).

(iii) Two people came in. They were selling coke.

Here the continuation "perhaps there were others ... " is good, so *they* cannot be referring to all the men who came in. It refers to just the two men the speaker was talking about in the first sentence. Noun phrases like *two people* are said to support "non-maximal reference anaphora." It might be argued that a specific (referential) interpretation of *two people* is what enables this reading; *more than/at least two people* does not seem to have a comparable interpretation.

We may formalize this referential interpretation by saying that *two men* (on this reading) denotes a principal filter; those two men who the speaker has in mind. The smallest live-on set of such a principal filter is smaller than the set of men: it contains just the two relevant individuals. The intersection of this set with VP is just the two-man set. So the formalization in (69) extends to this case and thus the example may speak in favor of a principal filter interpretation of *two men* (among other readings).

(71)* Compare the following sentences: (i) *A dog/every dog bit two women (you know, my neighbors)* and (ii) *A dog/every dog bit two or more/more than three women.* Do they have both a subject wide scope and an inverse, object wide scope reading? Set up a hypothesis that explains the data.

(i) can easily have both readings. (ii) does not easily have an object wide scope reading.

In (70) we have seen that specific indefinites may be regarded as denoting principal filters. In (i), the phrase *you know, my neighbors* suggests that we are dealing with a specific indefinite, too. The following descriptive hypothesis suggests itself. When the direct object receives a "specific" interpretation, it takes scope over the subject; when it does not, it cannot. *Two men* can be specific, because on one reading it denotes a principal filter; *two or more/more than three women* have no such readings.

(72)* Examine now what readings *Every prof assigned more than two readings to three students* **and** *Every prof assigned three readings to more than two students* **have. Do your findings make you change the hypothesis in (71)?**

The trick here is that we now have three quantifiers! The hypothesis in (71) can be checked by asking whether *three N* can take intermediate scope, that is, inverse scope inside the VP and still vary with the subject. Since the hierarchical order of VP-internal complements is a matter of debate, we check two sentences: in at least one of them *three N* must be taking inverse scope if it scopes highest inside the VP. So, are the following readings possible?

every prof > three students > more than two readings

and

every prof > three readings > more than two students

If yes, then on this construal *three students* and *three readings* do not denote principal filters. If they did, their witnesses could not vary with the individual professors. Also, if the given reading is available, then *three N* need not denote a principal filter in order to take inverse scope over the c-commanding *more than two N.*

The judgment seems to be that the critical reading is available. So the hypothesis in (71) is refuted. This example does not refute the assumption that *two men* can denote a principal filter. What it shows is that denoting a principal filter is not necessary for taking inverse scope.

This conclusion makes one want to go back and check if denoting a principal filter is strictly necessary for *two people* to support anaphora as in (iii) above. For instance, the following modified context is useful: *Every policeman reported that the following happened at 6 p.m. in the building he was watching. Two people entered. They were selling coke. Perhaps there were others who did the same ...* Indeed, it seems possible for the pairs to vary with the policemen and still support anaphora in the same way. This indicates that defining the antecedent as $SL(GQ) \cap VP$ is not a sufficiently general solution. (Indeed, a radically different treatment for the type of *two people* is proposed in Discourse Representation Theory.)

(73) Is the smallest live-on set of (a) an increasing, (b) a decreasing, (c) a non-monotonic quantifier an element of that quantifier?

The quantifiers $[\![$*at least one man*$]\!]$, $[\![$*fewer than two men*$]\!]$ and $[\![$*exactly one man*$]\!]$ have the same smallest live-on set: $[\![$*man*$]\!]$. *At least one man is a man* is true, *Fewer than two men are men* is false, and *Exactly one man is a man* is also false in the model (25), since we have three men who are all men. We see that its smallest live-on set may be too big to be an element of a decreasing or a non-monotonic quantifier. In the case of the increasing ones, we can be sure we cannot get into trouble: if, say, $\{a\} \in [\![$*at least one man*$]\!]$, then every superset of $\{a\}$ is.

REFERENCES

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2):159–219.

Beghelli, Filippo, Dorit Ben-Shalom, and Anna Szabolcsi. 1996. Variation, distributivity, and the illusion of branching. In this volume.

Keenan, Edward, and Dag Westerståhl. 1994. Generalized quantifiers in linguistics and logic. In Johan van Benthem and Alice ter Meulen (Eds.), *Handbook of Logic and Linguistics*. Amsterdam: Elsevier. Forthcoming.

Landman, Fred. 1991. *Structures for Semantics*. Dordrecht: Kluwer.

2

VARIATION, DISTRIBUTIVITY, AND THE ILLUSION OF BRANCHING*

Filippo Beghelli, Dorit Ben-Shalom, and Anna Szabolcsi

Department of Linguistics
UCLA

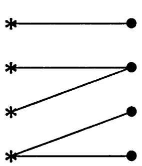
A well-known observation is that (1) has a reading on which *every building* scopes over *a fireman,* but (2) does not:

(1) A fireman checks the safety of every building.

(2) A fireman imagined that every building was unsafe.

The way to justify this claim is to point out that in (1), but not in (2), firemen can vary with buildings. For instance, (1) but not (2) is true in the following situation. In the diagram below, the four $*$'s represent all the buildings, and the $\bullet$'s firemen:

(3)



What is a precise way of saying what we did in drawing this diagram?

This question is the point of departure for the first part of this paper, which may be regarded as an extension of the Backgrounds chapter. We show, in rather informal terms, how witness sets can be useful in both explicating some basic intuitions about scope and understanding how particular denotational semantic differences between noun phrases affect their abilities to bear out

*This paper is based on "When do subjects and objects exhibit a branching reading?," read at WCCFL XII in April 1993 and at the CSLI conference on Logic and Language in June 1993. It presents the same results as the version in the WCCFL proceedings but spells the arguments out in more detail. We are grateful to Frans Zwarts for substantial discussions at the initial stage of the project, and to an anonymous Kluwer reviewer for comments. This research was partially supported by NSF Grant #SBR 9222501.

A. Szabolcsi (ed.), Ways of Scope Taking, 29–69.

certain scopal patterns. More generally, we suggest that the usual notion of scope needs to be factored into variation, distributivity, and maximality. This part lays some groundwork for several of the subsequent chapters and is thus of interest to all readers.

The second part shows that, already in this initial raw form, the above insights can be applied to make a novel claim concerning the availability of so-called branching readings. In logical terms, a branching reading can be defined for any sentence with a subject and a direct object. However, speakers of English accept only a fraction of these readings, so the question arises how the data can be predicted from the meanings of the participating quantifiers and the syntactic structure of the sentence. We propose that thinking about the behavior of quantifiers along the lines introduced in the first part leads to a simple answer to this question.[1]

1 THE INGREDIENTS OF SCOPE

1.1 Witnesses and variation

Recall the definition of a witness set from the Backgrounds chapter (exx. 37, 40):

(4) a. A set W is a witness of a GQ iff $W \in GQ$ and $W \subseteq SL(GQ)$, where $SL(GQ)$ is the smallest set the GQ lives on.

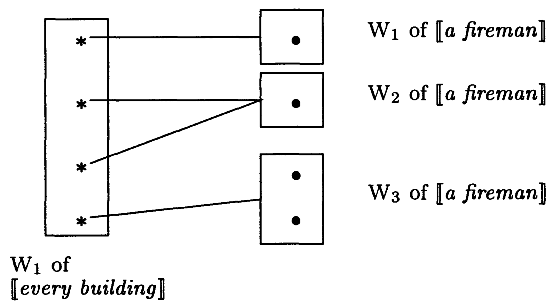
b. A GQ lives on a set of individuals A if, for any set of individuals X,

$$X \in GQ \quad \text{iff} \quad (X \cap A) \in GQ.$$

For instance, a witness set of the GQ denoted by *every building* is any set that contains every building and no non-building, and a witness set of the GQ denoted by *a fireman* is any set that contains at least one fireman and no non-firemen. The contents of (3) can now be described as in (5), and the general strategy, as in (6):

[1] According to the theory of generalized quantifiers, the term "quantifier" refers to the set of properties denoted by a noun phrase, and not to the noun phrase itself. In this paper we try to adhere to this norm. However, sometimes this would make the text pedantic and complicated. In these cases, we apply the term to the noun phrase as well.

(5)



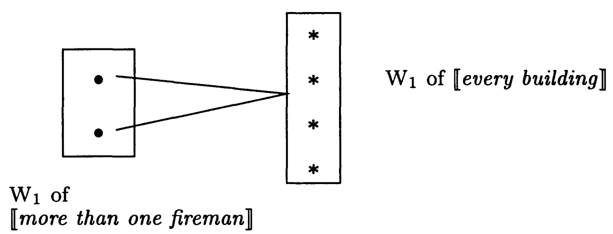
(6) To construct a situation that verifies the asymmetrical scope reading F > G, pick a witness W_i of the wide scope quantifier F. Using the relation denoted by the predicate, associate with each element of W_i a possibly different witness W_j of the narrow scope quantifier G.

We are now ready to study various limitations that this approach highlights.

One situation in which (7) is true is (8):

(7) More than one fireman checks every building (subject wide scope)

(8)



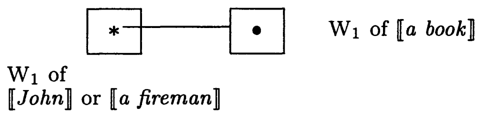
(8) contains only one witness associated with the narrow scope quantifier, and this is not an accidental property of the situation we are considering. It follows from the very meaning of *every building.* There can be only one set that contains every building and no non-building: the set of buildings itself.

In GQ theoretical terms, ⟦*every building*⟧ is a principal filter: it has a unique witness set (Backgrounds ex. 44). The same holds for ⟦*the (two) men*⟧ and ⟦*Andy and Carl*⟧, for instance. We now see that principal filters cannot exhibit variation (referential dependency) even in narrow scope position. Consequently, while variation is an important factor in our notion of scope, exhibiting variation and taking narrow scope cannot be identified.

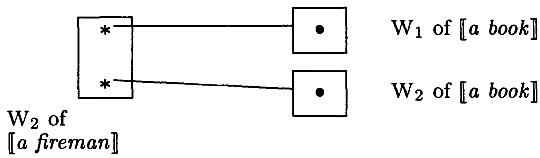
To see the complementary case, consider:

(9) John / A fireman read a book.

(10) a.



b.



In (10a), again, we have only one witness for the narrow scope quantifier, but here, this cannot be blamed on the meaning of *a book.* The unicity of the book set to be considered is forced by the fact that the witness of the wide scope quantifier has only one element. Whenever the relevant witness of the wide scope quantifier is a singleton, it is unable to induce variation (there is nothing to vary with), even if the narrow scope quantifier itself might be capable of exhibiting variation (as ⟦*a book*⟧ is).

What quantifiers fail to induce variation? The GQ ⟦*John*⟧ has no non-singleton witnesses at all: note that the smallest set ⟦*John*⟧ lives on is {john}. But ⟦*a fireman*⟧ has larger witnesses. How shall we judge the situation depicted in (10b), where the book-sets vary with the firemen? We propose that this variation is irrelevant, because the truth of the sentence is established already before we get to consider the second fireman and his book. Thus we may say that the quantifiers that cannot induce relevant variation are the ones whose minimal witnesses are singletons. (A minimal witness is one that ceases to be a witness if you take away even one element of it.)

1.2 Distributivity

We should hasten to add that having a non-singleton minimal witness is just a necessary, not a sufficient condition for a quantifier to induce variation. The pertinent data have been observed more or less independently by various scholars in the literature.[2]

(11) Two firemen read four books.

This sentence has a run-of-the-mill subject wide scope reading. But, unless the subject is accented in a particular way, it does not easily have a comparable object wide scope reading. Namely, it is easy to construe the four books as

[2]Farkas (1981, 1996), Verkuyl (1988), Kamp and Reyle (1993), Ruys (1993), Reinhart (1995), Kratzer (1995), Beghelli and Stowell (1994), and possibly others.

referentially independent of the two firemen (which is a precondition for the object wide scope reading), but not the firemen as varying with the books. Similarly,

(12) A fireman imagined that two buildings were unsafe.

Here *two buildings* may be read *de re,* i.e., the sentence may be interpreted as entailing the existence of two buildings. Nevertheless, even on this reading, firemen cannot vary with buildings. One way to express this is to say that *four books* in (11) and *two buildings* in (12) can take wide scope, but not distributive wide scope. Interestingly, the latter observation extends to universals, as in (2), repeated here:

(2) A fireman imagined that every building was unsafe.

Just as in (12), the existence of the buildings need not be a figment of a fireman's imagination, but even when *every building* is read *de re,* it cannot induce variation in the firemen. This observation is interesting for the following reason. Theories of plurals standardly assume that the scope of plural noun phrases needs to be factored into the scope of the existential closure applied to the set (or, plural individual) variable introduced by the NP and the scope of a distributive operator. On the other hand, the fact that universals do not induce variation in higher clauses has been taken to mean, plainly, that they are scopally trapped in their own clause. The parallelism of the data in the two domains suggests, instead, that distributivity needs to be factored out in both.

Thus it seems that the phenomenon of scope needs to be broken down, at least, into variation and distributivity. The present paper will merely capitalize on this basic observation and does not develop an appropriate novel approach to scope; some of the papers in this volume (Szabolcsi 1996, Beghelli and Stowell 1996, Beghelli 1996, Farkas 1996) will make several steps in that direction.

Given the above factorization, is the notion "scope" still useful? We propose to retain it in a primarily syntactic sense, in part to facilitate the comparison of our claims with those of others. By two quantifiers standing in an asymmetric scope relation we mean that the syntax of the logical or natural language under consideration has given one of the quantifiers the best possible chance to induce variation in the other. Thus, on one analysis, *a fireman* in (1) will be said to take asymmetric wide scope over *every building,* even though this particular choice of quantifiers cannot give rise to variation, but *every building* in (2) will not be said to take scope over *a fireman.*

1.3 An application: When does order matter?

The strategy in (6) has a significant logical limitation, to which we turn shortly. But before that, we can use the above considerations to answer the following simple question:

(13) Consider (i)–(iv). When are the two quantifier orders equivalent? (As usual, $\exists_2 x[g(x)]$ abbreviates 'there are x_1, x_2, such that they are distinct and $g(x_1)$, $g(x_2)$').

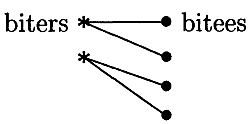
(i)	$\forall x \exists y[f(x,y)]$	and	$\exists y \forall x[f(x,y)]$
(ii)	$\exists x \exists y[f(x,y)]$	and	$\exists y \exists x[f(x,y)]$
(iii)	$\forall x \forall y[f(x,y)]$	and	$\forall y \forall x[f(x,y)]$
(iv)	$\exists_2 x \exists_2 y[f(x,y)]$	and	$\exists_2 y \exists_2 x[f(x,y)]$

Everybody knows, of course, that the two orders in (i) are not equivalent but the ones in (ii) and (iii) are. In view of this, one may be tempted to jump to the conclusion that the two orders are equivalent when the two quantifiers are identical, thus predicting that (iv) falls together with (ii) and (iii). But a moment of reflection shows that this is wrong. Consider the following linguistic instantiation of the formulae in (iv). The same noun *dog* is used throughout so that the restrictions can be ignored.

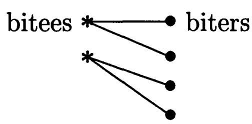
(14) a. (At least) Two dogs bit (at least) two dogs (subject wide scope)

b. (At least) Two dogs were bitten by (at least) two dogs (subject wide scope)

(15) a.



b.



Thus we must abandon the idea that the answer lies with the identity of the quantifiers. Instead, it seems the answer lies with variation.

The order of the quantifiers matters when one order gives rise to a different pattern of variation than the other. This may obtain when one order gives rise to variation and the other does not, or when both do but differently. Let us now briefly consider each of the four cases. To make talking about the examples easier, (i)–(iii) will also be paraphrased in the manner of (iv) above.

(16) a. Every dog bit a dog (subject wide scope)

b. A dog was bitten by every dog (subject wide scope)

(16a) is the best case for variation: $[\![$*every dog*$]\!]$ can induce variation and $[\![$*a dog*$]\!]$ can exhibit variation, because (unless the universe is accidentally too small) the former has a non-singleton minimal witness and the latter has more than one witness. (16b) is the worst case: $[\![$*a dog*$]\!]$ cannot induce relevant variation and $[\![$*every dog*$]\!]$ cannot vary. So the two orders will differ.

(17) a. A dog bit a dog (subject wide scope)

b. A dog was bitten by a dog (subject wide scope)

In (17a) as well as (17b), the wide scope quantifiers cannot induce relevant variation (the fact that the narrow scope ones might be able to exhibit variation does not come into play). So order makes no difference.

(18) a. Every dog bit every dog (subject wide scope)

b. Every dog was bitten by every dog (subject wide scope)

In (18a) as well as (18b), the narrow scope quantifiers cannot exhibit variation (the fact that the wide scope ones might be able to induce variation does not come into play). So again, order makes no difference, but for a different reason than in (17).

Returning to (14a, b), already spelled out and depicted above, ⟦*two dogs*⟧ can both induce and exhibit variation. In the (a) situation we may end up with two biters and four bitees, while in (b) with two bitees and four biters. Order makes a difference, despite the identity of the quantifiers.

This example indicates that although the behavior of the plain universal and existential quantifiers properly falls under a larger generalization, they are somewhat misleadingly special and thus it is dangerous to base intuitions solely on their behavior.

But, as has been mentioned above, the set of quantifiers considered above has still been quite limited in a crucial respect. This is to what we turn now.

1.4 Maximality

Take the following pair:

(19) a. Exactly one man saw exactly one woman (subject wide scope)

b. Exactly one woman was seen by exactly one man (subject wide scope)

Applying the above considerations to (19) we predict that (19a) and (19b) are logically equivalent, since ⟦*exactly one (wo)man*⟧ has only singleton minimal witnesses. But it is easy to see that the two readings are in fact independent! (In considering the situations below, the reader is invited to focus on the relevant subject wide scope readings, which are undoubtedly available, whether or not they are the intuitively most salient.) In (20), we simply outline situations in which one reading is true and the other is false and are not using witness sets:

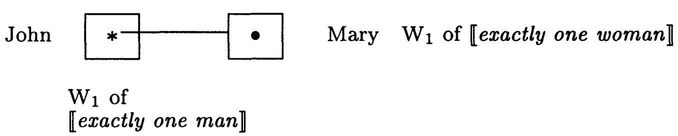
(20) a. John saw {Mary} (19a) true, (19b) false
Bill saw {Mary, Susie}
Peter saw {Judy, Claire}
No one else saw no one else

b. John saw {Mary} (19b) true, (19a) false
Bill saw {Mary, Susie, Judy}
Peter saw {Judy}
No one else saw no one else

In (20a), (19a) is true because John is the only man who saw just one woman, and (19b) is false because Susie, Judy, and Claire were all seen by just one man. In (20b), (19b) is true because only Susie was seen by just one man; (19a) is false because both John and Peter saw just one woman.

If we had tried to use the witness sets method outlined in (6), we would have failed miserably. In all the earlier cases, this method safely guaranteed that the sentences under consideration are true in the situation constructed. We might have embedded those situations in arbitrarily larger ones without any adverse effect. Not so in the present case. Consider:

(21)

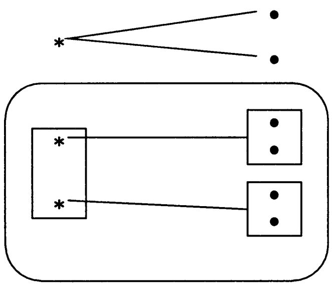


Both (19a) and (19b) are true here—but only if we guarantee that there are no more pairs in the man_saw_woman relation. If, for instance, (21) is embedded in (20a) or (20b), the truth values change dramatically.

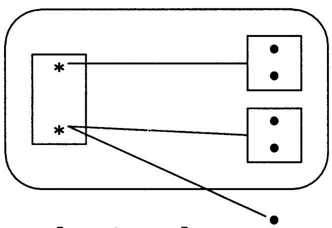
This in fact was to be expected. Up till now, we have restricted our attention to monotonically increasing quantifiers. As was seen in the Backgrounds chapter, for a quantifier to be increasing means, precisely, that whenever a sentence including it is established as true in some situation, it will remain true in arbitrary enlargements of that situation. And precisely this property is absent from decreasing or non-monotonic quantifiers, since both impose a maximality condition on the relevant situations. *Exactly one (wo)man* denotes a non-monotonic quantifier.

It is easy to see that the witness sets method runs afoul of decreasing and non-monotonic quantifiers in wide as well as narrow scope positions. Below, sentences are paired with situations whose encircled parts are constructed using the witness sets method. The sentences are all false in the larger situations.

(22) Exactly two/less than three firemen read two books.



(23) Two firemen read exactly two/less than three books.



1.5 Consequences for scope taking

One consequence is that our proposal above concerning when the order of two quantifiers matters holds only for pairs of increasing quantifiers. There are of course some cases even in the non-increasing domain where we get equivalences, e.g. John saw no man iff no man was seen by John, and exactly one man saw Judy iff Judy was seen by exactly one man. But giving a recipe for the general case becomes a more complicated matter.

Another, and more important, consequence pertains to the mechanisms of scope taking. We proposed a method for constructing situations that verify asymmetrical scopal readings: pick a witness of the wide scope quantifier and let the relation denoted by the predicate associate a possibly different witness of the narrow scope quantifier with each of its elements. We observed then that the viability of this method is limited to increasing quantifiers. This observation might indicate that the witness sets method is worthless. Alternatively, it might simply show that there is an empirically relevant intuition concerning how scopal readings are calculated or verified that pertains to one set of quantifiers but not to others. In other words, if the witness sets method indeed captures an empirically relevant intuition concerning the examples that it is applicable to, then scope taking cannot be a uniform phenomenon: decreasing and non-monotonic quantifiers must work in a way that is different from how increasing quantifiers do.

Several papers in this volume will make the empirical argument that different classes of natural linguistic quantifiers acquire their scope through different syntactico-semantic mechanisms. More specifically, it will be observed that there are two larger classes of quantifiers with markedly different scopal behavior. One of the classes contains only increasing quantifiers, while the other lumps together the decreasing and the non-monotonic items, along with some increasing ones. This indicates that monotonicity properties alone do not determine scopal behavior, but they do play a major role. And indeed, it will be argued that the manipulation of witness sets is insightful in connection with the behavior of the first class of quantifiers.

Specifically, Beghelli and Stowell (1996) argue that QPs belonging to the first class have designated landing sites in Logical Form (the specifiers of RefP, DistP, and ShareP, each associated with a distributive operator in a different way), while QPs belonging to the second class do not: they occupy the appropriate case positions. Szabolcsi (1996) proposes a connection between this syntax and Discourse Representation Theory, and discusses the relevance of monotonicity properties in detail.

The second part of the present paper does not yet pursue this syntactic, or representational, line; it remains within the realm of denotational semantics. We will be concerned with how factoring scope into variation, distributivity, and maximality makes it possible to predict what subject-object pairs speakers of English accept as supporting branching readings. First we give the gist of the analysis and then go on to present the details in more formal terms.

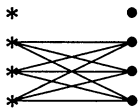
2 BRANCHING: AN INDEPENDENT READING?

2.1 The problem

Branching quantification in English was first studied by Hintikka (1974), Fauconnier (1975) and Barwise (1979). Their typical examples involve conjoined noun phrases, a reciprocal predicate, and some particle like *all.* (24) and (25) come from Barwise (1979, pp. 61–62):

(24) More than half of the dots and more than half of the stars are all linked by lines.

(25)



Two hallmarks of the configurations that make a branching reading true are independence (cf. the stars do not vary with the dots, and vice versa) and full connection (cf. the relevant stars are each connected to all the relevant dots). (Sher 1990 calls this the "each-all" version of branching.)

Branching readings are produced by a specific polyadic interpretation schema. This schema is not taken to be the contribution of any of the lexical items in the sentence; it is added to the derivation over and above the surface syntactically justified ingredients (in some theories, it is the interpretation of the Logical Form operation absorption). Specific proposals concerning the branching schema will be discussed in Sections 2.7, 2.8, and 2.10.1.

In addition to conjoined noun phrases, subject-object pairs may also support a branching reading. Many speakers even find this the preferred interpretation of certain sentences (see Gil 1982). E.g.,

(26) Three dogs bit two men.
'There is a set D of three dogs and a set M of two men, and each member of D bit each member of M'

But do all subject-object pairs support a branching reading? From a logical point of view (Sher 1990), there is no reason why they should not. However, Liu (1990, 1992) found that the availability of branching in English is severely limited. For instance, no such reading is attributed to (27a, b):

(27) a. Every dog bit two or more men.
b. No dog bit fewer than five men.

In this paper we are only interested in data involving plain subject-object pairs—that is, cases where, in distinction to Barwise's example, one quantifier phrase (QP) is structurally more prominent than the other, and no item like *all* is floating around. The question is this:

(28) What subject-object pairs support a branching reading?
Can the availability of branching be predicted from the meanings of the subject and object quantifier phrases and the syntactic structure of the sentence?

Interestingly, our quest leads to a reductionist answer:

(29) Predictions concerning when subject-object pairs support a branching reading can be made and indeed, come for free, if no special mechanism,

syntactic or logical, is assumed to create that reading. Each of the attested branching readings is logically equivalent to some other reading of the sentence that we want to derive anyway.

(30) Specifically, the branching readings of plain SVO sentences can be seen as special cases of either (A) scopally asymmetrical or (B) cumulative readings.

This result lends support to the suggestion by May (1989) and others that whatever branching readings are available in natural language are to be derived compositionally, relying on the contribution of adverbs like *all,* and appeal to a non-lexicalized branching schema is never necessary. We will make the following claims:

(31) A branching reading (of type A) is available exactly when the following conditions obtain at the same time:

   i. The meanings of the quantifier phrases preclude variation in the given configuration. This guarantees that the relevant two sets are independent.
   ii. The relation denoted by the verb is distributive (in the sense that it is strictly between individuals and not between groups). This, together with the fact that one of the quantifiers is assigned scope over the other, guarantees that the two sets are fully connected.
   iii. The nature of the quantifiers is such that the maximality condition on branching is met.

The discussion is organized as follows. Section 2.2 reviews Liu's branching data that this paper seeks to explain. Then the argument that most of the observed branching readings are logically equivalent to scopally asymmetrical readings is presented in two steps: Sections 2.3 through 2.5 present the intuitive core, and Sections 2.6 through 2.9 the formalism. Section 2.10 evaluates the results against alternative definitions of branching and against empirical data. Finally, Section 2.11 discusses the one case in which the branching reading is to be eliminated in favor of a cumulative reading, and concludes by raising the question whether the absence of genuine branching is an accidental gap in the semantics of English.

2.2 Liu's generalization

Liu (1990, 1992) conducted a careful empirical study concerning the scope and dependency behavior of noun phrases (NPs) in English. Her observation that NPs differ significantly in their ability to support inverse scope has inspired several of the papers in this volume. Beyond this, she also used her data to formulate a generalization concerning the availability of branching readings.

To begin with, Liu classifies noun phrases (NPs) according to their behavior in subject and object positions:

(32) a. Non-specific NPs: (i) Can depend on other NPs for scope interpretation, and (ii) Cannot easily make the subject scope dependent when they are in object position.

b. G(eneralized)-specific NPs: All the rest. (i) Cannot be dependent on others and/or (ii) Can easily make the subject dependent when in object position.

For instance, *few books* is non-specific:

(33) At least two men read few books

a. 'At least two men read few books, possibly different ones'
few books can be scope dependent

b. * 'Few books are such that at least two (possibly different) men read them'
few books as O cannot make S scope dependent

On the other hand, *every book* is G-specific:

(34) At least two men read every book

a. * 'At least two men read every book, possibly different ones'
every book cannot be scope dependent

b. 'Every book is such that at least two (possibly different) men read it'
every book as O can make S scope dependent

Note that the classification concerns NPs, and not NP denotations. Thus a NP will qualify as G-specific if it has at least one reading on which it can induce scope dependency in the subject while in object position. Using these criteria, Liu classifies NPs as follows.

(35) Non-specific NPs:
at least two N, more than two N, between two and five N, exactly two N, few N, fewer than two N, no N, neither N

(36) G-specific NPs:
all the N, every N, each N, most of the N, a majority of the N, some N, a (certain) N, the N, both N, one/two/three (of the) N

Turning to branching, Liu offers the following empirical generalization:

(37) a. When both NPs of a basic transitive sentence are G-specific, the sentence has a branching reading.

b. When one NP is G-specific and the other is non-specific, the sentence may or may not have a branching reading.

c. When both NPs are non-specific, the sentence has no branching reading.

For example:

(38) Two men read every book — can be branching

(39) a. At least one man read every book — can be branching

b. Two men read few books — cannot be branching

(40) At least one man read few books — cannot be branching

But why do we encounter any restrictions, and specifically these restrictions? Liu does not offer an explanation, formal or informal. However, there is something striking about her finding that whether an NP can participate in a branching reading correlates precisely with whether this NP can be scopally dependent and whether it can induce inverse scopal dependency. Assuming that these are indeed the relevant terms, the following question arises: What is it about branching that requires the NPs that support it to have particular scopal properties? The fact that the branching reading requires that the sets associated with the two quantifiers be chosen independently suggests a track to follow.

2.3 Independence and full connection: a first approximation

One crucial characteristic of branching is that the sets of individuals the quantifiers talk about are independent: there is no variation. This contrasts with the prototypical cases of asymmetric scope, where either the subject or the object induces variation in the other.

In Part I, we have seen, however, that there are particular choices of quantifiers with which scopal asymmetry cannot amount to variation. To recap, two prominent cases in the increasing domain are when (i) the wide scope QP is of the sort *John* or *a fireman* or (ii) the narrow scope quantifier is of the sort *every building, the (two) buildings,* or *Andy and Carl.* This raises the possibility that in the plain subject-verb-object cases that Liu examined, branching readings are but special cases of scopally asymmetrical readings. This is the insight that this paper explores.

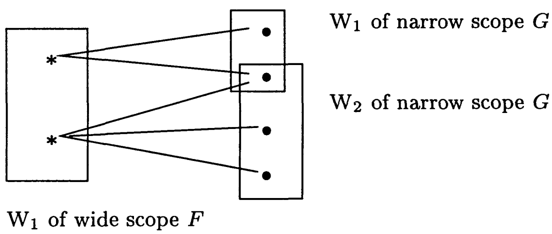
A second crucial characteristic of branching is that the two sets are fully connected. Are two independently chosen witness sets necessarily fully connected by the relation denoted by the predicate? Obviously not. Cases of cumulative and collective quantification are counterexamples (Scha 1981):

(41) (At least) two firemen put out (at least) three fires. 'Altogether two firemen put out fires, and altogether three fires were put out by them'

(42) Two firemen put out every fire. 'Two firemen as a collective put out every fire'

However, if one quantifier takes wide scope over the other and the relation between them is strictly distributive, then full connection is automatic. To see this, consider the kind of diagram that the method outlined in (6) produces:

(43)



This is the general case, where witnesses of G vary with the elements of some witness of F. Here the fact that F takes scope distributively over G entails that there is a witness of F such that each member of it is linked to each member of a possibly different witness of G. But we are considering special cases where the relevant witnesses for G are identical. Hence each member of F's witness is linked to each member of G's witness, which amounts to full connection.

In sum, we have shown that in the intuitively most accessible cases, an asymmetric scope relation involving a distributive predicate and particular choices of quantifiers inescapably yield a reading that is equivalent to a branching reading. Below we demonstrate that similar equivalences exist in other, intuitively less accessible cases as well. So, the question arises whether there are convincing cases of branching left without an asymmetric equivalent. We argue that there is only one type left, which, however, is known to have a cumulative equivalent. We conclude that the branching reading is never a genuine, separate reading of plain SVO sentences. Correspondingly, precise predictions concerning what pairs of quantifiers support "branching" come from establishing exactly when a scopal or a cumulative reading is equivalent to a branching one.

The Sections 2.6 and onward are concerned with substantiating this claim in more formal terms, also taking the maximality condition into account. Prior to that, however, we examine two issues to increase the initial plausibility of the enterprise.

2.4 Linear order: The role of scope restrictions

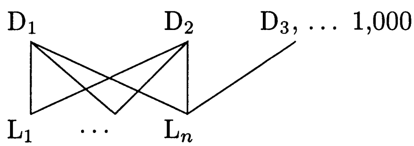
By every logician's definition, the polyadic branching schema is indifferent to the order in which the quantifiers occur in the sentence. Thus, if (44) has a branching reading, (45) is predicted to have one, too:

(44) More than one but fewer than six dogs bit every lion.

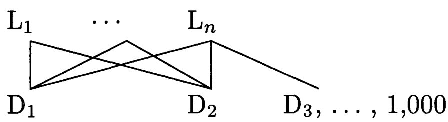
(45) Every lion bit more than one but fewer than six dogs.

Consider the two diagrams in (46), which are mirror images of each other, corresponding to the fact that (44) and (45) only differ in that subject and object are interchanged. Applying the appropriate definition (Sher's) to the quantifiers denoted by *every lion* and *more than one but fewer than six dogs,* (44) is predicted to have a reading which is true in (46a), and (45) to have a reading true in (46b).

(46) a.



b.



The prediction is not borne out: (44) is true of (46a), but (45) has no reading on which it is true of (46b). The fact that L_n is linked to 1,000 dogs does not affect (44) but falsifies (45).

The fact that switching the subject and the object affects the availability of the branching reading indicates that there is something fundamentally wrong with deriving that reading in a way that is inherently insensitive to the (linear, or c-command) order of the two quantifiers.

Our proposal, on the other hand, accounts for the contrastive behavior of (44) and (45) in a natural way. Recall that we are proposing that alleged branching readings are in fact special cases of others: in the present case, the asymmetric scopal reading is a possible candidate. What we find is that (44) has a relevant scopal reading that is true in (46a), but (45) has no scopal reading true in (46b).

Specifically, *every lion* denotes a principal filter, and *more than one but fewer than six dogs* a non-negative quantifier. As we shall show in (72), this combination guarantees that the subject wide scope reading of (44), 'there are more than 1 but less than 6 dogs such that each bit every lion,' is equivalent to a branching reading. (44) also has the O > S reading, 'every lion was bitten by a possibly different set of more than 1 but less than 6 dogs,' which is not branching and is irrelevant now.

In the case of (45), the branching reading would be equivalent to the O > S reading. But (45) simply has no O > S reading! The fact that modified numeral QPs like *more than one but fewer than six dogs* in object position do not take scope over the subject exemplifies one of the standard restrictions observed by Liu (1990):

(47) Every lion bit more than one but fewer than six dogs.

a. 'every lion bit a possibly different set of more than 1 but less than 6 dogs'

b. * 'there are more than 1 but less than 6 dogs such that each was bitten by every lion'

The fact that the absence of a particular scopal reading correlates with the absence of the logically equivalent branching reading confirms that the branching reading has no independent source: it is an epiphenomenon.

2.5 Bare indefinites

To cover Liu's core data, some new assumptions need to be made concerning bare (= non-modified) indefinites. As Liu observes, bare indefinites pattern with universals and definites in supporting a branching reading:

(48) a. Two or more kids climbed every tree. "branching" ok

b. Two or more kids climbed three trees. "branching" ok

c. Two or more kids climbed five or more trees. no "branching"

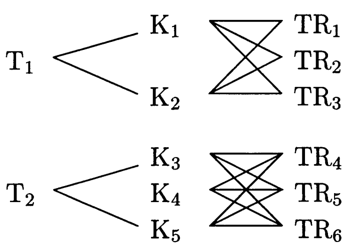
The easiest account of these facts is to postulate a principal filter reading for bare indefinites. This reading comes closest to Fodor and Sag's (1982) notion of a referential indefinite. Similarly to the case with a definite, the quantifier *three trees* on this reading talks about a set consisting of three trees that we "have in mind." This notion of referentiality is conceptually distinct from Enç's (1991) specificity, for instance. For Enç's purposes, *three trees* or even *any three* *of the trees* is specific if we know what superset we are drawing from. On the other hand, *three trees* denotes a principal filter if the three trees themselves are fixed. (Specificity in Enç's sense may pragmatically increase the noun phrase's ability to denote a principal filter.)

This account is not entirely correct, however. Consider the following (Cormack and Kempson's 1991 remarks point in the same direction):

(49) Every teacher saw that two or more kids climbed three trees.

On the account just proposed, (48b) has a branching reading when *two or more kids* takes wide scope but *three trees* denotes a principal filter. This predicts that when the complement clause in (49) has a branching reading, the kids may vary with the teachers but the trees may not. That is, all the teachers must have seen climbings of the same three trees, although they may have seen different kids climb them. But this prediction is false. For instance, the sentence may describe a situation in which each teacher saw two or more of his own pupils climb three of his own trees:

(49′)



It is not necessary for the trees not to vary with anything, which is what the principal filter interpretation requires. It suffices if the trees do not vary with the kids; and of course the kids must not vary with the trees, either. Notice that the standard treatment of *three trees* cannot possibly yield this result. If the sentence is assigned an $S > O$ reading, the trees will vary with the kids; and if it is assigned an $O > S$ reading, the kids will vary with the trees. What we need, intuitively, is a "relative principal filter" interpretation for the narrow scope quantifier.

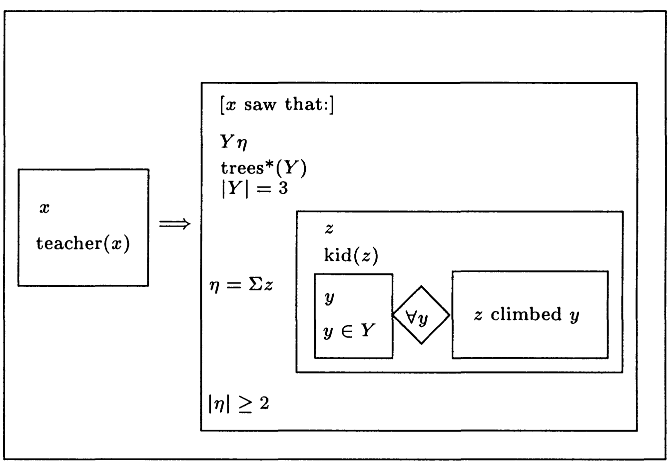
As Martin Honcoop (p.c.) points out to us, a straightforward way to obtain "relative principal filters" is to assume that *three trees* here is a principal filter denoter which, however, contains a phonetically null bound variable pronoun, so that (49) has an interpretation comparable to *Every teacher saw two or more kids climb the three trees in his yard.* But the same insight may be captured without making this particular syntactic claim.

We may adopt some basic assumptions of Discourse Representation Theory as in Kamp and Reyle (1993). Syntax proper and semantics proper are mediated by a level of discourse representations. A bare indefinite introduces a set (or a plural individual) referent. This referent may be placed either into the universe of the DRS that corresponds to the indefinite's place in syntactic structure or into the universe of any superordinate DRS. Finally, representations make explicit when a predicate is distributive with respect to a particular argument slot. Below is a Kamp and Reyle style representation of (49), simplified by not spelling out the contribution of *x saw that.*

In (50), the referent Y (and, as required by Kamp and Reyle, its associated conditions) is not introduced into the default DRS which contains *z climbed y* but, instead, into a superordinate one. Notice that introducing Y into a superordinate DRS does not assign wide scope to it over η in the traditional sense: kids do not vary with trees. This is due to the fact that the choice of the set (plural individual) Y is dissociated from distributivity, and hence from variation. This is unlike traditional generalized quantifier theory, where the two cannot be dissociated.

In this framework there is no need to postulate a separate principal filter reading for a bare indefinite. When the set (group) referent of a bare indefinite is introduced at least as high as the referent of some noun phrase NP that is syntactically more prominent and thus has already been processed by the DRS construction rules, the indefinite behaves like a principal filter with respect to that NP; when it is introduced into the main DRS, it behaves like a principal filter par excellence.

(50)



According to the analysis in (50), *three trees* functions as a "relative principal filter." It is not a true principal filter because its referent Y is not in the outermost box. But it has a fixed referent in the right hand side box in which *two or more kids* is processed, and this is all we are interested in for the purposes of branching.

To summarize, it turns out that the notion of a (genuine) principal filter is somewhat too demanding for our purposes; a relative principal filter suffices. In the interest of simplicity, however, we will continue to phrase our discussion of the conditions for branching in terms of principal filters, without adding the qualification "relative" all the time.

Another important question that arises here is whether we are only invoking DRT in order to solve the problem bare indefinites pose for branching. The answer is No. Beghelli and Stowell (1996) and Beghelli (1996) propose an empirical theory of quantifier scope, and Szabolcsi (1996) argues that their treatment of bare indefinites is essentially equivalent to adopting those assumptions of DRT that we appealed to above. This means that the main claim of our paper remains in effect: Each of the attested branching readings is logically equivalent to some other reading of the sentence that we want to derive anyway. Also, the claim that branching readings in plain SVO sentences can be seen as special cases of either scopally asymmetrical or cumulative readings remains true. We just need to add the qualification that the proper treatment of scope itself needs to go beyond traditional generalized quantifier theory.

Note, finally, that there is an even more radically semantic approach to scope in general and to problem this section has been concerned with in particular, namely, the one proposed in Farkas (1996) in terms of evaluation indices.

2.6 Independence in the general case

Let us begin by spelling out some of the reasoning in Part I in more precise terms. Variation can be formulated as in (51). (This specific formulation was suggested to us by F. Moltmann.) If the two quantifiers of a sentence are F and G and the relation denoted by the verb is R, $(F > G)(R)$ is the wide scope F reading.

(51) $(F > G)(R)$ is capable of exhibiting variation if it is not the case that in every model where $(F > G)(R)$ is true, the following holds:

For every witness w_1 of F, for every $x, z \in w_1$,
for every witness w_2 of G, for every $y, v \in w_2$,

$$((\langle x, y\rangle \in R \text{ and } \langle z, v\rangle \in R) \rightarrow ((\langle x, v\rangle \in R \text{ and} \langle z, y\rangle \in R)$$

In words: When there is variation, it need not be the case that whatever books one fireman read are the same as whatever books other firemen, if any, read. (Of course, a model may be too small to bear out potential variation.)

Spelling out the condition in (51), we derive the condition in (52):

(52) For $F > G$ to exhibit variation, there must be a model where $(F > G)(R)$ is true, but

$$(\langle x, y\rangle \in R \text{ and } \langle z, v\rangle \in R) \not\rightarrow (\langle x, v\rangle \in R \text{ and } \langle z, y\rangle \in R)$$

i.e., we need both
(a) $x \neq z$ within the same witness of F, and
(b) $y \neq v$ that distinguish two witnesses of G.

A quantifier F that cannot induce variation is one that never has two distinct elements x and z in its witness. This obtains when F has a unique witness and it is empty, e.g. *no dog*; or when F's witnesses are all singletons, e.g. *John, this man, John and no one else,* and *exactly one man.* A combination occurs with *fewer than two men.*

Note though that we only need to exclude variation that is "relevant" in view of the meaning of the quantifier. F's like *a man, some man, at least one man* have witnesses with more than one element but, since these quantifiers are increasing, the extra elements never make a difference. We can redefine the range of harmless quantifiers as those that have only singleton witnesses or are increasing and have singleton minimal witnesses. (Note that we cannot in general restate (52a) in terms of minimal witnesses. This would let all decreasing quantifiers in, since they all have the empty set as their minimal witness.)

A quantifier G that cannot vary is one that does not have two distinct witnesses. This obtains when G has a unique witness. This unique witness may be empty, as with *no man,* or non-empty, as with *John and Mary and no one else* on the one hand and with *John and Mary, every man* and *the(se) men* on the other. In these latter core cases, G "talks about" some fixed individuals, a notion neatly formalizable using the concept of a principal filter.

(53) The quantifier G is a principal filter iff it is of the form $\lambda P[A \subseteq P]$, with A non-empty, i.e., the properties (sets) that are elements of G are the supersets of a particular set A. A, which is also the unique witness of G, is called its generator set.

Names, universally quantified NPs, semantic definites, and their conjunctions are well-known principal filters. The quantifier *John and Mary* talks about the set {john, mary}. *Every man* talks about the set of men.[3] *These two men* talks about a set consisting of the two men we are pointing at, e.g., {peter, frank}. (Deictic *these two men* resembles pronouns in that its interpretation depends on the context (assignment, pointing), but in each context it talks about a unique set of individuals.)

[3] In some models *every man* may have an empty witness. We may choose to ignore these.

Prior to going further, let us note that in developing our argument it is crucial that we used witness sets, rather than elements, of quantifiers. Recall that an element of a quantifier may contain entities that do not belong to the smallest live-on set, i.e. the restrictor, of the quantifier. E.g., an element of *every dog* may contain cats and fire engines. Observe now that neither the ability to induce variation nor the ability to vary can be sensibly captured in terms of elements. Say, we have *Two lions bit every dog.* Suppose L_1 bit every dog and an old cat, while L_2 bit every dog and a young cat. Then the pertinent elements of *every dog* vary with the lions. This type of variation however is an artifact of the use of elements: it is never linguistically relevant.

2.7 The equivalence of scopal asymmetry and branching: the increasing case

When R is distributive, $(F > G)(R)$ with no variation exhibits "independence and full connection." Does this yield logical equivalence with branching? It depends on what exactly our definition of branching is. We will write the branching reading as $(F \times G)(R)$.

The definition of $(F \times G)(R)$ with two monotonic increasing generalized quantifiers is easy: it really involves nothing but independence and full connection. Technically, the latter means that the sets that are linked form a cross-product.

(54) If X and Y are properties (= sets), their cross-product $X \times Y$ is the set of all pairs $\langle x, y\rangle$ such that x is an element of X and y is an element of Y.

(55) Branching, MON↑–MON↑ (Barwise 1979):
For A and B that are monotone increasing quantifiers, $(A \times B)(R)$ is defined as $\exists X \exists Y[X \in A \,\&\, Y \in B \,\&\, X \times Y \subseteq R]$

Read: There are two sets, X and Y, such that X is an element of the generalized quantifier A, Y is an element of B, and the cross-product of X and Y is contained in the relation R denoted by the verb.

In view of our considerations above, it is easy to see that within this domain, $(F \times G)(R)$ is equivalent to $(F > G)(R)$ whenever F's minimal witnesses are singletons or G has a unique witness. For example:

(56) A dog bit three or more men. [S > O]
'There is a set X that contains a dog and a set Y that contains three or more men, and each element of X bit each element of Y'

(57) Three or more dogs bit five $\text{men}_{\text{filter}}$. [S > O]
'There is a set X that contains three or more dogs and a set Y that

contains five particular men, and each element of X bit each element of Y'

It is interesting to mention here that Westerståhl (1992) proves the following theorem for finite models:

(58) When Q_1 and Q_2 are MON↑ and ISOM, $\begin{matrix} Q_1 \\ Q_2 \end{matrix} \!\!>$ is equivalent to $Q_1 Q_2$ iff $Q_1 = \exists$ or $Q_2 = \forall$.

(Read: when Q_1 and Q_2 are increasing and have isomorphy (= topic-neutrality, quantity), the branching reading is equivalent to the wide scope Q_1 reading iff Q_1 is the existential or Q_2 is the universal quantifier.)

Notice that the reasoning that we presented above offers a simple intuition for why Westerståhl's theorem holds. Among the quantifiers that Westerståhl chooses to consider, the existential is the one with singleton minimal witnesses and the universal is the one with a unique witness. This explanation is interesting because it suggests how (58) generalizes to other quantifiers.

2.8 A general definition of branching

There is full agreement in the literature that (55) captures what the branching of two increasing quantifiers means. But once we turn to other quantifiers, we find disagreement concerning both for what cases branching can be defined and how it should be defined. The disagreement has two kinds of source: technical difficulties involved in providing a general definition and intuitive differences in how some cases should be evaluated.

Barwise (1979) defines the branching of two monotonic decreasing generalized quantifiers as follows:

(59) Branching, MON↓–MON↓ (Barwise 1979):
For A and B that are monotone decreasing quantifiers, $(A \times B)(R)$ is defined as $\exists X \exists Y [X \in A \,\&\, Y \in B \,\&\, (R \cap (A \times B)) \subseteq X \times Y]$

(60) Fewer than ten dots and fewer than six stars are all linked by lines.

(59) interprets (60) as follows: we have a set X containing fewer than ten dots and a set Y containing fewer than six stars, and whatever dots and stars are linked by lines are pairs drawn from X and Y.

The difference between (55) and (59) is due to the increasing versus decreasing nature of the quantifiers involved. It is exactly parallel to the difference in how *Three men walk* and *Fewer than ten men walk* can be expressed using formulae that begin with "There is a set X ... ":

(61) Three men walk
$\exists X[X \in \text{THREE\_MEN} \;\&\; X \subseteq (\text{WALK} \cap \text{MAN})]$

(62) Fewer than ten men walk
$\exists X[X \in \text{FEWER\_THAN\_TEN\_MEN} \;\&\; (\text{WALK} \cap \text{MAN}) \subseteq X]$

The increasing (61) merely states that the three men are walking men; it allows any number of further men to walk. The decreasing (62) requires that the fewer than ten men be all the walking men that there are—we call this a maximality condition. *Exactly two men walk* with a non-monotonic quantifier also requires a maximality condition; it would be expressed using $(\text{WALK} \cap \text{MAN}) = X$.[4]

Of course, the meanings of these sentences can be formalized in other ways as well, not beginning with "There is a set X … ." On the other hand, the branching readings must be formalized in this way. Beginning the definitions with "There are two sets, X and Y, … " ensures that the sets are chosen independently.

Going further, Westerståhl (1987) attributes to van Benthem a branching schema for certain non-monotonic quantifiers (naturally, with $X \times Y = R$) and himself proposes a general schema that applies to all continuous quantifiers.

Recall that we are interested in the definition of branching because we wish to examine the equivalence of scopal and branching readings in full generality. This means that we need a single definition, or a battery of definitions, that applies to all nine monotonicity combinations of two quantifiers. But the fact that the three monotonicity schemata differ in having $\subseteq$, $\supseteq$ or $=$ between $X \times Y$ and the restricted R already indicates that the problem is not trivial. The first attempt to overcome the difficulties is quite recent: Sher (1990).

Sher's definition is fully general and consists of two parts: the first part is essentially identical to the definition for two increasing quantifiers, and the second part imposes a maximality condition in order to take care of the decreasing and the non-monotonic cases. Her original version is as follows (P_1 and P_2 are the common nouns of the two NPs):

(63) $(Q_1 \times Q_2)(R)$ is defined as
$(\exists X)(\exists Y)[(Q_1 x)Xx \;\&\; (Q_2 y)Yy \;\&\; X \times Y \subseteq R \;\&$
$(\forall X')(\forall Y')(X \times Y \subseteq X' \times Y' \subseteq R \subseteq P_1 \times P_2 \rightarrow X \times Y = X' \times Y')]$

Since this is the definition we will use, let us consider it in some detail. First of all, we will modify it slightly:

(64) Branching (a slight modification of Sher 1990
Let F and G be generalized quantifiers whose smallest live-on sets are f

[4] In the increasing case it makes no difference whether we write $X \subseteq \text{WALK}$ or $X \subseteq (\text{WALK} \cap \text{MAN})$. Similarly, in (55) we might have written $X \times Y \subseteq (R \cap (A \times B))$; the restriction was omitted for the sake of simplicity.

and g, respectively, and R a restricted relation such that $\mathrm{Dom}(R) \subseteq f$ and $\mathrm{Ran}(R) \subseteq g$.[5] The branching reading $(F \times G)(R)$ is true iff
$\exists X \exists Y[X \in F \,\&\, Y \in G \,\&\, X \times Y \subseteq R \,\&$
$\forall X' \forall Y'[X \times Y \subseteq X' \times Y' \subseteq R \Rightarrow X \times Y = X' \times Y']]$.

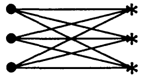
The working of (64) can be exemplified as follows:

(65) Exactly two dots and more than two stars are all linked by lines.
'There is a set X containing exactly two dots and a set Y containing more than two stars such that (i) $X \times Y \subseteq R$, i.e., each dot in X is linked to each star in Y and (ii) $X \times Y$ is not part of any bigger $X' \times Y'$ in the dot_links_star relation'

(64) differs from (63) in both a formal and a (minor) substantial respect. The formal difference is that (64) is stated using the formalism of generalized quantifier theory. The substantial difference is that in (63) the relation R is restricted by the sets P_1 and P_2 denoted by the common nouns of the relevant two noun phrases, while in (64) it is restricted by the smallest live-on sets of the corresponding generalized quantifiers. Since in most cases the common noun sets and the smallest live-on sets are identical, the change may seem insignificant (although natural: in a semantic definition, we use a semantic, not a syntactic, notion). However, a noun phrase like *John* contains no common noun but $[\![John]\!]$ has a smallest live-on set (in this case, {john}), so (64) becomes applicable to it. Furthermore, in the case of principal filters like *these two men* the smallest live-on set (for instance, {john, bill}) is smaller than the common noun set (MAN). We will see that here our semantic definition gives the desired results.

Let us see how (64) handles the non-monotonicity of *exactly two dots* in (65). The schema first guarantees that two appropriate sets are chosen independently and their members are fully connected. If we stopped here, however, (65) would be accepted as true in the following situation (if there are altogether three stars):

(66)



Condition (ii) excludes this. It requires that the cross-product that we use to verify (i) be the largest such, in the sense that it must not be properly contained in a larger cross-product within the same relation.

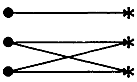
Does the maximality condition affect the increasing quantifier *more than two stars*? It does not. If in fact there were ten stars in the cross-product

[5]If F is the denotation of the object, and G of the subject, then R is the (restricted) converse of the relation denoted by the verb.

and we chose a Y that contains only, say, three stars, the sentence would be predicted to be false; but nothing prevents us from choosing Y to contain ten stars in the first place. In general, it follows from the notion of increasingness that it is insensitive to this kind of maximality condition. Hence, when both quantifiers are increasing, Sher's and Barwise's definitions coincide.

Sher differs from Barwise on decreasing quantifiers, however. While Barwise's (59) defines a reading that is equivalent to what has come to be called a cumulative reading (Scha 1981), Sher's definition gives a different reading.

(67) Fewer than three dots and fewer than three stars are all linked.



Barwise's definition renders the sentence false in this situation. We have two independent cross-products; we have to choose one of them. But whichever we choose does not exhaust the dot_links_star relation: the relation contains the pairs coming from the other cross-product, too. Sher's definition, on the other hand, renders the sentence as true: it only requires one $X \times Y$ that has the desired size and it not part of any bigger $X' \times Y'$. And there are even two such.

In the same way, Sher's definition renders both of the following sentences with non-monotonic quantifiers true in the above situation:

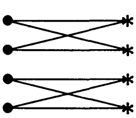
(68) Exactly one dot and exactly one star are (all) linked.

(69) Exactly two dots and exactly two stars are all linked.

Spaan (1992) accepts Sher's treatment of (67), however, he does not share Sher's intuition that (68) and (69) can be simultaneously true. He proposes that maximality should be defined not in terms of subsets but in terms of cardinality. His definition forces us to choose the biggest cross-product, in this case, the one involving two dots and two stars. This makes (69) true and (68) false.

On the other hand, Spaan explicitly agrees with Sher that (67) is true in the following situation, where all the independent cross-products have the desired size:

(70)



But Schein (1993, Ch. 12) argues that there are various linguistic examples that require a cross-product and are thus reasonably expected to fall under the heading branching and are false here, contra Sher's (and Spaan's) contention. For instance:[6]

(71) a. Fewer than three dots are linked to fewer than three stars, pairwise completely.

b. Exactly two dots are linked to exactly two stars, pairwise completely.

In sum, we have noted that in certain cases Sher's results differ from what Barwise or Spaan or Schein would find desirable. If they are correct, her (64) is somewhat too permissive. Can we still take (64) to be our etalon of branching when we examine plain SVO sentences? We believe that we can. Recall that we will be interested in what scopal readings of plain SVO sentences are equivalent to branching readings. As Section 2.10.1 will demonstrate, it turns out that in the cases where the four authors differ there are no equivalences anyway, wherefore the differences are immaterial to us.[7] There is one relevant difference, namely, Sher's treatment of the branching of two non-monotonic quantifiers differs from van Benthem's. In Section 2.11 we will submit that here Sher is correct.

2.9 The general equivalence of branching and scopal asymmetry

The main novelty of Sher's definition is that branching has three components: independence, full connection, and a separate maximality condition. The main theme of this section is how this third factor narrows down the range of scopal readings that are equivalent to a branching one.

In this paper we do not attempt to offer a beautiful formal result concerning the branching/scopal equivalence. Instead, we offer a kind of catalogue of the cases of equivalence whose correctness is easy to check using the basic definitions. We state it as a sufficient condition, but we conjecture that it is both sufficient and necessary.

(72) Equivalence of (64)-branching and scopal asymmetry:

(i) If either F or G is a principal filter generated by a singleton, then $(F > G)(R)$ and $(F \times G)(R)$ are equivalent irrespective of what the other quantifier is. In all other cases there is some qualification on the other quantifier to ensure the equivalence:

[6] Schein himself does not propose a definition of branching.

[7] We therefore opt for Sher's definition, which is the simplest and most general; Spaan's is restricted to the cases where X and Y are not empty.

(ii) If F is upward monotonic with singleton minimal witnesses, equivalence holds if G is upward monotonic.

(iii) If G is a principal filter, equivalence holds if F is non-negative.

(iv) If G has a unique empty witness, equivalence holds if F is a principal filter.

We will now systematically employ the notions of variation and maximality in giving a hint of why (72) is true.

In the examples the quantifier that guarantees no variation will be printed in bold face; our question is what restrictions need to be imposed on the other quantifier.

The cases in which the wide scope quantifier F cannot induce variation include the following:

(73) a. MON$\uparrow$ with singleton minimal witnesses:

filter: *John, this man, one man*$_{\text{filter}}$
non-filter: *at least one man, one or more men*

b. $-$MON with singleton (minimal) witnesses: *exactly one man*

c. MON$\downarrow$ with a unique empty witness: *no man*

d. MON$\downarrow$ with singleton or empty witnesses: *fewer than two men*

When F is a principal filter generated by a singleton set, maximality is automatically guaranteed, because the relation R is restricted to F's smallest live-on set, which is the singleton itself. Thus we are not allowed to consider pairs beyond those that have the unique element of the singleton as one of their members.

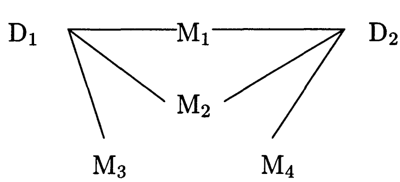
(74) **Fido** bit few/exactly two/at least two men.

This is the first case where our modification of Sher's original (63) makes a difference. There are several further cases to follow.

When F is an increasing singleton though not a filter, no restriction needs to be imposed on G as long as it is upward monotonic. G may be able to vary, because the variation F induces is irrelevant. But consider the following:

(75) a. **At least one dog** bit exactly two men.

b. **At least one dog** bit fewer than three men.



(75a, b) are true here on the S $\times$ O reading: there is a maximal cross-product containing exactly two/fewer than three men. But the S $>$ O reading is false, since individually, neither dog bit exactly two/fewer than three men.

When F is a non-monotonic singleton, equivalence fails unless G is a principal filter:

(76) a. **Exactly one dog** bit at least two men.

b. **Exactly one dog** bit John and Bill.

$$\begin{tikzcd} D_1 \arrow[r, dash] \arrow[dr, dash] & M_1 = \text{John} & D_2 \arrow[r, dash] \arrow[dr, dash] & M_3 \\ & M_2 = \text{Bill} & & M_4 \end{tikzcd}$$

Here (76a) is false on the S $>$ O reading but true on the S $\times$ O reading since, as we have seen, Sher allows us to ignore the existence of other independent cross-products. This kind of problem does not arise for (76b), since R is restricted to the set {john, bill}.

When F is decreasing and does not induce variation, equivalence holds only when G is a singleton filter, as in *No/fewer than two dogs bit* **John.**

(77) a. **Fewer than two dogs** bit John and Bill.

b. **Fewer than two dogs** bit few men.

c. **Fewer than two dogs** bit at least two men.

If in fact exactly one dog bit the said men, we already know from the previous example that only (77a) with a filter has a chance. But if no dog bit John and Bill, we can fail again. See the discussion of (79).

Notice that if someone chooses to interpret *fewer than two men* as 'fewer than two, but not zero, men,' then (s)he regards it synonymous with *exactly one man,* that is, the quantifier is non-monotonic, so the reasoning for (76) applies.

Let us now turn to the cases where the narrow scope quantifier G cannot vary:

(78) a. MON$\uparrow$ with unique witness = principal filter: *John, John and Mary, the(se) men, every man, two men*$_{\text{filter}}$

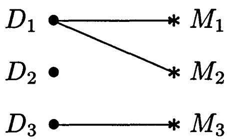
b. $-$MON with unique witness: *John and Mary and no one else*

c. MON$\downarrow$ with unique witness: *no man*

That G is a principal filter does not guarantee equivalence in general. E.g.,

(79) a. No dog bit **every man.**

b. Fewer than three dogs bit **every man.**



Both sentences are true here on the S > O reading, but not on the S × O reading. The reason is that $\emptyset \times \{M_1, M_2, M_3\} = \emptyset$ would only be maximal in the dog_bit_man relation if the latter were empty, too: here it is a subset of $\{D_1\} \times \{M_1, M_2\}$.

One way to avoid this problem is to make G a singleton filter. *No/fewer than three dogs bit* **John** is safe, because the whole generator set of the principal filter is affected uniformly. (A similar effect would arise if G were a group.) Another possibility is to require F to have no empty witness, i.e, to make F non-negative: either non-monotonic or increasing.

(80) a. Few but not zero dogs bit **every man.**

b. Exactly two dogs bit **every man.**

c. At least two dogs bit **every man.**

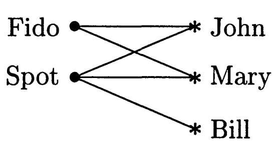
The reasoning that shows that (80a, b) yield equivalence is parallel to that for (76b), and the fact that (80c) with an increasing F yields equivalence is already familiar.

When G has a unique witness but is non-monotonic, equivalence obtains only when F is a singleton filter:

(81) a. Fido and Spot bit **John and Mary and no one else.**

b. Fido bit **John and Mary and no one else.**

c. Spot bit **John and Mary and no one else.**



(81a) is false here on the S > O reading but true on the S × O reading. The fact that Spot bit Bill, too, matters for S > O but not for S × O: given that Fido didn't bite Bill, there is no larger cross-product. On the other hand, (81b) is true and (81c) is false on both readings.

Finally, if G has a unique empty witness, F needs to be a principal filter:

(82) a. Few dogs, if any, bit **no man.**

b. Exactly two dogs bit **no man.**

c. At least two dogs bit **no man.**

d. Fido and Spot bit **no man.**

Suppose S > O is true. Then the set of dogs who bit no man is a witness W of F. In all four cases $W \times \emptyset = \emptyset$, but in (81a, b, c) this is not necessarily maximal in the dog_bit_man relation, which can happen to be non-empty. In in (81d), on the other hand the restricted relation is {fido, spot}_bit_man, which must be empty if S > O is true.

With this we have concluded the discussion of the cases summarized in (72).

2.10 Evaluation of alternatives

In this section we compare our proposal with three different kinds of alternatives. First, we argue that choosing Sher's definition of branching did not distort our picture. Second, we show that our results compare well with Liu's. Third, we comment on certain variations in speakers' judgments.

2.10.1 Sher versus Barwise/Spaan/Schein

Our first task is to justify the choice of using Sher's definition of branching, despite its divergence from Barwise's, Spaan's, or Schein's claims at various points. Here we are not claiming that Sher is more correct than the others (although she may be); what we are claiming is that the divergences make almost no difference in connection with the scopal/branching equivalence.

The typical cases where Sher's predictions differ from those of the others involve cases in which the quantifiers are decreasing or non-monotonic, and there is more than one cross-product in the model, cf. the discussion above of (67), (68), and (71). We argue that these cases make no difference since the corresponding SVO sentences, below, would not exhibit a scopal/ branching equivalence, anyway.

(68′)–(71′) Exactly two dots are linked to exactly two stars.

(67′)–(71′) Fewer than three dots are linked to fewer than three stars.

The reason is precisely that the S > O reading always allows variation here: different dots may be linked to different sets of stars. In the non-monotonic case, the S > O reading does not require for there to be even a single cross-product of the desired size. In the decreasing case the S > O reading may be true without any one of the cross-products being exhaustive of the relation (Barwise) or without the biggest cross-product (Spaan), or even any maximal one (Sher) being small enough.

Likewise, it is important to point out that the fact that the special branching schema "overgenerates" cannot be blamed on Sher's innovations: returning to even the most conservative definition, such as Barwise's, would not eliminate the empirical problem.

(83) contains two upward monotonic quantifiers. As has been mentioned, everybody's branching schema covers this case and interprets it in the same way (a). However, the branching paraphrase is not equivalent to either of the scopally asymmetrical paraphrases (b) and (c), whence our proposal predicts that (83) has no branching reading.

(83) At least two dogs bit more than two men.

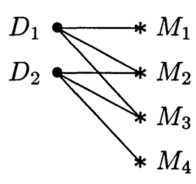
a. 'There is a set containing at least two dogs and there is a set containing more than two men, and each of these dogs bit each of these men' $= \mathrm{S} \times \mathrm{O}$

b. 'There are at least two dogs each of which bit a possibly different set of more than two men' $= \mathrm{S} > \mathrm{O}$

c. 'There are more than two men each of which was bitten by a possibly different set of at least two dogs' $= \mathrm{O} > \mathrm{S}$

Consider the truth of (83), and the truth of the paraphrases, in situation (84):

(84)



(83b) is true in (84), while (83a) and (83c) are false. The sentence itself, we claim, is just true here. If this is correct, then (83) lacks both the $\mathrm{S} \times \mathrm{O}$ reading and the (non-equivalent) $\mathrm{O} > \mathrm{S}$ reading.

But there are two branching cases proposed in the literature that we do not derive:

(85) No one loves no one.

(86) Exactly five men love exactly six women.

We return to these in Section 2.11.

2.10.2 The asymmetric/branching equivalence versus Liu

How do our results compare with Liu's findings? Let us begin by listing a sample of sentences which, according to (72i–iv), should be perceived as having a branching reading à la (64). The quantifiers that guarantee independence are printed in bold face; recall that restrictions on the others are due to maximality. For simplicity's sake, the examples rely on the S > O reading.

(87) **SINGLETON FILTER** > ANYTHING

Fido **A dog**$_{\text{filter}}$	bit	every man few, if any, men more than two but fewer than six men John and Bill and no one else

(88) ANYTHING > **SINGLETON FILTER:**

Every dog Few, if any, dogs More than two but fewer than six dogs Fido and Spot and no other creature	bit	**John** **a man**$_{\text{filter}}$

(89) **UPWARD SINGLETON** > UPWARD

At least one dog	bit	every man more than six men

(90) NON-NEGATIVE > **PRINCIPAL FILTER:**

More than six dogs Exactly one dog Few but not zero dogs	bit	**every man** **John and Bill** **two dogs**$_{\text{filter}}$

(91) PRINCIPAL FILTER > **UNIQUE EMPTY WITNESS:**

The dogs Fido and Snoopy Two dogs$_{\text{filter}}$	bit	**no man**

According to Liu, an SVO sentence is certain to have a branching reading if both NPs are G-specific; it may or may not have one if one NP is G-specific and the other is not (here Liu offers no generalization); and it never has one if both NPs are non-specific. The set of Liu's G-specific NPs is coextensive with those that have at least one (absolute or relative) principal filter interpretation in our terms.

Assuming this correspondence, our results give a fair approximation. All cases with two principal filters are ruled in, and only a single case of two non-filters is. The following cases deserve to be commented on.

First, consider the one filter plus one non-filter cases. The "non-negative > principal filter" class coincides with Liu's more detailed findings. The only cases that may raise an eyebrow are those where a filter can be combined with a decreasing quantifier (see 87, 88, and 91 above). We believe that not excluding mixed cases from the very beginning makes our approach more convincing; it is in fact insightful that only such a restricted set of further cases exhibits equivalence.

Next, consider the two non-filter cases, which are of the type *At least one dog bit more than six men.* Contra Liu, we are predicting that this kind of sentence does have a branching interpretation. But this is inescapable, since the S > O reading of this sentence is logically equivalent to a branching one according to everybody's definition of branching.

These cases might indicate the existence of a gap between the logician's definition of branching and the pretheoretical intuition that speakers apply when judging whether a sentence has a branching reading. The former is purely denotational; the latter seems somewhat representational. Specifically, speakers seem to prefer cases where one or both noun phrases introduce a discourse referent corresponding to a witness set, according to some version of DRT or according to Szabolcsi (1996).

In any event, note that we are not arguing that all these sentences have a genuine braching reading. On the contrary, we are arguing that in all SVO sentences branching is a mere illusion, due to the fact that an independently available reading is logically equivalent to what the branching one would be.

Finally, the following case may seem problematic for our proposal:

(92) Most (of the) students read two books.

As Liu explains, this sentence has one S > O reading on which pairs of books vary with students. On the other hand, for many speakers it lacks the standard O > S reading on which each of the two books was read by a different majority of the students. Instead, it has a classical branching reading: each member of some fixed majority of students read each member of some fixed set of books. (92) may thus seem to call for irreducible branching (see also Keenan 1996). This conclusion is not inescapable, however. Note that *two books* can denote a filter and thus the branching reading is equivalent to an S > O reading, just as in the examples reviewed earlier. The one thing that is peculiar about (92) is that although *most (of the) students* is in general not a filter, in subject position it cannot become dependent on the object. This fact requires an explanation but need not affect the present proposal.[8]

[8]There is one possible explanation of this fact that we are aware of: Honcoop's (1994). Following Ben-Shalom (1993), he suggests that inverse scope is calculated by a binary quantifier. Departing from Ben-Shalom, he proposes that the wide scope quantifier contributes

2.10.3 Variation in judgments

Speakers' judgments may vary with respect to certain examples. One source of variation is what noun phrases they are willing to interpret as principal filters. Everyone seems to have a filter interpretation for noun phrases with bare numerals like *two men,* and no one for, e.g. *few men* or *more than one but fewer than six men.* On the other hand, speakers differ somewhat as to whether they interpret *at least n* phrases and partitives as filters or not. Another source of variation is that some may interpret *few men* as non-negative 'few but not zero men,' at least in certain contexts. Such differences will affect whether *At least two men saw* **at least six (of the) dogs** and *Few men saw* **every dog** are classified with (90), for instance, but not the global predictions. Our predictions concern noun phrases interpreted as such-and-such quantifiers, not noun phrases as potentially ambiguous syntactic units.

Since these interpretation options play a role in E-type anaphora and inverse scope, we expect the same individual preferences to show up in those domains. For instance, the acceptability of Evans's (1980) famous example, quoted in (93a), seems to hinge on one's ability to interpret *few senators* in a non-decreasing fashion, as the implausibility of the modified (93b) version shows.

(93) a. Few senators admire Kennedy, and they are very junior.

b. ?* Few senators, if any, admire Kennedy, and they are very junior.

Or, consider the following descriptive generalization (based on our earlier work and on the results of the previous section): In the default case, if a direct object noun phrase *OBJ* is interpreted as a principal filter, this ensures that (i) the sentence has an O > S reading and (ii) if S is non-negative, its S > O reading is equivalent to a branching reading. So if a speaker accepts, say, *Every man saw OBJ* on the branching reading, (s)he is expected to accept it on the O > S reading as well.

2.11 Cumulative readings and conclusion

Approaching the end of our journey, we must ask whether the cases we have accounted for so far cover all "branching readings" that arise in plain SVO sentences. Above, we have in fact anticipated the answer: they do not.

its unique witness and the narrow scope quantifier a Skolem function to this binary quantifier. He suggests that if the narrow scope quantifier is proportional, the Skolem function would need to be implausibly complicated and thus predicts that only intersective quantifiers "skolemize." NB It may seem that any theory that accommodates the fact that direct object *two books* does not take distributive wide scope over the subject will take care of the problem of subject *most*. That is true in the case of (92), but not in general, for the same rigidity effect is exhibited with *every book* in the object position.

Consider (94), together with its interpretation according to (64):

(94) No one likes no one — branching, à la Sher
$\exists X \exists Y[\text{no-one}(X)\ \&\ \text{no-one}(Y)\ \&\ X \times Y \subseteq \text{man\_like\_man}\ \&$
$\forall X' \forall Y'[X \times Y \subseteq X' \times Y' \subseteq \text{man\_like\_man} \Rightarrow X \times Y = X' \times Y']]$

This is a bona fide reading of (94). But it is truth conditionally independent of the only scopal reading, 'Everyone likes someone.'

The unloving world reading of (94) constitutes the only unquestionable exception to the claim that existing branching readings of SVO sentences are equivalent to an existing scopal reading. As has been observed by Sher (1990) and Zwarts (1992), however, the reading in (94) is equivalent to a so-called cumulative reading (Scha 1981, Schein 1993). Cumulative quantifiers are also scopally independent, but the connection between the sets is weaker than in branching: it is enough for every element of X to be connected to some element of Y, and for every element of Y to be connected to some element of X. Following Zwarts's suggestion, we note that when both sets X and Y are empty, the requirements 'for every x, there is a y' and 'for every y, there is an x' are vacuously satisfied:

(95) No one likes no one — cumulative, à la Scha
'Exactly zero humans like humans, and exactly zero humans are liked by humans'

Thus it seems quite natural to regard this reading as a special case of cumulative quantification.[9]

The last case to consider is (96). Van Benthem defines a branching reading for two non-monotonic quantifiers of the *exactly*-type. Applied to (96′), this is not equivalent to any scopal reading, but it also differs from Sher's branching reading and Scha's cumulative reading of (96).[10]

(96) Exactly two dogs bit no more and no fewer than three men.

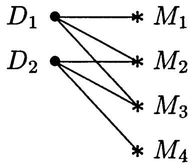
(96′) $\exists X \exists Y[\text{ex. two\_dogs}(X)\ \&\ \text{ex. three\_men}(Y)\ \&\ \text{dog\_bit\_man} = X \times Y]$

(96′) is false in all of the following situations. The S > O reading of (96) would be true in (97a), its cumulative reading in (97b), and its Sher-style branching reading in (97c). As compared to (97c), (96′) excludes the presence of D_3.

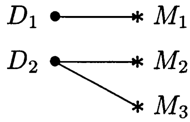
[9] Or of resumptive quantification, cf. May (1989).

[10] In the example, we replaced the second *exactly*-phrase with a ***no more and no fewer***-phrase, because we felt that the judgment is clearer in this way. The change cannot make a difference from van Benthem's point of view, since the two phrases are equivalent.

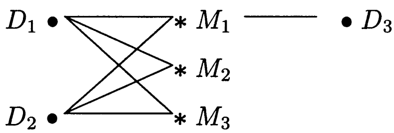
(97) a.



b.



c.



If empirical judgments confirm that (96′) exists as a separate reading, then it alone necessitates the use of a specific branching schema, all our previous arguments notwithstanding. Our own judgment is that (96′) is stronger than what (96) ever requires. (We believe that (96) just has an S > O reading and a cumulative one.)

With these observations, we take it that the data are compatible with our general claim:

(98) Simple SVO sentences have a branching reading only when that reading is logically equivalent to another, independently justified reading, namely, a scopally asymmetrical or a cumulative one.

This conclusion is based strictly on the workings of a particular sentence type and thus says nothing about the necessity of a special branching schema in others. However, consonant suggestions have been made about other sentence types. May (1989), Krifka (1991), Schein (1993), and van der Does and Verkuyl (1996), each of whom examines different data than we do, suggest either that alleged branching cases are due to contextual factors or that branching is to be derived compositionally, relying on the contribution of conjunction or of adverbs like *all,* and an independent branching schema is never necessary. Our result can be seen as adding one more piece to the puzzle. But since no one has pulled all these fragmentary results together, it is in fact not yet known whether they actually cover all the empirically relevant cases.

So let us assume, for the sake of the argument, that an independent branching schema is indeed not attested in natural language. Then the interesting question is whether this is an "accidental gap." As of date we are not able to answer this question, but let us indicate the beginnings of how we believe it might be approached.

Branching quantification is effected by a polyadic (here: binary) quantifier: one that satisfies more than one argument slot of the predicate at the same time. On the syntactic side, this means that the usual syntactic structure of the sentence is not respected; say, "absorption" needs to be invoked. On the semantic side, we note that polyadic quantifiers come in two varieties: those whose working is reducible to that of a sequence of (unary) generalized quantifiers, in which case polyadicity is semantically inessential, and those that are irreducible in this sense. In the latter case, the quantifier's schema acts like a "word" in that it makes a totally idiosyncratic contribution to the meaning of the sentence. The difference is that words are overt parts of the sentence, while polyadic quantifier schemata are not. This makes it plausible that learnable polyadic schemata must be tied to particular phrases like *all* (or, in other cases where a polyadic analysis has been proposed, to phrases like *between them, except, different ... different ...* , etc.),[11] or to particular constructions like coordination.[12] Other instances are presumably severely limited, judging from the fact that the meanings of sentences are by and large quite predictable. We expect that future research will identify quite strict constraints on the polyadic quantifiers of natural language. (See Ben-Shalom 1993 and Honcoop 1994 for some preliminary speculations.) At that point, it will be feasible to determine whether the branching quantifier schema is a possible but accidentally unattested entity, or it in fact violates some constraint and its absence can thus be predicted.

REFERENCES

Barwise, Jon. 1979. On branching quantifiers in English. *Journal of Philosophical Logic* 8(1):47–80.

[11]Peters and Kim (1995) show that reciprocals support a set of distinct interpretations. We note that these cover essentially the same spectrum that Sher (1990) attributes to various forms of independent quantification. The possibility arises that one important source of these independent readings is an overt or covert reciprocal. The investigation of this issue may help complete the puzzle.

[12]Or, cumulative readings in Hungarian require a special word order configuration. Compare cumulative (i) with S > O (ii):

(i) *Kevés ember (mindig) keveset végez*
few man (always) little-acc accomplishes
'Few people accomplish little (between them)'

(ii) *Kevés ember végez keveset*
few man accomplishes little-acc
'Few people are such that they accomplish little'

(But note that both van der Does 1992 and Schein 1993 argue against the polyadicity of cumulative readings.)

Beghelli, Filippo. 1996. The syntax of distributivity and pair-list questions. In this volume.

Beghelli, Filippo, and Tim Stowell. 1994. The direction of quantifier movement. *GLOW Newsletter* 32:56–57.

Beghelli, Filippo, and Tim Stowell. 1996. Distributivity and negation. In this volume.

Ben-Shalom, Dorit. 1993. Object wide scope and semantic trees. In Lahiri (Ed.), *Proceedings of SALT III.*

Cormack, Annabel, and Ruth Kempson. 1991. On specificity. In Jay L. Garfield and Murray Kiteley (Eds.), *Meaning and Truth*, 546–582. New York: Paragon House.

van der Does, Jaap. 1992. *Applied Quantifier Logics.* PhD thesis, University of Amsterdam.

van der Does, Jaap, and Henk Verkuyl. 1996. Quantification and predication. In Kees van Deemter and Stanley Peters (Eds.), *Semantic Ambiguity and Underspecification.* Stanford: CSLI.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22(1):1–25.

Evans, Gareth. 1980. Pronouns. *Linguistic Inquiry* 11(2):337–362.

Farkas, Donka. 1981. Quantifier scope and syntactic islands. In Roberta Hendrik et al. (Eds.), *Papers from the Seventeenth Regional Meeting, Chicago Linguistics Society*, 59–66.

Farkas, Donka. 1996. Evaluation indices and scope. In this volume.

Fauconnier, Gilles. 1975. Polarity and the scale principle. In Robin Grossman et al. (Eds.), *Papers from the Eleventh Regional Meeting, Chicago Linguistics Society*, 188–199.

Fodor, Janet, and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5(3):355–398.

Gil, David. 1982. Quantifier scope, linguistic variation, and natural language semantics. *Linguistics and Philosophy* 5(4):419–472.

Hintikka, Jaakko. 1974. Quantifiers and quantification theory. *Linguistic Inquiry* 5(1):153–177.

Honcoop, Martin. 1994. Inverse scope and Skolem functions. Unpublished manuscript, University of Leiden.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Keenan, Edward. 1996. Further beyond the Frege boundary. In Jaap van der Does and Jan van Eijk (Eds.), *Quantifiers, Logic and Language*, 179–201. Stanford: CSLI.

Kratzer, Angelika. 1995. Scope or pseudo-scope? are there wide scope indefinites? In Rothstein (Ed.), *Events and Grammar.* Kluwer. Forthcoming.

Krifka, Manfred. 1991. Definite Noun-Phrases are not quantifiers. *Linguistic Inquiry* 23(2):156–163.

Liu, Feng-hsi. 1990. *Scope Dependency in English and Chinese.* PhD thesis, UCLA.

Liu, Feng-hsi. 1992. Branching quantification and scope independence. In Jaap van der Does and Jan van Eijck (Eds.), *Generalized Quantifier Theory and Applications*, 297–331. Chicago: CSLI/University of Chicago Press.

May, Robert. 1989. Interpreting Logical Form. *Linguistics and Philosophy* 12(4):387–437.

Peters, Stanley, and Yookyung Kim. 1995. Reciprocals: Parametrized polyadic quantification. Presented at the 4th CSLI Workshop on Logic, Language and Computation, June 2-4.

Reinhart, Tanya. 1995. Interface strategies. Unpublished manuscript, Utrecht University.

Ruys, Eddy Godehard. 1993. *The Scope of Indefinites.* PhD thesis, Research Institute for Language and Speech, Utrecht.

Scha, Remko. 1981. Distributive, collective, and cumulative quantification. In Jeroen Groenendijk et al. (Eds.), *Formal Methods in the Study of Language*, 483–512. Amsterdam: Mathematisch Centrum.

Schein, Barry. 1993. *Plurals and Events.* Cambridge: MIT Press.

Sher, Gila. 1990. Ways of branching quantifiers. *Linguistics and Philosophy* 13(4):393–422.

Spaan, Martin. 1992. Parallel quantification. Unpublished manuscript, University of Amsterdam.

Szabolcsi, Anna. 1996. Strategies for scope taking. In this volume.

Verkuyl, Henk. 1988. Aspectual asymmetry and quantification. In Veronika Ehrich and Heinz Vater (Eds.), *Temporalsemantik: Beiträge zur Linguistik der Zeitreferenz*, 220–259. Tübingen: Niemeyer.

Westerståhl, Dag. 1987. Branching generalized quantifiers and natural language. In Peter Gärdenfors (Ed.), *Generalized Quantifiers*, 269–299. Dordrecht: Reidel.

Westerståhl, Dag. 1992. Iterated quantifiers. ITLI Prepublications.

Zwarts, Frans. 1992. Issues in quantification. Talk presented at the Workshop on Logic and Language, Stanford University.

3

DISTRIBUTIVITY AND NEGATION: THE SYNTAX OF *EACH* AND *EVERY**

Filippo Beghelli and Tim Stowell

Department of Linguistics
UCLA

1 INTRODUCTION

This paper is concerned with the syntax and semantics of quantifier scope construal, focussing on the distributive quantifiers *every* and *each*, and their interaction with negation. Our discussion is based on the theory of the syntax of quantifier scope developed more fully in Beghelli and Stowell (1994) and in Beghelli (1995).

The quantifier *every* has traditionally been analyzed in natural language semantics as the quantifier $\forall$, familiar from classical logic. We will show that *every* is more complex than this; a number of observations on its logico-semantic behavior lend plausibility to the view that *every* exhibits a kind of quantificational variability characteristic of licensed and bound elements. The quantifier *each* has been analyzed as a wide-scope variant of *every*, which is supposedly used in order to disambiguate between pairs of possible scope construals. We will show that the distinction between *every* and *each* is more properly characterized in terms of an intrinsic distinction between optional and obligatory distributivity. The effects of this distinction are often masked, however, by the effects of the syntactic mechanisms by which these notions are expressed in the grammar of natural languages, as we will see.

The paper is organized as follows. In Section 2, we introduce the general theory of scope and quantifier types on which the rest of the paper is based. In Section 3, we discuss the syntax of distributivity, concentrating on the distinctive behavior of QPs headed by *every* and *each*, which we refer to as Distributive-Universal QPs (DQPs). In Section 4, we examine the scopal interactions of DQPs with negation, bringing to light certain distinctive properties of these QPs, and highlighting some surprising differences between *every* and

*This research was supported in part by NSF grant #SBR 9222501 (Beghelli) and by grants from the Academic Senate of UCLA (Stowell).

A. Szabolcsi (ed.), Ways of Scope Taking, 71–107.

each. In Section 5, we discuss other differences between *every* and *each*, which we will use to explain the differential behavior that they exhibit with respect to negation.

2 TARGET SCOPE POSITIONS FOR QP-TYPES

2.1 Scope uniformity

Our analysis of *every* and *each* is formulated within the overall theory of quantifier scope developed in Beghelli and Stowell (1994) and in Beghelli (1995). We present here a sketch of that proposal; the reader is referred to those works for further discussion. We adopt two central assumptions of the standard theory of quantifier scope in generative grammar. First, quantifier scope is determined by c-command relations holding at the level of Logical Form (LF); second, Quantifier Phrases (QPs) are assigned scope by undergoing movement to their scope positions in the derivation of the LF representations.

However, we reject one central assumption that has guided virtually all previous work on scope, namely that all QPs have the same scope possibilities. This can be stated in terms of QUANTIFIER RAISING (QR), as in (1):

(1) **The Uniformity of Quantifier Scope Assignment (Scope Uniformity)**
Quantifier Raising (QR) applies uniformly to all QPs. Neither QR nor any particular QP is landing-site selective; in principle, any QP can be adjoined to any (non-argument) XP.

In this respect, we depart from the standard account in May (1977, 1985), as well as from refinements of it in Aoun and Li (1989, 1993), and Hornstein (1995).

The reason why Scope Uniformity cannot be maintained is empirical: different QP-types have correspondingly different scope possibilities. Some of the evidence for this conclusion is reviewed below.[1]

May (1977, 1985) assumes that pairs of subject and object QPs are typically scopally ambiguous, and concludes that all QPs normally undergo movement from their (S-structure) Case positions to distinct scope positions at LF. In other words, he assumes that Case positions never serve as scope positions for QPs. On the other hand, Hornstein (1995) proposes that every link in the

[1] Our approach builds on that of various authors, notably Kroch (1979) and Liu (1990), both of whom observe that quantifier scope is not uniform, in the sense that individual quantifiers differ from each other in their ability to take inverse scope. Our work builds, in part, on proposals in Beghelli (1993), Ruys (1993) and Beghelli et al. (1996), among others.

A-chain of a given QP is a possible scope position for that QP—including both the Case position occupied by the QP at Spell-Out and its θ-position.

In this study, we propose a hybrid theory, incorporating aspects of both May's and Hornstein's approaches.[2] The central innovative aspect of the system developed here is that it draws distinctions among various QP-types; whereas *certain* QP-types may take scope in their Case positions (remaining *in situ* at LF), other QP-types must move to distinct LF scope positions reserved for them. Moreover, there are further distinctions among those QP-types that must undergo movement, in the sense that each type has a designated LF scope position defined in the hierarchical phrase structure of the clause.

2.2 QP types

Although it is possible, *a priori*, to draw many distinctions among various QP-types, we believe that—in a first approximation—the syntax of quantifier scope can be adequately captured by recognizing five major classes of QP-types. Our classification incorporates insights of Szabolcsi (1994, 1996). The reader is especially referred to the latter paper, where the relation with our proposal is discussed at length.

QP-Types

a. **Interrogative QPs (WhQPs)**. These are familiar Wh-phrases such as *what, which man*, etc. We adopt the standard convention of attributing a [+Wh] feature to these QPs, encoding their interrogative force.

b. **Negative QPs (NQPs)**. These are QPs such as *nobody, no man*, etc. (In this group belong also French n-words such as *personne* 'nobody,' and possibly Italian/Spanish n-words such as *nessuno/nadie* 'nobody,' which sometimes require an overt negative element to license them.) We assume that these QPs bear a feature [+Neg].

c. **Distributive-Universal QPs (DQPs)**. These are QPs headed by *every* and *each*, which occur only with singular nouns. We attribute to them, in a first approximation, a distributive feature [+Dist(ributive)] (we will revise this assumption in Section 5, where we will attribute to *each* an intrinsic feature of distributivity [+Dist], leaving *every* underspecified for [Dist] and specified merely for universality [+Univ]). Both *each*-QPs and *every*-QPs are usually interpreted as both universal and distributive.

d. **Counting QPs (CQPs)**. These include decreasing QPs with determiners like *few, fewer than five, at most six, ...* and generally cardinality

[2] The hybrid claim that some quantifiers undergo scopal movement, while others do not, was put forth in Beghelli (1993).

expressions built by modified numerals (e.g., *more than five, between six and nine, more (students) than (teachers), ...*). The characteristic semantic property of these QPs is that they count individuals with a given property, have very local scope (take scope essentially *in situ*) and resist specific interpretations.

e. **Group-Denoting QPs (GQPs).** To this large class belong indefinite QPs headed by *a, some, several,* bare-numeral QPs like *one student, three students, ...* , and definite QPs like *the students.* The fundamental property of GQPs is that they denote *groups*, including plural individuals. Even leaving aside their referential reading (the type of epistemic specificity discussed first by Fodor and Sag 1982), GQPs can easily be construed as taking widest scope within their clause, though they might be c-commanded by other scopal elements. We maintain that this capacity for wide scope derives from their ability to introduce group referents. (Another property of GQPs that derives from this is that they support collective interpretations in contexts where DQPs require a distributive construal.) Indefinite and Bare-numeral GQPs can also support readings where they have very local scope, behaving like CQPs. We factor out such readings (exhibited by some of the members of this class) in terms of an ambiguity between a GQP and CQP reading.

2.3 Logical functions associated with QP-types

On the basis of this typology, we identify the following logical functions and relative LF positions where they are satisfied.

Scope positions for QP types

a. WhQPs take scope in the Spec of CP, where they assume their interrogative force by virtue of their [+Wh] feature being checked via Spec-Head agreement with the question operator Q.

b. NQPs take scope in the Spec of NegP, where their [+Neg] feature is checked via Spec-Head agreement with the (silent) Neg^0 head, as in Zanuttini (1991) and Moritz and Valois (1994). Clausal negation with *not*, which we assume involves negative quantification over eventualities or situations, is licensed in the same way. [3]

c. DQPs headed by *each* and *every* normally move to the Spec position of the Distributive-Universal category DistP, where they undergo Spec-head

[3]In other words, we assume, with Krifka (1989) and Schein (1993), that the correct logical translation of a negative sentence like *John didn't come* is not $\neg(come(j))$, but rather $no{:}e[come(e) \wedge Agent(e,j)]$ there are no events of coming where John is the agent.'

agreement with the Distributive-Universal head $Dist^0$, resulting in their characteristic interpretation. We will also suggest, however, that *every* can occur in other LF-positions as well, under certain circumstances; details are given in Section 4 and 5.

d. GQPs may select one of several distinct scope positions, resulting in the different interpretations that they receive:

   (i) GQPs that are referentially independent normally occupy the Spec of RefP position (located above CP), where they fulfill the function of (logical) subject of predication, and are interpreted with widest scope relative to other scope-bearing elements in their clause.

   (ii) A lower LF position, accessible by GQPs headed by an indefinite or a bare numeral, as well as QPs containing an externally bound variable, is the Spec of ShareP, which we locate just below DistP.[4] GQPs scoping in this position are interpreted with "dependent" specific reference, in the particular sense of specificity developed by Diesing (1990, 1992), i.e. ranging over individuals whose existence is presupposed. (This allows for a kind of narrow-scope specific reading, discussed below.) Whereas specific *indefinite* GQPs can occupy either the Spec of ShareP or the Spec of RefP position, specific *definite* GQPs must normally take scope in the Spec of RefP of that clause, and are scopally independent within it.

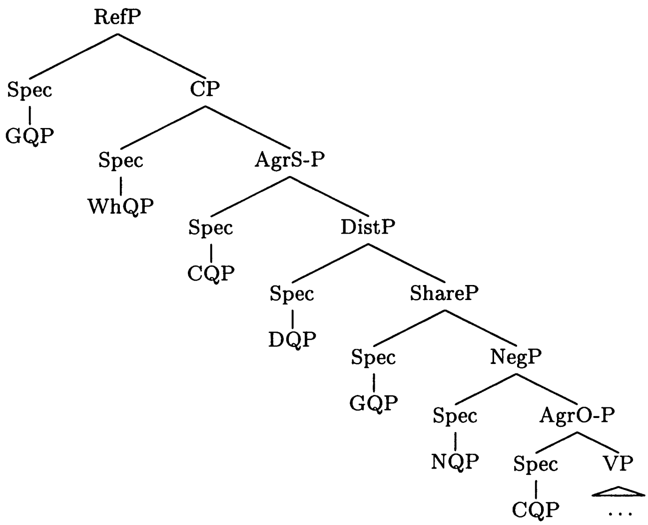
   (iii) Indefinite or bare-numeral GQPs may also take scope in their Case positions (i.e. *in-situ*), where they are interpreted non-specifically, like CQPs.

e. CQPs cannot ordinarily be interpreted as specific. Therefore they are interpreted in their Case positions and take scope *in-situ*. For a discussion of the properties of CQPs, the reader is referred to Szabolcsi (1996).

The relative scope positions of our five QP-types, based on their location in the functional structure of the clause, are given in (2):

[4]Definite QPs containing externally bound pronouns may also move to the Spec of ShareP, though we will not consider such cases here.

(2)



Given the well-known lack of island effects with definite and specific indefinite GQPs—which, like indexical pronouns and names, can have a *de re* construal even when they are embedded within islands—it has often been suggested that a wide-scope referential (*de re*) construal does not depend on movement. We will not be concerned here with the issue of how referential readings (cf. Fodor and Sag 1982) of indefinite QPs should be generated. We refer the reader to Kratzer (1995) for a recent proposal.

We assume that true GQPs become associated with an existential operator over a restricted variable, ranging over witness sets of the GQP.[5] This proposal seems to us essentially similar to that contained in Reinhart (1995), where the existential operator ranges over choice functions[6] (cf. Abusch 1994, Beghelli 1993, Beghelli 1995, Ruys 1993, etc. for further discussion).

Readers who are ill-at-ease with the postulation of functional categories will probably react with some skepticism to our claim that they play a central role in the syntax of quantifier scope assignment. We have several answers to this type of objection. First, with respect to the scope positions for WhQPs and NQPs, we are adding nothing new. Second, it is possible that our Spec of RefP position can be identified with the Topic position, and it is well known that topics undergo overt movement in many languages. (Our use of an LF landing site for GQPs forces us to adopt a somewhat broader notion of the "topic" function than what corresponds to the English Topic position, but many QP-types, including downward-entailing QPs, are forbidden from moving there.)

[5] We are grateful to Anna Szabolcsi for originally suggesting this idea to us.

[6] Kratzer (1995) develops Reinhart's suggestion to deal with the puzzle of island violations with referential indefinites.

Third, our Spec of ShareP position may correspond to the position of Diesing's scrambled narrow scope presuppositional QPs, though we make it accessible only to existential QPs of this type. Fourth, DQPs move overtly in some languages, to a position that we believe is none other than our Spec of DistP, as we will show below.

2.4 Scope and feature-checking

In the system that we propose, the movement of DQPs and GQPs to their scope position is driven by the need to check features that are associated to their QP-types. We will therefore refer to our proposal as a checking theory of scope assignment. We will return later on in this paper to the precise characterization of some of these features (in particular, to the different featural specification of *every* vs. *each*). Here we simply wish to present the overall picture, and evaluate some of its consequences.

Membership in any of the QP-types listed in Section 2.2 is indicated by a number of syntactic properties, some of which have been mentioned there. These properties are morphologically encoded in the determiner position of the DP or QP: this is obvious in the case of WhQPs and NQPs, as they bear *Wh*- and *n*-markings, but it arguably holds for other QP-types as well.

Thus, the determiners of DQPs (*each, every*) have what we may call *e*-morphology. Morphological markings (the presence of un-modified numerals, (in)definite article, etc.) distinguish the various subtypes of GQPs, and CQPs are characterized by the presence of modified numerals. These morphological specifications are not inherently different from the usual ones (agreement, case marking, etc.). We propose that they represent the syntactic encoding of logico-semantic features.

What is special with these, we propose, is that they carry logico-semantic features. WhQPs check their [+Wh] features through Spec-Head agreement with a Wh-operator hosted in C^0, and NQPs check [+Neg] in Spec of NegP, under agreement with the Neg-operator in Neg^0. Let us assume that a similar process obtains with the other QP-types. Feature-checking may appear to be more complex with the latter than it is with the former, but we are interested in pursuing the hypothesis that the process is essentially the same.

Our basic assumption is that DQPs need to check their [+Dist] features under agreement with a distributive operator (which we can indicate as $\forall$) hosted in $Dist^0$, whereas GQPs need to check group reference ([+group ref]) with an existential operator-head ($\exists$). Existential operator-heads occur in both $Share^0$ and Ref^0. The hierarchy in (2) thus corresponds to a hierarchy of operators. We claim that one of the basic roles served by the functional hierarchy of the clause is to encode the structural order in which semantic information is processed.

This gives the basic idea of what we think is going on in the process of scope assignment: scope is simply the by-product of agreement processes. Within this overall scenario, individual sub-types of QPs (and possibly individual quantifiers) realize additional features. GQPs are not, as a class, assigned a unique landing site: though definites typically take scope in Spec of RefP, numerals and indefinites can move to either RefP or ShareP. Extending the logic of our analysis, we suggest that when a GQP is endowed with an extra feature that marks it as the logical subject of predication, it will be driven to move up to (Spec of) RefP; otherwise it will remain in ShareP. If an indefinite GQP lacks the feature [+group ref] altogether, it behaves like a CQP, i.e. it goes no further than its Case position at LF). Unlike DQPs and GQPs, we assume CQPs do not have syntactically relevant features to check.

On a somewhat more technical level, we assume that scope positions can be reached either directly, through (leftward/upward) movement, or by (rightward/downward) reconstruction to a lower link in the chain of the QP. There is no principled difference between movement and reconstruction: each QP-chain is associated with one scope position, defined as the unique link which is compatible with the featural specification of the QP.[7]

2.5 The checking theory of scope versus other approaches

As noted above, the Checking Theory of scope that we develop here is in some respects a hybrid of May's theory (May 1977, 1985), which holds that all QPs undergo LF movement to their scope positions, and Hornstein's (1995) theory, which holds that quantifier scope is based strictly on chains formed by the movement of QPs to their Case positions in AgrSP and AgrOP. Our theory differs from these approaches in three important respects, however.

In assuming that only certain types of QPs undergo "QR" to a (non-Case) scope position, the Checking theory differs from May's theory, which holds that *all* QPs undergo QR at LF, and also from Hornstein's theory, which assumes that *none* of them do. More fundamentally, the Checking Theory is sensitive to the inherent semantic type of the QPs involved. First, certain QP-types must undergo LF movement from their Case positions, whereas others do not. Second, the Checking Theory provides targeted scope positions for each QP-type that does move; just as Wh-QPs and NQPs have targeted scope positions in the Spec of CP and NegP respectively, so DQPs headed by *every* or *each,* definite GQPs, indefinite GQPs, and CQPs have targeted scope positions too.

[7] Of course, this theory requires a suitable notion of Minimality to regulate movement. We do not explore this matter here; the reader is referred to Beghelli (1995) for a particular proposal in this direction.

These distinctive aspects of the Checking Theory of scope are motivated by the central empirical point that we wish to make, namely, that scopal ambiguity for pairs of clausemate quantifiers is much more restricted than has traditionally been assumed in the literature on quantifier scope. We are not referring here to the trivial observation that the discourse context may provide information that allows deductive reasoning to eliminate certain scope construals as unlikely or impossible; rather, we maintain that for certain combinations of quantifier-types, the *grammar* simply excludes certain logically possible scope construals. (In order to recognize this point, it is necessary to abstract away from the effects of discourse-related factors associated with Focus and Contrastive Topic intonation.)

We now turn our attention to the empirical generalizations that our theory captures. We begin by discussing the scopal behavior of indefinite GQPs, in terms of their interaction with DQPs and NQPs (including clausal negation). Next, in Sections 3 and 4, we examine DQPs and their scopal interactions with negation. In each case, one might object that May's or Hornstein's approach could account for the relevant data more simply, without invoking special functional projections for individual QP-types. To this objection, our reply is that the main strength of our approach lies in its ability to account for a range of data involving quantifier scope construals that are *not* ambiguous, where either of the alternative approaches would fail to distinguish in the appropriate way among different QP-types.

Independently of these factors, we believe that the extra complexity inherent in assuming a differentiated account of QP-types is compensated for by its being more theoretically uniform at a higher level. Our approach extends to all QP-types the basic analytical logic that has long been assumed for WhQPs, and more recently, for NQPs as well (cf. Zanuttini 1991, Moritz and Valois 1994).

Finally, we should draw attention to another general feature of the Checking Theory of scope developed here, which follows from the notion of targetted scope positions: the traditional notion that LF movement is typically optional can be dispensed with. Given that QP-types are endowed with certain intrinsic features, they *must* move to those scope positions where the features in question can be licensed.

2.6 Empirical justification

We have stressed that the fundamental motivation for our approach is empirical. We will now review some of the empirical justification for the rich structural representation that we hypothesize. We concentrate on interactions between clausemate QPs surfacing in subject and object positions, where one of the QPs is an indefinite GQP. We present only some of the relevant data in this section; further data will be considered in later parts of this paper. Scopal interactions between DQPs and negation (including both clausal negation and NQPs) are considered in Section 3; scopal interactions involving WhQPs are discussed extensively in Beghelli (1996). Furthermore, we will make only passing references, in discussing the predictions of our theory, to the scopal behavior of CQPs, since they bear only tangentially on the focus of the present paper; the reader is referred to Beghelli and Stowell (1994) and Beghelli (1995).

2.6.1 Clause-internal scopal asymmetries

We begin our empirical discussion by enumerating three predictions implied by the hierarchy of positions in (2):

(3) a. A WhQP should always take wide scope with respect to any other QP in their clause, other than GQPs when these are assigned scope in Spec of RefP.

b. A GQP should be scopally ambiguous with respect to a clausemate DQP, depending on whether the GQP moves to Spec of RefP or to Spec of ShareP.

c. A GQP object should be scopally higher than clausal negation, owing to the fact that it takes scope in Spec of ShareP or Spec of RefP–except in the case mentioned above where an indefinite or bare-numeral GQP remains in its Case position (Spec of AgrO-P) and receives a counting interpretation; cf. (diii) in 2.3. A GQP subject should always take wide scope with respect to clausal negation and/or a clausemate NQP.

d. A CQP in object position should never be able to take inverse scope over a GQP or DQP occurring in subject position.

Let us now see the empirical status of these predictions, and how they follow from our assumptions. Some of the predictions in (3) are, of course, familiar facts from the literature. For instance, (3a)—that WH-QPs take widest scope—is widely assumed, and we are essentially following a long tradition here.[8] Prediction (3b)—that clausemate GQP/DQP pairs are scopally ambiguous—is also a familiar fact, exemplified in paradigms such as (4):

(4) a. Every/Each student read two books.

b. Two students read every/each book.

In each case, the indefinite GQP headed by *two* can be construed either inside or outside the scope of the DQP headed by *every/each.*

[8]There is one apparent counterexample to the claim that WhQPs scope higher than QPs with *every, each, few, ...* , involving so-called pair-list readings of certain QP-types in certain syntactic positions; these are discussed in Beghelli (1996).

Our account of (4) does not differ empirically from the classical QR-based theory advanced by May (1977), although it derives the scopal ambiguity in a different way. The classical theory of May (1977) captures the ambiguity as a result of QR being free to apply sequentially, in either order, to both QPs. Either QP may adjoin to S, creating a higher S-node; then the other QP will adjoin to the higher S-node, taking wider scope than the QP that moved first. Since either QP can be the first to move, two LF-configurations are possible, resulting in the ambiguity. (This analysis could be translated into a Minimalist framework, by allowing both QPs to adjoin to AgrS-P, or by allowing one to adjoin to AgrS-P, and the other to adjoin to some other functional category, such as TP.)

In contrast, the Checking Theory of scope that we are advocating here must claim that the DQP will always end up in the same LF scope position, namely in the Specifier position of the Distributive Phrase (Spec of DistP). Hence the scopal ambiguity must arise in some other way. We suggest that it arises because indefinite GQPs have an ambiguous quantifier type, making more than one LF position available to them; in fact, we suggest that they have *four* possible LF landing sites. One of these—Spec of RefP—is superior to the DQP's position in Spec of DistP; another—Spec of ShareP—is inferior to it. The other two positions are both Case positions (Spec of AgrS-P and Spec of AgrO-P, for subjects and objects, respectively); of these, the latter is inferior to the LF position of the DQP, while the former is superior to it.

Consider now (4b), where an indefinite QP occurs in the subject Case position (Spec of AgrS-P) at Spell-Out, and a DQP occurs in the object position. Since the DQP must move to the Spec of DistP position, which is inferior to the Case position of the subject, a narrow scope construal of the subject will be possible only if the subject reconstructs to a scope position *lower* than Spec of DistP. For the GQP subject in (4b), a narrow scope construal of the subject must involve its reconstructing to the Spec of ShareP position, since it cannot reconstruct to the Spec of AgrO-P. (The possibility of its reconstructing to its θ-position is discussed below.)

The reader may wonder how the Checking Theory of scope can account for sentences containing two DQPs, such as *Each boy read every book* or *Every professor gave every student an A*. If DQPs headed by *each* and *every* have a unique LF landing site, then one might expect that a given sentence could contain only one of them. The analytical problem posed by such examples is no different in principle from that posed by multiple Wh-questions or by sentences containing multiple NQPs, e.g., in languages exhibiting "negative harmony" such as Spanish. For such cases, we follow a long tradition in assuming that the Spec positions of scopal categories can be multiply filled, either because there may be more than one specifier for the same projection, or through a process of absorption applying to quantifiers of the same logical type.

The first prediction in (3c)—that *indefinite GQP objects can take inverse scope over negation*—is also a familiar fact, based on examples like (5a, b):

(5) a. The students didn't read two/some books.
    b. No student read two/some books.

The second prediction in (3c)—the possibility of a *narrow-scope* construal for an indefinite GQP object, as in (5a), follows from our proposal that some (e.g., bare-numeral) GQPs can be interpreted as CQPs and remain in their Case positions at LF, as in 2.2 and 2.3.

Empirical support for the third prediction in in (3c)—that *indefinite GQP subjects must take scope over negation*—is less widely recognized, though it is supported by (6a, b):

(6) a. Two/some students didn't read this book.
    b. Two/some students read no books.

Assuming that the LF scope position of both clausal negation and NQPs is located at the NegP level, the *possibility* of a wide-scope construal of indefinite GQP subjects and objects is expected, given that indefinite GQPs have two possible LF landing sites above NegP in (2)—Spec of ShareP and Spec of RefP. (The distinction between these two positions is not obvious in examples like (5) and (6), and may appear at this stage to be an artifact of our account of (4); however, we will provide justification for this shortly.)

However, the GQP subjects in (6) apparently *must* take wide scope relative to negation, suggesting that there is no position below the scope domain of the negative operator (in Spec of NegP) that these subject GQPs can reconstruct to. Our hierarchical arrangement of scope positions provides an account of this, in the spirit of Hornstein (1995). Unlike an object GQP, whose Case position (Spec of AgrO-P) lies *within* the scope of negation, a subject GQP would have to reconstruct to a position within VP in order to derive a narrow-scope construal relative to negation, since the subject Case position (AgrS-P) is too high up. (Reconstruction to the Spec of ShareP can derive a narrow scope construal relative to a distributive operator in DistP, but it is not low enough to produce a narrow scope construal relative to negation.)

Thus, there is only one way in which a narrow-scope construal of a subject GQP relative to negation might be derived: by reconstruction of the subject GQP to its original θ-position below NegP. Evidently this option must be excluded. A natural way of deriving this result would be to assume that every quantifier phrase must syntactically bind a trace as a variable in the LF representation. (Though the semantic basis for such an assumption is not obvious, we will assume nevertheless that such a condition holds, on LF representations, at least.)[9] Then reconstruction of a GQP—or another quantifier phrase—to its original θ-position would be excluded, since there would be no trace in a lower position for the GQP to bind.

Simple indefinites (singular indefinites with the article *a/an* and bare plurals) in subject position do seem to be capable of reconstructing below NegP, however, as in (7):

(7) a. A student didn't write this book.

b. Students didn't write this book.

Furthermore, as is well known, simple indefinites and bare plurals can routinely be bound by generic operators and adverbs of quantification, whereas numerals and *some* do not show this type of variability. We can provide an explanation for the difference between (6) and (7) if we follow much recent work[10] in assuming that simple indefinites and bare plurals are actually restricted variables which can be unselectively bound by a variety of external quantifiers, including negative quantifiers. This will allow them to reconstruct into a θ-position because, being variables and not quantifier phrases, they do not need to bind variables themselves. Nor do they need to be checked with an operator-head in Spec,ShareP or Spec,RefP for existential quantification, because they are unselectively bound. Hence the contrast between (6) and (7).

Lastly, we should point out that the introduction of a special type for CQPs is motivated by a basic asymmetry in subject-object scope interactions. Whereas both DQPs and GQPs can, when in object position, take wide scope over a subject GQP (though not in the same way—cf. Section 3), CQPs are not able to take inverse scope:

(8) a. Some/one of the students visited more than two girls.

b. Some/one of the students visited few(er than three) girls.

c. Every student visited more/fewer than three girls.

In neither of (8a, b, c) can the object QP take scope over the subject (at least if normal intonation is employed). For example, we cannot construe (8a) to mean that for more than two girls, it is the case that some student, or one of the students, visited her.

This is derived directly under our analysis, since an object CQP cannot scope higher than Spec of AgrO-P, and a subject GQP, as seen above, cannot

[9]This assumption was not made in Beghelli (1995), where it was proposed that CQPs are allowed to reconstruct in their theta-positions. Some of the empirical problems handled by this latter, less restrictive view, remain as open questions in the solution suggested in the present paper.

[10]Heim (1982), Kratzer (1988), and Diesing (1988, 1990)

reconstruct lower than Spec of ShareP. (Nor can a subject DQP reconstruct below Spec of DistP.) Our assumptions about the local scope of CQPs are further confirmed by the observation that these QPs only support a *de dicto* reading when they are complements of intensional predicates:

(9) Someone wanted to visit more than two professors.

2.6.2 Cross-linguistic evidence

As a second argument for the Checking Theory of scope, we cite empirical evidence from surface constituent order in a number of languages, supporting our contention that there are distinctive scope positions defined in the phrase structure of the clause for DQPs and (particular construals of) GQPs. The paradigmatic case of one-to-one correlation between surface order and scope seems to be Hungarian, a language known to 'wear LF on its sleeve.' Szabolcsi (1996) presents striking evidence in support of the Checking Theory, by showing that, in Hungarian, a hierarchy of positions essentially similar to (2) governs the surface order of QPs. In this language, GQPs, DQPs, and CQPs move in the overt syntax to their specified scope positions in the hierarchy of functional projections in (2).

With respect to DQPs, Kinyalolo (1990) has shown that, in the Bantu language KiLega, universally quantified noun phrases that are obligatorily distributive must undergo overt leftward movement in the visible syntax.[11] We interpret this as evidence that KiLega requires DQPs to be spelled out in Spec of DistP, just as English requires (most) WhQPs to be spelled out in Spec of CP. Similarly, Khalaily (1995) shows that the Palestinian Arabic counterparts of our DQPs must undergo leftward movement in the overt syntax in a parallel fashion; he argues that Palestinian Arabic exhibits an overt counterpart to our LF movement to Spec of DistP, a conclusion that we concur with.

Further cross-linguistic evidence comes from the recent literature on scrambling in Hindi (Mahajan 1990) and various Germanic languages (cf. Kratzer 1988 and Diesing 1990, among others). A number of proposals have suggested that specific construals of indefinites are necessarily associated with (overt) leftward movement out of VP. Though the exact location of the landing site of scrambling is still being debated, we believe that the position that we identify as Spec of ShareP is a common landing site for scrambling. We will not develop this point here, however, since this would take us too far afield.

[11]It is significant to note that in KiLega universal terms that are not obligatorily construed with distributivity do not move leftward. In other words, only the quantifier corresponding to *each, every* triggers movement; the *all* quantifier does not. The latter is distinguished from the former in that it supports collective readings. Cf. Beghelli (1995) for discussion.

2.7 Semantic assumptions

Thus far, we have sketched out a theory of quantifier scope based on the typology of QPs listed in 2.2, the fixed scope domains ordained by the clause structure in (2), and the assortment of assumptions in Section 2.3 about where the individual QP-types can occur in LF. Before concluding this introductory overview of our theoretical approach, we should make our assumptions concerning the semantic underpinnings of our proposal explicit. We appeal chiefly to the theory outlined in Szabolcsi (1996). Szabolcsi's proposal is a development of the core tenets of Discourse Representation Theory (DRT). We give the following as a minimal set of hypotheses on which our approach rests:

(10) a. Following Szabolcsi's (1996) modification of standard DRT, we assume that GQPs introduce discourse referents in the form of restricted group variables. Such variables correspond to the minimal *witness set* of the QP in generalized quantifier theory parlance. Thus, a GQP like *two men* introduces a variable X ranging over sets containing two men and no non-men. We have suggested above that the variable introduced by a GQP must be checked with an existential operator-head that can only arise in two positions, as laid out in Section 2.3: Ref^0 and Share^0. Only simple indefinites and bare plurals act as plain variables.

b. Following standard DRT, we assume that CQPs are interpreted as generalized quantifiers. Because they do not introduce discourse referents (=variables), they do not undergo movement in LF above and beyond Case-driven movement.

c. We depart from standard DRT (and follow Szabolcsi 1996) in assuming that DQPs also introduce discourse referents, albeit of a different type than GQPs. Whereas GQPs introduce individual variables (whether singular or plural individuals—the term *group* covers both)—DQPs introduce *set* variables, which are again restricted variables ranging over witness sets of the quantifier. In Section 5 we discuss how the set variable introduced by a DQP gets bound, and by which operator. (Note, for clarification, that the distributive operator $\forall$ hosted in Dist^0 does not bind the set variable; this operator applies at a different level, that of the elements of the set.)

3 SCOPE AND DISTRIBUTIVITY

3.1 Varieties of scope judgements

Scope judgments involving quantifiers are commonly based on three types of interpretations, and related intuitions. The first type of intuition, usually invoked in assessing the interaction of existential quantifiers with a variety of logical operators (including negation and various intensional predicates), concerns existence presuppositions, as in (11):

(11) a. John wants to marry a Canadian princess.

b. John didn't marry a Canadian princess.

If the existentially quantified indefinite QP falls under the scope of *want* or *not*, then the speaker need not be committed to the existence of any Canadian princess; on the other hand, if the QP scopes over the logical operator, then the speaker is committed to the existence of one such individual. This sort of intuition will not concern us in this section.

A second type of intuition involves in scope interactions with negation and other downward-entailing operators. Consider, for example, the scopal interaction between negation and an existential or universal quantifier, as in (12):

(12) a. John didn't read a book.

b. John didn't read every book.

In these examples, the preferred reading is for negation to scope over the existential QP in (12a) and over the universal QP in (12b); however, the existential quantifier is also free to scope over negation in (12a), whereas in (12b), the universally quantified object can scope over negation only if it is focussed. In these examples, the primary basis for the scope judgements involves the interaction of the logical operator *not* with the logical operators $\forall$ and $\exists$.

A third type of intuition, commonly associated with indefinite QPs interacting with a variety of other QP-types, concerns distributivity. If a given QP_1 takes scope over an indefinite QP_2, then QP_1 is usually understood to distribute over QP_2. On the other hand, if QP_1 fails to take scope over QP_2, then distribution fails. Consider (13):

(13) a. Every boy read two books.

b. Five boys read two books.

If the indefinite GQP object falls under the scope of the subject QP, the total number of books involved is potentially much greater than two; the quantity associated with the existentially quantified GQP object is multiplied by the value of the other QP, so that (13b) can describe the reading of up to ten books, the total number depending on whether some of the boys might have accidentally read the same books. In such a distributive reading, we will describe the wide-scope QP as the DISTRIBUTOR and the narrow-scope indefinite as the DISTRIBUTEE or DISTRIBUTED SHARE (Choe 1987). If *two books* does *not* fall within the scope of the other QP, then distribution fails, and the sentence only describes the reading of a total of two books.

3.2 Collective and distributive construals

These intuitions about distribution rely on the possibility that the noun phrase serving as the distributed share is capable of referential variation, e.g. that each boy read a different pair of books in (13) (cf. Beghelli et al. 1996). In a situation where the boys happened to read the same set of books, as in *Five boys read all the books,* the DP *all the books* cannot serve as distributed share in the relevant sense, since there is no possibility of referential variation. Hence, narrow scope readings of some DP-types, including universals, cannot be accessed by intuitions of distributivity. We will assume, however, in agreement with Beghelli et al. (1996), that when a definite DP or DQP lands at LF in a position lower than that of another QP_2, it does take narrow scope with respect to QP_2.

The type of distributivity illustrated in these examples involves an overt indefinite GQP serving as the distributed share. In other cases, distributivity seems to involve distribution of events or agentive functions; this type of distributive reading is often contrasted with a so-called COLLECTIVE reading. Consider (14):

(14) John and Bill visited Mary.

On the distributive reading, John and Bill are each agents of distinct events involving visits to Mary; on the collective reading, John and Bill act together as joint agents of a single visiting event. We will assimilate this collective/distributive distinction to the paradigm in (13) by assuming that there is a covert existential quantifier over events in (14), as suggested by Davidson (1967), Kratzer (1988) and many others; if this existential quantifier falls under the scope of the subject GQP, then a distributive reading results; if the covert existential quantifier takes broad scope, then distribution fails, and a collective interpretation results.

We have labeled the QPs headed by *each* and *every* as Distributive-Universal QPs (DQPs)—*distributive*, because they must usually serve as distributors, and *universal* because they are usually understood to have the force of universal quantification.[12] The universal force that these QPs typically convey is illustrated in (15):

(15) a. All the boys visited Mary at six o'clock.
b. Every boy visited Mary at six o'clock.
c. Each boy visited Mary at six o'clock.

Suppose that the set of boys being quantified over consists of Tom, Dick, and Harry; then these sentences are all true if Tom, Dick, and Harry all visited Mary at six o'clock; they are all false if any one of the boys failed to visit Mary at six o'clock.

The distributive nature of *each* and *every*—as opposed to *all*—can be illustrated by considering contrasts such as the following:

(16) a. The Pope looked at all the members of his flock.
b. The Pope looked at every member of his flock.
c. The Pope looked at each member of his flock.

(17) a. All the boys surrounded the fort.
b. ? Every boy surrounded the fort.
c. ? Each boy surrounded the fort.

In (16), the universally quantified objects all allow for a distributive construal, where the object QP serves as a distributor and a looking event serves as a distributed share; but only *all* allows for a collective construal, where distribution fails and there is a single looking-event. Thus, in (16a), the Pope might have looked at the assembled multitude with a single glance, but in (16b) and (16c), he must have looked individually at each and every member of his flock.

In (17), the predicate *surround* requires an event with a semantically plural agent; this requires a collective (nondistributive) construal of the subject QP, which must denote a (plural) group. Such a construal is possible with a universal QP headed by *all* in (17a), but not with a DQP headed by *every* or *each* in (17b, c); the DQPs force a distributive construal, where a surrounding-event serves as the distributed share, attributed individually to each member of the set of boys, a reading that is incompatible with the semantics of the predicate.

Another property distinguishing *each* and *every* from *all* is grammatical number: *each* and *every* are grammatically singular, combining with morphologically singular NPs and binding singular pronouns as variables:

[12] Actually, the situation is apparently somewhat different in Hungarian, where several non-universal QP occur overtly in what appears to be the Spec of DistP position; cf. Szabolcsi (1996).

(18) a. All the $boys_i$ said $they_i$ were tired.

b. Every boy_i said he_i was tired.

c. Each boy_i said he_i was tired.

We believe this property is related to their strong distributive behavior; we will return to this point in Section 5 (cf. also Beghelli 1995).

Summarizing our discussion thus far, we have reviewed two ways in which a distributed share can be provided to set up distribution. The first involves an overt indefinite GQP functioning as a distributed share for another QP; the second involves a covert existential quantifier over events functioning as a distributed share, on a distributive (non-collective) event construal. We have also seen two contexts where DQPs force distributive (non-collective) event construals in configurations where universally quantified QPs headed by *all* and other QP-types allow a nondistributive (collective) construal. Henceforth, we will refer to *each* and *every* as STRONG DISTRIBUTIVE quantifiers.

3.3 Other diagnostics of strong distributivity

So far, we have not shown that DQPs headed by *every* or *each* differ from QPs headed by *all* (or from other types of QPs, for that matter) with respect to distribution over overt indefinite GQPs. At first glance, they appear not to. For example, in (19), the indefinite object GQPs seem to be allowed to function as distributed shares for various types of subject QPs:

(19) a. Tom, Dick, and Harry read two books about India.

b. Three boys read two books about India.

c. All the boys read two books about India.

d. Every boy read two books about India.

e. Each boy read two books about India.

(For many speakers, the DQP subject headed by *each* in (19e) seems to favor a distributive construal over the indefinite object somewhat more strongly than the other subject QPs do, but this does not appear to be an absolute requirement.) Thus, while *each* and *every* may be more strongly distributive than GQPs (including those headed by *all*) with respect to covert event quantification, such a distinction does not seem to be justified when an overt indefinite GQP functions as the distributed share.

This conclusion turns out to be premature, however. Recall that there are two possible LF scope positions below the Spec of DistP for GQP objects: the Spec of ShareP and the Spec of AgrO-P. (We have already suggested that GQPs may remain in their Case positions at LF, and that when they do so, they have the counting interpretation characteristic of CQPs; this assumption was necessary in order to account for the fact that GQP objects are free to scope under negation.) Thus, it is possible that the GQP objects in (19) are actually occurring in Spec of AgrO-P rather than in the Spec of ShareP.

An interesting difference between DQPs and other QP-types emerges when we consider structures involving singular indefinite QPs modified by the adjective *different,* which functions as an unambiguous marker of true *distributed share* status.[13] Only QPs headed by *every, each* can enforce a distributive reading when they take scope over *a different N.* The following examples illustrate this:

(20) a. Every boy read a different book.
  b. Each (of the) boy(s) read a different book.
  c. * All the boys read a different book.
  d. * The boys read a different book.
  e. * Five boys read a different book.

DQPs also differ from GQPs with inverse scope construals. Whereas DQP objects headed by *each* or *every* can assume the distributor function, other QP-types, including GQPs headed by *all*, cannot:

(21) a. A (different) boy read every book.
  b. A (different) boy read each book.
  c. * A (different) boy read all the books.
  d. * A (different) boy read Ulysses and Dubliners.
  e. * A (different) boy read two books.

In (21c–e), the subject GQPs may not be construed as distributed shares, and *different* must be understood to mean 'different from some other boy mentioned previously in the discourse,' whereas in (21a, b), the subjects can be so construed, and *different* can be understood to differentiate among the referents of the distributed share. In addition, in sentences like (19b) the distributive reading 'there are two books about India, such that for each one, two (possibly different) boy read it' is not generally available, as noted by Kamp and Reyle (1993), Ruys (1993) and references cited therein.

Actually, examples where an inverse distributive reading appears to be available with GQPs have been quoted in the literature. In this vein, Reinhart (1995) cites the well-known *American flag* example noted originally by Hirschbuehler (1982), to which we may add a more benign floral example:

[13]Items like *a different N* also have an anaphoric reading: 'an N which is not identical to the one mentioned before.' This reading is irrelevant here.

(22) a. An American flag was hanging in front of two buildings.
b. Blossoms sprang out of two rosebushes.

Such examples, however, rely crucially on the special properties of simple indefinites and bare plurals, which—as noted above in the discussion of (7)—are allowed to reconstruct to their VP-internal thematic positions. The inverse distributive readings disappear with different choices for the subject GQP:

(23) a. Five guards stood in front of two buildings.
b. Three blossoms sprang out of two rosebushes.

Neither of these cases seem to readily allow for a distributive interpretation—which our approach correctly predicts, since reconstruction to the original VP-internal thematic position is precluded for these GQP subjects, as we have already seen. In contrast to such cases, any indefinite GQP can serve as the distributed share when a DQP headed by *every* or *each* functions as the distributor, as observed above. Consequently, we believe that Hirshbuehler's type of example is simply reflective of the special reconstructive abilities of simple indefinites and bare plurals, and does not undermine our distinction between DQPs and GQPs with respect to inverse distributive construals. This suggests that our distinction between strong distributivity with DQPs headed by *every* or *each* and the type of distributivity exhibited by other QP-types does extend to distribution over overt indefinite GQPs after all. The contrast between (19) and (21) also suggests that this distinction is syntactically based, insofar as it is sensitive to c-command relations holding between the two QPs.

We are now in a position to relate the strong distributivity exhibited by DQPs to the syntactic structure for quantifier scope that we introduced in Section 2. We suggested there that DQPs always move from their Case positions to the Spec of DistP at LF, and we pointed out that this movement seems to take place in the visible syntax (before Spell-Out) in some languages. We also suggested that when a DQP scopes over a clausemate GQP, the GQP normally occurs in the Spec of ShareP. We now propose to exploit this structure to characterize the difference between strong distributivity (associated with DQPs) and the type of distributivity exhibited by other QP-types.

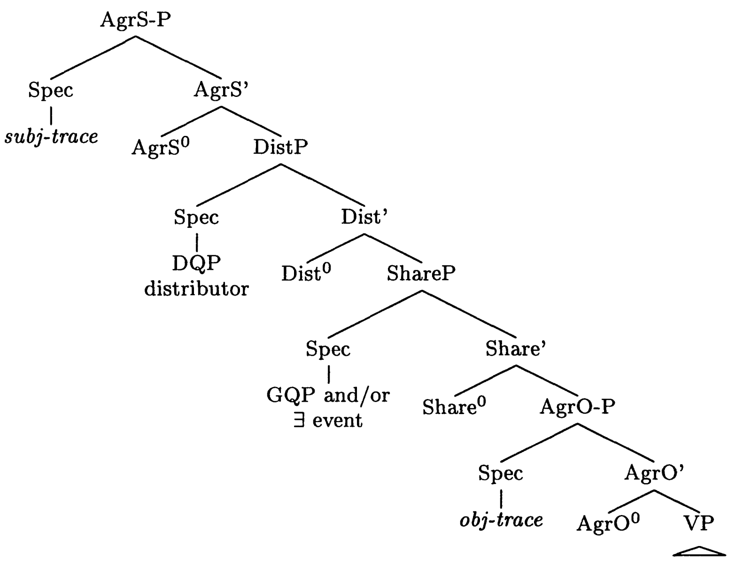
Strong distributivity seems to have three characteristic diagnostic properties:

(24) **Strong Distributivity**

a. DQPs headed by *each/every* are Strong Distributors.
b. Strong Distributivity is obligatory.
c. Strong Distributivity can arise under an inverse scope construal, e.g., where the distributee is in Spec of AgrSP and the distributor is in Spec of AgrOP.

Let us now review our assumptions about scope assignment with DQPs. Suppose that DQPs bear an intrinsic feature of (strong) distributivity [+Dist]. As discussed in Section 2, this feature must be checked in the same way that features such as [+Wh] and [+Neg] must: under Spec-Head agreement with a functional head. Thus [+Dist] DQPs must appear in the Spec of DistP at LF in order for their distributive feature to be licensed. The Dist^0 head selects as its complement a functional category containing the QP corresponding to the distributed share. This functional category, which we label ShareP, requires an existentially quantified indefinite GQP (the distributed share) to occur in its Spec position, just as NegP and [+Wh] CP require NQPs and WhQPs to occur in their Spec positions. When a DQP takes distributive scope over an indefinite GQP, the indefinite moves to Spec of ShareP at LF; when there is no overt indefinite and the GQP simply forces a distributive (non-collective) event construal, a covert quantifier over events moves to the Spec of ShareP. The complement of Share^0 contains the Verb Phrase and various lower-level functional projections (including NegP and AgrOP):

(25)



Thus, a chain of syntactic dependencies captures the strong distributive nature of DQPs. Our account captures the characteristic properties of Strong Distributivity in (24); (24a, b) follow from the mechanism of feature-checking, and (24c) follows from the fact that Spec of DistP and Spec of ShareP are possible LF landing sites for DQPs and indefinite GQPs, respectively.

Let us see how our system works with a few simple examples:

(26) a. Every boy visited Mary at six o'clock. [15b]

b. The Pope looked at each member of his flock. [16c]

c. Each boy read two books about India. [19e]

d. A (different) boy read every book. [21a]

In every case, the DQP headed by *each/every* must move to Spec of DistP, where its [+Dist] feature is checked. This requires the presence of an active Dist head, just as the movement of a WhQP to Spec of CP requires the presence of an active [+Wh] Comp head. The active Dist head selects a ShareP with a Share head that licenses (and requires) an existential QP in Spec of ShareP, by the familiar feature-checking mechanism.

We mentioned above that a number of recent studies (Schein 1993, Higginbotham 1985, Kratzer 1988, Diesing 1988, 1990) have adopted Davidson's (1967) proposal for the existence of an event argument and proposed a syntactic position for it.[14] We wish to adopt the proposal that event arguments are syntactically realized, but in a modified form; we suggest that this argument position occurs VP-internally, and that it functions as a θ-position of the usual sort, i.e., as a syntactic position in which overt and covert QPs may originate (cf. Stowell 1991). Adverbial QPs ranging over events such as *rarely, never, always* originate there; the same is true of the WhQP *whether* and the NQP *not* (clausal negation), and the covert existential event QP $\exists$. Just as *whether* and *not* move to their scope positions in Spec of CP and NegP respectively to have their quantificational features checked, so $\exists$ moves to the Spec of ShareP.

In (26a, b) there is no overt GQP, so the covert existential quantifier over events must move to the Spec of ShareP, forcing a distributive (non-collective) construal. In (26c, d) there is an overt indefinite, which is free to move into the Spec of ShareP, resulting in distribution over books in (26c) and over boys in (26d). These overt indefinites are also free to move to the Spec of RefP instead, resulting in a wide scope construal, in which case the event quantifier must move to the Spec of ShareP.

Our analysis implies that the covert event QP $\exists$ does not need to move to the Spec of ShareP if there is an overt indefinite GQP that can move there instead, as in (26c, d). This does not seem to be correct, however; it seems that the event quantifier is always forced to move to Spec of ShareP, since it is virtually impossible to construe a DQP as taking distributive scope over an overt indefinite with a collective (nondistributive) event construal. We are not

[14]There is some disagreement on whether this argument position is realized for all types of predicates, or just for stage-level (or possibly just for eventive) predicates; we will assume that the position exists for all types of predicates, but that in the case of individual level predicates, it cannot be existentially quantified: it can only be (semi-)generically quantified.

certain whether the latter observation is a fact that the syntax of LF should try to account for, or whether it is a fact about the ontology of permissible event-types; for concretenes, we will assume the latter view, but we will not try to resolve this issue here.

We have provided a syntactic account of Strong Distributivity, but so far we have not attempted to explain the type of distributivity associated with non-DQP distributors. We have seen that the latter type of distributivity, which we will refer to as WEAK DISTRIBUTIVITY or PSEUDO-DISTRIBUTIVITY, has the following characteristic properties:

(27) **Pseudo-Distributivity (Weak Distributivity)**

a. Plural definite and indefinite GQPs (including QPs headed by *all*) are Pseudo-distributors.

b. Pseudo-distributivity is optional.

c. Pseudo-distributivity cannot arise under an inverse scope construal, e.g., where the distributee is in Spec of AgrS-P and the distributor is in Spec of AgrO-P.

Property (27c) suggests that Pseudo-distributivity does not make use of distributor movement to a targetted scope position such as Spec of DistP *per se*; otherwise, we would expect that any QP-type that can trigger Pseudo-distributivity should be able to do so regardless of where it originates within the clause. We will not provide an explicit account of Pseudo-distributivity here; the reader is referred to Beghelli (1996) for detailed discussion. We will simply sketch the essentials of the proposal given there. Pseudo-distributivity arises through the agency of a covert distributive element corresponding to floated *each* (cf. Roberts 1987). Like its overt counterpart, silent *each* is optionally generated between AgrS-P and AgrO-P. Pseudo-distributivity is supported if silent *each* is c-commanded by (the trace of) the GQP that acts as (pseudo-)distributor, and c-commands the LF position of the QP that functions as (pseudo-)distributee. In the case where the distributed share is an indefinite GQP object, the lower scope position in question may be the Spec of AgrO-P.

4 STRONG DISTRIBUTIVITY AND NEGATION

In Section 2, we outlined the basic scope interactions exhibited by definite GQPs in relation to both DQPs and NQPs (including so-called clausal negation). Thus far, however, we have avoided any discussion of the scopal interaction between DQPs and NQPs. Our structure in (2) suggests that we should expect DQPs to uniformly take scope over NQPs, since the Spec of DistP (the target scope position of DQPs) asymmetrically c-commands the Spec of NegP (the target scope position of NQPs).

The facts of DQP/NQP scope interactions with negation are much more complex than this, however. It turns out that DQP subjects behave differently from DQP objects, and, to make matters worse, *each*-DQPs behave differently from *every*-DQPs. We will concentrate on structures involving clausal negation marked by the particle *not*), which we have analyzed as an NQP that originates in the *Event* argument position and moves to the Spec of NegP (like any other NQP) to have its negative feature checked at LF. Since the same analytical logic extends to other types of NQP such as *nothing*, *no man*, etc., we will not discuss them explicitly here, in order to keep the discussion to a manageable length.

It turns out that DQPs, far from scoping comfortably above negation, seem to be awkward or ungrammatical with it in most cases; in the one example where they seem to coexist happily (29a), negation scopes over the DQP, rather than vice-versa:

(28) a. ?? Every boy didn't leave.
b. ?? Each boy didn't leave.

(29) a. John didn't read every book.
b. ?? John didn't read each book.

Before proceeding further, we should comment briefly on the status of our judgments, since they depart from what is generally assumed about such data. Our judgments are based on a neutral, non-focussed intonation; if the DQP or the negated verb is focussed, these examples become grammatical, with distinct (and generally unambiguous) scope construals. We assume that these focussed readings have distinct LF representations associated with them, but we will say nothing further about them here; we are interested in explaining the marginal status of the non-focussed readings.

The Checking Theory of DQP licensing, combined with our account of event quantification, accounts directly for these data, with the exception of (29a), which we discuss further below. In each case, the DQP should be forced to move to the Spec of DistP, activating Dist^0 and its complement ShareP. But there is no existential QP available in any of these examples to occupy the Spec of ShareP and satisfy the checking requirements of its head. None of these sentences contain any overt indefinite GQPs, and in every case the event variable is bound by the (cliticized) event-NQP *n't*—or its null counterpart, if *n't* is really the head of NegP—so there cannot be a covert existential event-QP, either. (There is only one *Event* argument position available, and it is impossible for two distinct QPs to originate there, just as it is impossible for two distinct QPs to originate in any other argument position.) Since there is no indefinite QP that can move to the Spec of ShareP, the checking requirements of the head of ShareP cannot be satisfied, and the Checking Theory predicts that all of these examples to be excluded. This yields the desired result in every case except (29a), to which we return below.

When the DQPs in (28)–(29) are replaced by (definite) universally quantified GQPs headed by *all*, the results are fully grammatical:

(30) a. All the boys didn't leave.
 b. John didn't read all the books.

These examples seem to behave like the examples involving scopal interactions between indefinite GQPs and negation discussed in Section 2: the subject GQPs must scope over negation—at least on the neutral intonation—while the objects are scopally ambiguous. These examples can thus be assimilated to the treatment of GQPs given earlier. We account for the difference between *each/every* and *all* by assigning QPs headed by *all* to the type of GQPs, with the proviso that the Spec of ShareP position is unavailable to these universally quantified GQPs for reasons already discussed. (Only QPs that are capable of referential variation may occur there, i.e. indefinites and definites containing free variables.) The decision to treat *all* as the head of a GQP also fits in with its ability to occur as the subject of collective predicates, as discussed in connection with examples (16)–(17).

The data in (28)–(29), as well as the contrast between (30) and (28)–(29) provides strong support for our approach to Strong Distributivity, as well as our distinction between Strong Distributivity and Pseudo-distributivity. But although our treatment of Strong Distributivity correctly excludes (28a, b) and (29b), these examples do not show that DistP should be placed above NegP, as in our proposed structure, rather than beneath it. In fact, we would predict the same result if NegP were placed higher than DistP; since the NQP *not* would still originate in the *Event* argument position and bind its trace there as a variable, which ought to prevent the covert existential event-QP $\exists$ from originating there as well.

The crucial evidence for our relative hierarchical placement of DistP and NegP comes from sentences similar to those in (28) and (29), but with an overt indefinite GQP, as in (31) and (32):

(31) a. Every boy didn't read one book.
 b. Each boy didn't read one book.

(32) a. One boy didn't read every book.
 b. One boy didn't read each book.

The first thing to observe about these examples is that they are markedly better than their counterparts in (28a, b) and (29b). Our account of strong distributivity predicts this; although the presence of *n't* precludes an existential event-QP, there is an overt indefinite GQP that can move to the Spec of ShareP at LF, thus satisfying the requirements of the activated Share^0 head. Moreover, (31a, b) and (32b) have precisely the scope readings that we expect to find, given our structure in (2): on the preferred reading, the indefinite GQP scopes over negation and under the DQP headed by *every* or *each*: thus, (31a) translates as 'for every boy, there is one book that he didn't read.' The crucial point is that the grammatical scope construal has the DQP and the indefinite GQP both scoping above negation, supporting our hierarchical placement of DistP and ShareP relative to NegP.

The only problematical example in this paradigm is (32a); here, the *every*-DQP seems to be unable to scope over negation, even though there is an indefinite GQP subject available, which should be able to move to ShareP. Our Checking Theory of scope, as outlined thus far, fails to capture this. (32a) is problematical in the same way that (29a) is; the DQP seems to be forced to scope *under* negation, even though we would tend to expect it to have the opposite scope relation, at least if it behaved like *each*. We will discuss both (29a) and (32a) in Section 5.

At this point, we would like to comment on the significance of the data that we have been looking at for our general approach to quantifier scope. In Section 2, we observed that some QP-types support inverse scope construals, while others do not; in Section 3, we saw that only a subset of the former group of QP-types support inverse distributive scope construals (namely, DQPs). In this section, we have seen that even DQPs disallow any scope construal over negation unless they also distribute over an overt indefinite, which must itself scope over negation. We have also seen that universally quantified GQPs headed by *all*, which (unlike DQPs) cannot take inverse distributive scope over subject GQPs, apparently can take inverse nondistributive scope over negation. Such facts are virtually impossible to account for in terms of traditional treatments of quantifier scope, or, indeed, in terms of any theory that does not recognize distinctions among various QP-types in terms their scopal behavior. It is also interesting to note, inter alia, that in (31a, b) and (32b), the inverse scope construal of the object QPs relative to negation represent the *only* grammatical scope construals for these sentences (on the neutral intonation); this should come as a surprise to anyone who might still maintain that inverse scope construals are only marginally available, and that surface c-command relations are the basis of scope construals.

5 *EVERY* VERSUS *EACH*

5.1 Distributive *Each* and universal *Every*

In our introductory remarks, we mentioned that *each* has sometimes been characterized as a variant of *every*, which allows (or requires) a wide scope construal where *every* does not. Thus Fodor and Sag (1982) describe *each* as "a quantifier that favors wide scope." Based on our discussion thus far, it is evident that we are inclined to seek an account for the distinctive behavior of *each* that goes beyond the statement of a predisposition towards wide scope.

In fact, *each* and *every* exhibit a number of other differences, which collectively suggest that *every*, unlike *each*, can receive a non-distributive universal construal in certain configurations, behaving essentially like *all*. We believe that these differences are related to those discussed in Section 4 (ex. (32)) involving scope interactions between DQP objects and negation, where *each*-DQPs were well-behaved from the perspective of our theory, whereas *every*-DQPs seemed to behave more like GQPs headed by *all*.

As a point of departure, we point to two well-known differences between *each* and *every* that both indicate a more uniformly distributive character of *each*. First, *each*, unlike *every*, occurs in Quantifier Float constructions, which provide unambiguous distributive construals for sentences with GQP subjects, where a collective construal would otherwise be possible. In such cases, *each* arguably occupies the Spec of DistP position (cf. Sportiche 1988, Beghelli 1995). Second, *each*, but not *every*, occurs in Binominal *Each* constructions, which also have a strong distributive interpretation (cf. Safir and Stowell 1989, Beghelli 1995). Although we will not discuss either of these constructions here, the fact that they both occur with *each*, rather than with *every*, does tend to suggest that *each*, rather than *every* is the canonical distributive quantifier in English. To our knowledge, there is no distributive construction that makes the cut in the opposite way.

A third difference between *each* and *every* concerns collective universal construals of DQPs headed by *every* in examples such as the following:

(33) a. It took all the boys to lift the piano.
b. It took every boy to lift the piano.
c. * It took each boy to lift the piano.

Although DQPs headed by *every*, like those headed by *each*, normally force a distributive (non-collective) construal, as we saw above, this requirement seems to be relaxed in contexts such as that in (33). While we do not have an explanation to offer for why the requirement should be relaxed in this construction, the distinction between *each* and *every* that it reveals suggests that, in at least one context, *every* can serve as a non-distributive universal quantifier.

The fourth difference between *each* and *every* concerns modification by *almost.* This particle can qualify any quantifier or numeral designating a fixed quantity that is understood as the end point of a scale, including universal quantifiers like *every* and *all*; but it cannot combine with *each*:

(34) a. One boy ate almost twenty apples.
b. One boy has eaten almost nothing.
c. One boy ate almost all the apples.
d. One boy ate almost every apple.
e. * One boy ate almost each apple.

This suggests that *all* and *every*—but not *each*—can designate the end point of a scale, here the full set of apples. Note that the ungrammaticality of (34e) cannot be due to a failure of distributivity, since the DQP should be free to distribute over the indefinite subject.

A fifth difference concerns modification of universal and proportional quantifiers by the particle *not.* Whereas *not* can combine with a variety of proportional quantifiers, including *more/less (than) n, many,* or with *every* and *all*, it cannot combine with *each*:

(35) a. Not more than ten boys ate an ice-cream cone.
b. ? Not ten boys ate an ice-cream cone.
c. Not many boys ate an ice-cream cone.
d. Not all the boys ate an ice-cream cone.
e. Not every boy ate an ice-cream cone.
f. * Not each boy ate an ice-cream cone.

Although this test groups *every* with *all*, rather than with *each*, it is not obvious what underlying semantic property is being diagnosed here. (The marginal status of the bare numeral example in (35b) suggests that a proportional function of the quantifier may be relevant, but (35a) seems to have a non-proportional construal.) In any event, it seems reasonable to assume that *every* has a core function of pure universality that *each* lacks.

While none of the differences between *each* and *every* enumerated in this section provides the basis for a coherent analysis of either the syntax or the semantics of these two quantifiers, they all point towards the conclusion that *every* is fundamentally more like a canonical universal quantifier than *each* is, and conversely that *each* is fundamentally more like a pure distributive operator than *every.*

5.2 *Every* and unselective binding

A further difference between *each* and *every* pertains to the fact that *every*-DQPs can be construed generically, whereas *each*-DQPs cannot:

(36) a. Every dog has a tail.

b. Each dog has a tail.

Example (36a) can be construed as a claim about dogs in general, whereas (36b) must be construed as claim about a particular set of dogs previously mentioned in the discourse. In a similar vein, Gil (1992), citing the paradigm in (37)–(38), observes that *each*-DQPs pattern with definite GQPs (in our terms), whereas *every*-DQPs pattern with generically construed GQPs headed by *all*:

(37) After devoting the last three decades to a study of lexical semantics, George made a startling discovery.

a. Every language has over twenty color words.

b. All languages have over twenty color words.

c. ? Each language has over twenty color words.

d. ? The languages have over twenty color words.

(38) George has just discovered ten hitherto-unknown languages in the Papua New Guinea highlands.

a. ? Every language has over twenty color words.

b. ? All languages have over twenty color words.

c. Each language has over twenty color words.

d. The languages have over twenty color words.

Gil accounts for this by attributing to *each* a feature [+Definite], which *every* is supposed to lack: "while for *every*, the domain of quantification is free, for *each* it is contextually determined." (p. 20).

While this description of the contrasts in (37)–(38) seems to be more or less correct, it is not the case that *every*-DQPs must always be construed generically. Consider (39):

(39) Emma and Anna found lots of beautiful shells on the beach.

a. They examined each shell carefully.

b. They examined every shell carefully.

c. They examined all the shells carefully.

d. ? They examined all shells carefully.

Here, the *every*-DQP seems to be construed as definite, quantifying over a contextually determined set in just the same way as the *each*-DQP and the definite GQP *all the shells*, in contrast to the generically construed QP *all shells* in (39d). The same is true of all of the *every*-DQPs discussed in Sections 1-3. Thus, the generic construal of the *every*-DQP in (37a) and (38a) seems to be a function of the particular syntactic context in which it occurs, which imposes a generic construal on simple indefinites headed by *a* in much the same fashion:

(40) a. A man (usually) parts his hair on the left. (Generic)

b. Arby met a man at the conference. (Existential/Specific)

The variable interpretation of the indefinites and bare plurals in contexts such as (40) led Heim (1982) and Kratzer (1988) to conclude that indefinites and bare plurals function syntactically as (restricted) variables rather than as true QPs; these variables are supposed to be bound by external unselective quantifiers. The relevant quantifiers are a null generic (weakly universal) quantifier GEN taking clausal scope in (40a) and Heim's existential closure operator in (40b), to which Diesing (1988) assigns VP-level scope, and which we have analyzed as originating in the VP-internal *Event* argument position, and taking scope at the ShareP level.

If we now apply the same reasoning to the data in (36)–(39), we are led to the surprising conclusion that DQPs headed by *every* are variables, rather than true QPs. This *prima facie* surprising result is reminiscent of an observation due to Groenendijk and Stokhof (1993), who note quantificational variability effects with examples like the following:

(41) For the most part, John knows which book every student bought.

Here *every* seems to be interpreted more like *most* than like either *all* or *each*, suggesting, perhaps, that when *every* seems to behave like *each*, it may be exhibiting a similar type of unselective binding effect. Let us now consider how this might be possible, bearing in mind that we need to preserve the obvious fact that *every* is a kind of universal quantifier.

When *every*-DQPs occur in generic contexts, they are interpreted as though they were universal-generic QPs (just like indefinites in the same environments) because they contain restricted variables (ranging over sets) bound by a silent generic quantifier. The meaning that we want to assign to examples like (36a) *Every dog has a tail* under this analysis is thus something like 'in the default situation s where X is the set of all dogs in s, all members of X have a tail.' When *every* occurs in a context associated with reference to a single situation-time, it acquires its contextualized universal-distributive reading, presumably because it is bound by an analogous silent definite quantifier. Thus, a sentence like *Every boy lifted the piano* would be translated along the lines of 'there is a (particular) past situation s, a set X of all boys in s, such that all the members of X lifted the piano.'

Of course, this idea raises the issue of how a GQP headed by *every* can be analyzed as a universal variable. The theory presented in Section 2 allows us to account for this. We have assumed, with Szabolcsi (1996), that *every* and *each* introduce discourse referents, in the form of set variables. The set variable of *each*, we will now assume, must be bound by a definite operator—as required by its definiteness features, which we have reviewed above. On the other hand, the set variable introduced by *every* can be bound by other operators as well, including GEN.

On its normal (strongly) distributive use that shows up in non-generic past-tense contexts, *every* seems to be interpreted identically to *each*. At this point, one might ask exactly what kind of operator it is that licenses this canonical use of DQPs. The most obvious candidate is the existential quantifier over events. But this option is precluded for us, if this quantifier must appear in Spec of ShareP and the DQP headed by *every* or *each* must appear in the Spec of DistP. Another possible candidate is the silent (definite or indefinite) existential quantifier ranging over situation-times proposed by Stowell (1993). This quantifier is an existential counterpart of GEN; it is introduced as the internal argument of a Tense predicate heading the category TP.[15] We have not attempted to locate TP within the hierarchy of functional projections in (2), but it seems reasonable to suppose that it lies below AgrS and DistP. If so, it would be free to move to the Spec of RefP and act as the binder for *every, each.*

5.3 Scope interactions with negation, revisited

Now let us return to a consideration of the puzzling facts concerning object *every*-QPs in sentences containing clausal negation, repeated here:

(29) a. John didn't read every book. (NOT $>$ $\forall$)

b. ?? John didn't read each book.

(32) a. One boy didn't read every book. (NOT $>$ $\forall$)

b. One boy didn't read each book. (EACH $>$ ONE $>$ NOT)

[15] More precisely, according to Stowell's proposal, this existential quantifier originates as the Specifier of the time-denoting category ZP, which serves as the internal argument of a Tense predicate such as PAST. A tense predicate is a dyadic predicate of temporal ordering, which relates an event-time or situation-time (denoted by its internal argument) to a reference-time (denoted by its external argument).

Recall that the *each*-DQPs in these examples are well-behaved, from the perspective of our theory of distribution; it is the *every*-DQPs that are problematic. These *every*-DQPs should be required to move to the Spec of DistP, just like the *each*-DQPs are. (This requirement on *each*-DQPs is responsible for the scope construal in (32b), and for the ungrammaticality of (29b).)

Example (29a) is surprising because it shows that *every*-DQPs are not always required to move to the Spec of DistP above NegP; if they were, (29a) would be as odd as (29b). Example (32a) is even more surprising, because it shows that *every*-DQP objects are not just *allowed* to remain under NegP; here, they are actually *required* to do so. In (32a), the failure of the *every*-DQP to move above negation to the Spec of DistP cannot be attributed to the lack of an indefinite within the clause to satisfy the requirements of ShareP; evidently some other factor is at work here, inhibiting movement of the *every*-DQP to the Spec of DistP.

We would now like to relate these facts to some of the other properties of *every* discussed in this section. The essential idea is that *every*-QPs introduce a set variable, which gets bound by negation when the *every*-QP occurs in its scope. Developing this idea further, it seems plausible to assume that the set variable introduced by *every* must be bound by the closest potential binder available. Since negation is closer to it in the hierarchy of functional projections than the existential quantifier over times in the complement of Tense, it is the closest potential binder and will bind the set variable of *every.* We can then say that *every* fails to be interpreted as scoping over negation in (29a) and (32a) because the set variable that it introduces must be bound by the closest unselective quantifier it can find, and the NQP that ends up in the Spec of NegP serves this role. Thus, (29a) would receive an interpretation roughly along the following lines: 'there is no situation s and set X of (all) books in s, such that John read (every member of) X at s.'

A coherent picture is finally beginning to emerge: whereas *each* is a true distributive QP, *every* is not. Moreover, *every* exhibits some degree of quantificational variability, in the sense that its set variable can be bound by negative and generic operators. We have presented a possible account of this behavior, on the basis of the semantic justification of QP-types originally proposed by Szabolcsi (1996), cf. Section 2.7.

There are, however, two crucial facts about *each* and *every* that we still have to clarify. First, we must account for the fact that *each* is obligatorily distributive, whereas *every* is only optionally so. Second, if *every* is not inherently distributive, i.e., if it is really an unusual kind of universal, then we must explain why it differs from *all* in exhibiting Strong Distributivity in contexts such as those discussed in Section 3.

We would like to suggest that the solution to these problems lies in the featural specification of *every* and *each.* Both *every*- and *each*-QPs have access to Spec of DistP because they are singular, and this is a pre-condition for the distributive operator in $Dist^0$ to apply to them. On the other hand, *all* is plural, and hence does not have access to DistP.[16]

Each-QPs are endowed with a [+Distributive] feature, which *must* be checked in Spec of DistP; *every*-QPs, on the other hand, are underspecified for [Distributive]. Accordingly, *every*-QPs move to Spec of DistP only when their set variable is not bound by a lower operator, such as negation, which would then be the closer binder. When no negative operator intervenes, the set variable of *every* is bound by the existential quantifier over situation-times that has raised to Spec of RefP.

5.4 Concluding remarks

In this study, we have drawn attention to previously unrecognized scope interactions involving *each, every*, negation, and various types of indefinite QPs. We have suggested that these can most naturally be accounted for under the assumption that various quantifier types, such as DQPs and NQPs, are associated with fixed scope positions defined in the hierarchical pharase structure of the clause (DistP and NegP, respectively). We have also drawn distinctions among various types of (in)definite and numeral QPs (GQPs and CQPs), and proposed that these too have certain dedicated scope positions in the functional structure of the clause, though a greater amount of scopal freedom is allowed with these.

In addition, we have claimed that a number of otherwise puzzling differences between *each* and *every* can most readily be explained by extending to QPs headed by *every* the Heim-Kamp notion that NPs that have been traditionally considered purely quantificational in fact introduce variables ("discourse referents" in DRT parlance), and by assuming that such variables can be bound by certain external operators. This, we have argued, yields in some cases additional meanings and scope positions beyond the fixed ones that we have suggested at the outset.

REFERENCES

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics* 2(2):83–135.

Aoun, Joseph, and Audrey Li. 1989. Constituency and scope. *Linguistic Inquiry*

[16]The singular agreement property of every-QPs presumably forces distributive predication even when they do not move to Spec of DistP, but are bound by negative or generic operators. These however, would be cases of Pseudo-distributivity: i.e., we assume that a silent distributor is inserted; cf. Beghelli (1995).

20(2):141–172.

Aoun, Joseph, and Audrey Li. 1993. *The Syntax of Scope.* Cambridge: MIT Press.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In Amy J. Schafer (Ed.), *Proceedings of the North East Linguistic Society 23, Volume One*, 65–80. Amherst: GLSA Publications, University of Massachussetts.

Beghelli, Filippo. 1995. *The Phrase Structure of Quantifier Scope.* PhD thesis, University of California, Los Angeles.

Beghelli, Filippo. 1996. The syntax of distributivity and pair-list questions. In this volume.

Beghelli, Filippo, Dorit Ben-Shalom, and Anna Szabolcsi. 1996. Variation, distributivity, and the illusion of branching. In this volume.

Beghelli, Filippo, and Tim Stowell. 1994. The direction of quantifier movement. *GLOW Newsletter* 32:56–57.

Choe, Jae-Woong. 1987. *Anti-Quantifiers and a Theory of Distributivity.* PhD thesis, University of Massachusetts, Amherst.

Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (Ed.), *The Logic of Decision and Action*, 81–95. University of Pittsburgh Press.

Diesing, Molly. 1988. Bare plural subjects and the stage/individual level contrast. In Manfred Krifka (Ed.), *Genericity in Natural Language.* University of Tübingen.

Diesing, Molly. 1990. *The Syntactic Roots of Semantic Partition.* PhD thesis, University of Massachusetts, Amherst.

Diesing, Molly. 1992. *Indefinites.* Cambridge: MIT Press.

Fodor, Janet, and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5(3):355–398.

Gil, David. 1992. Scopal quantifiers: Some universals of lexical ef fability. In Michel Kefer and Johan van der Auwera (Eds.), *Meaning and Grammar: Cross-linguistic Perspectives*, 303–345. Mouton de Gruyter.

Groenendijk, Jeroen, and Martin Stokhof. 1993. Interrogatives and Adverbs of Quantification. In Katalin Bimbó and András Máté (Eds.), *Proceedings of the 4th Symposium on Logic and Language*, 1–29, Budapest. Akadémiai Kiadó.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* PhD thesis, University of Massachusetts, Amherst.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16(4):547–593.

Hirschbühler, Paul. 1982. VP-deletion and across-the-board quantifier scope. In James Pustejovsky and Peter Sells (Eds.), *Proceedings of NELS 12*, 132–139. GLSA Publications, University of Massachussetts, Amherst.

Hornstein, Norbert. 1995. *The Grammar of Logical Form from GB to Minimalism.* London: Blackwell.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Khalaily, Shamir. 1995. QR and the minimalist theory of syntax: The case of universally and negative quantified expressions in Palestinian Arabic. Unpublished manuscript, University of Leiden.

Kinyalolo, Kasangati. 1990. *Syntactic Dependencies and the Spec–Head Agreement Hypothesis in KiLega.* PhD thesis, University of California, Los Angeles.

Kratzer, Angelika. 1988. Stage–level and individual–level predicates. In Manfred Krifka (Ed.), *Genericity in Natural Language*, 247–284. University of T/"ubingen.

Kratzer, Angelika. 1995. Scope or pseudoscope? Are there wide-scope indefinites? In Susan Rothstein (Ed.), *Events and Grammar.* Dordrecht: Kluwer. Forthcoming.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch (Ed.), *Semantics and Contextual Expression*, 75–115. Dordrecht: Foris.

Kroch, Antony. 1979. *The Semantics of Scope in English.* Garland.

Liu, Feng-Hsi. 1990. *Scope and Dependency in English and in Chinese.* PhD thesis, University of California, Los Angeles.

Mahajan, Anoop. 1990. *The A/A' Distinction and Movement Theory.* PhD thesis, MIT.

May, Robert. 1977. *The Grammar of Quantification.* PhD thesis, MIT.

May, Robert. 1985. *Logical Form.* Cambridge: MIT Press.

Moritz, Luc, and Daniel Valois. 1994. LF pied–piping and Specifier–Head agreement. *Linguistic Inquiry* 25(4):667–707.

Reinhart, Tanya. 1995. Interface strategies. Unpublished manuscript, University of Utrecht.

Roberts, Craige. 1987. *Modal Subordination, Anaphora, and Distributivity.* PhD thesis, University of Massachusetts, Amherst.

Ruys, Eddy. 1993. *The Scope of Indefinites.* PhD thesis, Research Institute for Language and Speech, Utrecht.

Safir, Ken, and Tim Stowell. 1989. Binomial *each.* In *NELS 18, volume 2*, 426–450. GLSA Publications, University of Massachussetts, Amherst.

Schein, Barry. 1993. *Plurals and Events.* Cambridge: MIT Press.

Sportiche, Dominique. 1988. A Theory of Floating Quantifiers and Its Corollaries for Constituent Structure. *Linguistic Inquiry* 19(3):425–449.

Stowell, Tim. 1991. The alignment of arguments in Adjective Phrases. In Susan Rothstein (Ed.), *Perspectives on Phrase Structure: Heads and Licensing (Syntax and Semantics, Volume 25)*, 105–135. San Diego: Academic Press.

Stowell, Tim. 1993. The syntax of tense. Unpublished manuscript, UCLA.

Szabolcsi, Anna. 1994. Quantifiers in pair–list readings and the non–uniformity of quantification. In Paul Dekker and Martin Stokhof (Eds.), *Proceedings of the Ninth Amsterdam Colloquium*, 646–664. ILLC/Department of Philosophy, University of Amsterdam.

Szabolcsi, Anna. 1996. Strategies for scope taking. In this volume.

Zanuttini, Rafaella. 1991. *Syntactic Properties of Sentential Negation: A Comparative Study of Romance Languages.* PhD thesis, University of Pennsylvania.

4

STRATEGIES FOR SCOPE TAKING*

Anna Szabolcsi

Department of Linguistics
UCLA

1 OVERVIEW

Standard theories of scope are semantically blind. They employ a single logico-syntactic rule of scope assignment (quantifying in, Quantifier Raising, storage, or type change, etc.) which roughly speaking "prefixes" an expression α to a domain D and thereby assigns scope to it over D, irrespective of what α means, and irrespective of what operator β may occur in D:

(1) The semantically blind rule of scope assignment:

$$\alpha[_D \ldots \beta \ldots] \Rightarrow \alpha \text{ scopes over } \beta$$

There are two basic ways in which (1) turns out to be incorrect: the resulting interpretation may be incoherent, or the resulting interpretation may be coherent but not available for the string it is assigned to.

Szabolcsi and Zwarts (1993) focus on the first case. Take a version of (1) that is assumed to operate in surface syntax: *wh*-fronting. In a sizable class of cases, called "weak island violations," this rule yields unacceptable results. For instance:

(2) a. Who do you think that I mentioned this rumor to?
b. Who do you regret that I mentioned this rumor to?
c. Who didn't you mention this rumor to?

(3) a. How do you think that I solved this problem?

*This paper has been presented in talks and in class in Budapest, at MIT, at UCLA, at CSSP 95, and at the Tenth Amsterdam Colloquium; I thank the audiences of these presentations for comments. I am grateful to Michael Brody, Diana Cresti, Carmen Dobrovie-Sorin, Donka Farkas, Irene Heim, László Kálmán, Yoad Winter and, most of all, Dorit Ben-Shalom for discussion. This research was partially supported by NSF grant #SBR 9222501.

A. Szabolcsi (ed.), Ways of Scope Taking, 109–154.

b. * How do you regret that I solved this problem?

c. * How didn't you solve this problem?

(4) a. Who do you think that I got the ring I am wearing from?

b. * Who do you regret that I got the ring I am wearing from?

c. * Who didn't you get the ring that you are wearing from?

Szabolcsi and Zwarts submit that the violation is semantic in nature. *How* in (3b, c) and *who* in (4b, c) ought to scope over domains D that they are unable to. The reason is that manners and collectives are elements of proper join semi-lattices. Szabolcsi and Zwarts argue that the computation of the denotation of a factive context requires taking meets, and that of the negative context, complements. Since these operations are not defined in join semi-lattices, manners and collectives cannot scope over such contexts. For the moment, let it suffice that the $\alpha > \beta$ scope relation, pace (1), is not semantically unconstrained.

To illustrate the second case, which the present paper is concerned with, consider the fact that quantifiers in English often scope over operators that are higher in the surface syntactic hierarchy. These cases are attributed to the covert operation of (1). This account predicts that all quantifiers α interact uniformly with all operators β. But they do not. E.g., some but not all direct objects can scope over the subject (5), and some but not all direct objects can scope over negation (6):[1]

(5) a. Three referees read every abstract.
"every $N >$ three N"

b. Three referees read few abstracts.
*"few $N >$ three N"

(6) a. John didn't read many abstracts.
"many $N >$ not"

b. John didn't read few abstracts.
*"few $N >$ not"

It turns out that these contrasts have to do with semantics, too; however, they pertain to the syntax/semantics interface, rather than pure semantics. That is, the starred examples are not incoherent; simply, the given form cannot carry

[1]The scope interpretation that matches surface hierarchy often outshines the one that does not. Y. Winter (p.c.) suggests that in checking whether the latter, inverse reading is possible, it is useful to test examples where the primary reading is pragmatically dispreferred. This procedure lets real inverse readings shine without creating the false impression that all inverse readings are possible: some examples will just end up nonsensical.

the intended meaning. Proof is that the same α's are able to scope over the same β's in English when they are originally higher in syntactic structure (7) or when they acquire such a higher position via overt fronting (8):

(7) a. Few referees read three abstracts.
"few $N >$ three N"

b. Few women didn't like John.
"few $N >$ not"

(8) Few men did no one/every woman/two women like.
"few $N >$ no N/ every N/ two N"

Examples comparable to (8) are in fact standard in Hungarian, a language that disambiguates scope in surface structure (see below).

It does not seem desirable to develop a theory that maintains the omnivorous rule (1) and supplements it with a variety of filters on its overt or covert application. Such a strategy would simply not be explanatory. Instead, I argue for an approach that is as constructive as possible. This constructive methodology is in the same spirit as the combinatory categorial approach to syntax in Szabolcsi (1992) and references cited therein, although the results to be discussed in this chapter are entirely independent of categorial grammar.

The assumption is that "quantification" involves a variety of distinct, semantically conditioned processes. Each kind of expression participates in those processes that suit its particular semantic properties. Thus the heuristic principle is this:

(9) What range of quantifiers actually participates in a given process is suggestive of exactly what that process consists in.

Based on data in Liu (1990, 1992), proposals how to devise semantically conditioned specialized scopal mechanisms were first made in Ben-Shalom (1993) and Beghelli (1993). A both empirically and theoretically more fully developed version of the latter is Beghelli and Stowell (1994, 1996) and Beghelli (1995).

In this paper I first summarize those features of Ben-Shalom's semantic proposal that will be important in the core discussion. I proceed to reviewing certain aspects of Beghelli and Stowell's syntactic theory, and suggest that data from Hungarian, a language that "wears its LF on its sleeve," provide specific empirical support for them. Then I propose that Beghelli and Stowell's LF, especially in the light of some of the Hungarian data, can be quite directly mapped onto somewhat modified Kamp and Reyle (1993) style Discourse Representations.[2] The main concrete modification to be proposed pertains to

[2]Potentially, other dynamic theories could be used, too. Kamp and Reyle's is special in that it happens to include significant work on plurals, as opposed to Heim's (1982) File

widening the class of discourse referents. Finally, the Hungarian data will be shown to provide evidence that the denotational semantics of the noun phrase delimits, but does not determine, whether it introduces a discourse referent.

2 CONSTRUCTIVE APPROACHES TO DIFFERENTIAL SCOPE TAKING

2.1 Ben-Shalom (1993)

Ben-Shalom restricts her attention to a representative subset of the data in Liu (1990) that do not involve partitives.[3] Some features of her proposal that are directly relevant to the present paper are as follows. Consider (10) and (11):

(10) Three referees read every abstract.

(11) Three referees read fewer than five abstracts.

The standard way to calculate the object wide scope, O > S reading of (10) is to form the set of things read by three referees and check whether every abstract is in that set. But if the formation of this set, which is not the denotation of a surface syntactic constituent of the sentence, is a freely available option, then it can be used in calculating an O > S reading for (11), too. This is the standard assumption in the literature. However, (11) does not readily admit an O > S reading. This suggests that the O > S reading of (10) is not calculated in the above mentioned way, either. Rather, it must be calculated in some alternative way that is available when the intended wide scope quantifier is, say, *every abstract* but not when it is, say, *fewer than five abstracts.*

Ben-Shalom proposes that inverse scope is effected by a specific binary quantifier [O > S].

(12) If S and O are generalized quantifiers and R is the relation denoted by a transitive verb, the binary quantifier $[O > S]$ is defined to operate as follows:

For every $a \in A$, $S(R(a))$,
where A is some set determined by O.

Change semantics. The intuition my analysis is based on relies on the representational character of DRT; it remains to be seen whether DPL-style reincarnations of DRT would be equally suited to this purpose.

[3]Liu's generalizations are reviewed in Section 2.2 of Beghelli, Ben-Shalom, and Szabolcsi (1996).

$\lambda x[S(R(x))]$ is the property denoted by the subject+verb segment of the sentence; in the examples at hand, it is the property of being read by three referees. Informally, (12) says, "Grab a set A determined by the quantifier denoted by the object and check, for every element a of this set, whether it has the property that three referees read it." (The fact that Ben-Shalom formulates her proposal using a binary quantifier is immaterial for our present concerns, so it will not be dwelt on.)

Let us underline the procedural difference between the standard calculation of scope and the one encoded by $[O > S]$. The difference is twofold. On the standard account, we construct the set denoted by $\lambda x[S(R(x))]$ and let O operate on it. Using $[O > S]$, this set does not need to be constructed and O is not a predicate operator. Instead, O contributes a domain of entities, each of which is checked for the property $\lambda x[S(R(x))]$.[4]

The binary quantifier $[O > S]$ works most straightforwardly when O is a principal filter, because a principal filter determines a unique set, called its generator, within its restrictor. The unique set $[\![$*every man*$]\!]$ determines is the set of men; the unique set $[\![$*John and Bill*$]\!]$ determines is the set {john, bill}, etc. When O is just monotone increasing, it determines several suitable sets (in a big enough model), called its witnesses, so the operation of $[O > S]$ is less simple but still perfectly viable. But when O is monotone decreasing or non-monotonic, it does not determine any suitable set on its own. As is explained in detail in Chapter 1, the truth conditions of *Fewer than six men walk* or *Exactly six men walk* cannot be specified as "There is a set A consisting of fewer than/exactly six men such that each $a \in A$ walks." Hence $[O > S]$ is inapplicable to non-increasing quantifiers.

According to Ben-Shalom, $[O > S]$ captures the empirical facts correctly because the best inverse scope takers in English are indeed principal filters. In the discussion below I will consider a wider range of quantifiers in a wider range of contexts, and propose a somewhat similar account of them, exploiting the fact that the strategy "Grab a witness set and check its elements for property P" generalizes exactly to the increasing quantifiers.

The discussion of Beghelli and Stowell's proposal will make clear that, however insightful Ben-Shalom's proposal is, the overall picture of scope interaction is more complex than Liu's pioneering work suggested. Two important factors are (i) the need to factor out the contribution of distributivity and (ii) the fact that the possibility of inverse scope depends, not only on the choice of the wide

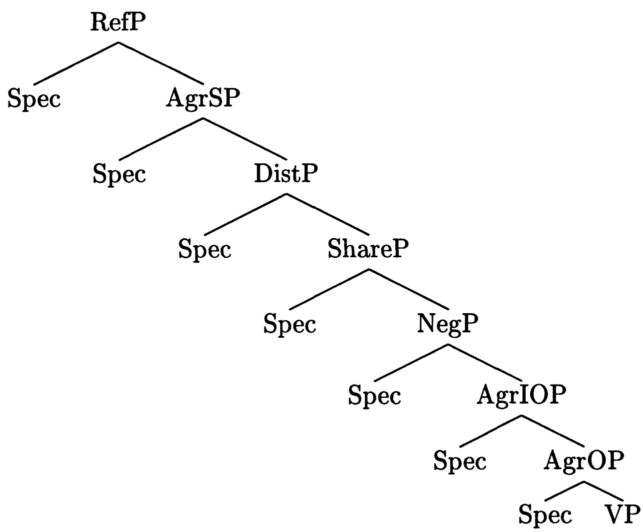
[4]It might be objected that checking whether an entity has property $\lambda x[S(R(x))]$ involves checking whether it is in the corresponding set, but this is not really so. To use a mathematical example, we may not be able to construct the set of prime numbers, but we may well be able to determine whether a given number is a prime, by examining what its divisors are. This example also reveals that the checking procedure may be intensional and/or invoke inferential processes. I thank Ed Keenan for discussion on this issue.

scope taker but, sometimes, also on the choice of the narrow scope taker. Thus the account requires a more complex set of assumptions.

2.2 Beghelli and Stowell (1994, 1996)[5]

Like Ben-Shalom, Beghelli and Stowell dispense with Quantifier Raising, an omnivorous movement rule without a specific landing site, and propose that Logical Form in English includes, among others, the following hierarchy of functional projections. Abbreviations: RefP = Referential Phrase, AgrSP = Subject Agreement Phrase, DistP = Distributive Phrase, ShareP = Distributed Share Phrase, NegP = Negative Phrase, AgrIOP = Indirect Object Agreement Phrase, AgrOP = Direct Object Agreement Phrase, VP = Verb Phrase.

(13)



Each type of quantifier acquires its scope by moving into the specifier of that functional projection which suits its semantic and/or morphological properties. When the sentence contains more than one quantifier that needs to land in a particular specifier, that position is filled multiply and its content undergoes absorption. Some important options are as follows.

Definites (*the two men*) move to the specifier of RefP, and distributive universals (*every man*) to the specifier of DistP. The head of DistP, a distributive operator, selects for a ShareP complement, which can accommodate either an indefinite (*two (of the) men*) or an existential quantifier over events. Indefinites may alternatively move to the specifier of RefP.

[5] See Stabler (1996) for a reformulation of Beghelli and Stowell's syntax in computationally preferable terms.

Modified numerals (*more than six men, fewer than six men, exactly six men,* and indefinites whose noun is destressed) do not move to either RefP, DistP, or ShareP. They just move to the appropriate agreement specifier positions to receive Case. The fact that modified numeral subjects easily take widest scope follows from the fact that AgrSP in English happens to be higher than DistP and ShareP. On the other hand, indirect and direct object modified numerals happen to have their agreement positions quite low in the structure, and they scope accordingly.[6]

Scope relations arise in two ways. They may simply follow from the hierarchy specified in (13). For instance, an indefinite direct object may scope above a universal subject by moving into RefP, which happens to be above DistP:

(14) a. Every man read two of the books

b. $[_{\text{RefP}}$ two of the books $[_{\text{DistP}}$ every man ...]]

Or, the inverse reading of *Two of the men read every book* comes about by moving *every book* to DistP and *two of the men* to ShareP.

Inverse scope may also be due to reconstruction: a phrase can be lowered into the position(s) of its trace, typically, into its VP-internal position.[7] The simplest assumption is that any kind of lowering is restricted to undoing semantically insignificant movement, i.e. an expression can be lowered from its Case position but not from RefP, DistP, or ShareP. For instance,

(15) a. More than three men read every book

b. $[_{\text{AgrSP}}$ more than three men$_1$ $[_{\text{DistP}}$ every book ... $[_{\text{VP}}$ ... t_1 ...]]]

The converse is not possible: *Every man read more than three books* does not receive an inverse scope interpretation. *Every man* cannot undo its presence in DistP and reconstruct into a VP-internal position below AgrOP:

(16) a. Every man read more than three books

b. $[_{\text{AgrSP}}$ t_1 $[_{\text{DistP}}$ every man$_1$ $[_{\text{ShareP}}$ $\exists e$ $[_{\text{AgrOP}}$ > 3 books $[_{\text{VP}}$ ... t_1 ...]]]]]

There is a slight difference between (16) and *More than three men read more than six books.*

[6]Definites, universals, and bare indefinites also pass through their own agreement positions for Case reasons. Since DistP and ShareP are lower than AgrSP, subjects must undergo some kind of lowering when targeting these positions. Various ways to execute this are discussed in Beghelli (1995).

[7]I base this part of the overview on Beghelli (1995), who considers the modified numerals data in greater detail than Beghelli and Stowell.

(17) a. More than three men read more than six books

b. [$_{\text{AgrSP}}$ > 3 men$_1$ [$_{\text{AgrOP}}$ > 6 books ... [$_{\text{VP}}$ ... t_1 ...]]]

Here inverse scope is very difficult but, in contrast to (16), can be forced by context. Since *more than three men* as a subject can in general reconstruct into its VP-internal position, this is predicted. (The marginality of reconstruction when the object is also a modified numeral calls for an independent account.)

Definites and bare indefinites do not move to DistP even when they are interpreted distributively; instead, their distributive interpretation is due to a silent operator comparable to floated *each*. Beghelli and Stowell call this "pseudo-distributivity." Silent *each* can apparently occur below AgrSP, ShareP, AgrIOP, and AgrOP, but not below RefP. This captures the fact that even when direct object *three books* moves to RefP and is therefore referentially independent of subject *two of the men,* it cannot make the latter referentially dependent, since there is no distributive operator between the two positions.

(18) a. Two of the men read three books

b. [$_{\text{RefP}}$ three books [$_{\text{AgrSP}}$ two of the men$_1$ [$_{\text{ShareP}}$ t_1 [...]]]]

On the other hand, in the structure below the property of having read three of the books can be distributed over *two men,* because the latter has a trace in AgrSP associated with silent *each*:

(19) a. Two men read three of the books

b. [$_{\text{RefP}}$ two men$_1$ [$_{\text{AgrSP}}$ t_1 EACH [$_{\text{ShareP}}$ three of the books [...]]]]

Similarly, the direct object in RefP can distribute over a subject that reconstructs into VP, because it has a trace in AgrOP and AgrOP may have silent *each* associated with it.

In sum, the distributivity of universals is due to a separate distributive operator (Dist) and, similarly, the distributivity of definites and bare indefinites is due to a separate distributive operator (silent *each*). Once the distributive key and the distributive operator are separated, they can move separately. This possibility is made crucial use of. *Every man* and *(the) two men* are allowed to move upward unboundedly to a higher RefP, but the corresponding distributive operators, being heads or adverbs, stay put. Thus it is predicted that (20) has a *de re* reading, where every woman or two particular women have the property of there being more than six men who think that the women will fall in love with them; but the men cannot vary with the women, as this property does not distribute:

(20) More than six men imagine that every woman/two women will fall in love with them.

The fact that Dist and *each* do not move up, together with the fact that the QP's landing site in the higher clause, RefP, is itself not associated with a distributive operator, underlies the traditional observation that "QR is clause-bounded."

3 CLAIMS TO BE MADE

Below I will examine Hungarian data in the light of Beghelli and Stowell and make the following main claims:

(21) Hungarian distinguishes scope positions in its surface syntax that are highly reminiscent of those postulated by Beghelli and Stowell for Logical Form in English.

(22) Some noun phrases can occur in only one of the above scope positions, but others can occur in more than one, and their interpretations vary accordingly.

(23) It is known that the presuppositional versus existential interpretation of noun phrases may be a function of their position. Hungarian is shown to exhibit similar positional distinctions in a new dimension, distributivity.

(24) Scope taking mechanisms fall into two broad categories. In the one case, the noun phrase introduces a "logical" subject of predication (not identical to a grammatical subject, i.e. a nominative). In the other, it performs a counting operation on an independently defined predicate denotation.

(25) The above distinction is not a purely denotational one, instead, it is representational/procedural. It is reminiscent of the basic insight of DRT. Introducing a logical subject of predication can be assimilated to introducing a discourse referent. Anaphora facts will motivate a revision of what items introduce discourse referents and the distinction of two kinds of referents: individuals (atomic or plural) and sets.

(26) In general, the logical forms Beghelli and Stowell derive for English sentences can be seen as direct instructions for constructing DRS's.

4 SCOPE POSITIONS IN HUNGARIAN

4.1 Hungarian surface structure disambiguates scope

Hungarian has come to be known as a language that "wears its LF on its sleeve." A substantial body of work by Hunyadi, Kenesei, É. Kiss, Szabolcsi, and others since the early eighties has established that surface order and intonation disambiguate scope.[8] For instance, the following sentences are unambiguous; the scopal order of quantifiers matches their left-to-right order.[9]

(27) a. Sok ember mindenkit felhívott.
many man everyone-acc up-called
'Many men phoned everyone'
= many men > everyone

b. Mindenkit sok ember felhívott.
everyone-acc many man up-called
'Many men phoned everyone'
= everyone > many men

(28) a. Hatnál több ember hívott fel mindenkit.
six-than more man called up everyone-acc
'More than six men phoned everyone'
= more than 6 men > everyone

b. Mindenkit hatnál több ember hívott fel.
everyone-acc six-than more man called up
'More than six men phoned everyone'
= everyone > more than 6 men

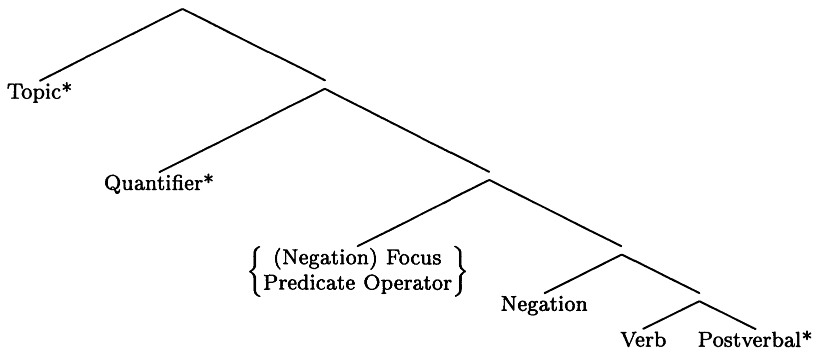
More precisely, it is their occurrence in specific syntactic positions that defines the quantifiers' scope. Simple syntactic tests distinguish the positions in (29), which I label with the pretheoretical names that have by now become more or less traditional; I coined the speaking name Predicate Operator for one subtype of what is traditionally called Focus. As usual, the * indicates that the given position may be filled multiply:[10]

[8] The Appendix will show that there are in fact significant exceptions in the postverbal field. But this does not affect the argument in the bulk of the paper, which pertains to preverbal DPs.

[9] For simplicity's sake, in this paper I will only consider cases in which the postverbal universal is unstressed. It is agreed, following É. Kiss (1987), that the alternative, heavy stressed option involves stylistic postposing in Phonetic Form.

[10] Topics are flatly intoned and not contrastive; contrastive topics (paraphrasable by "as for ... ") have a scooped intonation, must be followed by some operator, and are analyzed by É. Kiss (1987) as instances of Left Dislocation. In this paper I am not concerned with Left Dislocation, so even the position is omitted from the diagram.

(29)



The fact that left-to-right order determines scopal order follows from (30). For recent discussions, see É. Kiss (1991, 1994).

(30) In Hungarian, operators c-command their scope at S-structure (where c-command is defined in terms of first branching node).

Typically, a Hungarian sentence with n scope-bearing DPs will have n or $n-1$ in the preverbal field, so that their scopes are indeed disambiguated by surface order. The postverbal field is assumed to have a flat structure. It is rare but possible to have more than one scope-bearing DP postverbally; what their relative interpretation is is an interesting question which I will return to in the Appendix.

Some of the diagnostics of which position a DP occupies in the preverbal field are as follows:

(31)
a. Topics, but not other preverbal items, can be followed by sentential adverbials like *tegnap* 'yesterday.'
b. When a Topic or Quantifier precedes a non-negated finite verb that has a prefix, the prefix is in proclitic position.
c. When a Focus or Predicate Operator precedes a non-negated finite verb that has a prefix, the prefix occurs postverbally.[11]
d. A sequence of Quantifiers cannot be broken by a non-Quantifier.
e. A DP in Focus receives an exclusion-by-identification interpretation; a DP in Predicate Operator does not.

4.2 A parallelism with Beghelli and Stowell's LF

I argue that the extent to which Hungarian surface structure reveals the syntax of scope is even greater than has been thought. In general, it demonstrates that QPs are not simply lined up in the desired scopal order but occupy

[11]That is to say, the finite V moves into a functional head that is higher than the position of the prefix.

specific positions. And in particular, the traditionally distinguished positions correspond quite closely to the specifier positions of the functional categories in Beghelli and Stowell's (13). For the time being, I ignore the postverbal field.

(32) Hungarian

Topic	$\approx$	Spec, RefP
Quantifier	$\approx$	Spec, DistP
Focus (with indefs.)	$\approx$	Spec, ShareP
Predicate Operator	$\approx$	Spec, AgrP/VP

This parallelism is supported by data that pertain to (i) exactly what noun phrases occur in each position, and (ii) what kind of interpretation they receive there.

Some restrictions on the occurrence of DPs in these positions are well-known. E.g. a Topic must be specific, and universals do not occur in Topic or Focus (this latter fact was first observed in Szabolcsi 1980, p. 66). However, no systematic investigation of these matters has been carried out to date. In what follows I examine a representative sample. The data are summarized in (33) on the next page. Note that many DPs occur in more than one position; as we shall see, their interpretations vary accordingly.

Let us see how the distribution of DPs supports the parallelism in (32).

Proper names, definites, and those indefinites that take widest scope in their own clause are placed into [Spec, RefP] in Beghelli and Stowell. The Hungarian counterparts, when preverbal, occur in Topic.

Distributive universals are placed into [Spec, DistP] in Beghelli and Stowell. The Hungarian counterparts, when preverbal, occur in Quantifier position.

Bare indefinites that scope under distributive universals are placed into [Spec, ShareP] in Beghelli and Stowell. The Hungarian counterparts can occur in Focus with a comparable scope interpretation.

Modified numerals, which do not readily take inverse scope in English are placed into [Spec, AgrP] or [Spec, VP] in Beghelli and Stowell. The same holds for indefinites whose N is destressed and whose numeral is interpreted as "exactly *n*." The (relevant) Hungarian counterparts cannot occur higher than the Predicate Operator position.[12]

[12] If a constituent of DP is set into contrast, the whole DP is pied piped to Focus. This option is irrelevant to us and is not indicated in the table.

(33)

	Topic	Quantifier	Focus	PredOp	Post-V
a legtöbb fiú 'most of the boys'	+				+
valamely fiú/bizonyos fiúk 'some boy(s)'	+				+
Péter, Péter és Mária 'Peter,' 'P and M'	+		+		+
a fiú(k) 'the boy(s)'	+		+		+
hat fiú 'six boys'	+		+	+@@	+
sok fiú 'many boys'	+	+		+@@	+
minden fiú 'every boy'		+			+
valamennyi fiú 'each boy'		+			+
még Péter is 'even Peter'		+			+
hat fiú is 'even/as many as six boys'		+			+
Péter is 'Peter, too'		+			+
semelyik fiú (neg. concord) 'none of the boys'		+			+
legalább hat fiú 'at least six boys'		+			+
több, mint hat fiú 'more than six boys(1)'		+		+	+
hatnál több fiú 'more than six boys(2)'				+	+#
pontosan hat fiú 'exactly six boys'				+	+#
kevés fiú 'few boys'				+	+#
kevesebb, mint hat fiú				+	+#
hatnál kevesebb fiú 'less than six boys(1,2)'				+	+#
legfeljebb hat fiú 'at most six boys'				+	+#
fiú(k) 'boy(s), existential'				+	+#

@@ With the noun destressed

Only if PredOp/Focus is filled or V is negated

In view of the above data as well as in anticipation of the discussion below, it seems justified to refer at least to Hungarian Topic as (spec of) HRefP and Hungarian Quantifier as (spec of) HDistP. On the other hand, I will retain the labels Focus and PredOp since here, it seems, the pertinent similarities are functional and the residual differences are significant. ShareP, unlike Focus, does not host definites; PredOp, unlike AgrP, is not Case-related, etc.

Apart from the fact that scopal movement can be visible, the crucial respect in which Hungarian differs from English is that Hungarian has no agreement (Case) positions mixed with the scope positions in the preverbal field, whence scope relations are independent of the argument hierarchy. In the Appendix I outline an analysis of Hungarian sentence structure that, among other things, captures the observations above.

5 OUTLINE OF THE ANALYSIS

In what follows, I will focus on the positions HRefP, HDistP, and PredOp, and argue that their inhabitants contribute to the interpretation of the sentence as summarized in (34) through (36). (Focus is omitted, because it has an obvious additional semantic function that is irrelevant to the present concerns.) I formulate my claims with respect to Hungarian and will argue for them using Hungarian data, but recall that I believe that, modulo the obvious cross-linguistic differences, these data are supportive of Beghelli and Stowell's approach and my claims are intended to hold of their logical forms, too. In fact, some of these claims are incorporated into Beghelli and Stowell (1994, 1996).[13]

(34) DPs that occur both in HRefP and Focus, as well as *valamely/bizonyos N* 'some N(s)' that only occur in HRefP, contribute an individual to the interpretation of the sentence, i.e., an atomic or a plural individual (the atoms of) which correspond(s) to the element(s) of a minimal witness set of the DP.[14] This individual serves as a logical subject of predication. Predication may be distributive or collective, depending on the nature of the predicate.

(35) A DP in HDistP contributes a set to the interpretation of the sentence, i.e., a witness set. This set serves as a logical subject of predication mediated by a distributive operator.

(36) A DP in PredOp does not contribute an entity to the interpretation of the sentence and does not serve as a logical subject of predication. It performs a counting operation on the property denoted by the rest of the sentence. If that predicate is distributive and thus denotes a set, the DP

[13] *A legtöb fiú* 'most (of the) boys' and *fiú(k)* 'boy(s), existential' are not included in my three categories. Their analysis goes beyond the scope of this paper.

[14] A witness set of a generalized quantifier GQ is a set that is (i) an element of GQ, and (ii) a subset of the smallest set GQ lives on. E.g. a witness set of $[\![\textit{two men}]\!]$ is a set containing two men and no non-men. See Chapter 1 for discussion.

counts its elements. If that predicate is collective and thus has plural individuals in its denotation, the DP counts the atoms. The result of counting may even be compared to the cardinality of the common noun set, i.e. the DP's determiner need not be intersective.

The basic distinction that I wish to make is between DP denotations that contribute an entity as a target of predication and DP denotations that operate on the denotation of the predicate in the manner of generalized quantifiers. Such a distinction seems straightforward between names, definites and bare indefinites on the one hand and modified numerals on the other.[15] Distributive quantifiers might seem to side naturally with the latter group, but I claim they indeed side with the former and end up as one subspecies in the "subject of predication" category. This is what the proposals in (34) through (36) attempt to capture.

It seems to me that a natural framework for expressing these proposals is a version of the Discourse Representation Theory expounded in Kamp and Reyle (1993). The claim that some DPs serve as logical subjects of predication should translate as the claim that they introduce discourse referents. Following Kamp and Reyle (1993, p. 168), by "introduces a discourse referent" I mean that the rule processing the DP introduces a referent either into the universe of the very DRS to which it is applied or into the universe of a superordinate DRS. Thanks to such referents, these noun phrases support non-maximal reference anaphora. This contrasts with rules that take care of quantifiers; these place a discourse referent into a newly created subordinate DRS (introduce duplex conditions). These latter noun phrases only support maximal reference anaphora (constructing an antecedent for a subsequent pronoun involves the abstraction operation that intersects the denotations of the first and the second arguments of the determiner).[16]

Kamp and Reyle stipulate that when a DP "introduces a discourse referent" then, at the point of introduction, it is associated with all and only the conditions that come from material inside the DP. That is, even if a referent is introduced into a superordinate DRS, it will never be divorced from its DP-internal conditions. This needs to be stipulated, because Kamp and Reyle's discourse referents are plain variables ranging over the whole universe, and DP-internal conditions are represented as predicated of them. In contrast, in (34) and (35) I assume that a referent is a sorted variable that is ab ovo restricted to ranging over (plural individuals formed from minimal) witness sets

[15] The claim that HRefP serves as a logical subject of predication squares entirely with É. Kiss's (1992, 1994) analysis of Hungarian, although she makes no comparable claims about the other positions.

[16] The distinction between maximal and non-maximal reference anaphora is illustrated and examined in Problems (69)–(72) of Chapter 1.

of the generalized quantifier denoted by the DP. E.g., the discourse referent introduced by *two men* is a variable over plural individuals made up of two men. Since a witness set, by definition, is of the right "size" and contains only entities drawn from the determiner's restriction, the inseparability of the referent from the information that comes from the DP follows without further stipulation.

Note that this proposal differs from the usual notion of restricted quantification, which relies on the (smallest) set the GQ lives on, i.e. its common noun set, rather than a witness.

Kamp and Reyle's stipulation in fact takes care of a problem discussed in Abusch (1994) and Reinhart (1995). The example comes from Heim (1982): *If a cat likes a friend of mine, I always give it to her.* On the intended interpretation, a *friend of mine* is to be construed as having wide scope. But if only existential closure is outside the conditional and the predicate *friend of mine* is in the antecedent, the sentence will be incorrectly verified by any model where there is someone who is not a friend of mine. Abusch (1994) proposes a specific syntactic mechanism to percolate the predicate up to the quantifier. Reinhart (1995) invokes choice functions in the interpretation of indefinites. My own proposal is highly compatible with Reinhart's, given that the value of her choice function is exactly my witness set. Reinhart (1995) and Winter (1996) show how to obtain those choice functions compositionally; their procedure might be adopted by the present theory.

The behavior of DPs that occur in HRefP and Focus (the latter the functional counterpart of Beghelli and Stowell's ShareP) is straightforwardly derivable from the properties Kamp and Reyle attribute to set denoter referents (singular or plural individuals, in present terms). What DRT gains from Beghelli and Stowell, in turn, is a characterization of distributivity that is empirically more precise and less stipulative. Recall from 2.2 that silent *each* is claimed to behave much like its overt counterpart, whose behavior is governed by well-studied principles of syntax.

Let us assume, then, in general that the DRS construction algorithm does not take the simple phrase structures used in Kamp and Reyle as input but, rather, its operation is directly determined by the kind of Logical Form Beghelli and Stowell's analysis assigns to the sentence. This will have clear advantages in connection with the treatment of inverse scope. Kamp and Reyle comment on the fact that not all noun phrases can take inverse scope, but eventually they opt for the stipulation that a syntactically lower noun phrase may be processed before a syntactically higher one, which is equivalent to assuming an unconstrained QR. Beghelli and Stowell's theory eliminates QR and replaces it with an articulated syntactic theory of where each type of noun phrase ends up at LF. Their LF now specifies the correct orders in which to process noun phrases.

But there are reasons for more substantial modifications of DRT. These have to do with the behavior of DPs in HDistP, see (35), in comparison with those in PredOp, see (36). I will argue that the inhabitants of HDistP, universals among them, are construed as targets of (obligatorily distributive) predication. This claim will be supported by showing that (i) they support only distributive readings and (ii) they introduce discourse referents, although not exactly the same kind as inhabitants of HRefP. Only the inhabitants of PredOp, which are all "counters," operate on predicate denotations in the manner of generalized quantifiers.[17]

I believe that the picture that we are led to is a generalization of Ben-Shalom's (1993) insight. Recall from 2.1 that, restricting her attention to the calculation of inverse scope, Ben-Shalom argued that there is a procedural difference in the evaluation of sentences involving names, definites, specific indefinites, and universals on the one hand and those involving modified numerals on the other. In the former case, she proposes to start out with a set determined by the quantifier and check its members for some property. In the latter case, she proposes to directly tackle the predicate's denotation. In present terms, the difference is precisely that the former act as subjects of predication and the latter as predicate operators.

Pursuing the DRT analogy, these observations amount to adding a procedural flavor to DRT, in the following sense. DPs that introduce discourse referents do not only differ from others in how they support anaphora, which is largely a matter of logical syntax. They also differ at the interface between DRSs and the model theory, because the verification of the truth of sentences containing them is carried out using different procedures.

This procedural intuition may be reminiscent of Brentano and Marty's distinction between categorical versus thetic judgments, revived in Kuroda (1972), Sasse (1987), and Ladusaw (1994). At present I am not in a position to judge how far a deeper parallelism might go, but this issue certainly merits further investigation, since it may tie together formal and informal lines of research. (One obvious difference is that the present proposal is concerned strictly with the contribution of particular DPs, not with whole sentences/judgments.) Likewise, the "subject of predication" and the "predicate operator" types of verification procedures may be relevant in connection with the construction of mental models, in a sense similar to Webber (1979) and Crain and Hamburger

[17]To avoid misunderstanding, notice that I am using the notion of a generalized quantifier in two different senses in this paper: in a denotational sense and in a representational/procedural sense. From a denotational perspective all noun phrases denote generalized quantifiers (sets of predicate denotations). This remains true whatever further considerations are invoked; hence I am free to appeal to notions like witness sets and monotonicity. From a representational/procedural perspective, only a subset of the noun phrases operate directly on predicate denotations: those that do not introduce a referent (logical subject of predication).

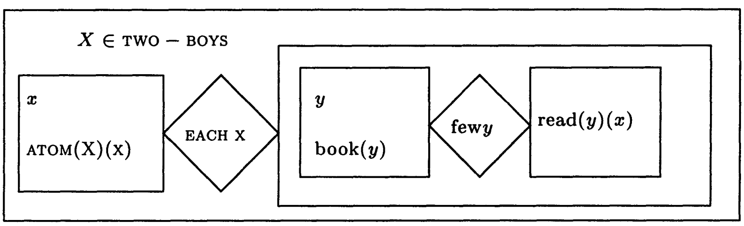
(1992). Finally, the two modes of operation recall the "look-up" versus "compute" distinction in Szabolcsi and Zwarts (1993). But developing a broader procedural theory that subsumes these goes beyond the scope of this paper.

In concrete terms, I will be arguing that the Beghelli and Stowell-style logical forms in (37) and (39) correspond to discourse representations as in (38) and (40), respectively.[18]

(38) is much like in Kamp and Reyle. The differences are (i) that X is now understood as a variable over plural individuals, not sets, and (ii) X is a restricted (sorted) variable. I will use the following notational convention: $X \in DP$ is a variable ranging over plural individuals whose atoms are the elements of some minimal witness set of $[\![DP]\!]$. I represent *few books* simply in terms of a duplex condition. Note that the cardinal and the proportional readings behave alike from the present perspective. EACH is Beghelli and Stowell's silent *each.*

(37) $[_{\text{RefP}}$ Two boys$_1$ $[_{\text{AgrSP}}$ t_1 EACH read$_2$ $[_{\text{AgrOP}}$ few books$_3$ $[_{\text{VP}}$ t_1 t_2 $t_3]]]]$

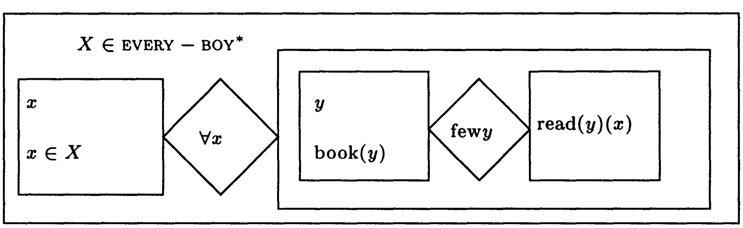
(38)



(40) involves *every boy* that, according to my proposal, introduces a set referent. Notation: $X \in DP^*$ is a variable ranging over witness sets of $[\![DP]\!]$, and $\forall$ is the distributive operator Dist.

(39) $[_{\text{AgrSP}}$ t_1 read$_2$ $[_{\text{DistP}}$ every boy$_1$ Dist $[_{\text{AgrOP}}$ few books$_3$ $[_{\text{VP}}$ t_1 t_2 $t_3]]]]$

(40)



[18]The explanation of why referents in HRefP are based on minimal witnesses while those in HDistP are plain witnesses is given in Section 8.3.

This replaces a "tripartite" structure in Kamp and Reyle.[19]

With these general considerations in mind, let us turn to the justification of (34) through (36), with reference to Hungarian.

6 DISTRIBUTIVE AND COLLECTIVE READINGS

6.1 Distributivity in HDistP

The reason why the Hungarian Quantifier position deserves the label HDistP is that all DPs occurring there are strictly distributive. (Although we get distributive readings elsewhere, too, as will be discussed below.)

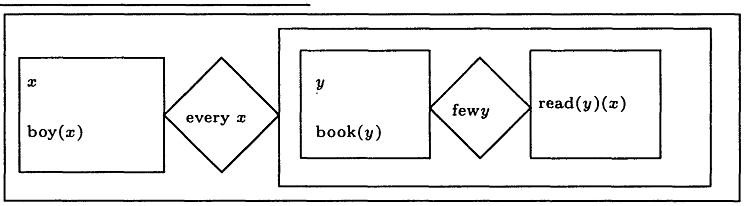
Some DPs occur only in HDistP and not in the other three distinguished positions. Universals, *minden fiú* 'every boy' and *valamennyi fiú* 'each boy' are the paradigmatic cases. But all *is* 'also, even' phrases are like universals in that they are barred from HRefP, Focus and PredOp.[20] For their distributivity, consider:

(41) Kati is fel-emelte az asztalt.
Kati also up-lifted the table-acc
'Kati lifted up the table, too'

This sentence cannot mean that along with others, Kati was a member of a collective that lifted up the table. It can only mean that Kati lifted the table on her own, and someone else did too.

(42) Hat fiú is fel-emelte az asztalt.
six boy even up-lifted the table-acc
'As many as six boys lifted up the table'

Here the contribution of *is* 'even' is essentially scalar: *hat* ... *is* means that six is considered many. Nevertheless, while the same sentence without *is* may well have a collective reading, (42) may only mean that there were as many as six individual table liftings.

19 

20 It may be interesting to mention that Hunyadi (1981) explains the identical surface distribution of *is* 'also, even' phrases and universals with reference to the fact that the morpheme *is* derives from the conjunction *és* and universals semantically reduce to conjunction. Similar relations have been in the focus of much recent work directed at Japanese and Korean.

But the most interesting new facts involve the observation that some noun phrases may occur in more than one position, and their interpretation varies accordingly.

Consider first telic predicates that can be either distributive or collective. (43) shows that names, definites and bare indefinites (the DPs that occur both in HRefP and in Focus) support either reading. DPs in HDistP do not support a collective reading at all. Finally, DPs in PredOp support an unmarked distributive reading of the sentence as well as a marked collective one, which has the flavor "It took as many/few as n boys to VP."

In the examples below the first DP is one that occurs only in the given position and the second is one that occurs in different positions with varying interpretations.

(43) a. Kati és Mari / Két fiú fel-emelte az asztalt. HRefP
'Kati and Mari / 'Two boys lifted up the table'
OK *lifting: collective*

b. Minden fiú / Több, mint hat fiú fel-emelte az asztalt. HDistP
'Every boy / 'More than six boys lifted up the table'
* *lifting: collective*

c. Kevesebb, mint hat fiú / Több, mint hat fiú emelte fel az asztalt. PredOp
'Less than six boys / 'More than six boys lifted up the table'
OK *lifting: "it took n"-collective*

Similar results are obtained with purely non-distributive telic predicates: "once only" predicates. Notice that here the distributive interpretation is out, no matter what the subject is: the same sand castle cannot be destroyed more than once (I mark this with #). See Szabolcsi and Zwarts (1993, Section 5) for some discussion.

(44) a. Kati és Mari le-rombolta a homokvárat. HRefP
'Kati and Mari tore down the sand castle'
OK *destruction: collective*
\# *destruction: distributive*

b. Minden fiú / Több, mint hat fiú   le-rombolta a homokvárat.   HDistP
'Every boy / 'More than six boys   tore down the sand castle'
* *destruction: collective*
\# *destruction: distributive*

c. Kevesebb, mint hat fiú / Több, mint hat fiú   rombolta le a homokvárat.   PredOp
'Less than six boys / 'More than six boys   tore down the sand castle'
OK *destruction: "it took n"-collective*
\# *destruction: distributive*

On the other hand, there are other non-distributive predicates like *surround* where even the "it took n" flavor is absent, and modified numerals in PredOp support an unmarked collective interpretation of the sentence. I suspect that this difference, which otherwise plays no role in my analysis and will not be investigated further, is due to the stativity of the predicate. (As for the choice of the verb, note that *surround* differs from *gather,* for instance, in that (i) if a plurality of entities surround something (in one layer), then no subset of them surrounds it, but (ii) a single entity may surround something by forming a full circle on its own.)

(45) a. Az X birtok és az Y birtok körül-öleli a kastélyt.   HRefP
'Estate X and estate Y surround the castle'
OK *surrounding: collective*
OK *surrounding: concentric circles*

b. Minden birtok / Több, mint hat birtok / Sok birtok   körül-öleli a kastélyt.   HDistP
'Every estate / 'More than six estates / 'Many estates   surround the castle'
* *surrounding: collective*
OK *surrounding: concentric circles*

c. Kevesebb, mint hat birtok
Több, mint hat birtok öleli körül a kastélyt. PredOp
Sok birtok
'Less than six estates
'More than six estates surround the castle'
'Many estates
OK *surrounding: collective*
OK *surrounding: concentric circles*

The behavior of DPs in Quantifier position fully supports the idea that this position is analogous to [Spec, DistP]. Not only do the Hungarian counterparts of *every boy* and *each boy* occur in this position, but a variety of further DPs do, too. And while the latter can support collective readings elsewhere, in this position they only support distributive readings.[21]

However, the following question presents itself: Do the collective or distributive readings arise in the same manner in all three positions?

6.2 Two types of collective readings: HRefP and PredOp

In the foregoing discussion I was careful to use a wording according to which a DP "supports a collective/distributive reading of the sentence." The reason is that I wished to remain entirely neutral as to what role this DP specifically plays in the formation of such a reading. I argue that in every one of the three positions that we are considering the DPs play a somewhat different role.

First consider the contrast between collective interpretations supported by DPs in HRefP versus DPs in PredOp:

(46) a. Ez a hat fiú fel-emelte az asztalt. HRefP
'These six boys lifted up the table (together)'

b. Ez a hat birtok körül-öleli a kastélyt.
'These six estates surround the castle (together)'

(47) a. Több/kevesebb, mint hat fiú emelte fel az asztalt. PredOp
'It took more/less than six boys to lift up the table (together)'

b. Több/kevesebb, mint hat birtok öleli körül a kastélyt.
'More/Less than six estates surround the castle (together)'

Following Kamp and Reyle (1993), I propose that in (46) the subject introduces a plural individual referent and 'lifted up the table' is predicated of it

[21] These data are clear counterexamples to Gil's (1995, p. 326) Universal 1: "If a quantifier is distributive-key, it is also universal."

collectively. More precisely, Kamp and Reyle treat bare indefinites as "set denoters," although they note that these sets are in one-to-one correspondence to plural individuals and the plural individual view is intuitively preferable. I am switching to plural individuals on the technical level, too, reserving the option of having set referents for another kind of DP.

In Kamp and Reyle's theory, collective predication is the only way to obtain a collective interpretation for the sentence, and in fact, they do not discuss convincing examples that would force one to think otherwise. But the examples in (47) are such. The subjects do not introduce a discourse referent either in a technical sense (see the anaphora facts below) or in an intuitive sense. The sentences in (47) are in no way "about" some boys or estates. Thus I claim that these sentences receive their collective interpretation in a different way. Namely, it is the predicate that denotes a group, as opposed to a set of individuals, and what the DP does is to count the atoms of this group. E.g.,

(48) 'The collective that surrounds the castle and consists of estates has more/less than six atoms'

Thus the sentences in (47) have a collective interpretation but their subject DPs are not interpreted collectively.[22]

So, in line with Kamp and Reyle, I assume that DPs in HRefP/Focus denote plural individuals that can be subjects of collective or distributive predication, while DPs in PredOp are counters. In distinction to Kamp and Reyle, however, I assume that the latter can count either the elements of a set, or the atoms of a group, whichever the predicate they operate on denotes. This takes care of (46) versus (47).

7 TWO TYPES OF DISCOURSE REFERENTS

In this section I discuss various aspects of (35), i.e. the claim that DPs in HDistP introduce a set referent.

[22]In English, some of the counting quantifiers have a variant that introduces a plural individual. This is claimed in Groenendijk and Stokhof (1984) and corroborated by S. Spellmire (p.c.). Thus, we have,

Some more/fewer than six men lifted the table [collectively].

The suspicion might arise that the English counterparts of the Hungarian examples only work with these variants (with the determiner *some* possibly "suppressed"). Notice, however, that *Few estates surround this castle* clearly differs in meaning from *A/*Some few estates surround this castle* and yet, is impeccable. Thus the phenomenon cannot be reduced to the subject introducing a plural referent. I should add that corresponding Hungarian DPs in PredOp do not allow for the plural construal at all.

7.1 No plural individual referent in HDistP

Let us turn to anaphora facts that establish whether a DP introduces a plural individual referent. In Kamp and Reyle, the most important mark of DPs that introduce a plural referent is that they can antecede a collective subject pronoun even when the latter is inside their own distributive predicate, see (50)–(52) below. Here is why this is the test case. In cross-sentential anaphora like *Many boys came. They were curious,* the pronoun constructs an antecedent for itself using the restrictor 'boy' and the predicate 'came.' But a pronoun located inside a predicate cannot use that same predicate in constructing an antecedent for itself. It can only corefer with a previously introduced discourse referent. And since we want a collective interpretation for the pronoun, the discourse referent it corefers with must be a plural individual, too.

It turns out that the Hungarian data are even easier to judge than the English. In Hungarian, DPs that contain a numeral are themselves in the singular and, alongside with conjunctions of singulars, trigger singular agreement on the predicate:

(49)	John és Bill Két ügyvéd Sok ügyvéd Hatnál több ügyvéd	titkárnőt vett fel/vett fel titkárnőt.
	'John and Bill two lawyer many lawyer more than six lawyer	secretary-acc hired{3sg}'

In cross-sentential anaphora, all these DPs antecede plural pronouns. When however they c-command a (possibly non-overt) pronoun, a singular pronoun receives a bound individual variable reading, while a plural pronoun receives a coreferential reading. Given this morphological distinction, all that needs to be judged in Hungarian is whether a DP can be linked to a plural pronoun in Kamp and Reyle's diagnostic context. For transparency, I replicate the Hungarian pronouns in the translations:

(50)	John és Bill Két ügyvéd	olyan titkárnőt vett fel, akivel előbb	elbeszélget-ett{3sg} elbeszélget-tek{3pl}
	'John and Bill two lawyers	hired a secretary that	he had interviewed they had interviewed'

If {3sg}, interview distributive;
if {3pl}, interview can (must?) be collective.

(51)	Minden Sok ügyvéd	olyan titkárnőt vett fel, akivel előbb	elbeszélget-ett{3sg} * elbeszélget-tek{3pl}
	'Every lawyer many lawyers	hired a secretary that	he had interviewed * they had interviewed'

If {3sg}, interview distributive;
if {3pl}, example *

(52)	Hatnál kevesebb Sok ügyvéd	vett fel olyan titkárnőt, akivel	elbeszélget-ett{3sg} * elbeszélget-tek{3pl}
	'Less than six lawyers many lawyers	hired a secretary that	he had interviewed * they had interviewed'

If {3sg}, interview distributive;
if {3pl}, example *

We see that the demarcation line lies exactly where Kamp and Reyle place it in English on the basis of judging the available interpretations. Only in the case of DPs that occur in HRefP/Focus can the plural pronoun be linked to the DP itself, cf. (50). In (51)–(52), with DPs that occur in HDistP and PredOp, respectively, the plural pronoun may at best pick up DP's smallest live-on set or be interpreted deictically.

7.2 Essential quantifiers and distributivity

The fact that DPs in HDistP are never linked to a plural pronoun in this context might suggest that they are interpreted in essentially the same way as those in PredOp, namely, as generalized quantifiers. The difference would consist in the first type having distributivity built into their definition.

This correlation is interesting, because Partee (1995, p. 564) conjectures (extending a claim in Gil 1989, 1995) that all essentially quantificational DPs are distributive. To make Partee's point perhaps even stronger, let me reinterpret "essentially quantificational" as those DPs whose determiner is not purely intersective and which cannot be taken to denote (atomic or plural) individuals, either. *Every N* and proportionals are essentially quantificational. Furthermore, non-individual denoting DPs whose restrictor is presupposed not to be empty are essentially quantificational. The reason is that a presupposition that pertains to only one argument of the determiner prevents the determiner from being symmetrical (and hence intersective).

In fact, Hungarian offers further subtle confirmation of Partee's hypothesis. Consider the PredOp data discussed in (47). If *több/kevesebb, mint hat N* is replaced by *az N-ek közül több/kevesebb, mint hat* 'more/fewer than six among the Ns,' the closest we can get to a partitive in Hungarian, the collective readings disappear.

(53) a. Több/kevesebb, mint hat fiú emelte fel az asztalt.
OK 'It took more/less than six boys to lift up the table (together)'

b. A fiúk közül több/kevesebb, mint hat emelte fel az asztalt.
'More/fewer than six among the boys lifted up the table, individually'

Similarly, if we have *sok* 'many' or *kevés* 'few' in PredOp and they are interpreted proportionally, the collective readings disappear. We may say that both changes result in essentially quantificational DPs.

Now, it is possible to maintain that all DPs in HDistP are essentially quantificational in this slightly modified sense. Recall what we have here: *every N, many N, at least/more than n N,* and *also, even* phrases. Crucially, it is not counter-intuitive to say that when *több, mint hat fiú* 'more than five boys' occurs in HDistP, we presuppose that there are boys. Maybe we are even thinking of boys drawn from a known superset of individuals, that is, the phrase may be specific in Enç's (1991) sense.

If all DPs in HDistP have semantic properties that make them essentially quantificational, then the fact that they are invariably distributive may simply follow from Partee's generalization.

7.3 Set referents in HDistP

It seems now that both the anaphora facts and the distributivity facts concerning HDistP correlate with the inhabitant DPs being essentially quantificational. If essentially quantificational DPs are automatically to be analyzed as having a "tripartite" structure, then such an analysis seems very well motivated. I submit, however, that there are other facts that receive a natural explanation if we assume that these DPs introduce a discourse referent of some sort, and the same facts remain mysterious on the "tripartite" analysis.

The Hungarian data are critical in developing this argument. The reason is that the diagnostics of introducing a discourse referent have to do with non-maximal reference anaphora and referential variation. According to Beghelli and Stowell, in English only universals reside in DistP. But a universal has a unique witness that is identical to its restrictor (= smallest live-on) set. Therefore, maximal reference anaphora (computed by intersecting the restrictor and the predicate sets) and non-maximal anaphora to some witness set come out the same. Likewise, universals will not exhibit referential variation, however they may be entered in the DRS. Therefore, the behavior of universals is compatible with more than one analysis. To see what properties the syntactic position *per se* has, we would need to test non-maximal anaphora on a DP with witnesses distinct from the restrictor, and referential variation on a DP with more than one witness. In Hungarian, DPs like 'many men' and 'more than five men' occur in the same HDistP position as 'every man,' thus the relevant tests can be performed. Furthermore, since the same DPs occur in PredOp, too, minimal pairs can be formed to isolate the properties present only in HDistP.

It should be clear that my factual claims below concern the behavior of Hungarian DPs, and it is for students of English to decide whether *many men* and *more than five men* exhibit similar behavior. Now two questions arise. Is it possible at all for me not to predict that these English DPs behave analogously? It is, because I show in Section 9 that denotational semantics delimits, but does not determine, a DP's actual mode(s) of operation. Hence the fact that a Hungarian DP is denotationally equivalent to some English DP does not entail that they operate identically. But what is the crosslinguistic significance of the Hungarian facts then? Since I have argued for a global analogy between HDistP and English DistP on the one hand and PredOp and English AgrP on the other, the Hungarian data may offer an insight into the way DPs in these positions operate, even if the items that occur in those positions are not exactly the same.

Consider first the following contrast in the behavior of *több, mint hat diákunk* 'more than six students of ours' in HDistP versus PredOp, with respect to a variant of the "others" test, cf. Problems (69)–(70) in Chapter 1. Imagine two teachers in the process of correcting the exams of a large class. When they are done with some of the exams, the exchange in (54a) is felicitous, while the one in (54b) is not.

(54) a. Több, mint hat diákunk félreértette a kérdést.
Lehet, hogy még másokat is találsz.
'More than six of our students (**HDistP**) misunderstood the question.
Maybe you will find others, too'

b. Több, mint hat diákunk értette félre a kérdést .
* Lehet, hogy még másokat is találsz.
'More than six of our students (**PredOp**) misunderstood the question.
* Maybe you will find others, too'

When 'more than six of our students' is in HDistP, as in (54a), the dialog is perfectly coherent. The first teacher's remark is unambiguously about a particular set of more than six students. The second teacher's remark means that there may be students outside this set who also misunderstood the question. In contrast, when 'more than six of our students' is in PredOp, the first teacher's remark can only mean that the number of students who misunderstood the question is greater than six. This cannot be followed by a remark about the "others." To begin with, this interpretation does not present a set of individuals in comparison with whom certain individuals may be "others." Moreover, however the exams yet to be corrected will turn out, they will not change the fact that the overall number of those who misunderstood the question is greater than six.

I conclude that the DP in HDistP introduces a set that is salient enough for anaphora to build on. This set is a witness of the generalized quantifier denoted by the DP. But a DP in PredOp crucially does not support this kind of anaphora, because it does not talk about individuals at all.

The details of the interpretations of the complement subjects below point to the same conclusion quite unambiguously:

(55) a. Legalább két elemző úgy gondolja, hogy több, mint hat hazug igazat mond.
'At least two analysts think that more than six liars (**HDistP**) are truthful'

b. Legalább két elemző úgy gondolja, hogy több, mint hat hazug mond igazat.
'At least two analysts think that more than six liars (**PredOp**) are truthful'

Farkas (1996) argues that the descriptive content (DC) of any noun phrase may be evaluated with respect to the worlds introduced by superordinate clauses; in the present case, this entails that whatever determiner the complement subject might have, the entities referred to may be liars in the speaker's world, not in the analysts' worlds. This in fact does not follow from the present proposal and thus, if correct, the mechanism Farkas proposes needs to be incorporated. On the other hand, there is a difference between the possible interpretations of (55a,b) that goes beyond what the evaluation of the DC explains.

Namely, (55a) can mean that there is a fixed set of more than six liars such that a fixed set of at least two analysts think that they are truthful. That is, on this reading the liars and the analysts are chosen independently. In contradistinction to this, in (55b) it may at best be a coincidence if the liars the analysts think to be truthful are identical; there is no reading that guarantees it. This difference between (55a) and (55b) follows straightforwardly if we assume that the DP in HDistP introduces a referent corresponding to a witness (a set of more than six liars), but the DP in PredOp merely counts how many liars each analyst thinks are truthful. The fact that the liars can be chosen independently of the analysts in (55a) follows from the assumptions concerning discourse referents: they may be introduced into either the current DRS box or into any superordinate box. And the fact that the analysts nevertheless do not become dependent follows from the fact that the distributive operator invariably gets stuck in its base position. (These square with other proposals that Farkas makes.) No mechanism with a comparable effect is available to DPs that do not introduce a referent, cf. (55b).

With these, I take it to be established that DPs in HDistP, in distinction to PredOp, introduce discourse referents.

We are now faced with the residual question of why, then, these DPs fail to support anaphora in (51). We may stipulate that coreference in the strict sense involves a relation between a pronoun and an expression denoting an individual, atomic or plural. Then one (natural) difference between bare indefinites like *hat fiú* 'six boys' and inhabitants of HDistP is that the referent that the former introduces is an individual but the referent that the latter introduces is a set. As was noted above, such a distinction can be accommodated in Kamp and Reyle's framework with a minimal modification.

This stipulation may be beneficial in explaining why, according to Beghelli and Stowell, bare indefinites never move to [Spec, DistP] and thus need to receive their distributive interpretation in a different way. We may correlate the feature that is checked in DistP with introducing a set, not an individual, referent.

How should universals in DistP and HDistP be analyzed, then? Recall that because they denote principal filters, they conform happily to both the referent and the tripartite analyses. By default, we want to treat them in the same way as the other, more discriminating inhabitants of the same syntactic position, i.e., using discourse referents.

It turns out that this analysis is the only one compatible with Stowell and Beghelli's independent claims. In general, they argue that distributivity is a separate factor even in the case of universals; what remains, then, is a set. More specifically, they discuss the following two types of data:

(56) John didn't read every book.

(57) What did every boy read?

The notable property of (56) is that, on normal intonation, it only allows a reading where *not* takes scope over *every book.* The notable property of (57) is that it has a pair-list reading. Beghelli and Stowell (1996) and Beghelli (1996) analyze both cases by assuming that the universal acts as a variable bound by some operator (the negation or the question operator). Details aside, this would make no sense on the usual interpretation of universals, but it makes good sense if the universal introduces a set referent, since that is a bindable variable in DRT terms. [23]

[23]Incidentally, the result that universals may be bound is not unique to this analysis; dynamic semantics can produce the same, as observed by Groenendijk and Stokhof (1993).

8 THE SUBJECT OF PREDICATION MODE OF OPERATION

8.1 Grab a witness and predicate distributively

Let us now see what the proposed analysis really is.

There is a sharp intuitive difference between Hungarian sentences that have HDistP or PredOp filled, even when there is no truth conditional difference. DPs that occur in both positions are especially instructive in this regard.

(58) Tegnap sok diákunk meg-betegedett. HDistP
yesterday many student-1pl pfx-sickened
'There is a set of many students of ours such that each fell ill yesterday'

(59) Tegnap sok diákunk betegedett meg. PredOp
yesterday many student-1pl sickened pfx
'The students of ours who fell ill yesterday were many'

The examples are chosen in such a way that, due to the possessive construction, they are both "presuppositional" and the 'many' phrases are interpretable as proportional in both cases. If this is so, then there is no standardly known reason for the sentences in (58) and (59) to be perceived as not meaning the same. But that is the perception; no native speaker would be tempted to say otherwise, even though they might not be able to explicate the difference. This is something to account for.

My account is that in (58) we take a set of students and claim that each of them fell ill. In (59), we take those who fell ill, and count our students among them.

The semantic analysis of HDistP that I am advocating is a generalization of Ben-Shalom's (1993) proposal for inverse scope and Chierchia's (1993) proposal for pair-list readings, which is based on Groenendijk and Stokhof's (1984). As was reviewed above, Ben-Shalom assumes that inverse scope is effected by a binary quantifier whose working can be illustrated as follows:

(60) a. Three referees read every/two abstract(s)
b. for every $x \in A$, three referees read x
where A is a witness set of the quantifier *every/two abstract(s)*

Chierchia assumes that pair-list readings are effected by a binary quantifier whose working can for present purposes be simplified as follows:[24]

[24] In Szabolcsi (1996a) I argue against using (61) as the general representation of pair-list readings, because it does not fit the full range of quantifiers that support pair-list; but here I appeal to (61) for an insight to be applied to a crucially restricted set of examples. See specifically Sections 3.1 and 5 in Szabolcsi (1996a).

(61) a. What did every/two boy(s) read?

b. for every $x \in A$, what did x read
where A is a witness set of the quantifier *every/two boy(s)*

That is, in both cases the quantifier that takes inverse scope or induces a pair-list reading is said to contribute a set to the interpretation of the sentence, associated with a separate distributive operation "every $x \in A$." These authors assume that this behavior of the quantifier is "unusual:" it obtains specifically in the inverse scope or pair-list context. My proposal differs from theirs in that I am assuming that offering up a witness to distributive predication is how quantifiers in HDistP always operate.

To illustrate with an English example, I am assuming that *Every referee read three abstracts,* on its direct S > O reading is also calculated in the manner of (62b), rather than (62c); whether (62b) is thought to involve a binary quantifier is immaterial:

(62) a. Every referee read three abstracts

b. for every $x \in A$, x read three abstracts
where A is a (=the) witness set of the quantifier *every referee*

c. every(referee)(read three abstracts)

It is worth emphasizing that the word "every" in (62b) stands for the distributive operator and in (62c) for the actual determiner. Thus the following Hungarian example makes the contrast more transparent, perhaps:

(63) a. Több, mint hat fiú el-ment. HDistP
more than six boy away-went

b. for every $x \in A$, x left
where A is a witness of 'more than six boys'

c. more-than-six(boy)(left)

8.2 The increasingness constraint

At this point it is crucial to go back to the data in (33) and observe a peculiar fact about the distribution of DPs:[25]

(64) Both HRefP and HDistP accommodate only increasing quantifiers. All decreasing and non-monotonic quantifiers are confined to PredOp.

[25]HDistP accommodates ***semelyik fiú*** 'none of the boys' and ***Péter sem*** 'Peter either,' which seem to contradict the increasingness claim. But Szabolcsi (1981) argued that ***semelyik fiú*** is just the negative concord form of ***minden fiú*** 'every boy;' similar claims have been made about negative concord in Italian by Haegeman and Zanuttini (1990). Similarly, ***Péter sem*** is the negative concord form of ***Péter is*** 'Peter also.' So these are not counterexamples. All genuinely decreasing quantifiers, as well as the non-monotonic ones, occur in PredOp.

This fact calls for an explanation. What kind of an explanation shall it be? Recall the heuristic formulated in (9) and used in various chapters of this book:

(65) What range of quantifiers actually participates in a given process is suggestive of exactly what that process consists in.

In the light of (65), (64) suggests that DPs in both HRefP and HDistP are interpreted in a way that is only available to increasing quantifiers. My analysis above has exactly this property. DPs in both HRefP and HDistP have been argued to put up a witness as a logical subject of predication, and this is possible only when the DP is increasing. Consider the following fact (see Section 2.1 as well as Chapter 1):[26]

(66) If Det is increasing, but not if it is decreasing or non-monotonic,
$\textsc{det}(N)(P) = \exists A$, A a witness of $\textsc{det}(N)$, $\quad \forall x \in A, \quad Px$

The left hand side is the standard (generalized quantifier theoretic, or "tripartite") specification of the truth conditions. The right hand side is the analysis I am proposing. (66) says that the proposed analysis yields the correct truth conditions if and only if the quantifier is increasing. In the spirit of (65), the analysis predicts the increasingness constraint.

On the other hand, the standard GQ theoretic or "tripartite" analysis of the inhabitants of HDistP would yield logically correct results for all quantifiers. Hence the assumption that DPs in HDistP operate in that manner would not be able to explain the constraint. It would predict that the inhabitants of HDistP are as heterogeneous as those of PredOp.[27]

8.3 Witnesses and minimal witnesses

Recall from (38) and (40) that referents in (H)RefP are claimed to be based on minimal witnesses, but referents in (H)DistP on plain, not necessarily minimal, witnesses. This choice has to do with two factors: collective readings and anaphora.[28]

Consider first (43a), 'Two boys lifted up the table.' A witness set of $[\![two\ boys]\!]$ is any set that contains two boys and no non-boys. It may therefore be a set that contains, say, four boys. But if the table was lifted up by a collective

[26]Logically speaking, Det also needs to be conservative and have extension, but all natural language determiners are thought to have these properties, so they will not discriminate between potential empirical cases.

[27]It may be possible to give a pragmatic account of the facts behind (64), as is suggested in Kadmon (1987). I believe, however, that such an account would involve developing a major theory that shifts the borderline between semantics and pragmatics in a fundamental way. As no one to my knowledge has laid out such a theory, for the time being its benefits cannot be taken for granted.

[28]I thank Y. Winter for discussion on these matters.

of four boys, then (43a) is not true. Similarly, if the example contained a disjunction, 'John or Bill lifted up the table,' a witness set of ⟦*John or Bill*⟧ would be {j, b}—but the sentence would be false in a situation where the collective comprising both John and Bill did the lifting. Thus for collective readings we need plural individuals based on minimal witnesses: just two boys in the first case, just John or just Bill in the second.

When the same DPs participate in distributive readings, the choice between minimal and non-minimal witnesses does not make a truth-conditional difference, because the quantifiers in (H)RefP are all monotonically increasing: 'there is a set of just two boys each of whom is tall' allows for there being a larger set with four tall boys and is therefore the same as 'there is a set of at least two boys each of whom is tall.' But anaphora facts confirm that the referent introduced by *two boys* is one with just two boys.

(67) Two boys came in. They were tired.

While the first sentence is compatible with four boys coming in, the pronoun in the second appears to refer to just those two boys that we singled out. In sum, it is justified to assume that referents introduced in (H)RefP are plural individuals based on minimal witnesses of the quantifier, irrespective of whether they are subjects of distributive or collective predication.

The situation is different in HDistP. Here the anaphora facts alone are decisive. Quantifiers in HDistP are always subjects of distributive predication and they are all monotonically increasing. Hence it makes no truth-conditional difference whether we operate with minimal or non-minimal witnesses. But consider anaphora. The critical example is (54a).

(54)a. Több, mint hat diákunk félreértette a kérdést.
Lehet, hogy még másokat is találsz.
'More than six of our students (**HDistP**) misunderstood the question.
Maybe you will find others, too'

Recall that here *mások* 'others' was claimed to refer to students who fall outside a particular set. Now, a minimal witness of ⟦*more than six students*⟧ has exactly seven students. The question is, are we forced to construe the first sentence to be about exactly seven students? No. This discourse is just as fine if the actual number of the students talked about is eight or nine. But then the referent introduced in HDistP must be any witness, not a minimal witness, of the quantifier.

8.4 Essential quantification, again

In Section 7.2 I pointed out that the obligatorily distributive interpretation of DPs in HDistP falls under a slightly modified version of Partee's (1995) generalization. Namely, all inhabitants of HDistP are essentially quantificational in the sense that they do not denote (singular or plural) individuals and their determiners are non-intersective (universal, or proportional, or at least presuppositional). Partee conjectures that all essentially quantificational DPs are distributive.

On the present account, inhabitants of HDistP introduce a set referent and are associated with a distributive operator, the head of the functional projection. This account is weaker than one based on Partee's generalization might be, since distributivity is not linked to any other semantic property of the noun phrase. On the other hand, Partee's generalization is a descriptive, not a theoretical one; for the time being, it is not known why the entailment might hold. Note also that even if essential quantifiers are all distributive, not all distributive quantifiers are essentially quantificational. Not only do we have distributive readings for sentences with *hat fiú* 'six boys' that denotes a plural individual, but distributive readings with purely cardinal *sok fiú* 'many boys' and *hatnál több fiú* 'more than six boys' in PredOp are also impeccable. Furthermore, *a legtöbb fiú* 'most of the boys' is an inherently proportional and (in my judgment) invariably distributive quantifier in Hungarian, but it resides in HRefP and not in HDistP. That is to say, distributive readings plainly cut across the positions HRefP, HDistP, and PredOp. My conclusion is that the correlation between distributivity and certain semantic properties is an open question for the time being; it is to be hoped that its explanation will shed more light on the nature of (H)DistP as well.

What remains to be accounted for on my analysis is the observation, made in Section 7.2, that DPs in HRefP and HDistP are presuppositional in some sense. As Ben-Shalom (p.c.) points out, this may follow from the fact that if there is no non-empty witness to serve as the subject of predication, predication will not be just false but will not even take place.

In fact, this reasoning prompts us to modify the usual assumption concerning exactly what is presupposed in presuppositional DPs. The usual assumption is that the determiner's restrictor is presupposed to be non-empty. But while this assumption may be sufficient to explain the absence of presuppositional DPs from existential contexts, it does not seem sufficient to do justice to the felicity conditions of the pertinent sentences. Consider the following in the context "In the history of the Vatican, ... ":

(68)	Hat lengyel pápa Több, mint öt lengyel pápa	könyvet írt.
	six Polish Popes (**HRefP**) more than five Polish Popes (**HDistP**)	book-acc wrote

These examples do not seem more felicitous in 1995, when the restrictor (the set of Polish Popes) is non-empty than they would have been fifty years earlier.

When the above DPs operate in the subject of predication mode, they appear to presuppose that at least six Polish Popes have existed in history (who then may or may not have written books). That is, exactly as the present analysis predicts, it seems that the existence of a non-empty witness, and not that of a non-empty restrictor, is presupposed.[29]

9 THE ROLE OF DENOTATIONAL SEMANTIC PROPERTIES: IMPORTANT BUT LIMITED

Both classical DRT and my modified version of it propose a non-uniform treatment of noun phrases: some are said to introduce discourse referents and others to operate on predicate denotations. An obvious question to ask is to what extent the denotational semantic properties of each noun phrase determine in what mode it will operate.

I argued above that there is at least one crucial respect in which denotational semantics plays a delimiting role: unless an explicit maximality condition is added, only monotonically increasing quantifiers allow for the paraphrase 'There exists a set or plural individual such that...' Thus only increasing quantifiers can have discourse referents corresponding to them. And indeed, it was observed that HRefP and HDistP accommodate only increasing quantifiers. Below, I will point out a somewhat similar constraint in connection with PredOp.

It would be very interesting, then, to be able to show that a DP's mode of operation is fully determined by its denotational semantic properties. Unfortunately, this does not seem possible. In fact, even at the present stage of the research, the Hungarian data seem to indicate, quite unambiguously, that the enterprise is hopeless. In other words, parallel to the fact that the difference between the proposed modes of operation is not purely denotational, the

[29] It may be observed that Diesing (1992) proposes to account for a somewhat similar intuition concerning the specific versus non-specific interpretations of bare and modified indefinites. Apart from the interpretation of presuppositionality, some of the crucial respects in which her proposal differs from the one developed here are as follows. (i) She assimilates specific (presuppositional) indefinites to restricted quantifiers and (ii) she assumes that non-specific indefinites always introduce variables captured by an existential closure operator.

Many of the observations motivating my analysis can be seen as reasons for rejecting Diesing's. Ad (i), treating specific indefinites as quantificational prevents her theory from accounting for the data that motivate Kamp and Reyle to assume that these DPs introduce plural individual discourse referents. In fact, Diesing's only empirical argument for the quantificational analysis comes from antecedent contained deletion. However, if any bit of Beghelli and Stowell's theory of LF is correct, then the fact that we observe some LF movement does not in itself allow us to diagnose that movement as QR and the participating DP as a "quantifier." Ad (ii), the assumption that all non-specific indefinites are variables captured by existential closure, irrespective of whether they are monotonic increasing, decreasing, or non-monotonic, gives logically incorrect results, as was argued above.

conditions for a DP to operate in a given mode are not purely denotational, either. This seems like an important, and in fact natural, conclusion.[30] Let us see some of the relevant data.

First of all, we have seen that the same noun phrase may occur in more than one distinguished position in Hungarian and, accordingly, operate in more than one way. For instance, DPs like ***több, mint hat fiú*** 'more than six boys' can occur either in HDistP or in PredOp. Or, ***sok fiú*** 'many boys' can occur in HRefP or HDistP or PredOp, with the same proportional interpretation. Thus, there can be no one-to-one correspondence between denotational semantic properties and modes of operation.

More strikingly, we can point to cases where two denotationally equivalent DPs behave differently. For instance, the determiner 'more than six' has two versions. The (a) version is analytic (syntactic comparison), the (b) version is synthetic (morphological comparison). Now, the former occurs either in HDistP or in PredOp, but the latter only in PredOp:

(69) a. Több, mint hat fiú ment el/el-ment. PredOp/HDistP
more than six boy went away/away-went

b. Hatnál több fiú ment el/??el-ment. PredOp
six-than more boy went away/away-went

I see no independent semantic difference between the two versions, which indicates that the lack of ambiguity in the synthetic version is idiosyncratic.

Similarly, ***legalább hét fiú*** 'at least seven boys' does not, according to my own judgment, occur in PredOp, although logically equivalent 'more than six boys' has a variant that does. This, again, seems like an accidental gap.

In sum, an increasing DP that is in principle capable of supporting a discourse referent may or may not actually do so, on one or any of its uses.

Note a cross-lingustic consequence. If two denotationally equivalent Hungarian DPs do not need to operate identically, then a Hungarian DP and its English "counterpart" do not necessarily do so, either: it is an empirical question whether they do.

Let us now turn to the question whether and how occurrence in PredOp is constrained. PredOp does not care about monotonicity: it hosts increasing, decreasing, and non-monotonic quantifiers. On the other hand, it is remarkable that ***minden fiú*** 'every boy' and ***a legtöbb fiú*** 'most (of the) boys' do not occur there; the former is confined to HDistP and the latter to HRefP. What excludes them? The fact that they have non-intersective determiners cannot be the reason, for instance. (53) already demonstrated that a quantifier in PredOp

[30] At the present stage of research, the noun phrase's choice among the denotationally speaking available options seems arbitrary. It is to be hoped that further research will identify the critical factors, whatever they might be.

may well be partitive or proportional. Likewise, decreasing ***kevés fiú*** 'few boys' is invariably in PredOp, whether proportional or intersective. Furthermore, we are faced with another idiosyncracy here. According to the textbook analysis, ***most of the boys*** is equivalent to ***more than 50% of the boys***(or pick whatever larger figure you prefer). But, as can be expected on the basis of the data reviewed earlier, ***a fiúknak több, mint 50 százaléka*** 'more than 50% of the boys' does occur in PredOp.

The descriptive generalization I offered in (36) was that DPs in PredOp perform a specific operation on predicate denotations: they count. The absence of 'every boy' is natural then: it surely is not a counter. The fact that 'most of the boys,' in distinction to 'more than 50% of the boys' is excluded indicates that being a "counter" is in part a representational/procedural notion, too.

Interestingly, Hungarian word order is not the only empirical domain that sets these two DPs apart. Consider binominal *each* and existential sentences with a coda in English; two well-studied constructions, whose accounts in the literature are standardly in denotational semantic terms:[31]

(70) a. * The professors met most of the boys each.
b. The professors met more than fifty per cent of the boys each.

(71) a. * There will be most of the boys in the yard.
b. There will be more than fifty per cent of the boys in the yard.

Sutton (1993), whose work is the source of the first datum concluded, somewhat desperately, that these contrasts eliminate the possibility for a denotational semantic characterization of what DPs work with binominal *each.* She proposed that what all the good examples have in common is that they are "counters;" a proposal reinforced by ***The professors met one/*a boy each.*** While the general theory in the present paper does not immediately explain why specifically counters need to be involved in (70), I hope to have substantiated that this type of non-denotational conclusion need not be that desperate.

10 APPENDIX ON HUNGARIAN

In this Appendix, I wish to address two issues pertaining to Hungarian that may be necessary for the reader to make good use of the data presented. One concerns the presentation of (29), the global structure of a Hungarian sentence, in current syntactic terms. The other, with which I start, is this:

[31] Comorovski (1995) argues that partitives with a strong determiner may occur in presentational *there*-contexts when they are not anaphoric. This finer qualification will still not distinguish between ***most of the*** and ***more than n% of the.***

(72) What positions do postverbal DPs occupy and what are their scope options?

All literature on Hungarian agrees that postverbal DPs scope under preverbal ones (for two exceptions, see fns. 9 and 34). What has never been seriously examined, to my knowledge, is what scopal options postverbal DPs have within their own domain. Given that the postverbal field is assumed to have a flat structure, É. Kiss's general proposal makes either of the following two predictions:

(73) a. If operators in Hungarian c-command their scope at S-structure (in terms of first branching node c-command), then quantifiers in the postverbal field can be interpreted in either order.

b. If operators in Hungarian precede and c-command their scope at S-structure, then quantifiers in the postverbal field are interpreted in left-to-right order.

The reason why these predictions have not been scrutinized, I believe, is that having more than one scopal expression in the postverbal field is not usual and the judgments are rather difficult. (Since Hungarian goes out of its way to provide means to disambiguate scope, the postverbal field is not the domain of choice for scope interaction.) But if we now look at the postverbal field with the moral of Stowell and Beghelli's work on English in mind, we can construct critical data that are quite straightforward to judge. Such examples involve plural definites, universals, and modified numerals, especially decreasing ones.

The choice of 'a Tuesday' for Focus allows us to control for the possibility that a postverbal quantifier scopes out of the postverbal field; if the Tuesdays do not vary, scope interaction is confined to the postverbal field, which is what we are interested in.

(74) a. Egy keddi napon harapta meg hatnál több kutya
a Tuesday day-on bit pfx six-than more dog
Katit és Marit.
Kati-acc and Mari-acc
'It was on a Tuesday that more than six dogs bit Kati and Mari
OK (a Tuesday >) more than six dogs > Kati and Mari'
OK (a Tuesday >) Kati and Mari > more than six dogs'

b. Egy keddi napon harapott meg hatnál több kutya minden fiút.
'It was on a Tuesday that more than six dogs bit every boy'
OK (a Tuesday >) more than six dogs > every boy
OK (a Tuesday >) every boy > more than six dogs

c. Egy keddi napon harapott meg hatnál több kutya kevés fiút.
'It was on a Tuesday that more than six dogs bit few boys'

OK (a Tuesday >) more than six dogs > few boys
?? (a Tuesday >) few boys > more than six dogs

d. Egy keddi napon harapott meg minden kutya kevés fiút.
'It was on a Tuesday that every dog bit few boys'
OK (a Tuesday >) every dog > few boys
* (a Tuesday >) few boys > every dog

What we find is essentially the same pattern as in English. 'Kati and Mari' and 'every boy' easily take inverse scope over a modified numeral. With great difficulty, 'few boys' can take inverse scope over another modified numeral. But it is unthinkable for 'few boys' to take inverse scope over a universal.[32]

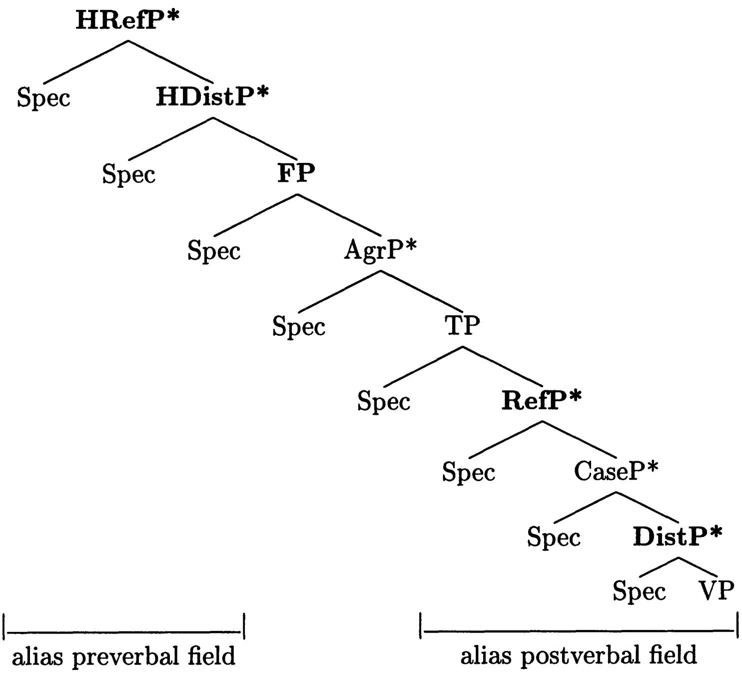
These facts are inconsistent with both (73a) and (73b). What this means is that scopal order in Hungarian is not fully determined by S-structure. The inverse scopal orders must be due to LF movement, by and large in the same way as in English.

This observation eliminates an alleged idiosyncracy of Hungarian. Since the preverbal positions are operator (A-bar) positions, it is quite natural for DPs that move there overtly to have their scope determined once for all. (The same holds for English DPs that undergo overt *wh* or negative fronting.) On the other hand, DPs in the postverbal field are thought to occupy argument (A) positions at S-structure, just like non-fronted DPs do in English. Thus postverbal Hungarian DPs can be expected to have their scope interpretation determined in essentially the same way as English DPs in A-position.

I am thus led to positing two "scopal fields" in Hungarian: the preverbal one, with landing sites for overt operator movement, and the postverbal one, with comparable landing sites for covert operator movement. The global structure that these are embedded in is as follows. NB the XPs generated under the Kleene star respect the binary branching constraint: they do not form flat substructures.

[32] I chose a subject-object word order to make the judgments simpler. It seems to me that the judgments are contingent merely on linear order, however.

(75)



As was argued in the foregoing sections, the preverbal field contains HRefP, HDistP, Focus and PredOp. For the sake of simplicity, I take the latter two to be alternative specifiers of FP.[33] The postverbal field contains RefP and DistP, but no FP. I assume that each (H)Dist head has an event quantifier as its share, albeit I do not posit SharePs all over the place. The linearly $n+1$th event quantifier quantifies over subevents of the linearly nth; the ultimate event variable resides in the VP.

In (75), the two fields are separated by a series of functional projections. In line with Brody (1990), I assume that the surface position of the verb is derived by fronting, i.e. by movement into a functional head which is not separated from the specifier of FP by any overt material. The details of the movement of the verb and of the verbal prefix (whose surface position serves to diagnose whether a DP is in FP or HDistP) are immaterial to our present concerns; see Szabolcsi (1996b).

DPs move out of VP to check their nominative, accusative, etc. features in the appropriate CaseP (only *pro* moves up to AgrP). They may stay in CaseP and end up postverbally in surface structure, or they may move on to one of the

[33]Or, there might be two distinct [+F] functional heads, one that hosts Focus and another that hosts PredOp; see also the discussion of (76). This possibility is explicitly allowed in the theory of Horvath (1995).

preverbal operator positions. At present, we are interested in the postverbal option.

CaseP's are generated in one cluster, in a random order. This accounts for the facts that the order of postverbal DPs is independent of grammatical function and that the linearly first can always take scope over the linearly second. In addition, CaseP's are flanked by RefP and DistP, LF movement into which follows the same mechanics that Beghelli and Stowell propose for English. Likewise, there is a possibility of reconstruction into VP. As in the discussion of Beghelli and Stowell at the outset, I assume that only semantically insignificant movement can be undone by reconstruction. Thus a DP that has moved to RefP or DistP cannot be reconstructed.[34]

These assumptions derive the data in (74) as follows. In (74a), the inverse reading is due to the movement of 'Kati and Mari' into RefP. In (74b) and (74c), the inverse readings are due to the reconstruction of 'more than six dogs' into VP; in the latter case the marginality of this reading will need an independent account, as in English. In (74d), the inverse reading is unquestionably out, because 'every dog' cannot reconstruct into VP.

The last question to touch on concerns postverbal counting quantifiers. (33), the table summarizing the distribution of DPs in the distinguished positions notes a peculiarity:

(76) A counter must occur in PredOp unless (i) there is already another counter in PredOp, or (ii) Focus is filled, or (iii) the verb is negated. Why?

Recall that PredOp is in complementary distribution with Focus before the finite verb stem. It differs from Focus in two ways. First, DPs in Focus receive an exhaustive interpretation, while DPs in PredOp do not receive any "extra" interpretation.[35] Second, DPs in Focus are negated directly, while DPs in PredOp are not:

(77)	Mari ment el. Mari went away 'It is Mary who left'	Nem Mari ment el. not Mari went away 'It is not Mari who left'
(78)	Kevés fiú ment el. few boy went away '(There are) Few boys (who) left'	Nem ment el kevés fiú. not went away few boys 'There aren't few boys who left'

[34]In addition, names, definites and "referential" indefinites that occur in Focus or in the postverbal RefP must reach the main DRS somehow; I remain agnostic on whether this is to have a syntactic reflex of some sort.

[35]Drawing from Kenesei (1986) and van Leusen and Kálmán (1993), Szabolcsi (1994) proposes that this contrast follows from the fact that the appropriate notion of exhaustivity, which has come to be called exclusion-by-identification, is defined only for singular or plural individuals. The inhabitants of Focus denote individuals but those of PredOp do not.

Given these differences, it was justified in the main text to distinguish between Focus and PredOp. This paid off in view of the functional parallelism between Beghelli and Stowell's ShareP and Focus with bare indefinites on the one hand, and Beghelli and Stowell's AgrP/VP positions and PredOp on the other. In this section, I am making the simplifying assumption that Focus and PredOp are the alternative specifiers of the same functional head with a [+F] feature. Now, the question is why counters exhibit the peculiar distribution noted in (33). I adopt a suggestion by M. Brody (1990; p.c.), who observes that the behavior of counters resembles that of *wh*-phrases in, say, English: they must check their [+F] feature overtly unless another item has checked its [+F] feature overtly. Counters that remain postverbal are analogous to *wh*-in-situ.

(79) a. $[_{[+F]P}$ Hatnál több lány] hívott fel kevés fiút.
six-than more girl called up few boy-acc
'The girls who phoned few boys were more than six'

b. * Felhívtam kevés fiút.
up-called-I few boy-acc
'I phoned few boys'

(80) a. Where did you buy what?

b. * You bought what?

Thus a syntactic condition analogous to the one governing the distribution of *wh*-phrases (the Wh-criterion) can be thought to account for the data.[36]

Finally, we must ask why modified numerals are [+F]. A simple, perhaps also simplistic, answer might be this. The DPs that can introduce discourse referents and serve as targets of predication are topics in some generalized sense. The DPs that cannot introduce discourse referents are bound to be part of the comment. [+F] is perhaps nothing else than "is part of the comment."[37]

[36]We may note, however, at least two relevant differences between the two domains. First, *wh*-in-situ may be located in a different clause than the overtly moved *wh*-phrase, while in-situ counting quantifiers must be clausemates to the overt checker of [+F]. Second, the postverbal counter does not by any means take scope in PredOp; it takes scope in situ. This is confirmed by the fact that another quantifier may scope between them. In the sentence below, 'everyone' unambiguously scopes over 'few jokes:'

Mari/Hatnál több fiú mesélt mindenkinek kevés viccet.
'It was Mary / There were more than six boys who told everybody few jokes'

[37]This view is consonant with the bipartite ⟨grounding, claim⟩ representations in Kálmán (1994). Kálmán argues that a [+F] constituent is part of the claim and the remnant of the grounding. I thank J. Horvath, L. Kálmán, and M. Brody for discussions on the feature [+F].

REFERENCES

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics* 2(2):83–135.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In Amy J. Schafer (Ed.), *Proceedings of the North East Linguistic Society 23, Volume One*, 65–80, Amherst. GLSA Publications, University of Massachussetts.

Beghelli, Filippo. 1995. *The Phrase Structure of Quantifier Scope.* PhD thesis, UCLA.

Beghelli, Filippo. 1996. The syntax of distributivity and pair-list questions. In this volume.

Beghelli, Filippo, Dorit Ben-Shalom, and Anna Szabolcsi. 1996. Variation, distributivity, and the illusion of branching. In this volume.

Beghelli, Filippo, and Tim Stowell. 1994. The direction of quantifier movement. *GLOW Newsletter* 32:56–57.

Beghelli, Filippo, and Tim Stowell. 1996. Distributivity and negation. In this volume.

Ben-Shalom, Dorit. 1993. Object wide scope and semantic trees. In Lahiri (Ed.), *Proceedings of SALT III.*

Brody, Michael. 1990. Remarks on the order of elements in the Hungarian focus field. In István Kenesei (Ed.), *Approaches to Hungarian, Volume 3*, 95–123. Szeged: JATE.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1(2):181–234.

Comorovski, Ileana. 1995. On quantifier strength. In Emmon Bach et al. (Eds.), *Quantification in Natural Languages*, 145–177. Dordrecht: Kluwer.

Crain, Stephen, and Henry Hamburger. 1992. Semantics, knowledge, and np modification. In Robert D. Levine (Ed.), *Formal Grammar: Theory and Implementation*, 372–401. Oxford University Press.

Diesing, Molly. 1992. *Indefinites.* Cambridge: MIT Press.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22(1):1–25.

Farkas, Donka. 1996. Evaluation indices and scope. In this volume.

Gil, David. 1989. Scopal quantifiers: Some universals and effability. Unpublished manuscript.

Gil, David. 1995. Universal quantifiers and distributivity. In Emmon Bach et al. (Eds.), *Quantification in Natural Languages, Volume 1*, 321–362. Dordrecht: Kluwer.

Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers.* PhD thesis, University of Amsterdam.

Groenendijk, Jeroen, and Martin Stokhof. 1993. Interrogatives and adverbs of quantification. In Katalin Bimbó and András Máté (Eds.), *Proceedings of the 4th Symposium on Logic and Language*, 1–29. Budapest: Akadémiai Kiadó.

Haegeman, Liliane, and Raffaela Zanuttini. 1990. Negative concord in West Flemish. Unpublished manuscript.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* PhD thesis, University of Massachussetts, Amherst.

Horvath, Julia. 1995. Structural focus, structural case, and the notion of feature-assignment. In Katalin É. Kiss (Ed.), *Discourse Configurational Languages*, 28–64. New York: Oxford University Press.

Hunyadi, László. 1981. *A nyelvi polaritás kifejezése a magyarban [The Expression of Linguistic Polarity in Hungarian].* PhD thesis, Hungarian Academy of Sciences.

Kadmon, Nirit. 1987. *On Unique and non-Unique Reference and Asymmetric Quantification.* PhD thesis, University of Massachussetts, Amherst.

Kálmán, László. 1994. Conditionals, quantification and bipartite meanings. In *Working Papers in the Theory of Grammar, Volume 1, Number 3.* Theoretical Linguistics Programme, Budapest University and Research Institute for Linguistics, Hungarian Academy of Sciences.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Kenesei, István. 1986. On the logic of Hungarian word order. In Werner Abraham and Sjaak de Meij (Eds.), *Topic, Focus, Configurationality*, 143–159. Amsterdam: John Benjamins.

É. Kiss, Katalin. 1987. *Configurationality in Hungarian.* Dordrecht: Kluwer.

É. Kiss, Katalin. 1991. Logical structure in linguistic structure: The case of Hungarian. In James Huang and Robert May (Eds.), *Logical Structure and Linguistic Structure*, 387–426. Dordrecht: Kluwer.

É. Kiss, Katalin. 1992. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11(1):85–120.

É. Kiss, Katalin. 1994. Sentence structure and word order. In Ferenc Kiefer and Katalin É. Kiss (Eds.), *Syntax and Semantics 27*, 1–90. San Diego: Academic Press.

Kuroda, S.-Y. 1972. The categorical and the thetic judgment: Evidence from Japanese syntax. *Foundations of Language* 9(2):153–185.

Ladusaw, William. 1994. Thetic and categorical, stage and individual, weak and strong. In Mandy Harvey and Lynn Santelmann (Eds.), *Proceedings from Semantics and Linguistic Theory IV*, 220–229, Ithaca. Cornell University Department of Modern Languages and Linguistics Publications.

van Leusen, Noor, and László Kálmán. 1993. The semantics of free focus. University of Amsterdam.

Liu, Feng-hsi. 1990. *Scope Dependency in English and Chinese.* PhD thesis, UCLA.

Liu, Feng-hsi. 1992. Branching quantification and scope independence. In Jaap van der Does and Jan van Eijck (Eds.), *Generalized Quantifier Theory and Applications*, 297–331. Chicago: CSLI/University of Chicago Press.

Partee, Barbara. 1995. Quantificational structures and compositionality. In Emmon Bach et al. (Eds.), *Quantification in Natural Languages*, 541–601. Dordrecht: Kluwer.

Reinhart, Tanya. 1995. Interface strategies. Unpublished manuscript, Utrecht University.

Sasse, Hans-Jürgen. 1987. The thetic/categorical distinction revisited. *Linguistics* 25(3):511–580.

Stabler, Edward. 1996. Computing quantifier scope. In this volume.

Sutton, Melody. 1993. Binomial each. Master's thesis, UCLA.

Szabolcsi, Anna. 1980. Az aktuális mondattagolás szemantikájához [On the semantics of functional sentence perspective]. *Nyelvtudományi Közlemények* 59–82.

Szabolcsi, Anna. 1981. The semantics of topic-focus articulation. In Jeroen Groenendijk et al. (Eds.), *Formal Methods in the Study of Language*, 513–540. Amsterdam: Mathematisch Centrum.

Szabolcsi, Anna. 1992. Combinatory Grammar and projection from the lexicon. In Ivan Sag and Anna Szabolcsi (Eds.), *Lexical Matters*, 241–268. Stanford: CSLI.

Szabolcsi, Anna. 1994. All quantifiers are not equal: the case of focus. In *Proceedings of the 5th Symposium on Logic and Language*, Vol. 42–43, 171–187. Acta Linguistica Hungarica.

Szabolcsi, Anna. 1996a. Quantifiers in pair-list readings. In this volume.

Szabolcsi, Anna. 1996b. Verb and prefix movement in Hungarian. Unpublished manuscript, UCLA.

Szabolcsi, Anna, and Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1(3):235–284. Reprinted in this volume.

Webber, Bonnie Lynn. 1979. *A Formal Approach to Discourse Anaphora*. New York: Garland.

Winter, Yoad. 1996. Choice functions and the scopal semantics of indefinites. Unpublished manuscript, OTS, Utrecht University.

5

COMPUTING QUANTIFIER SCOPE*

Edward P. Stabler

Department of Linguistics
UCLA

1 A PUZZLE

There is a quantifier scope ambiguity in (1). In addition to the preferred normal scope reading paraphrased in (1ns), this sentence has the inverse scope reading paraphrased in (1is):

(1) Some linguist speaks every language.

(1ns) There is some linguist x such that x speaks every language

(1is) For every language y, there is some linguist or other who speaks y

Liu (1990) and others point out that certain objects, such as those with decreasing denotations, do not allow an inverse scope reading, as in:

(2) Some linguist speaks at most 2 languages.

(2ns) Some linguist x is such that x speaks at most 2 languages

(2is) There are at most 2 languages y such that some linguist or other speaks those 2 languages y

(2is) is perfectly intelligible: it says that linguists speak at most 2 languages altogether. This does not seem to be available as an interpretation of (2). This is arguably not just a preference; sentence (2) just cannot be interpreted as (2is).

If it is true, as it seems, that a certain semantically identified collection of quantifiers does not allow inverse scope readings, how can this be explained? If there is a mechanism for "raising" quantifiers, how could a semantic property like decreasingness be relevant to whether the mechanism can apply? One might

*This work was inspired by stimulating discussions with Anna Szabolcsi, Filippo Beghelli, Fernando Pereira, and especially Dorit Ben-Shalom and Ed Keenan.

A. Szabolcsi (ed.), Ways of Scope Taking, 155–182.

think that the semantic values of the structures computed in any computational model of the language user could have no more relevance than the dollars do when a calculator computes your bank balance.[1] But the puzzle is resolved when we remember the most basic insight about computational systems: formal properties can reflect semantic properties. In a computational model of the language user, to make sense of generalizations like the one mentioned above, we need a theory in which the semantic generalizations are revealed to hold in virtue of the formal representation of quantifiers.

Suppose that we just subcategorize the quantifiers according to their scoping behavior, thereby distinguishing them syntactically. While this allows scoping distinctions to be captured formally in a computational approach, it still misses something in the generalization noted above. That generalization does not merely identify the distinctive scoping behavior of certain classes of quantifiers, but also relates that behavior to the semantic values of the quantifiers in these classes. Putting the matter this way, it is easy to see how the remainder of the puzzle must be resolved. The way to capture this semantic generalization is to elaborate the representational account of scope in such a way that the very representational features that determine the scope of quantifiers also determine fundamental aspects of their inferential role in the language, and hence fundamental aspects of their meaning such as decreasingness. This paper provides a preliminary account of this sort.

Providing such an account has become more challenging in some recent derivational approaches to syntax. If some lexical item has a syntactic requirement which is met in the course of a derivation, there may be no need to assume that the requirement is in some significant way still present in the derived structure. A "checked" or fulfilled requirement that has no further role may be regarded as a deleted syntactic feature. This perspective on syntactic derivation, according to which key features of the structures are deleted in the course of a derivation, apparently conflicts with the idea that derived syntactic structure is the "output" of linguistic analysis and the "input" to later cognitive processes. This tension between what is required for the syntax and what is required for semantic interpretation and later cognitive processes is quite clear in recent debates about the role of chains, for example. Brody (1995) argues that because chains are needed for interpretation, we should assume

[1] This seems to be exactly what Chomsky (1995a) is puzzled about when he insists that semantic properties cannot be relevant to an account of a language user, any more than an intentional relation to the earth is really relevant when we say "the meteorite is aiming for the earth." A certain puzzlement about the role of semantic properties in linguistics is natural, and a resolution of the puzzle is sketched here. It is no surprise that there are no significant generalizations about meteorite trajectories stated in terms of what they are aiming for, and hence no comparable puzzle about semantic generalizations in astronomy. Meteorites do not represent targets to themselves the way language users represent things to themselves and each other.

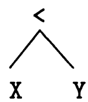
they are available in the syntax, and hence an additional derivational notion of movement becomes unnecessary. But this argument is unsound. If chains are needed for interpretation but not needed to condition the syntactic derivation, the appropriate conclusion is that interpretation depends on aspects of linguistic structure to which the syntax is insensitive. This sort of perspective on semantic values is familiar in categorial grammar and certain other traditions, and is quite natural in transformational grammar as well.

To illustrate how this whole picture could work, a very simple illustrative syntax is formalized in §2, incorporating a simple version of the theory of quantifiers proposed in Beghelli and Stowell (1996) and Szabolcsi (1996). Following Chomsky (1995b), the structures defined in this simple syntax have some properties that are syntactically relevant, and others that are semantically relevant. A preliminary compositional semantics is provided for the structures with semantically relevant properties; and sound inference patterns are defined in which the very features that determine scoping options thereby determine inferential role. That is, for this simple language, we have a purely formal account that conforms to semantic generalizations of the kind mentioned above. In §3 some basic assumptions of this approach are identified and contrasted with alternative approaches.

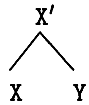
2 A SIMPLE GRAMMAR

A simple example will illustrate how the various pieces of a story about human language could go together to explain semantic generalizations about quantifier scope. This example is not intended to be an adequate representation of English, or even of just those English-like constructions that it generates. But it will be clear in this fragment how syntax, semantics and inference are related in such a way as to allow quantifiers to be used in semantically appropriate ways by a computational system. The proposal is that some such relation is found in human language, and that this is how the semantic generalizations about quantifier scope will ultimately find their explanation. Relevant aspects of the fragment will be discussed in a more general way in §3.

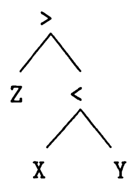
Before presenting the grammar, we informally sketch the basic assumptions. Following Koopman (1994) and Sportiche (1995) we assume that all syntactic requirements must be satisfied either by head movement or by an appropriate relation between a head and a specifier. Following Chomsky (1995b) we assume that all these requirements are encoded as properties of lexical heads. To indicate that a head X has a certain complement Y, rather than "percolating" just some of the features of the head X, we assume that all the features of the head X "project over" those of Y. To depict this in a tree, we will use the notation



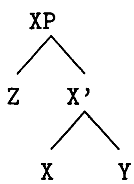
The more traditional X-bar structure here would be something like the following:



A head `X` and complement `Y` may combine with a specifier `Z` to yield a structure like the following:



The traditional X-bar structure here would be something like this:



With these new simpler structures, it is a trivial matter to find the head of any projection; at each internal node one simply goes down the "lesser" branch, the one "pointed to," until reaching the collection of features which is the head. A maximal constituent is either the root of a complete structure or else a node that does not project over its sister. We will assume that all branching is binary, and that the order of heads, complements and specifiers is uniformly the one just depicted: specifier, head, complement.

To implement a simple account of movement, LF movement, and reconstruction, the language will contain, in addition to simple constructions like (3i) for the determiner `some`, constructions in which this element has been split into (3ii) its phonetic features `/some/`, (3iii) its interpretable features `(some)`, and (3iv) its bare categorial structure:

(3)

(i)	(ii)	(iii)	(iv)
d: [=n]	d: [=n]	d: [=n]	d: [=n]
|	|	|	
some	/some/	(some)	

(The feature =n will be explained just below.) Rather than lowering material in reconstruction after having raised it, we aim to achieve the same results by allowing movements to split a category into these various components along its chain. This makes a "one pass" computation of the syntactic structure possible.

To indicate that a head x has an unfilled selection requirement, an unfilled "receptor" for some category y, we will use the feature =y. The head some that selects a noun will have the feature =n. This feature will be licensed by "incorporating" the categorial feature n of the selected constituent by head movement, where this incorporation is possible only in a strictly local configuration. All head movements will be "covert;" that is, no phonetic material will be moved when these relations are established. When n is incorporated into a head with the feature category =n (together with any other requirements of the selected head), we will indicate that this receptor has been filled by removing the feature =n and deleting the category n.[2] If the moved verb has any interpretable and phonetic features, they will be left behind.

Turning to quantifier phrases, we will adopt a simplified version of the basic ideas of Beghelli and Stowell (1996) and Szabolcsi (1996).[3] The following four categories of determiners are distinguished:

(4) negative determiners (no),
distributive and universal determiners (each, every),
group denoting determiners (the, some, a, one, three, ...),
counting determiners (few, fewer than 5, more than 6, ...).

Furthermore, we assume the special functional categories ref, dist and foc (or share), which provide specifier positions that distinguish among these quantifiers. In simple clauses with transitive verbs, we order these categories with respect to the complementizer c, tense t, and the verb v as follows:

c (ref) (dist) (foc) t v.

Beghelli and Stowell (1996) cite work from Szabolcsi, Kinyalolo and others to support the claim that quantifiers of (roughly) these 4 types are restricted to

[2] The similar grammars explored in Stabler (1996) distinguish "weak" selection features =x from "strong" selection features =X which trigger overt head movement. For present purposes, we ignore overt head movement.

[3] The reader is referred to these sources for the motivations and details of the sort of syntactic theory of quantifiers considered here.

different surface positions in other languages; but in English the argument for the various specifier positions is less direct. In any case, this account will be approximated in our simple language fragment, and the syntactic properties of objects formed with negative and counting determiners will accordingly reflect their inability to take inverse scope. More precisely, in the fragment defined below, as in the Beghelli and Stowell (1996) account of human languages, we have the following:

(5) (i) negative quantifiers are interpreted in the specifier of `neg`

(ii) distributive and universal quantifiers are interpreted in the specifier of `dist`

(iii) group denoting quantifiers can interpreted in the specifiers of `ref`, `foc` or in their case positions

(iv) counting determiners are interpreted in their case positions

In the simple language defined here, we assume that the case positions in a simple clause are the specifier of `t` and the specifier of `v`. We will not treat Beghelli and Stowell's (1996) interesting suggestions about quantification over events, or the differences between `each`, `every` and `all`, or the full account of negation. There seems to be no obstacle to elaborating the fragment presented here to encompass these ideas.

2.1 Syntax

We begin as in Keenan and Stabler (1994), letting the syntax comprise four sets: a vocabulary V, a set of categories Cat, a lexicon Lex which is a set of expressions which are structures built up from V and Cat, and a set of structure-building rules $\mathcal{F}$ which map sequences of expressions to other expressions. For any such grammar $G = \langle V, Cat, Lex, \mathcal{F} \rangle$, the language $L(G)$ is the closure of Lex under $\mathcal{F}$. That is, the language is all the structures that can be built from the lexicon using the structure building operations that the grammar provides.

We define the **vocabulary** V as the union of three sets:

(6) a. $PI = \{$ `some, less, than, one, no, every, linguist, sentence, speaks, believes, is` $\}$

b. $P = \{/\mathtt{x}/ \mid \mathtt{x} \in \mathtt{PI}\}$ (phonetic features)

c. $I = \{(\mathtt{x}) \mid \mathtt{x} \in \mathtt{PI}\}$ (interpreted features)

d. $V = (PI \cup P \cup I)$

A lexical item may include some element of PI in a structure together with syntactic features, where we distinguish among these, the basic categories and the others:

(7) a. *base* = {v,d,n,c,case,t,neg,ref,dist,foc}. (basic categories)

b. *features* = $\{\texttt{=x} \mid \texttt{x} \in \textit{base}\} \cup$ (selection requirement)
$\{\texttt{+x} \mid \texttt{x} \in \textit{base}\} \cup$ (weak assignment of x to specifier)
$\{\texttt{+X} \mid \texttt{x} \in \textit{base}\} \cup$ (strong assignment of x to specifier)
$\{\texttt{-x} \mid \texttt{x} \in \textit{base}\}$ (requirement of x)

Finally, projections of any head are labeled with either > or < as indicated above.

The lexicon is a finite set of trees, lexical heads. We can think of each lexical head as a sequence of features, but for convenience we will separate the categorial feature, the phonetic and interpretable features, and the other features, depicting them in the following way:

```
        category:[other features]
                     |
  phonetic and interpretable features
```

If the category feature is absent, we will put a 0 in that position. If the phonetic or interpretable or both are absent, we may have just (interpretable material) or /phonetic material/ or 0. If there are no other syntactic features, we have the empty list [] of other features.

We now present some elements of the **lexicon**. Here are four determiners and two nouns, but we allow the determiner some to be syntactically 3 ways ambiguous:

```
 d:[=n]      d:[=n,-foc]      d:[=n,-ref]      d:[=n]
   |              |                |              |
  some           some             some      more than one


d:[=n,-neg]    d:[=n,-dist]    n:[-case]    n:[-case]
    |               |              |            |
    no            every         student     sentence
```

We have transitive verbs like speaks. To introduce a simple recursion we allow believes to select a c object and a d subject. To facilitate the discussion of semantic properties in §2.3, we include is as a transitive verb, ignoring the special syntactic properties of this verb.

```
v:[+case,=d,=d]    v:[+case,=c,=d]    v:[+case,=d,=d]
        |                  |                  |
     speaks            believes              is
```

Finally, functional categories provide the glue that ties the lexical elements together. These elements have no phonetic features:

```
c:[=ref]      c:[=dist]      c:[=foc]      c:[=t]
   |              |              |             |
   0              0              0             0

ref:[=dist,+ref]      ref:[=foc,+ref]      ref:[=t,+ref]
      |                     |                    |
    (ref)                 (ref)                (ref)

dist:[=foc,+dist]     dist:[=t,+dist]      foc:[=t,+foc]
      |                     |                    |
   (dist)                (dist)                (foc)

t:[=neg,+CASE]        t:[=v,+CASE]         neg:[=v,+neg]
      |                     |                    |
     (t)                   (t)                 (neg)
```

As will become clear, these lexical entries do not provide for sentences with more than one universal or distributive quantifier. In a language where all universal quantifiers are covertly fronted, we could allow a lexical entries for `dist` to recursively select another `dist`. In a language where only one quantifier is fronted overtly, we could let `agr` select `dist` with a strong `+DIST` feature, while letting the recursive form select `dist` with a weak `+dist` feature. We leave this and other similar elaborations aside for the present.

Roughly following Chomsky (1995b), we have two basic structure building relations, *merge* and *move*, but we break *move* up into 3 different functions, according to whether there is movement of a complete category with its phonetic features and interpretable structure, a movement of just the interpretable structure (leaving phonetic properties behind), or a movement of just phonetic features. These operations are called *movepi, movei, and movep*, respectively. So our set of **structure building functions** is

$$\mathcal{F} = \{merge, movepi, movei, movep\}.$$

We now define each of these structure building operations in turn. Example applications will be presented in a derivation below.

(8) The function *merge* combines two expressions, in response to a selection feature `=x`. It is restricted to two cases: (i) a head with `=x` merges with

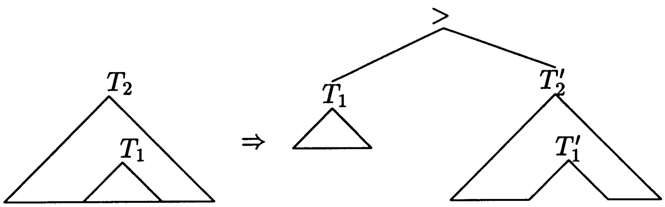


Figure 1 *movepi* raises subtree T_1, replacing it by empty subtree T_1'

a constituent of category x on its right (a complement), deleting both of these features; (ii) a nonhead with =x merges with a category x on its left (a specifier), deleting both features and checking any other +y, -y pairs. For the moment we will assume that in both cases the x head incorporates all the features of the other category, except when v incorporates d it leaves behind all features of the form -y.

(9) The function *movepi* is similar to what is usually called XP substitution. Following Chomsky (1995b) in conceiving of it now as a structure building operation, it applies to a single constituent whose head has a "strong "+X feature, moving the closest -x proper subconstituent to its specifier position as shown in Figure 1. We will not allow this operation to move a constituent out of a position where it must be interpreted. That is, to get the results in (5i)-(5iv), this function can move a -case subtree T_1 unless it is the specifier of dist, foc, ref or neg. So let's say that a subtree T_1 is **movable**, $\mu(T_1)$, if, and only if it is not the specifier of dist, foc, ref or neg, and it has at least one of the features -case, -dist, -foc, -ref, -neg. Furthermore, whenever *movepi* or *movei* move the interpretable part of a constituent that has no -case feature, moving it to check some other feature -x, the movement leaves behind a special interpreted feature $\blacktriangleleft_x$.

(10) The second type of movement, *movei*, is triggered by a "weak" feature +x rather than a strong feature. This is covert XP movement, similar to *movepi* except that the phonetic features of the moved constituent are left behind. Only the syntactic and interpreted features move.

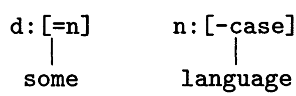
(11) What happens when movement is triggered by a strong feature +X but the only available -x constituent is not movable? As discussed just above, this happens for example in the case of a -neg subject which ends up in the specifier of neg with an unchecked +CASE feature. In this case, *movep* can apply, moving just the phonetic structure to the position of the strong feature where it must be pronounced.

This completes the specification of the entire grammar $G = \langle V, Cat, Lex, \mathcal{F} \rangle$. Formally, the language $L(G)$ is everything that can be derived from the lexical elements *Lex* using any of the five structure building functions in $\mathcal{F}$. Of course, we will typically be interested in derivations of well-formed clauses, that is, expressions with root c in which all structural requirements have been satisfied. Let $\Gamma(G)$ be this set of all expressions with root c in which all structural requirements have been satisfied.

2.1.1 A simple clause

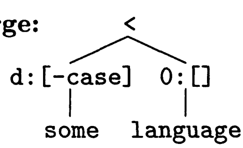
The grammar allows a number of derivations of well-formed clauses that would be pronounced as `some linguist speaks some language`. One of them interprets both arguments of the verb in their case positions. We step through the derivation of this simple structure first. We begin by drawing the components of the object from the lexicon:

step 0 lexicon:



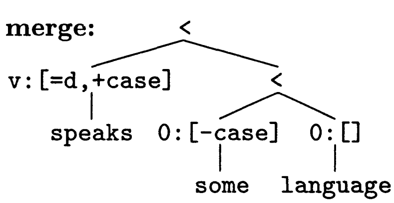
These components can be merged, deleting `=n` and the categorial feature `n`:

step 1 merge:



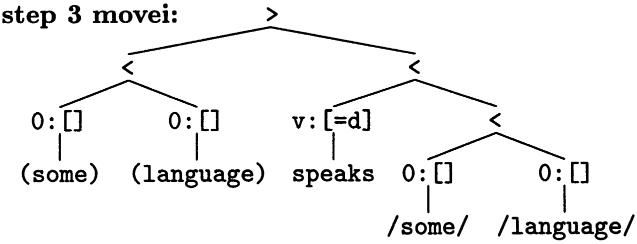
Then we merge the result with the lexical item `speaks`, deleting `=d` and the categorial feature `d`:

step 2 merge:



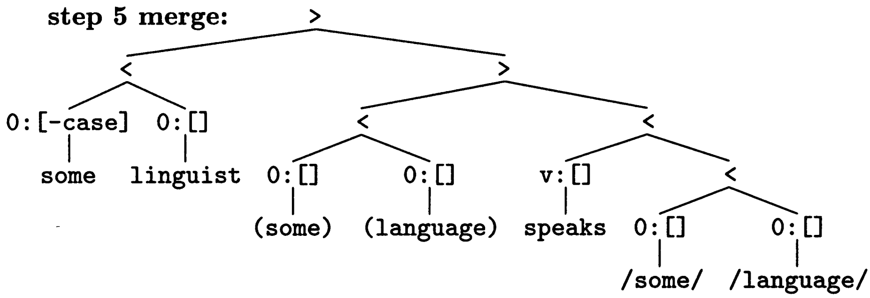
The result has `+case` and the derivation cannot proceed until this feature is discharged, so we covertly move the object to the specifier position, deleting `+case` and `-case`. This has the effect of splitting the interpretable and categorial features of the object from the phonetic features, the latter being left behind:

step 3 movei:



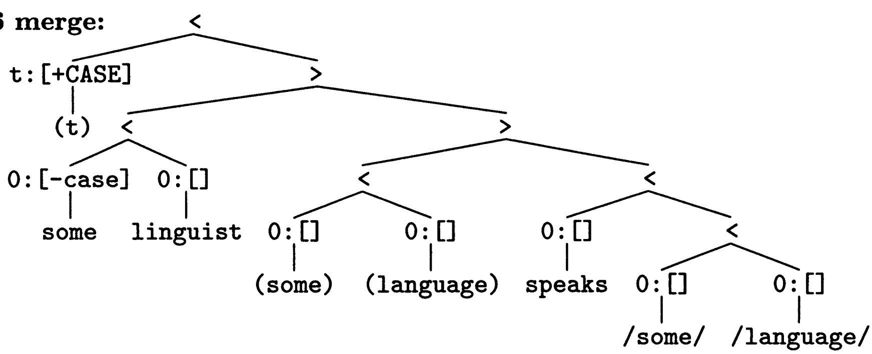
Step 4 of the derivation merges **some** with **linguist** to form the subject, using *merge* as was done in Step 1. The subject can then be merged with the result of Step 3 to give us a verbal projection that contains both the subject and the object:

step 5 merge:



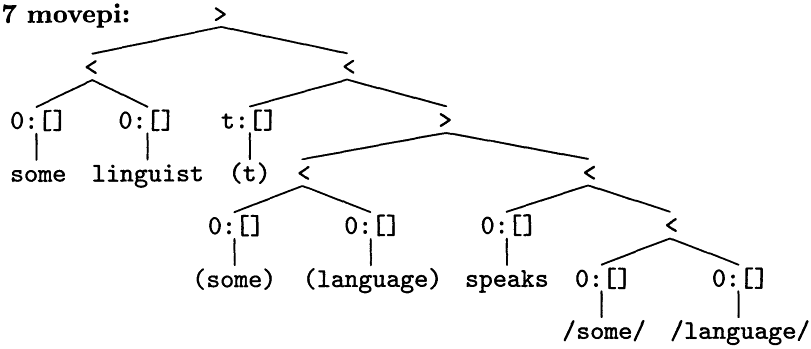
This structure can then be merged as the complement of the lexical t that selects it, deleting =v and v:

step 6 merge:



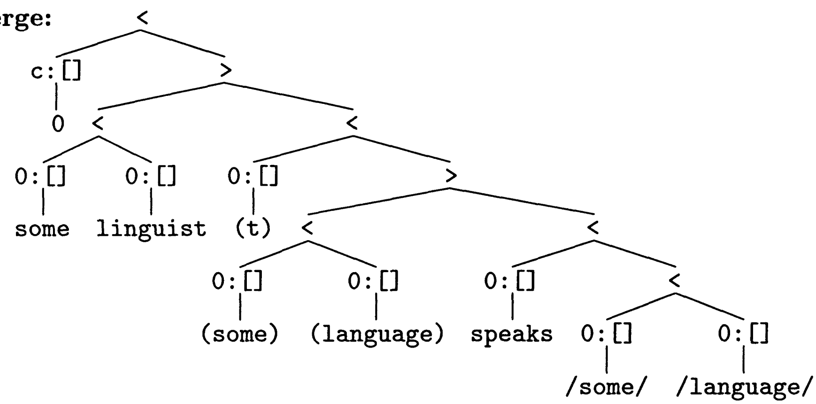
The head of this structure has a strong +CASE feature which must be assigned, so we overtly move the subject, deleting +CASE and -case:

step 7 movepi:

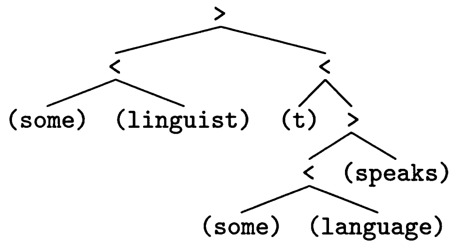


Finally, this whole structure can be taken as the complement of the lexical item `c` that selects it, deleting `=t` and `t`:

step 8 merge:



Notice that in this derived structure, there are no outstanding syntactic features except for the categorial feature `c`. Reading the phonetic material in order across the leaves of the tree we find `/some linguist speaks some language/`. Reading the interpretable constituents across the leaves of the tree we find: `(some linguist some language speaks)`. Actually, for PF and for LF, we require more structure than just the string. If we strip out just the parts of this last structure that are semantically relevant, we have the structure:



Leaving off the parentheses, the semantically relevant structure of the verb and its arguments is just

$[_{t}$`[some linguist] [[some language] speaks]].`

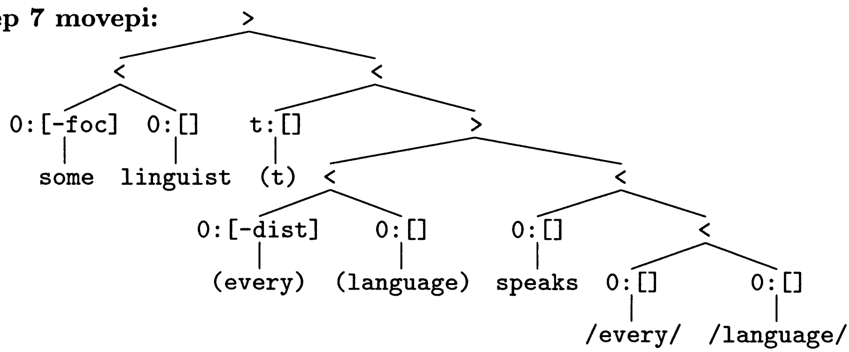
In the next section we consider a derivation in which this arrangement of semantic constituents is altered.

2.1.2 Inverse scope

The grammar allows a number of derivations of well-formed clauses that would be pronounced as `some linguist speaks every language`. In one of them, `every language` in `dist` scopes over `some linguist` in `foc`. This structure has the inverse scope reading (1is) discussed in the introduction. (We will discuss the interpretation of this structure in more detail in §2.3, below.)

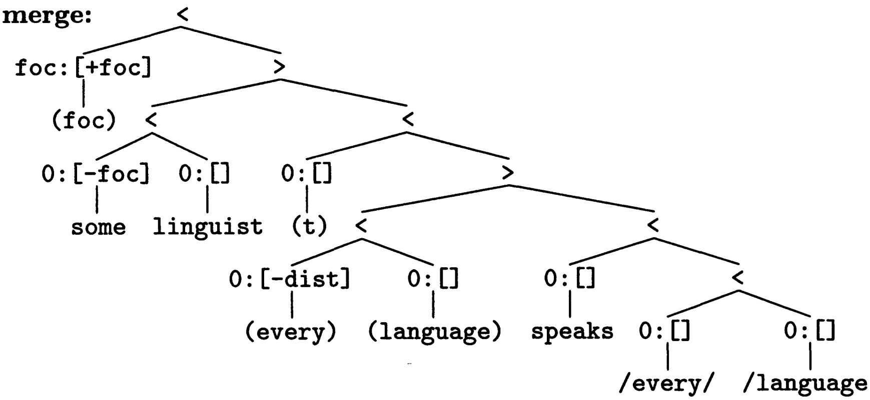
We begin the derivation with the same first seven steps as in the previous case, combining the `v` phrase with `t` and raising the subject to the specifier position. The resulting structure is just like the one we have at the same stage of the previous derivation except we have the additional features `-dist` and `-foc` which must be discharged.

step 7 movepi:

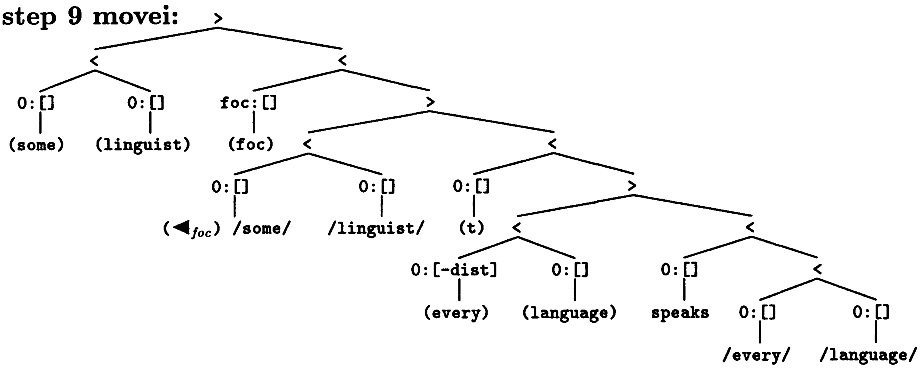


To discharge these additional features, a few more derivational steps are needed. First, we merge with `foc` in Step 8, and then raise the focused argument to specifier position in Step 9. Notice that Step 9 leaves behind the semantic feature $\blacktriangleleft_{foc}$, since the constituent being moved has already had its `-case` feature checked:

step 8 merge:

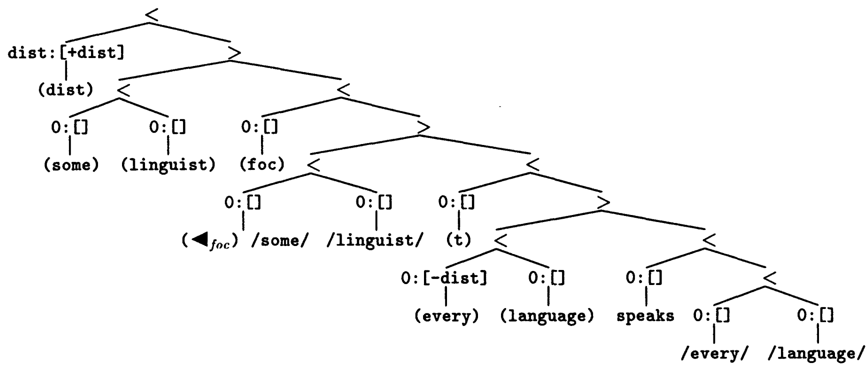


step 9 movei:

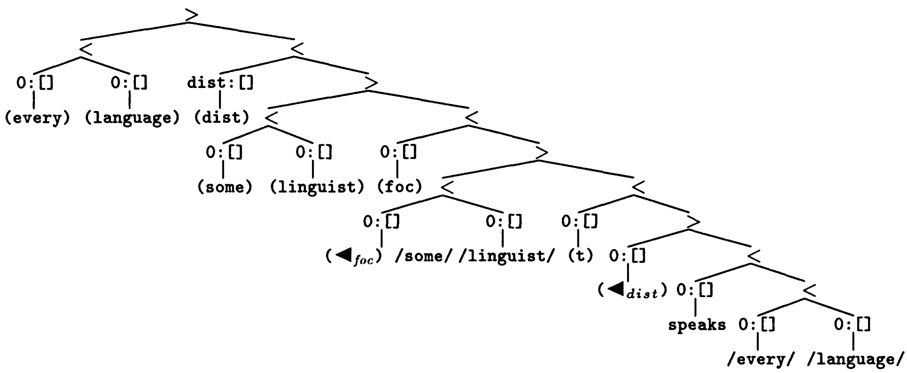


The structure that results at step 9 has only one syntactic feature left, namely the category `foc`. This feature can be deleted by merging it with the lexical item `dist` which selects it, in step 10. Then in step 11, the `+dist` feature of this new head pulls the interpreted features of the `-dist` object up to its specifier position to cancel `+dist`:

step 10 merge:

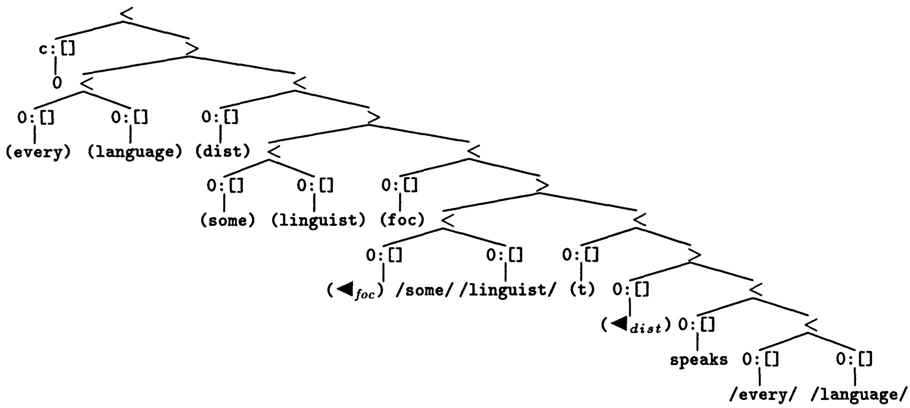


step 11 movei:



The phrase that results at step 11 again has just one syntactic feature, `dist`, which allows it to be the complement of a lexical `c` head in step 12, completing the derivation of a clause in which all syntactic features except `c` have been canceled:

step 12 merge:



The semantically relevant structure here, the LF, is

`[`$_{dist}$`every language [`$_{foc}$`some linguist [`$_{t}$◀$_{foc}$ ◀$_{dist}$ `speaks]]]`.

2.2 Basic properties

By derivations of the sort presented in the previous sections we can establish that $\Gamma(G)$ contains a range of structures of the sort we are interested in:

PROPOSITION 1

(a) [$_{\texttt{dist}}$`(every language)`[$_{\texttt{foc}}$`(some linguist)`[◀$_{foc}$◀$_{dist}$ `speaks`]]] $\in \Gamma(G)$

(b) [$_{\texttt{ref}}$`(some linguist)`[$_{\texttt{dist}}$`(every language)`[◀$_{ref}$◀$_{dist}$ `speaks`]]] $\in \Gamma(G)$.

(c) [$_{\texttt{dist}}$`(every language)`[`(some linguist)` ◀$_{dist}$ `speaks`]] $\in \Gamma(G)$.

Other, more general properties of $\Gamma(G)$ can be established from the definition of the grammar:

PROPOSITION 2

(a) There is no derivation of a c$\in \Gamma(G)$ in which the interpretable structure of the object `no language` appears higher than the interpretable structure of the subject `some linguist`.

(b) There is no derivation of a c$\in \Gamma(G)$ in which the interpretable structure of the object `less than one language` appears higher than the interpretable structure of the subject `some linguist`.

Proof: (a) follows immediately from the movability requirement μ that restricts the domains of the functions *movepi* and *movei*. These are the only functions that move interpretable structure, and they are blocked from moving any specifier of `neg`, but `no language` will have a `-neg` feature which can only be checked in that position.
(b) follows again from the movability requirements. Since `less than one language` will have its case feature checked immediately in object position, and since it has no `-dist`, `-foc`, `-ref` or `-neg` feature, its interpretable structure cannot be moved at all. □

In sum, we have captured in this fragment the basic structural distinctions set out in (5).

Although chains, sequences of constituents coindexed by numbers or variables, are not present at any point in our derivations, we can still talk about series of movements which would have been regarded as constituting a chain in earlier theories. It is clear that the following properties fall out of our definition of the grammar:

PROPOSITION 3 Trace is immobile.

Proof: This follows immediately from the fact that the movements defined here leave no features on the trace which could trigger any later syntactic operation. □

PROPOSITION 4 Chains are uniform.

Proof: There are no chains in the syntax, and no indexing of categories with numbers or elements of a model. However, we can look at the possible sequences of movements and see that various desired sorts of "uniformity" are guaranteed. For example, no constituent can move to more than one case position; no head can adjoin to another head and then move to a specifier position; and so on. These follow from our definitions of the movement operations, so that no additional uniformity requirement need be stipulated. □

Finally, it is a simple matter to compute the syntactic structures generated by the grammar G, as we see stepping through the derivation of the example structure above.[4] These structures determine quantifier scope, so what we have seen is how the computation of quantifier scope can proceed without reference to semantic values of any expressions. This leads us back to the original puzzle: haven't we missed something important? Before tackling this question, let's sketch a semantics for the fragment.

[4]It is not quite so simple to provide a reasonable account of how people might compute these structures incrementally, as they hear a sentence from beginning to end. This problem is explored in Stabler (1996).

2.3 Semantics

We will interpret the simple clauses in $\Gamma(G)$ extensionally, providing no account of sentences involving the verb `believe`. For simplicity we will also ignore the contribution of tense etc. However, it will be relevant whether an argument appears in the a projection of a head with a tense feature `(t)`, the tense phrase, or not, as we will see. The language $\Gamma(G)$ contains predicates which take various numbers of arguments. We use the notation q^j to indicate that q is a predicate that takes j arguments. As in standard extensional semantics, a predicate q^j is denotes a j-ary relation Q^j on the (nonempty) domain A. And as usual, we let nouns denote 1-ary relations.

We adapt to our purposes a formulation of semantic theory from Ben-Shalom (1996).[5] The basic elements of our models are infinite sequences s of objects from some nonempty domain, together with three special registers which will hold focused, referentially identified, or distributed elements. Since the contents of these registers are always in the sequences they are associated with, we can depict them as pointers to particular elements in the sequence, as for example in

$$s = (\overset{\substack{ref\\ \downarrow}}{s_0}, s_1, \overset{\substack{dist\\ \downarrow}}{s_3}, s_4, \overset{\substack{foc\\ \downarrow}}{s_5}, s_6 \ldots)$$

Here, the element in the first position is also the content of the *ref* register, and this element will be denoted *ref*(s). The contents of the third and fifth registers are similarly *dist*(s) and *foc*(s), respectively. We call a sequence together with these registers an r-sequence.

The result of extending r-sequence s by adding some $a \in A$ to the first position is denoted by as as usual. The result of extending s by adding some $a \in A$ to the first position and also marking the new element as the contents of the *ref* register will be denoted

$$\overset{\substack{ref\\ \downarrow}}{a}\ s$$

Similarly for extensions with elements that are marked *dist* and *foc*.

[5] Ben-Shalom (1996) notices that the standard model theory of propositional modal logic (as in, e.g. Goldblatt 1992) can be adapted to a language of predicates and generalized quantifiers by letting the points of the model be sequences and letting the generalized quantifiers correspond to modal operators. Compare Alechina (1995). The resulting semantic theory is remarkably simple and natural. For brevity, in defining the models below we mention just one accessibility relation R, namely extension, but the definition of $\models_s$ shows that we actually use different extensions according to the restrictors of the quantifiers. See Ben-Shalom (1996) for a more careful treatment.

With these conventions, we can let a model be a triple $\mathcal{M} = (S, R, V)$ where S is the set of infinite r-sequences of objects from domain A; R is a binary relation on S such that

sRs' iff $s' = as$ for some $a \in A$;

and V is a function from predicates q to subsets of S defined as follows:

$$s \in V(q) \text{ iff } \begin{cases} s_0 = s_1 & \text{when } q = \texttt{is}^2, \\ (s_0, \dots, s_{j-1}) \in \mathrm{Q}^j & \text{for any other } q = q^j \end{cases}$$

Then an r-sequence $s \in S$ in a model $\mathcal{M} = (S, R, V)$ satisfies an expression as follows:

$$\begin{array}{lll} \mathcal{M} \models_s q & \text{iff} & s \in V(q) \\ \mathcal{M} \models_s [_f\texttt{every } q\ \phi] & \text{iff} & \text{for every } a \in \mathrm{Q},\ \mathcal{M} \models \overset{f}{\underset{as}{\downarrow}} \phi \\ \mathcal{M} \models_s [_f\texttt{some } q\ \phi] & \text{iff} & \text{for some } a \in \mathrm{Q},\ \mathcal{M} \models \overset{f}{\underset{as}{\downarrow}} \phi \\ \mathcal{M} \models_s [_f\texttt{no } q\ \phi] & \text{iff} & \text{for no } a \in \mathrm{Q},\ \mathcal{M} \models \overset{f}{\underset{as}{\downarrow}} \phi \\ \mathcal{M} \models_s [_f\texttt{less than one } q\ \phi] & \text{iff} & \text{for less than one } a \in \mathrm{Q},\ \mathcal{M} \models \overset{f}{\underset{as}{\downarrow}} \phi \\ \mathcal{M} \models_s [\blacktriangleleft_f\ \phi] & \text{iff} & \text{for } a = f(s),\ \mathcal{M} \models_{as} \phi \end{array}$$

As usual, we say that an expression ϕ is verified by a model, $\mathcal{M} \models \phi$ if and only if for all $s \in S$, $\mathcal{M} \models_s \phi$. And an expression ϕ is valid in frame (S, R) just in case for all valuations V, $(S, R, V) \models \phi$.

Consider for example the expression

$$\phi = [_{dist}\texttt{every language}[_{foc}\texttt{some linguist}[_t\blacktriangleleft_{foc}\blacktriangleleft_{dist}\texttt{speaks}]]].$$

We stepped through the syntactic derivation of this expression in §2.1.2. Let's show that this sentence ϕ is true in a model $\mathcal{M}$ where there are two linguists $l0, l1$ and two languages $a0, a1$ where $l0$ speaks $a0$ and $l1$ speaks $a1$. That is, the universe A is $\{l0, l1, a0, a1\}$, so S is the set of infinite r-sequences of these elements. We have the relations

$\mathrm{LINGUIST} = \{l0, l1\}$
$\mathrm{LANGUAGE} = \{a0, a1\}$
$\mathrm{SPEAKS} = \{(a0, l0), (a1, l1)\}$

We adopt the convention of listing the elements of tuples in these relations in order of decreasing obliqueness, putting the subject last. Using this order for the arguments slightly simplifies the definition of the satisfaction relation, as will become clear.

Now consider any particular $s \in S$. We will show that for any such s, $\mathcal{M} \models_s \phi$. If s is some r-sequence

$$s = (\overset{\overset{ref}{\downarrow}}{a_0}, \overset{\overset{dist}{\downarrow}}{l_1}, \overset{\overset{foc}{\downarrow}}{l_1}, \dots)$$

The first three elements of this r-sequence are $ref(s) = a0$, $dist(s) = l1$ and $foc(s) = l1$. To show that our example is verified by this and every other r-sequence, the contents of the registers and the first positions in the sequence are irrelevant. Any other r-sequence will verify the sentence ϕ too.

To establish $\mathcal{M} \models_s \phi$, we simply step through the truth definition given above. First we see that

$$\begin{aligned}\mathcal{M} \models_s \phi \text{ iff } \quad & \text{for every } a \in \text{LANGUAGE},\\ & \mathcal{M} \models_{\underset{a}{\overset{dist}{\downarrow}}\, s} [_{foc}\texttt{some linguist}[_t \blacktriangleleft_{foc} \blacktriangleleft_{dist} \texttt{speaks}]].\end{aligned}$$

Notice that this step involves considering extending of the original r-sequence s to r-sequences as in which the *dist* register has also been reset to a. Using the truth definition again, we see that

$$\begin{aligned}\mathcal{M} \models_s \phi \text{ iff for every } a \in \text{LANGUAGE}, \quad & \text{and}\\ \text{for some } l \in \text{LINGUIST}, \quad & \mathcal{M} \models_{\underset{l}{\overset{foc}{\downarrow}}\,\underset{a}{\overset{dist}{\downarrow}}\, s} [_t \blacktriangleleft_{foc} \blacktriangleleft_{dist} \texttt{speaks}].\end{aligned}$$

The next two steps involve extending the r-sequences of interest without resetting any registers. Rather, they extend the r-sequence by copying the contents of the registers. First we extend with the contents of the *foc* register, then with the contents of the *dist* register:

$$\begin{aligned}\mathcal{M} \models_s \phi \text{ iff for every } a \in \text{LANGUAGE}, \quad & \text{and}\\ \text{for some } l \in \text{LINGUIST}, \quad & \mathcal{M} \models_{l\,\underset{l}{\overset{foc}{\downarrow}}\,\underset{a}{\overset{dist}{\downarrow}}\, s} [\blacktriangleleft_{dist} \texttt{speaks}].\end{aligned}$$

$$\begin{aligned}\mathcal{M} \models_s \phi \text{ iff for every } a \in \text{LANGUAGE}, \quad & \text{and}\\ \text{for some } l \in \text{LINGUIST}, \quad & \mathcal{M} \models_{a\,l\,\underset{l}{\overset{foc}{\downarrow}}\,\underset{a}{\overset{dist}{\downarrow}}\, s} \texttt{speaks}.\end{aligned}$$

Now we can see that the satisfaction relation does hold, since for every $a \in$ LANGUAGE there is some $l \in$ LINGUIST such that $allas \in V(\texttt{speaks})$. This is the case because for every $a \in$ LANGUAGE there is some $l \in$ LINGUIST, $(a, l) \in$ SPEAKS.

In this simple semantic theory, we see that the scope of a quantifier is determined by the syntactic position of its interpretable structure. We have constituents $\blacktriangleleft_{dist}$, $\blacktriangleleft_{foc}$, and $\blacktriangleleft_{ref}$, which are "bound" in some sense, but they are unlike variables in certain other respects. There are not infinitely many of them. Notice also that a clause in which no arguments have moved is perfectly interpretable. That is, there is no need to move just in order to produce a variable that can be bound by a quantifier. Furthermore, the positions of quantifiers is severely restricted by the syntax. Since we only have two decreasing quantifiers in the fragment, `no` and `less than one`, and since we know from Proposition 2 that when these quantifiers occur in object position their interpretable structures never move above the interpretable structure of the subject, it follows that decreasing objects will never have an inverse scope reading.

The puzzle about this is that the computation of syntactic structure, sketched in the previous section, determines quantifier scope and yet it makes no reference to the semantic values of the determiners. This should be puzzling, as we can see from the fact that, as far as the syntax is concerned, there is no reason not to expect a decreasing quantifier to have the feature `+dist` or `+ref`. The fact that there is no such quantifier in human languages would then be just an accident, but this is implausible. To avoid this impasse, the theory needs one additional ingredient.

2.4 Inference

The decreasing quantifiers can be characterized by their inferential behavior. Stating our inference rules over the structures in $\Gamma(G)$, and showing just the interpretable structures, a quantifier `D` is decreasing if the following inference is valid:

$$\frac{\begin{array}{c}\texttt{[D N1 V]}\\ \texttt{[}_{dist}\texttt{every N2 [}\blacktriangleleft_{dist}\texttt{ some N1 is]]}\end{array}}{\texttt{[D N2 V]}}$$

This observation suggests a solution to the puzzle with which we began. As noted in the introduction, we can assume that the very features of a determiner which determine scoping options also constrain the inferential role of the structures containing the determiner. Let's sketch how this might work.

Beginning with the easiest case, we can assume that we have the following inference rule for any determiner `D`,

$$\frac{\begin{array}{c}\texttt{[}_{neg}\texttt{ D N1 V]}\\ \texttt{[}_{dist}\texttt{every N2 [}\blacktriangleleft_{dist}\texttt{ some N1 is]]}\end{array}}{\texttt{[}_{neg}\texttt{D N2 V]}}$$

That is, any determiner which appears in the specifier position of `neg` will be one that licenses this inference. Any such determiner is decreasing. Once this rule is given, it is no longer an accident that negative quantifiers scope the way they do. A link between decreasingness and scope is established, in purely formal terms.

Notice that we do not say, if a determiner is decreasing then it may occur in the specifier of `neg`. That is not quite true, in the first place, since counting determiners may also be decreasing even though they cannot occur in the specifier of `neg`. But more importantly, this claim reverses the order of explanation. On the present account, we do not assume that the speaker has some "grasp" of the semantic value of the determiner first, and then decides where to put it in the syntactic structure. Rather, the speaker uses the determiner in a certain way, in the syntax according to the requirements specified in its features, and in inference. Constraints on its semantic value then follow.

The situation here is artificially simple, but we have secured the basic point, that the very features which determine syntactic properties can also determine inferential and semantic properties in theories like this one. Notice for example that although the only counting determiner in our fragment is decreasing, we could have had counting determiners like `more than one`, which are not decreasing. So in this case we will not be able to say that every counting determiner will license the characteristic inference of decreasing quantifiers. Rather, the particular kind of quantifier that appears in the counting quantifier positions will matter. This is just a case where formal distinctions must be drawn within the categories of items that can occupy the syntactic positions.

3 SOME ALTERNATIVES AND ELABORATIONS

3.1 Alternative explanations of scoping restrictions

There are many alternative approaches to the computation of quantifier scope. Most are easily assigned to one of the following categories:

Syntactic restructuring with variables: The approach offered here falls in this category, together with previous approaches to "quantifier raising" (QR). Other examples of this sort of approach are provided by Higginbotham and May (1981), Fiengo and May (1995), and others.

Syntactic restructuring without variables: Keenan (1987) shows how, if we allow the subject and verb to form an interpretable constituent which is, in turn, interpreted with the object, we can compute inverse scope without the use of bound variables.

Semantic restructuring with variables: It is also possible to compute the normal and inverse scope readings from a single syntactic structure by, in effect, storing and raising the embedded object quantifier in the computation of logical form. Examples of this sort of approach are provided by Cooper (1983), Pereira (1990), Dalrymple et al. (1994) and others.

Semantic restructuring without variables: It is possible to avoid variables on this approach too. Instead of raising the object quantifiers to bind a variable, the object quantifiers can be interpreted as combining with the verb to yield a functions from quantifiers to propositions, as in Nam (1991) and others.

The present approach differs from previous QR-based approaches in various respects: the classes of quantifiers distinguished by Beghelli and Stowell (1996) are distinctive, as are the assumptions about the structural positions available for these constituents. The representation of previous positions of arguments is also distinctive both syntactically (there are finitely many types of symbols $\blacktriangleleft_x$, not infinitely many variables), and semantically (the symbols $\blacktriangleleft_x$ are interpreted as propositions in a propositional calculus, not as terms in a predicate calculus). In a certain sense, there are no "variables," though the elements $\blacktriangleleft_x$ play a very similar role.

The basic properties of the present approach which do the work in solving the puzzle though are just these: the quantifiers fall into various syntactic categories with distinctive properties, and this variety also plays a role in determining inferential role and semantic value. So the semantic restructuring accounts face two difficulties. In the first place, there are empirical considerations of the sort adduced by Beghelli and Stowell (1996), Szabolcsi (1996) and other work in the same tradition. If it is true that the different classes of quantifiers are syntactically distinguished, with their surface positions related to different positions in linguistic structure by restricted movement relations of just the sort found elsewhere in syntax, then a syntactic approach to scoping phenomena is clearly preferable. Semantic approaches face a second kind of problem when we try to elaborate them to solve the puzzle of §1.

Let's briefly consider a recent proposal from Dalrymple et al. (1994) as a representative example.[6] This work assumes that a single syntactic structure (an "f-structure") like (12) is mapped into either of two alternative logical forms (13) or (14):

(12)
$$f\colon\begin{bmatrix}\textsc{pred} & \text{`speak'} \\ \textsc{tense} & \textsc{pres} \\ \textsc{subj} & g\colon\begin{bmatrix}\textsc{spec} & \text{`every'} \\ \textsc{pred} & \text{`linguist'}\end{bmatrix} \\ \textsc{obj} & h\colon\begin{bmatrix}\textsc{spec} & \text{`some'} \\ \textsc{pred} & \text{`language'}\end{bmatrix}\end{bmatrix}$$

(13) $f_\sigma \rightsquigarrow_t \mathit{every}(w, \mathit{linguist}(w), \mathit{some}(z, \mathit{language}(z), \mathit{speak}(w, z)))$

(14) $f_\sigma \rightsquigarrow_t \mathit{some}(z, \mathit{language}(z), \mathit{every}(w, \mathit{linguist}(w), \mathit{speak}(w, z)))$

The deduction of either logical form is made possible by providing lexical entries for such predicates that allow its arguments to be bound in either order. The predicate speak is associated with the higher order linear logic form (15) in a deductive system in which this form is mutually derivable with (16):

[6]The features of the Dalrymple et al. (1994) proposal that I discuss here are very basic, shared by many other approaches. The way the linear logic eliminates the need for "quantifier storage," the way certain logical forms are properly made unavailable – these interesting and distinctive features of the Dalrymple et al. (1994) proposal are not considered here.

(15) $\forall X.\, g_\sigma \rightsquigarrow X \multimap (\forall Y.\, h_\sigma \rightsquigarrow Y \multimap f_\sigma \rightsquigarrow \mathit{speak}(X, Y))$

(16) $\forall Y.\, h_\sigma \rightsquigarrow Y \multimap (\forall X.\, g_\sigma \rightsquigarrow X \multimap f_\sigma \rightsquigarrow \mathit{speak}(X, Y))$

Without going into the details, by binding the inner formula first, using the first of these forms leads to normal scope, while using the latter leads to inverse scope. Notice that the linear logic is used in the deduction of the logical forms, but the language of the logical forms themselves, shown in (13) and (14), does not contain linear connectives like $\multimap$ or the higher order variables X, Y.

This is a "semantic restructuring" account because the mechanisms used in the deduction of the alternative scopes – a linear logic in this case – are not the ones used in the deduction of the syntactic structure. How could such an account be elaborated to disallow the inverse scope reading when the object is a decreasing quantifier? Since one of the main points of the Dalrymple et al. (1994) approach is apparently to make either argument of the predicate equally available, it appears that a significant change will be required. Any account is going to have to reconstruct some sort of asymmetry between the two arguments of the verb, and then have some corresponding difference between the lexical entries for the various sorts of quantifiers so that derivations of the unwanted scopes are unavailable. Supposing that this could be done, there is another problem, which is to indicate a plausible link between the restrictions on the scoping options and the inferential roles of the quantifiers. This is likely to be difficult too, because in accounts like this one, the mechanisms which govern the calculation of the logical form do not appear in the logical form itself. They are eliminated. As noted above, the language used in the deduction of the logical forms is distinguished from the language of the logical forms themselves. Whatever restrictions there are on quantifier scoping are presumably realized in the former, whereas inference patterns are presumably defined over the latter. In the fragment discussed above, on the other hand, there is only one language, $\Gamma(G)$. It is syntactic form, the interpreted form, and the language of reasoning, all at once. In this language, it is easy to tie the grammatical features responsible for quantifier scoping to the features that license certain inference patterns.

3.2 Alternative explanations of the semantic generalizations

Barwise and Cooper (1981) note that "processing" a quantified statement `D N Pred` need not involve considering the generalized quantifier `D N`, which in a simple extensional treatment is a function from properties to truth values (or a set of properties). Such a function (or set) can be quite large, but rather than considering the whole function, it suffices, when the quantifier is monotone, to select a witness and check whether this witness stands in an appropriate relation to the property denoted by `Pred`. When the function is increasing, we check to see that every member of the witness is a member of the denotation of the predicate. Similar ideas are proposed by Ben-Shalom (1993) and Szabolcsi (1996).

These are clearly not accounts of what people do when they use the language, at least not in any direct sense. A witness or the generator of a principal filter can be very large too, infinite and undecidable, without any corresponding difficulty in the language users' use of the phrase in inference. Humans do not, of course, actually "check" each member of these sets for membership in another set, performing some computational step for each element of the set the way a semantic automaton might. Rather, we must draw our conclusions by deductive steps defined over representations of more abstract relations among whole sets, representations that are in at least this respect like our syntactic structures with their interpretable elements. Then, no surprise, the size of the sets being reasoned about does not generally have any bearing on the complexity of the reasoning.

So if the witness and generator based verification procedures are metaphorical, involving verification steps that are never really performed by the human language speaker, can we explain why these have seemed relevant in semantic theories for human languages? Yes. Inference methods can be regarded as justified by certain relations among the verification procedures associated with the sentences involved. For example, the direct verification, by checking all elements of the witness sets, of `every human speaks a language`, includes as a proper part all of the steps that would be required in the direct verification of `every linguist speaks a language`. That is why the inference of the latter from the former is justified. With such an intimate connection between these notions, it is not in the least surprising that insights about verification conditions are relevant to inferential roles, which are in turn tied to formal, syntactic properties of linguistic expressions in the computational model of the speaker-hearer.

4 CONCLUSIONS

There is a philosophical tradition which rejects the idea that pure thought is coded into language and then expressed, as if we could think the very thoughts we do without knowing anything of any human language, pure semantics. This view is rejected by Dummett (1993), for example,

> A view that might claim to represent common sense is that the primary function of language is to be used as an instrument of communication, and that, when so used, it operates as a code for thought. On

> this view, it is only because we happen to lack a faculty for directly transmitting thoughts from mind to mind that we are compelled to encode them in sounds or marks on paper ...
>
> The idea of a language as a code became untenable because a concept's coming to mind was not, by itself, an intelligible description of a mental event: thought *requires* a vehicle.

The conception of language and reasoning offered in this paper, according to which at least much of what we call reasoning is reasoning with the structures of our native language, fits with this rejection of the code conception of thought. We do not need to assume that there are two steps in reasoning: the formulation of pure thoughts, and then the bringing together of words to express them. A certain kind of bringing together of words, spoken or not, is thinking.

This view can also be contrasted with views according to which we have many languages: at least, the one we speak and the one we reason with. The objections to this sort of account, which we find in certain "semantic" accounts of quantifier scoping, have a different character. Here, it is argued that such views cannot provide the simplest explanations of semantic generalizations about syntax.

It is sometimes held that the meaning of an expression is uniquely determined by its inferential role. No such thing has been assumed here. But the connection between inferential role and meaning is especially clear in the case of quantifiers and logical constants. From this perspective, it is no surprise to find that the same formal structure that determines the scoping options of a quantifier – whatever, exactly, that structure is – can determine the inferential role of the quantifier. Consequently, semantic generalizations about the scoping options of quantifiers are not surprising, and we can make sense of how such generalizations can be true in a computational model of the language user.

REFERENCES

Alechina, Natasha. 1995. *Modal Quantifiers.* PhD thesis, University of Amsterdam.

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.

Beghelli, Filippo, and Tim Stowell. 1996. Distributivity and negation. In this volume.

Ben-Shalom, Dorit. 1993. Object wide scope and semantic trees. In *Procs. of the Workshop on Semantics and Linguistic Theory, SALT III.*

Ben-Shalom, Dorit. 1996. *Semantic Trees.* PhD thesis, UCLA. Available at http://128.97.8.34/.

Brody, Michael. 1995. *Lexico-Logical Form: A Radically Minimalist Theory.* Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1995a. Language and nature. *Mind* 104:1–61.

Chomsky, Noam. 1995b. *The Minimalist Program.* Cambridge, Massachusetts: MIT Press.

Cooper, Robin. 1983. *Quantification and Syntactic Theory.* Dordrecht: Reidel.

Dalrymple, Mary, John Lamping, Fernando Pereira, and Vijay Saraswat. 1994. A deductive account of quantification in LFG. In M. Kanazawa, C.J. Piñón, and H. de Swart (Eds.), *Quantifiers, Deduction and Context.* Stanford, California: Center for the Study of Language and Information.

Dummett, Michael. 1993. *The Seas of Language.* Oxford: Clarendon Press.

Fiengo, Robert, and Robert May. 1995. *Indices and Identity.* Cambridge, Massachusetts: MIT Press.

Goldblatt, Robert. 1992. *Logics of time and computation.* Stanford: CSLI/Stanford University.

Higginbotham, James, and Robert May. 1981. Questions, quantifiers and crossing. *The Linguistic Review* 1:41–79.

Keenan, Edward L. 1987. Unreducible n-ary quantifiers in natural language. In P. Gardenfors (Ed.), *Generalized Quantifiers.* Dordrecht: Reidel.

Keenan, Edward L., and Edward P. Stabler. 1994. Abstract syntax. In *Configurations*, Université du Québec à Montreal. Forthcoming. Available at http://128.97.8.34/.

Koopman, Hilda. 1994. Licensing heads. In David Lightfoot and Norbert Hornstein (Eds.), *Verb Movement*, 261–296. New York: Cambridge University Press.

Liu, Feng-hsi. 1990. *Scope Dependency in English and Chinese.* PhD thesis, Department of Linguistics, UCLA.

Nam, Seungho. 1991. Scope interpretation in nonconstituent coordination. In *Proceedings of the West Coast Conference on Formal Linguistics, WCCFL 10*, 337–348.

Pereira, Fernando C.N. 1990. Categorial semantics and scoping. *Computational Linguistics* 16:1–10.

Sportiche, Dominique. 1995. Sketch of a reductionist approach to syntactic variation and dependencies. In Hector Campos and Paula Kempchinsky (Eds.), *Evolution and Revolution in Linguistic Theory.* Washington: Georgetown University Press.

Stabler, Edward P. 1996. *Acquiring and parsing languages with movement.* Oxford: Basil Blackwell. Forthcoming. Draft available at http://128.97.8.34/.

Szabolcsi, Anna. 1996. Strategies for scope taking. In this volume.

6

EVALUATION INDICES AND SCOPE

Donka F. Farkas

Linguistics Board
UCSC

1 INTRODUCTION

In this paper I propose a theory of scope that is less structure driven than the traditional approach. The traditional view of scope is structural in the sense that the relative scope of two expressions is taken to be determined by their relative position at some level where hierarchical relations are encoded. More precisely, in this view, e_1 is in the scope of e_2 iff e_2 *commands* e_1 at the appropriate structural level. I take the term *command* in its generic sense here, meaning 'higher than' and leave the details of how to define its domain unspecified for now. Common to all varieties of command is that it is defined at the sentence level.

The structural approach to scope has immediate repercussions for syntax under the assumption that semantic structure is read off of syntactic structure; the original motivation for the syntactic level of LF was precisely the assumption that command relations in syntax determine semantic command relations and therefore scope. Under this assumption, e_1 is in the scope of e_2 iff e_2 commands e_1 at LF.

My aim here is to propose a theory in which the relative scope of two expressions is a matter of possible dependencies between indices, seen as Kaplan-style coordinates of evaluation. According to Kaplan (1979), expressions are evaluated relative to an index I, which is a sequence of coordinates including a world, w, a time, t, a three-dimensional location, p, as well as a coordinate for speaker and addressee. Coordinates of evaluation fix the parameters with respect to which the denotation of an expression is determined. In the terminology of this paper Kaplan's coordinates are called *indices of evaluation.*

The proposal builds on insights concerning temporal reference in Enç (1986), and is close in spirit to work in Discourse Representation Theory (DRT). The main difference between the indexical theory of scope to be presented here and

A. Szabolcsi (ed.), Ways of Scope Taking, 183–215.

the DRT approach found in Kamp and Reyle (1993) is that scope relations in DRT are encoded structurally at the level of Discourse Representation Structure: different scopal relation induce DRSs that differ structurally. In the view presented here the difference is one of content rather than structure.

Structural considerations will turn out to be relevant here too but the theory differs from the traditional one in the following two respects: (i) scope is seen as essentially discoursal; there is no attempt at reducing discoursal effects to sentential ones, as in the acommodation approach of Roberts (1989) and Poesio and Zucchi (1992); (ii) structural considerations at both the syntactic and the semantic level are seen as underdetermining scopal relations: structure determines when an expression *may* be in the scope of another, but not when it *must* be in its scope.

After outlining basic assumptions in Section 2, I turn to presenting the proposal in Section 3. Section 4 shows how the indexical theory presented here accounts for the scopal properties of indefinite and distributive noun phrases i.e., noun phrases that contribute the domain over which a distributive predication holds. (In the terminology of Gil 1995, distributive noun phrases are DIST.KEY.) The final section discusses briefly issues connected to discourse scope, and mentions some of the problems that remain open.

2 SCOPE OF WHAT WITH RESPECT TO WHAT

I am concerned here with the scope of noun phrases (henceforth DPs) with respect to intensional operators (modals, the conditional operator), intensional predicates and nouns (such as *believe, belief, dream*), and 'quantificational' DPs, i.e., DPs whose D(eterminer) contributes a proportional (strong) quantifier. I restrict my attention to DPs whose contribution to lf, the representation of semantic structure, includes at least a subscripted variable x_n, and a descriptive content, DC_n, in the form of a predicative expression on x_n. In addition, quantificational DPs, i.e., DPs whose D is a 'strong' quantifier such as *every, each* or *most,* induce a tripartite quantificational structure in which the quantifier is contributed by the determiner, and where the variable and the DC occur in the Restrictor. In this case the quantificational force of the DP is determined by the quantifier it contributes. I follow DRT and File Change Semantics in assuming that the quantificational force of non-quantificational DPs depends on their position in the semantic structure. I will come back below to the question of exactlv how expressions containing such free variables are to be interpreted.

Scope of the DC A descriptive condition DC_n on a variable x_n constrains the assignment of values to x_n to those assignments where the value meets the descriptive condition. One aspect of the issue of noun phrase scope concerns the question of determining the world in which this condition has to be met. Thus, (1a) has two readings depending on whether the variable introduced by the underlined noun phrase is to be a friend of mine in w, the world of evaluation of the whole sentence, or in w_J, the world according to John, introduced by the matrix predicate *believe.* Similarly, (1b) can be interpreted either with the DC evaluated with respect to w, or with the DC evaluated with respect to the worlds introduced and quantified over by the modal.

(1) a. John believes that *a friend of mine* is a crook.
 b. *A friend of mine* might be a crook.

When the world of evaluation is w, the DC is said to have wide scope with respect to the predicate and the modal; when the world of evaluation is that introduced by the predicate or the modal the DC will be said to have narrow scope with respect to them. Note that the issue of the scope of the DC has consequences for the question of where the referent of the noun phrase is supposed to exist, since under standard assumptions, for ordinary descriptions at least, if an individual satisfies a description at some world w, the individual must exist in w (or must have a counterpart there).

An account of DC scope based on Quantifier Raising (QR) has the noun phrase occur in a position commanding the intensional predicate or modal at LF in the wide scope reading, and it has the predicate or modal command the noun phrase at LF in the narrow scope reading. An account of DC scope in DRT would have the contribution of the DC entered in the main box in the wide scope reading, and it would have the DC entered in the subordinate box created as a consequence of the use of the intensional predicate or the modal in the narrow scope reading. In the account to be presented here, the structural properties of the two readings are identical. The difference lies in the choice of the modal evaluation index of the DC, a choice that is not fully determined by structural factors.

The problem of the scope of the DC of a noun phrase with respect to intensional predicates or operators is reminiscent of the issue of the temporal interpretation of the DC discussed in Enç (1986). Enç shows, based on examples such as (2),

(2) The fugitives are now in jail.

that the temporal reference of the DC of a noun phrase is in principle independent of the temporal reference of the main predication of which the noun phrase is an argument. In her analysis, the temporal coordinate (or index) of the DC can be set to any discoursally salient value. The account of the scope of the DC to be proposed here builds on this similarity.

Scope of the variable I turn now to another aspect of scope that arises in connection with quantification, to be referred to below as the scope of the variable. The issue concerns the question of whether there is co-variation of values assigned to some variable x with values assigned to some other variable y. The expression that contributes x will be said to be within the scope of the expression that contributes y, and to be dependent on y. In order for such co-variation to be possible, y must vary, i.e., its interpretation must involve the assignment of several values, and the interpretation of x must vary with the interpretation of y; x must therefore be a variable whose denotation varies at different assignments. Thus, under the assumption that proper names are rigid designators, if x were contributed by a proper name it could not vary with assignments of values for y. An expression may induce dependency iff its interpretation involves a set or group whose members may serve as evaluation indices to other expressions. Examples of such expressions are noun phrases denoting sets or groups, as well as modals, whose interpretation involves a set of worlds. Now an expression may be dependent iff its evaluation is allowed to vary, i.e., its interpretation may involve varying assignments to the variable it contributes.[1] Typical examples of such expressions are indefinite noun phrases. Expressions that may not vary are those whose values are fixed once and for all, such as proper names, or whose values are fixed relative to the coordinates of the speech act, such as deictic noun phrase. A noun phrase such as *every apple* may induce dependency because it introduces a set of evaluation parameters (one for each apple) that may serve as index values to other expressions, as will be seen in this section. Such an expression denotes an absolute principal filter iff the set of apples that forms the domain of quantification is fixed with respect to the sentence in which the noun phrase occurs, as well as the larger context. It is possible, however, to interpret a noun phrase of this form so as to involve quantification over a set of sets of apples, as in (3):

(3) Every child ate every fruit.

Here the most likely interpretation is one where *every fruit* is dependent on *every child*: it quantifies over a set of sets of fruits, one set *per* child. In this case the domain of quantification of the second quantifier depends/co-varies with the elements of the domain of the first. In order for this interpretation to arise the

[1] In the terms used in Szabolcsi (1996a) and Beghelli, Ben-Shalom and Szabolcsi (1996), an expression may induce variation iff it is interpreted as a generalized quantifier whose witnesses are of cardinality greater than 1; an expression may undergo co-variation iff it allows a non principal filter interpretation. Below I am concerned with the factors that give rise to these interpretations.

context must have provided information about the association of sets of fruits to children, or such information must be readily accommodatable. The noun phrase *every fruit* here no longer denotes an absolute principal filter. Trying to analyse this example as involving a hidden description containing a bound pronoun, on a par with Martin Honcoop's suggestion for relative principal filters mentioned in Beghelli, Ben-Shalom, and Szabolcsi (1996), describes the issue rather than solves it. Note that the dependent reading of *any* noun phrase has a paraphrase involving a description that contains a bound pronoun exactly because of the details of the semantics of the bound reading.

A similar argument can be constructed for showing that, in appropriate contexts, definite noun phrases such as *the red button* in (4),

(4) Every child pressed the red button.

may be dependent, and therefore do not have to denote an absolute princial filter. Here felicity conditions on the use of the definite determiner require the context to have provided information on the basis of which each child is associated with a (single) red button, or such information must be readily accommodatable. I therefore differ here from Beghelli, Ben-Shalom and Szabolcsi (1996) according to whom definite noun phrases, as well as noun phrases whose determiner is *every* are assumed to always have widest scope because they always denote principal filters. In my view noun phrases whose determiner is definite tend to denote absolute principal filters because of the familiarity condition accompanying the use of the definite article. This condition is most easily met when the referent of the noun phrase, i.e., the entity that is to serve as the value of the variable contributed by the noun phrase, is a single entity. The dependent, relative principal filter denoting reading arises when the context provides familiar single values associated to each value of the variable the definite noun phrase depends on. Noun phrases whose determiner is *every* denote an absolute principal filter when the domain from which the variable they contribute is to be given values is a single set. The dependent, relative principal filter denoting reading arises when for each value of the dependency inducing variable, the context provides a set for the variable bound by *every* to range over. For such a reading to arise, the quantificational structure contributed by the dependent noun phrase must be within the Restrictor or the Nuclear Scope of the quantificational structure contributed by the dependency creating expression. (For further discussion, see Farkas 1993, 1994.)

I turn now to outlining the relevant assumptions about quantification and the relevant details of nominal, modal, and situational quantification.

I assume a tripartite quantificational structure at the level of lf, where the Restrictor introduces a set of 'cases', presupposed to be non-empty, which form the domain of quantification. The view of quantification adopted here is therefore *restricted*; cases that do not satisfy the restrictor are irrelevant. Most generally, the lf of quantificational expressions is of the form in (5):

(5) Restrictor Q Nuclear Scope

Such an expression is true iff Q-many ways of making the Restrictor true are also ways of making the Nuclear Scope (NS) true. What the 'cases' in the domain of quantification are depends on the type of quantificational expression involved. The domain of quantification is a set of worlds in modal sentences and afactual (or non-indicative) conditionals such as those exemplified in (6), as well as in generic quantification if one adopts the view, most recently argued for in Condoravdi (1994), that genericity involves a modal dimension. The domain of quantification is a set of situations within a single world in adverbial quantification and factual (or indicative) conditionals of the type exemplified in (7), and it is a set of individuals in nominal quantification, exemplified in (8).

(6) a. John may be sick.
b. If John were sick, he'd be in bed.

(7) a. Sometimes, when John is hungry he is grouchy.
b. If John is hungry he is grouchy.

(8) Every student left.

In order to verify a quantificational claim one has to identify the set of relevant cases based on the information provided by the Restrictor, and then one has to check whether the possibly complex property expressed by the NS is true of the appropriate number of cases. Note that expressions whose evaluation co-varies with the cases that constitute the domain of quantification will be dependent on the expressions that introduce these cases. In (6b), (7a), and (8) there is an overt restrictive phrase: the antecedent in (6b) and (7b), the *when*-clause in (7a), and the semantic content of the subject noun phrase in (8). The context may further restrict the relevant set of cases to a salient subset of the cases identified by the overt restriction. In (6a) there is no overt restrictive phrase. Here the relevant set of cases is determined by the interpretation of the modal, as well as contextual factors. Below I restrict my attention to examples where there is an overt restrictive phrase.

Following DRT, I assume that both instances of existential closure proposed in Heim are dispensed with. Text level existential closure is rendered superfluous, as suggested in Heim (1982, Ch. 3), by the requirement that the text be true in the model with respect to which it is interpreted. If truth conditions are stated in terms of assignment functions, this reduces to the requirement that there be an assignment function that satisfies the conditions imposed by the DRS or lf constructed on the basis of the text. Text-level existential closure is thus replaced by existential quantification over assignment functions interpreting free variables. NS-level existential closure is rendered superfluous by the requirement that the NS be true of the appropriate number of cases introduced by the Restrictor: it reduces to the requirement that there be an assignment function that satisfies the conditions in the NS for Q-many cases that satisfy the Restrictor. NS-level existential closure is thus reduced to distributive existential quantification over assignment functions, where the key domain is identified by the Restrictor. These reductions are significant in so far as they follow from a definition of truth of expressions containing free variables.

I now turn to the relevant features of each type of quantification exemplified above.

2.1 Nominal quantification

Let us assume that expressions are interpreted with respect to a model $M = \langle W, U, V \rangle$, where W is a set of worlds, U is a set of individuals, and V is a set of valuation functions assigning intensions to constants.

It is important in what follows to be able to keep track of individuals across worlds. This can be done either by allowing the domains of worlds to overlap or by enriching the model with counterpart relations that connect individuals across worlds. I adopt the former option though nothing crucial depends on this choice.

The 'cases' that form the domain of quantification are the individuals in w (or the individuals in some contextually salient situation in w) that meet the DC condition of the quantificational DP. A way of making such a case true amounts to chosing an evaluation function that assigns to the variable contributed by the quantificational DP a value that meets the DC condition of the DP. The logical form of sentences involving nominal quantification is as in (9),

(9) Restrictor $\quad Qx_n \quad$ Nuclear Scope

where x_n is the variable contributed by the quantificational DP.[2] Truth conditions for an expression of this form are given in (10).

(10) A quantificational lf of the form in (9) is true in w wrt M iff there is an assignment function f such that there are Q-many assignment

[2]Whether the quantifier unselectively binds any or all other variables in the Restrictor is a question that I leave open here. Heim (1982) assumes unselective binding. In traditional analyses that treat indefinites as quantificational, the assumption is that such variables are bound by a narrow scope existential quantifier. To implement the latter solution in a Heimean framework one would need to introduce a Restrictor-level existential closure operation. The choice between these alternatives has repercussions with respect to the 'proportion problem'.

functions f_R that extend it with respect to x_n and that verify the lf in the Restrictor, such that each f_R has an extension f_{NS} which verifies the NS.

The notion of extension used here is defined in 11.

(11) An assignment function f_e extends f with respect to x_n iff f_e agrees with f on all variables with the possible exception of x_n.

Note that the elements in U that verify the lf in the Restrictor constitute the smallest live-on set of the generalized quantifier denoted by the noun phrase. If Q requires there to be more than one f_R the witness sets of the GQ are of cardinality greater than 1.

The lf of (8) is given in (12):

(12) Restrictor:	$\forall x_1$	Nuclear Scope:
x_1		*leave* $'(x_1)$
student $'(x_1)$		

The quantifier here binds x_1 because it is contributed by the DP that introduces x_1. The truth conditions of the lf in (12) are given in (13).

(13) The lf in (12) is true in w with respect to M iff there is some assignment function f with the following property: every assignment function f_R which extends f wrt x_1 such that $f_R(x_1) \in$ *V(student')* at w has the property of having a trivial extension f_{NS} such that $f_{NS}(x_1) \in$ *V(leave')* at w.

We assumed above that (8) involves quantification over all the students in w. The context may however restrict the domain to some salient subset of students.

The evaluation function f is referred to below as the *base* evaluation function. The interpretation of nominal quantification involves then, besides the base function, a set of functions F_R, which are the extensions of f wrt the variable(s) bound by the quantifier that verify the Restrictor, as well as a set of functions F_{NS}, which extend the functions in F_R and verify the Nuclear Scope.

The values of x_1 above are fixed relative to F_R because x_1 is the variable contributed by a quantificational DP. In examples that exhibit scope ambiguities, non-quantificational DPs contribute variables in the Restrictor or the Nuclear Scope. Variables in the Restrictor pose problems that lie beyond the scope of this paper so I restrict my attention here to the latter case, exemplified in (14):

(14) Every student speaks an Indo-European language.

In (14), the material in the Restrictor is the variable contributed by the subject noun phrase and its DC. The cases introduced by the Restrictor, which form the domain of quantification, are those individuals in w (or some contextually salient situation in w) who satisfy the Restrictor. The NS contains the material contributed by the indefinite, as well as the main predication. The sentence is true just in case the NS is true with respect to each case that satisfies the Restrictor, i.e., just in case the NS is true of every student.

The two interpretations of (14) that are known as the 'wide scope' and the 'narrow scope' reading of the indefinite result from the option of fixing the value of the variable it contributes by the base assignment function f (wide scope), or by the functions in the set F_{NS} (narrow scope). When the indefinite has wide scope, the variable it contributes is assigned a value by the base function, and this value is inherited by the functions in F_R and F_{NS}. When the indefinite has narrow scope, the functions in F_{NS} extend those in F_R with respect to the variable contributed by the indefinite. In the former case the indefinite is independent of the distributive quantificational DP, while in the latter, it is dependent on it. Under the wide scope reading the indefinite may denote a principal filter, while in the narrow scope case it may not.

In structural accounts of scope this difference corresponds to a difference in the structural position of the variable contributed by the indefinite. In Quantifier Raising (QR) accounts, the indefinite noun phrase is raised so as to command the quantificational noun phrase at LF to give the wide scope reading, while in the case of the narrow scope reading QR results in a configuration where the quantificational noun phrase commands the indefinite. The analysis in NP-preposing terms in Heim (1982) achieves the same result. In DRT the variable contributed by the indefinite is added to the main DRS box to give the wide scope reading, and to the NS box to give the narrow scope reading.

2.2 Modal quantification

Following Kratzer (1979, 1980), modal sentences such as (6) are assumed to involve quantification over worlds. The 'cases' that form the domain of quantification are a subset of W. In simple modal sentences, such as (6a), the worlds in the relevant subset are those in which the contextually supplied propositions in the modal base are true. In modal conditionals, the worlds in the domain of quantification are the set of worlds W_R such that the worlds in this set are closest to the base world w, and are such that the proposition expressed by the antecedent is true in them.

The lf of a modal conditional is a quantificational structure with the lf contributed by the antecedent, lf_a, in the Restrictor, and the lf contributed by the consequent, lf_q, in the NS. The quantifier binds the worlds in W_R:

(15) Restrictor: $\forall w_R$ Nuclear Scope:
lf_a lf_q

One makes the Restrictor true here by chosing a world $w_R \in W_R$. The truth conditions of (15) are given in (16).

(16) An lf of the form in (15) is true in w wrt M iff there is an assignment function f such that for every w_R that is closest to w such that w_R has the property that f verifies lf_a at w_R, f verifies lf_q at w_R.

(I will not be concerned below with the 'closeness' requirement.) The issue that arises now is that of the choice of world parameters for particular subparts of lf_a and lf_q. The available choices are the base world, w, and the worlds w_R in W_R. These choices are available because W_R has just been introduced, and because the base world needs no introduction. Note that because W_R is a set of worlds, it may give rise to co-variation. Evaluation of a subpart of the conditional with respect to the base world gives rise to the wide scope reading of that expression with respect to the conditional; evaluation of such a subpart with respect to W_R gives rise to the narrow scope reading of the expression with respect to the conditional. In the latter case, the evaluation of the expression will co-vary with the worlds in W_R. In QR-based accounts these choices correlate with different structural positions at LF: wide scope interpretation correlates with a position commanding the conditional, and narrow scope correlates with a position commanded by the conditional. In DRT the wide scope interpretation correlates with material entered in the main DRS box, and the narrow scope interpretation correlates with material entered in the antecedent or the consequent boxes. To exemplify, consider (17):

(17) If someone were here she'd help.

Under the wide scope reading of the indefinite in the antecedent, one has to find a value for the variable it contributes within the domain of the base world, w, such that in all worlds in w_R in which that person is here, she helps. Under the narrow scope reading of the indefinite, each world w_R such that there is an individual in that world who is here is such that that individual helps.

The ambiguity illustrated above concerned the choice of variable. In examples such as (18),

(18) If a friend of mine were here she'd help.

the additional question of what world the DC must be met in arises as well. That question involves the issue of whether the value assigned to the variable contributed by the indefinite has to be a friend of mine in w or in the worlds in W_R and we will come back to it below.

2.3 Quantification over situations

Sentences involving quantification over situations, exemplified in (7), will be treated on a par with modal quantification by assuming that the domain of a world w includes a set S_w of situations in w. Each world w defines its own extensional model M_w, such that $M_w = \langle S_w, U_w, V_w \rangle$, where U_w is the set of the individuals in w, and V_w assigns values to constants relative to the situations in S_w. Truth of an lf in a world w reduces to truth in a situation in w:

(19) An lf is true in w with respect to M iff there is a situation $s \in S_w$ such that the lf is true in s with respect to M_w.

Truth with respect to a situation is defined as follows:

(20) An lf is true in a situation s in a world w with respect to M_w iff there is an evaluation function f_w that verifies the lf at s, where f_w assigns values to variables relative to the situations in S_w.

Sentences involving quantification over situations receive a treatment parallel to modal quantification, where situations in a world play the role of worlds in a model. In modal conditionals the domain of quantification is worlds in W that are closest to the base world w, and that are such that the antecedent is true in them. In non-modal conditionals the domain of quantification is made up of situations in S_w that are closest to what is expected, and that are such that the antecedent is true in them. The lf of such a conditional is as in (21).

(21) Restictor: $\forall s_R$ Nuclear Scope:
lf_a lf_q

A way of making the Restrictor true now amounts to choosing a situation in S_w in which lf_a is true. Given (19), an lf of the form in (21) is true in w iff it is true in some $s \in S_w$. In order for (21) to be true in some $s \in S_w$ every way of making the Restrictor true (within the limits of what is expected) must be a way of making the Nuclear Scope true as well. The truth conditions are given in (22):

(22) An lf of the form in (21) is true in s wrt M_w iff there is an assignment function f_w with the following property: every minimal s_R that is closest to what is expected in w such that f_w verifies lf_a at s_R has an extension s'_R such that f_w verifies lf_q at s'_R.

A minimal situation in which an lf is true is a situation made up of only individuals and relations that satisfy the conditions in the lf; a situation s' is an extension of a situation s iff s' is part of s.

Sentences involving quantification over situations exhibit scope ambiguities with respect to the situations s_R. The ambiguity illustrated in (17) is parallelled in (23),

(23) If/When a boy he likes comes over, Johnny shows him his turtle.

where the indefinite in the antecedent may be interpreted with respect to the base situation, giving the wide scope reading, or with respect to each situation s_R, giving the narrow scope reading. In the latter reading, the interpretation of the indefinite co-varies with the situations that form the domain of quantification.

In QR-based approaches to scope, the wide scope reading of the indefinite with respect to the conditional is given by an LF where the indefinite has been raised to a position that commands the conditional; the narrow scope reading is given by an LF where the indefinite is commanded by the conditional. In DRT, the wide scope reading is given by a DRS where the contribution of the indefinite has been entered into the main DRS box; the narrow scope reading is given by a DRS where the contribution of the indefinite occurs within the antecedent box.

Common to all the cases of narrow scope variables discussed here is that they involved interpretations where the values assigned to the variable varied. The value of a narrow scope variable x_n in the NS of a nominal quantificational structure is determined by the set of functions F_{NS} because these functions extend the functions in F_R wrt x_n; the values assigned to x_n co-vary with the values assigned to the variable(s) bound by the quantifier. A narrow scope variable in the Restrictor or the NS of a modal quantificational structure is given values by a set of assignment functions whose modal indices range over the worlds quantified over by the modal. The values assigned to such a variable may vary from world to world. The same is true, mutatis mutandis, of narrow scope variables in expressions involving quantification over situations. The value assignments of narrow scope variables in quantificational expressions vary with values assigned to the variable bound by the quantifier. In all these cases the noun phrases introducing the dependent variables are not interpreted as principal filters. Under the wide scope readings, on the other hand, the noun phrases in question are interpreted by a single valuation function at a single world or situation, which is consistent with a principal filter evaluation.

2.4 Conclusion

It has been argued so far that the evaluation of noun phrases involves the following two possibly distinct issues relevant to scope: (i) the issue of the scope of the DC, and (ii) the issue of the scope of the variable. The former pertains to the world or situation in which the DC has to hold of the referent, while the latter pertains to the properties of the evaluation function that assigns values to the variable. Narrow scope variables in quantificational structures are assigned varying values. Such noun phrases will be said to have *non-rigid* reference independently of whether the variation is intensional or extensional.

The necessity of separating these two issues is made evident in quantificational structures, where a narrow scope variable has non-rigid reference while a wide scope one has rigid reference, independently of the scope of its DC. Note that in examples involving indefinites in the complements of intensional predicates such as *believe* or *think,* the wide scope interpretation of a variable could in principle be analyzed as the accommodation of the referent in the base world. Thus, for an example like (1a), one could claim that the variable has only a narrow scope reading, under which it must be assigned a value from the domain of w_J, the world introduced by the matrix predicate *believe.* The wide scope reading would then be the result of accommodating this referent in w_0, the world of the discourse, which is the world the matrix sentence is evaluated in. (For discussion of when this type of accommodation may occur, see McCawley 1981.) Under this view, widest scope would reduce to global accommodation and intermediate and narrowest scope would reduce to varieties of local accommodation. An accommodation-based approach is not available, however, for wide scope indefinites in quantificational structures, since in their case the wide scope reading involves rigid reference while the narrow scope reading involves non-rigid reference. The former therefore can not be reduced to accommodation of the referents of the latter.

In the next section a treatment of scope phenomena is presented that does not connect scopal variation to variation in the relative position of the relata in a scope relation and handles both quantificational and extensional contexts.

3 THE PROPOSAL

3.1 Evaluation indices

Following work in DRT, I assume that the basic components of lfs are variables and conditions placed on them. A fundamental distinction assumed here is between conditions contributed by the main predicative expression of a sentence and conditions contributed by the constituents that realize the arguments of the predicate as well as various adjuncts. The conditions on variables that will be of interest here are predicative conditions, i.e., conditions involving a logical predicate and its arguments.

Truth conditions are stated in terms of functions assigning intensions to variables, and functions in V assigning intensions to constants. The evaluation of both variables and predicative conditions may vary relative to world, situation, and temporal parameters. In what follows I will be concerned only with the former two. Variables may be interpreted with respect to a single assignment function at varying worlds or situations, as in the narrow scope indefinite cases discussed under modal and situational quantification above, or they may be interpreted with respect to varying assignment functions at a single world, as in the case of narrow scope indefinites discussed under nominal quantification. In all these cases the expressions contributing these variables are said to have non-rigid reference. Variables may also be interpreted with respect to a single assignment function at a single world, and a single situation, as in the case of widest scope indefinites. These indefinites have rigid reference. The parameters of variation for predicative conditions concern the worlds or situations at which the extension of V for the predicate in question has to be checked.

In evaluating lfs one has to fix which functions play a role in assigning values to variables, and which worlds or situations predicative expressions are evaluated with respect to. It is assumed here that these parameters of variation are indicated by evaluation indices. The evaluation index of a variable specifies the evaluation function (or set thereof) that is crucial in determining the relevant values of that variable; the evaluation index of a predicative condition whose predicate is P specifies the worlds or situations at which $V(P)$ has to be checked. Under this view, variation in scope concerns variation in the value of evaluation indices. The scope relation is defined as in (24):

(24) An expression e_1 is in the scope of an expression e_2 iff the value of an evaluation index of e_1 is set relative to e_2.

In (1a), repeated here as (25),

(25) John believes that *a friend of mine* is a crook.

the DC of the indefinite is within the scope of the intensional predicate iff the DC is to be evaluated at the world introduced by the matrix predicate. In the terminology used here, the modal index of the DC is set to the value introduced by the predicate. In (14), repeated here as (26),

(26) Every student speaks an Indo-European language.

the indefinite has narrow scope with respect to the universal iff the assignment function parameter of the indefinite is dependent on the assignment function evaluating the universal.

The lf of a simple sentence such as (27a) has the components in (27b-d), where the main predication, MP in (27c), is contributed by the main predicative expression of the sentence, the variable contributed by the indefinite noun phrase is given in (27b), and its DC, given in (27d), is contributed by the descriptive content of the noun phrase.

(27) a. A man left.

b. $x_1 \quad f(w_n)$

c. MP: $leave'(x_1)w$

d. DC_1: $man'(x_1)w_m$

The w's above are the values of the modal indices of the expressions they accompany. The modal index value of x_1, w_n, specifies that one has to consider the value f assigns to x_1 at w_n. The modal index value on the predicative condition in (27c) specifies that $f(x_1, w_n)$ must be in the extension of *leave'* at w. The truth conditions of (27a) under the lf in (27b-d) are given in (28).

(28) (27a) is true at w wrt M iff there is an assignment function f such that

(i) $f(x_1, w_n) \in V(leave', w)$ and

(ii) $f(x_1, w_n) \in V(man', w_m)$

A question that arises now is whether the modal index values of a variable, its DC, and the MP in which the variable is an argument may indeed be set to values that are independent of one another, as has been done in (27) for illustrative purposes. The discussion of the scope of the DC in the previous section has shown that the modal index value of the DC of a variable may at least in principle be independent of the modal index value of the MP in which the variable occurs. The discussion of the scope of the variable suggests that its index is also independent of the index of the MP. Examples which will be discussed below show that the index of a variable and its DC are at least partially independent of one another.

Obviously, (27), when uttered in a neutral context, is interpreted with all modal index values set to w_0, the world of the discourse. Alternative values for modal indices can be selected only if such values are made available in the discourse. Modal index values are made available by being introduced by modal expressions such as modal operators, intensional predicates (*believe, dream, want*), intensional nouns (*belief, dream*), or conditionals. The crucial assumption therefore is that evaluation index values are selected from a set of values made available by the context. In the case of modal indices, the world the discourse occurs in, w_0, is always in this set since it is always present in the context. Other worlds are made available by being introduced through the use of modals or intensional predicates and nouns.

These suggestions can be implemented in an an extended version of the view of context proposed in Stalnaker (1979). In this view the context c contains, besides the set p_c of propositions in the common ground, a set R_c of discourse referents, i.e., the file cards of Heim (1982), and a set I_c of evaluation index values. I_c contains the following proper subsets: W_c, a set of worlds in W,

S_c, a set of situations, T_c, a set of temporal reference points, and F_c, a set of evaluation functions. Just as the speaker and addressee are always present in R_c, the world in which the discourse occurs, w_0 is always present in W_c, and the time of the utterance is present in T_c.

The context change potential of utterances affects not only the common ground and R_c, but also I_c. The use of an intensional predicate or noun, a modal or a conditional has the effect of adding elements to W_c, and thereby adding to the set of accessible modal index values. Evaluation functions that figure in the truth conditions of an expression are added to F_c at the point when the relevant expression is being added to the context.

3.2 Free and bound indices

Illustrating with the modal indices of main predications, I show now that they come in two varieties, *free* and *bound.* If a modal index is free, any w in W_c is, in principle, an admissible value for that index.[3] If a modal index is bound, its value must be set to the world (or worlds) introduced by some particular expression.

The modal indices of the MPs of main clauses are free; their default value is w_0 but this choice may be overridden in favor of previously introduced worlds. Consider for instance the discourse in (29):

(29) I had an unpleasant dream last night.
The weather had turned unbearably hot and my room was not air-conditioned.

After the addition of the first sentence to the context the new context, c_1 contains a new set $W_{c_1} = W_c + w_d$, where w_d is a world that models my dream. The second sentence is interpreted with respect to the new context, c_1, and therefore it may be interpreted as being asserted of w_0, the world in which the discourse occurs, or as being asserted of w_d, the world just introduced. In the former case the modal index of the MP of the second sentence is w_0, and in the latter it is w_d. The second choice is possible only because w_d has just been introduced, and therefore w_d is now in W_{c_1}. Under the latter indexing the second sentence is, in discourse structural terms, an elaboration of the first.[4] The point of interest for present purposes is that the MP of a matrix clause may be indexed to a world other than w_0 in case such a world has been recently introduced, and the discourse can be coherently interpreted

[3] I am ignoring now certain constraints concerning (counter)-factuality imposed by the morphology.

[4] Note that in this account, modal subordination is independent of sentential subordination. This is an advantage, given that discoursal subordination may occur even in the absence of complement taking expressions, as in (i) below:

as being about that world. Under this account then, the structure of the interpretation of (29) under the two indexings is identical. What differs is the value of the modal index of the MP (and, most likely, the other elements) of the second sentence. The claim here is that a recently introduced world may serve as the value of a free index of an expression in subsequent discourse giving rise to discoursal modal subordination. Actual choices are, of course, limited by pragmatic and discourse coherence factors. Intensional predicates and nouns may have discourse scope under this account because the worlds they introduce may serve as values to modal indices of expressions occurring in subsequent discourse. This approach to modal subordination in discourse is different from that proposed by Roberts (1989), and Poesio and Zucchi (1992), where discoursal subordination is reduced to sentential subordination at the level of DRS by the use of structural accommodation, a mechanism that copies structure. For reasons of space, the two approaches will not be contrasted here.

Unlike the modal index of MPs of matrix clauses, the modal index of the MPs of complements of intensional predicates and nouns, as well as the modal indices of the MP in modal quantificational structures, are bound. Thus, the modal index of the complement clause in (30) is necessarily set to w_J, the world introduced by the main predicate, as indicated informally in (30) by the modal indices subscripted to the clausal brackets.

(30) [John believes [that Mary is sick]$_{w_J}$]$_{w_0}$

The dependency of the modal index of complements on the world introduced by the expression the complement is an argument of is parallelled in the temporal realm by cases where the time reference of a complement is dependent on the time reference of the matrix.

In modal conditionals the Restrictor introduces a set of worlds, and the modal index of the MPs in the antecedent and in the conditional are bound to these worlds. (The fact that the lf in the Restrictor serves to identify the worlds in question, while the lf in the NS is interpreted as predicating something of them is the result of the way tripartite quantificational structures are interpreted.)

The fact that the index of the MP of the complement of an intensional predicate or noun is bound by the world(s) introduced by that predicate or noun follows from the semantics of these constructions. The matrix expression introduces a world or set of worlds, and the complement expresses a claim about these worlds. The modal index of the MP of the complement is fixed by the truth conditions of the the matrix predicate. (I am not concerned here

(i) There's a good movie on tv tonight. It's about this girl from Hungary who visits her cousin in New York.

with the question of whether intensional predicates and nouns introduce single worlds or sets thereof.)

In quantificational structures, the index of the MP of the NS is bound by the 'cases' introduced in the Restrictor because of the role of the Restrictor and the NS in the interpretation of quantificational structures: the former sets up a set of 'cases' and the latter expresses a claim about these cases. Again, the interpretation rule of quantificational expressions fixes the index of the main predication in the NS to the cases introduced by the Restrictor.

So far then, it has been claimed that predicative conditions as well as variables have modal indices whose values specify the worlds in which the predicate must hold of its arguments (in the case of predicative conditions), and the worlds from whose domain the variable must be given values (in the case of variables). The modal index of the MP of matrix clauses is free; its value can be set to w_0, the world in which the discourse occurs, or to some other previously introduced world. The modal index of the MP of arguments of intensional predicates as well as that of MPs in quantificational structures are bound. In the former case the modal index of the MP of the argument must be set to the world or worlds introduced by the intensional expression in question; in the latter case the value of the index is determined by properties of the immediate tripartite structure the MP is part of. The value of the modal index of the MP in the Nuclear Scope is determined on the basis of the interpretation of the Restrictor. The value of a bound index is determined by properites of its local context: the expression contributing the bound index is an immediate constituent of the expression contributing the value of the binder.

It will be assumed below that an evaluation index is free unless it is bound. The only cases of bound indices that will be relevant below are cases where the index of a constituent is fixed by the semantics of the expression of which it is an immediate part.

3.3 Scope of the DC

The facts concerning the scope of the DC are straightforwardly accounted for under the assumption that the evaluation indices of DCs are free. The freedom of the modal indices of the DC of noun phrases parallels the freedom of their temporal index argued for in Enç (1986).[5] Under present assumptions this amounts to claiming that the modal index of the DC of a noun phrase can

[5]There may well be noun phrase types whose DCs are bound to the parameters of their main predications. These would be noun phrases such as bare plurals, that have necessarily narrow scope with respect to expressions in higher clauses. In the present framework, restrictions on the scopal properties of noun phrases amount to restrictions on their valuations. Such restrictions may affect both the functions that interpret the variable contributed by the noun phrase, and those that interpret the DC, and they may have overt morphological or syntactic reflexes. Thus, the main claim of Farkas (1985) amounts to requiring the DC

be set, at least in princple, to any world in W_c, where c is the current context. In particular, the value of this index may be the same as that of the index of the MP or it may be some other world that is salient in the context. (This is exactly the same as the behavior of temporal indices in Enç's analysis.) Note that under this approach the DC of a noun phrase e_1 has 'narrow scope' with respect to an intensional predicate or modal e_2 iff the modal index of the DC is given as value the worlds introduced by e_2. The DC has 'wide scope' with respect to a modal or an intensional predicate iff the value of its modal index is not set to the worlds introduced by them. The two readings of (1) would have their indices set to the values given in (31) and (32):

(31) [John believes [that [a friend of mine]$_{w_J}$ is a crook]$_{w_J}$]$_{w_0}$ (narrow scope)

(32) [John believes [that [a friend of mine]$_{w_0}$ is a crook]$_{w_J}$]$_{w_0}$ (wide scope)

The choice of w_J in (31) is available because the DP occurs within the complement of *believe,* which introduces w_J; the choice of w_0 in (32) is available because the world of the discourse needs no special introduction.

Under this view, the two readings have representations that are structurally identical; they differ in the values of the evaluation indices of the DC, and not in the structural position of the DC relative to the predicate *believe* at either the syntactic or the semantic level. The fact that the DC occurs within the argument of *believe* makes the choice in (31) possible but not necessary.

3.4 Scope of the variable

Recall that the question of the scope of the variable as defined here concerns the possible dependency between assignment functions for one variable x_n on varying assignments for another variable x_m. In case of variable dependency x_n is evaluated with respect to assignment functions that co-vary with the assignment functions evaluating x_m. Thus, in (33),

(33) Every student speaks an Indo-European language.

the two interpretations result from the option of evaluating the variable contributed by the indefinite by the base assignment function f, giving the wide scope reading, or by the set of functions F_{NS} which extend the functions in the set F_R, giving the narrow scope reading. In the former case the value of the variable that is relevant to the interpretation is that given to it by f because

contributed by subjunctive relative clauses to be interpreted by a function that ranges over a set of worlds.

neither the functions in F_R nor, crucially, the functions in F_{NS} extend f wrt this variable. In the latter case the relevant values of the variable are given by the functions in F_{NS} because now these functions extend those in F_R wrt this variable. In the narrow scope reading, the indefinite co-varies with the distributive, while in the wide scope reading it does not.

I propose that the function (or set thereof) which determines which values of a variable are relevant appears in lf as an evaluation index on the variable on a par with modal indices. The choices are determined by the position and role of the variable in lf, which in turn are dependent on morphological and surface structure considerations. The function index of a variable contributed by a quantificational DP is bound: it has to be the set of functions F_R. The function index of a variable contributed by a garden-variety indefinite is free. For a variable contributed by an indefinite in the NS the available options are (i) the base function f, which, just like the world of discourse, is always an option for a free function index, and (ii) the functions in the set F_{NS}. This latter option is available because the variable in question is in the NS in a nominal quantificational structure. The functions in F_R and F_{NS} are made available because they are introduced by the rule that interprets quantificational structures. I am assuming therefore that the base function f is always in the set of accessible evaluation functions F_c, and that at the point of evaluating the Restrictor, the set F_R is added to F_c, and at the point of evaluating the NS, F_{NS} is added and F_R is taken out.

The two options for (14) are represented as in (34a) and (34b):

(34) a. Restrictor: $\forall x_1$ Nuclear Scope:
x_1 F_R | x_2 f
DC_1: $student'(x_1)$ | DC_2: $\textit{I-E language}'(x_2)$
MP: $speak'(x_1, x_2)$

b. Restrictor: $\forall x_1$ Nuclear Scope:
x_1 F_R | x_2 F_{NS}
DC_1: $student'(x_1)$ | DC_2: $\textit{I-E language}'(x_2)$
MP: $speak'(x_1, x_2)$

The effect of having the value of the function index of x_2 be f in (34a) is the same as having x_2 be bound by text-level existential closure; the effect of having the value of the function index set to F_{NS} is the same as having the variable be bound by NS-level existential closure. The truth conditions of (34a) and (34b) are as in (35):

(35) a. The lf in (34a) is true in w wrt M iff there is an assignment function f such that every assignment function f_R that extends f wrt x_1 such that $f_R(x_1) \in V(student')$ has a (trivial) extension f_{NS} such that $f(x_2) \in V(\textit{I-E language}')$, and $\langle f_{NS}(x_1), f(x_2)\rangle \in V(speak')$.

b. The lf in (34b) is true in w wrt M iff there is an assignment function f such that every assignment function f_R that extends f wrt x_1 such that $f_R(x_1) \in V(\textit{student}')$ has an extension f_{NS} that extends f_R wrt x_2 such that $f_{NS}(x_2) \in V(\textit{I-E language}')$ and $\langle f_{NS}(x_1), f_{NS}(x_2)\rangle \in V(\textit{speak}')$.

(Modal indices have been ignored above.)

In both (35a) and (35b), $f_R(x_1) = f_{NS}(x_1)$; in (35a) $f(x_2) = f_R(x_2) = f_{NS}(x_2)$, while in (35b), the values assigned to x_2 by the functions in F_{NS} covary with the functions in F_R, and are possibly different from the values assigned to this variable by f and the functions in F_R.

It is crucial to note that the representations of the two readings are structurally identical; the only difference concerns the value of the function index on x_2. In structural approaches, the choice of evaluation function for a variable is unambiguously determined by the structural position of the DP contributing the variable at LF, or the structural position of the variable at the level of DRS. The same is true for the value of the modal parameter. In the approach proposed here the value of these parameters is encoded directly in the semantic representation in the form of values for evaluation indices. Structural considerations are relevant only in determining what options are available. The approach to scope developed here is closer in spirit to dynamic logic than to DRT since it makes no use of the structural notions made available in DRT. The question of how anaphora puzzles that initially motivated the structures proposed in DRT can be treated in the present approach is outside the scope of this paper.

In examples involving modal quantification the cases introduced by the Restrictor are worlds which must serve as values to the modal index of the MPs of both Restrictor and NS. The modal index of variables contributed by indefinites in either the Restrictor or NS is free and therefore these worlds may or may not serve as values of the modal index of variables in the Restrictor or the NS. If they do, the assignment function giving values to that variable will range over the worlds introduced by the Restrictor resulting in narrow scope readings. If they don't, the variable will be rigid with respect to these worlds resulting in wide scope readings. In the wide scope reading of the variable contributed by the indefinite in (36),

(36) John must write about a French philosopher.

the variable is interpreted by $f(w_0)$; in the narrow scope reading, the variable is interpreted by $f(W_D)$, where W_D are the worlds quantified over by the modal. In the latter case the indefinite refers non-rigidly since the extension of f at different worlds may vary. In the former case the indefinite refers rigidly.

3.5 Conclusion

In the approach proposed here wide and narrow scope readings for both the DC and the variable are distinguished by evaluation index values and not by differences in the hierarchical position of the relevant expressions at LF or DRT. Consequently, neither structural level needs to be tampered with in order to account for scope facts, by QR in the case of LF, and by allowing variables contributed by indefinites special freedom with respect to where they can be entered in the case of DRT.[6] The freedom of scope of variables and DCs contributed by garden variety indefinites is a result of the fact that their evaluation indices are free. This in turn is a consequence of the assumption that this is the default state of indices associated with ordinary indefinites. As mentioned in footnote 5, there are many types of noun phrases whose interpretation is constrained in various ways. Deictic noun phrases, for instance, must be interpreted by an evaluation function whose parameter is fixed to the context of utterance; this is why they are immune to variation, and therefore denote principal filters. Proper names are also immune to variation, and denote principal filters, because they behave like constants: all evaluation functions must assign the same value to the variable they contribute at all parameters of evaluation. Special indefinites, such as bare plurals, are constrained to have evaluation indices that are dependent on the indices of their main predication, which has the consequence of always giving them narrow scope with repsect to higher predicates. In this chapter I restrict my attention to garden variety indefinites. The next section points out some welcome consequences of the indexical approach regarding the scope domains of these indefinites, as well as of DPs whose D is *every*.

[6]For different proposals dispensing with (most) instances of QR, see Beghelli (1993) and Dobrovie-Sorin (1993). See also Beghelli and Stowell (1996) for a proposal in which scope is determined by a richly articulated LF built by movement driven by particular features on DPs. Here I am concerned with what the semantic import of these features would be, a matter that is crucial independently of whether one accepts this particular view of LF or not. In my approach the scope of garden-variety indefinites is free because they do not impose any restrictions on their evaluation functions apart from Heim's Novelty Condition, while DPs whose determiner is *every* affect scope the way they do because they contribute the Restrictor and Quantifier of a quantificational structure. In a feature driven movement approach, garden-variety indefinites have to be marked with a feature that allows them to move to a variety of positions, or not move at all, while universal distributive DPs have to be marked with a feature that forces them to move to particular position that ensures their interpretation as providing the key of a distributive predication.

4 SCOPE OF *EVERY* VS. SCOPE OF INDEFINITES

4.1 Data

I first review the empirical generalizations established in the literature and then show how they are accounted for under the present proposal. (For relevant discussion, see Farkas 1981, Fodor and Sag 1982, Ludlow and Neale 1991, and Abusch 1994.)

Based on the fact that (37a) has no reading in which officials co-vary with committee members, while (37b) has a reading in which a single high placed official is involved,

(37) a. A high placed official claimed that John talked to every member of the committee.
b. Every member of the committee claimed that John talked to a high placed official.

the following generalizations emerge:

A. *The scope of* **every** *and modal quantifiers is upward clause-bounded with respect to indefinites or other quantifiers.*

B. *The scope of indefinites is upward unlimited.*

The fact that modals pattern like *every* is shown by the interpretations of the examples in (38): the modal may not have scope over the variable contributed by the indefinite in (38a), *every* may not have scope over the modal in (38b), but the variable contributed by the indefinite may have scope over the modal in (38c).

(38) a. A student thinks that John must leave.
b. It is possible that every candidate will win.
c. It is possible that a candidate will win.

The first generalization prevents an indefinite from being dependent on a non-clause-mate modal or distributive that it commands, while the second generalization allows an indefinite to be independent even if commanded by a modal or a distributive. The issue here concerns the relation between the evaluation index values introduced by a quantificational structure and the evaluation index values of a variable contributed by an indefinite. An indefinite is within the scope of a DP whose determiner is *every* iff the evaluation index of the variable contributed by that indefinite is F_R or F_{NS}, where F_R and F_{NS} are the functions that interpret the Restrictor and the NS of the quantificational structure induced by *every*. Saying that *every member of the committee* in (37a) cannot have scope over the indefinite *a high placed official* amounts to the claim that the variable introduced by the indefinite cannot be indexed by the functions that interpret the Restrictor or the NS of the quantificational structure contributed by the quantificational DP and its sentence. Saying that the indefinite *a high placed official* in (37b) may have scope over the quantificational DP in the matrix amounts to the claim that the variable contributed by the indefinite does not *have* to be interpreted by F_{NS}.

In order to capture the generalization in A one has to prevent the evaluation index values introduced in a quantificational structure from being accessible as index values for variables contributed by expressions occurring in clauses that structurally command the quantificational expression. In order to capture the generalization in B one has to allow the evaluation index value of a variable contributed by an indefinite in the Restrictor or the NS to be independent of the values introduced by the quantificational structure.

A separate question concerns the scope domain of the DC of various types of indefinites and quantificational DPs. Here we will be concerned only with the scope domain of DCs of indefinite noun phrases and noun phrases whose D is *every* with respect to intensional predicates and operators. As mentioned above, the freedom of scope of the DC of a noun phrase may be constrained by the semantic nature of that noun phrase.

Now the DC of the indefinite in (39) can be assumed to hold either at w_0, or w_M, or w_J, where w_M and w_J are the worlds introduced by *think* and *believe* respectively:

(39) Mary thinks that John believes that a/every friend of mine is a spy.

This leads to the generalization in C:

C. *The scope of the DC of indefinites and DPs whose quantifier is* **every** *is upward unlimited.*

In order to capture this generalization the evaluation index value of the DC of an indefinite or an *every* DP must be allowed to be chosen from any previously introduced value.

4.2 Accounting for the generalizations A-C

Generalizations B and C follow from what has been said so far. The relevant assumptions are the following: (i) the DCs of garden variety indefinites and quantificational noun phrases are free; (ii) index values introduced in a matrix clause are accessible to expressions occurring in subordinate clauses. The first assumption is a natural one given that under the present approach the default case is for an index to be free. The freedom of the index of the DC of indefinite as well as definite noun phrases is independently needed to account for their interpretation in discourse; it is analogous to the freedom of their temporal indices discussed by Enç.

The second assumption is rooted in the notion that index accessibility involves a concept of priority that takes account of hierarchical structure. A view of priority according to which matrix clauses are prior to the clauses that serve as their arguments or to those that are arguments of their arguments is needed in accounting for anaphoric data as well. We will come back to the notions of accessibility and priority shortly.

The availability of widest scope readings in the present approach is the result of the base function and the base world being always present in I_c. The availability of intermediate scope readings is the result of having intermediate expressions introduce accessible index values. The preference for narrowest or widest scope readings occasionally noted in the literature may be due to the high degree of salience of the base world and base function, as well as that of the index values of the MP in which the variable is an argument.

The main advantage of this approach over an LF-based analysis is that the unlimited upward scope of DCs and variables contributed by indefinites is accounted for without having to posit unlimited QR (or NP-Preposing), and without having to assume long distance binding by text-level existential closure, as in Abusch (1994). This is a significant result since this type of unconstrained raising rule or binding relation has no analogue anywhere else in the grammar. In DRT, the unlimited upward scope of indefinites is captured by allowing the variable and its DC to be entered at the level of any superordinate box. The question this theory raises is why one cannot enter other types of information, such as main predications or various operators, at superordinate boxes as well.

Another welcome consequence of the approach proposed here is that it accounts straightforwardly for what turns out to be a paradox in structural acounts of scope, namely the fact that the DC of a distributive noun phrase is upward unlimited, while the scope of the distributive quantifier contributed by the noun phrase is upward clause-bounded. (See Farkas 1993 and Ludlow and Neale 1991 for discussion.) Consider (40).

(40) Mary thought that a witch claimed that every person in this room had contact with her.

Here *a witch* cannot co-vary with assignments given to the variable contributed by the distributive noun phrase. This is predicted by the present theory since the index contributed by the distributive is not accessible to the indefinite. The DC of the distributive, on the other hand, may have scope over both *claim* and *think.* This, again, is predicted since the scope of DCs is upward unlimited.

Note also that in the approach presented here, intermediate readings of indefinites are predicted and handled without any difficulty. A sentence of the type in (41), first discussed in Farkas (1981),

(41) Every library ordered every book that was written by a famous American linguist.

is correctly predicted to have a reading where the noun phrase *a famous American linguist* has scope over *every book* but is within the scope of *every library.* Under this reading the indefinite is interpreted by the set of functions in F_{NS} that interpret the NS of the quantificational structure contributed by *every linguist,* and not by those contributed by *every book.* In this case, *a famous American linguist* denotes a relative principal filter in the sense that its reference is rigid with respect to *every book* but is non-rigid with respect to *every library.* When this noun phrase has widest scope it is interpreted by the base evaluation function, and therefore it refers rigidly with respect to *every library* as well. When it has narrowest scope, it is interpreted by the functions in F'_{NS} that interpret the NS of the quantificational structure contributed by *every book.* A way of treating intermediate readings in DRT is given in Beghelli, Ben-Shalom and Szabolcsi (1996).

The present proposal predicts that the modal index of a variable and its DC may be set to different values. That this prediction is correct is illustrated by (40) above, as well as (42).

(42) I would be happy if somebody who is actually rich would have been poor.

(Similar cases are discussed in Farkas 1985 and Abusch 1994.) This example has a reading where the variable contributed by the indefinite is indexed to the worlds introduced by the Restrictor, thus allowing its value to vary from world to world, but the modal index of the DC of the indefinite must be set to w_0. Here then the variable has narrow scope with respect to the worlds introduced by the Restrictor, while its DC has wide scope with respect to them.[7]

The generalization in A can be captured by rendering the index values introduced in a subordinate clause inaccessible as index values of expressions

[7]There is a complication with examples such as

(i) John thinks that every friend of mine cheated.

which cannot be interpreted with the quantifier having wide scope and the description having narrow scope with respect to *think*: the sentence cannot be understood as quantifying over all and only those actual individuals who are friends of mine according to John, excluding those individuals who exist only in w_J and are friends of mine there. It may be that there is a constraint requiring the DC and its quantifier to share their modal index so that one doesn't have to look at several worlds when determining what constitutes a case for quantification.

in superordinate clauses. Crucial to the present account of limits on scopal domains is the way one defines accessibility.

Let us start with the most general definition in (43):

(43) A value v is accessible as a possible value to an index of an expression e_1 iff v has been introduced prior to e_1.

Let us assume now that at the discourse level priority is defined in terms of temporal sequencing. This allows any index value v contributed by an expression e_2 to be accessible to expressions in subsequent discourse. This assumption accounts for (29), as well as the other cases of modal subordination discussed in Farkas (1993). Constraints concerning discourse coherence and larger discourse structure will play a role in accounting for the various limits on the discoursal scope of various types of expressions.

At the level of complex sentences, let us assume that the constituents of a clause S_1 are prior to the constituents of a clause S_2 if S_1 c-commands S_2 in surface structure. Under these assumptions, the generalization in B is accounted for.

Finally, let us assume that clause-mates are simultaneous, i.e., that the index value introduced by an expression e_2 is accessible as a value to any expression e_1 that occurs in the same minimal clause as e_2. This predicts the possibility of 'inverse' scope in (44).[8]

---

[8]There are cases when command seems to matter even within a minimal clause. Thus (i) may not be interpreted with the indefinite within the scope of the distributive while (ii) can.

(i) John showed somebody/a student every picture.

(ii) John showed every picture to somebody/a student.

É. Kiss (p.c.) has also noted that the relative scope of clause-mate quantifiers is fixed by surface c-command as well. Thus, while the direct object in (iii) below may be interpreted as dependent on the subject, in the sense that the set of apples quantified over varies with each child, the opposite is not possible:

(iii) Every child ate every apple.

In the face of such facts one could either adopt a strict c-command based priority relation within the clause as well, and relax it only for the cases where inverse scope within a clause obtains, or one can adopt the permissive version suggested here and restrict it in the cases where inverse scope within a clause does not obtain. The former alternative owes an explanation for the instances when inverse scope is possible; the latter owes an explanation for those when it is not. Note also that in the present approach, the scope possibilities of various types of noun phrases depend directly on constraints on the functions that evaluate them. As mentioned before, proper names and deictics are not susceptible to variation, and therefore they will always have readings equivalent to widest scope, while those noun phrases whose indices must be bound by their MPs will always have narrowest scope readings; garden-variety indefinites on the other hand place no special conditions on their evaluation (apart from the novelty condition), which explains their scopal versatility. 'Cardinality' noun phrases such as *three books, three or more books,* and *at least twenty people* may be interpreted as introducing a set and giving information about its cardinality. When such noun phrases are within

(44) A proofreader read every paper.

Under these assumptions priority at the sentence level corresponds to c-command defined on clausal domains in surface structure.

The main advantage of the present approach to scope is that it accounts straightforwardly for the upwardly unlimited scope of DCs and indefinites both within a sentence and beyond it without having to resort to unbounded movement or unbounded binding. The unlimited scope of these expressions is the result of the availability of index values introduced in prior discourse. The selection of values for the evaluation indices of these expressions is reminiscent of antecedent selection for non-reflexive pronouns. In the case of expressions whose evaluation indices are bound, there is a local relation between the element introducing the value and the constituent whose index must be set to the value in question. For the cases discussed here the binding of the index is intimately connected to the interpretation of the construction the expression with the bound index is an immediate constituent of.

In approaches where scope is determined by hierarchical relations at LF the upward unlimited scope of DCs and indefinites can be accounted for either by

---

the scope of a quantifier, i.e., when they refer non-rigidly, because they are interpreted with respect to a set of evaluation functions, they give information about the cardinality of a set of sets. The preference of certain cardinality noun phrases for narrow scope when in non-subject position, noted originally by Liu (1990), might be due to the specific way in which information about cardinality is being given. Thus, 'imprecise cardinality' noun phrases such as ***at least 10 books, five or more books,*** etc., may tend to be interpreted as having narrow scope because they are the type of noun phrase one would use to describe a set of sets of varying sizes. I suggest therefore that the relative difficulty with which such noun phrases take wide scope over a quantified subject has to do with the type of cardinality information they are giving. Note that one cannot ban them from introducing a set, since they do just that in subject position in (iv),

(iv) More than fifteen/at least fifteen people came to the party.

One cannot ban them from taking inverse scope over a simple indefinite either, since (v) has a reading where the body guards co-vary with the officials,

(v) A body guard has been assigned to more than fifteen officials.

and finally, one cannot claim that they always have wide scope in subject position, because of examples like (vi):

(vi) More than fifteen students cannot fit in a classroom of this size.

The subject noun phrase here is most naturally interpreted as referring to the cardinality of a non-specific set. See Szabolcsi (1996b) for relevant discussion. Note that one cannot even ban such noun phrases from ever having inverse scope over a quantified noun phrase because of examples such as (vii).

(vii) Every member of our group wrote monthly letters to more than a thousand prisoners of conscience assigned to it by the central office.

Example (vii) has an interpretation according to which there is a set of more than a thousand prisoners of conscience such that every member of the group wrote monthly letters to each of those prisoners.

allowing unlimited QR (or NP-preposing), or by allowing text-level existential closure to bind a variable indefinitely lower, as proposed in Abusch (1994). Both alternatives involve an unbounded relation of a type that has no precedent elsewhere: either an unconstrained movement or an unconstrained binding relation between quantifier and variable.

Clause bounded quantifier scope follows in the present account from the assumption that quantifiers are not raised, or at least, not raised beyond their clause. The aspect of the proposal that permits an account of scope phenomena that does not rely on unbounded movement or unbounded binding is the conceptual separation of the issue of the scope of variables and their DCs from the issue of the scope of a quantifier. The latter problem concerns the availability of the index value introduced by the quantifier as a possible value for expressions occurring in previous or subsequent discourse. The scope of a quantifier is upward clause-bounded because the index value it introduces is not accessible to non-clausemate c-commanding expressions. The issue of the scope of variables and their DCs concerns the availability of index values introduced by quantifiers and other expressions to serve as values for expressions whose indices are free. The unlimited upward scope of such expressions follows from the assumption that index values in I_c, are accessible to expressions occurring in subsequent discourse.

5 CONCLUSION

5.1 Discourse scope

The approach to scope presented above predicts that intensional predicates, modals, and quantificational DPs have unlimited 'downward' scope: they introduce index values in I_c that are accessible for expressions occurring in lower subordinate clauses or in subsequent discourse. Discourse coherence factors, as well as morphological considerations may limit choices in certain cases but I will not be concerned with these issues here.

The phenomenon of modal subordination in discourse, mentioned briefly above, shows that this prediction is correct with respect to modal and situational indices. The phenomenon of 'telescoping,' discussed in Roberts (1989), and in more detail in Poesio and Zucchi (1992), shows that the prediction is correct with respect to function indices introduced by quantificational DPs as well. The phenomenon is exemplified in (45), due originally to Barbara Partee:

(45) Every candidate walked to the platform.
He shook hands with the dean and got his diploma.

In the present account discoursal subordination occurs when an expression (in this case the pronoun in the second sentence) is given an evaluation index introduced by an expression occurring in a sentence occurring in previous discourse (in this case the functions evaluating the variable contributed by the quantificational DP). The problems of accounting for limits on the discourse scope of various types of indices, and of contrasting this approach to Poesio and Zucchi's treatment remain beyond the scope of this paper.

5.2 Extra wide scope

It has been noted in the literature that *each* and less easily, *every,* may take scope over an indefinite in a higher clause.[9] Thus, at least for some speakers, the indefinite in the matrix clause in (46) may receive a non-rigid interpretation with respect to the cases identified by the Restrictor in the subordinate clause, namely the relevant invited speakers.

(46) A/some student made sure that each/every invited speaker had a ride.

For an account compatible with the theory of scope presented here, see Farkas and Giannakidou (1996). Here I will only make some empirical observations. First, while the scope of these distributives may cross one sentence boundary, it may not cross more than one. Thus, the matrix subject in (47) cannot be dependent on the cases introduced by the distributive noun phrase:

(47) A/some student made sure that John arranged that each/every invited speaker has a ride.

Second, it appears that the nature of the matrix predicate influences the possibility of the distributive to take extra wide scope. Thus, speakers who accept the relevant reading of (46) cannot interpret (48) with the distributive having extra wide scope:

(48) A/some journalist or other said that each/every candidate won.

Finally, *each* seems to take extra wide scope significantly more easily than *every.* Other quantifiers such as *all, several, most,* or cardinal numerals do not seem to allow for this possibility.

The extra wide reading of the distributive in (46) is clearly problematic for the present approach, under the assumption that higher clauses are prior relative to lower clauses. Given the empirical observations above, however, one should certainly not abandon the fundamental observation formulated in generalizations A and B above, i.e., one should not conclude that the scope

[9]Thanks to Ivan Sag for reminding me of these cases. See Ioup (1977) for relevant discussion.

of distributives, just like that of indefinites, is upward unbounded. I suspect that the correct solution will make crucial use of the emphatically distributive character of *each* and, to a lesser extent, *every,* and cannot be reached without a deeper understanding of the nature of distributivity and its interaction with quantification.[10] Progress in this area will also shed light, I think, on the scopal properties of various types of plural noun phrases as well as of determiner types that have not been included in the present discussion, such as cardinals and amount denoting determiners. My main goal here is to show that an indexical treatment of both DC and variable scope successfully handles some essential properties at the level of the discourse and the complex clause, and therefore that it is worth pursuing with respect to the many questions that remain open.

REFERENCES

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics* 2(2):83–135.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In Amy J. Schafer (Ed.), *Proceedings of the North East Linguistic Society 23, Volume One*, 65–80, Amherst. GLSA Publications, University of Massachussetts.

Beghelli, Filippo, Dorit Ben-Shalom, and Anna Szabolcsi. 1996. Variation, distributivity, and the illusion of branching. In this volume.

Beghelli, Filippo, and Tim Stowell. 1996. Distributivity and negation. In this volume.

Condoravdi, Cleo. 1994. *Descriptions in Context.* PhD thesis, Yale University.

Dobrovie-Sorin, Carmen. 1993. On the denotation and scope of indefinites. Unpublished manuscript, Centre National Research Scientifique, Paris.

Enç, Mürvet. 1986. Toward a referential analysis of temporal expressions. *Linguistics and Philosophy* 9(4):405–426.

Farkas, Donka. 1981. Quantifier scope and syntactic islands. In Roberta Hendrik et al. (Eds.), *Papers from the Seventeenth Regional Meeting, Chicago Linguistics Society*, 59–66.

[10]Another possibility would be to capitalize on the presupposed nature of the domain of these two quantifiers (discussed in Moltmann 1994). This, however, is less promising since it does not distinquish *each* and *every* from *most.* A promising direction to pursue is that of developing the distinctions drawn between various types of universal quantifiers in Vendler (1967) and McCawley (1977).

Farkas, Donka. 1985. *Intensional Descriptions and the Romance Subjunctive Mood.* New York: Garland.

Farkas, Donka. 1993. Modal anchoring and NP scope. Linguistics Research Center working paper LRC-93-08.

Farkas, Donka. 1994. Specificity and determiner reduplication in Hungarian. Paper presented at the Predication and Specification Workshop, Santa Cruz.

Farkas, Donka, and Anastasia Giannakidou. 1996. How clause-bounded is the scope of *every*? To appear in *Proceedings of Semantics and Linguistic Theory VI.*

Fodor, Janet, and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5(3):355–398.

Gil, David. 1995. Universal quantifiers and distributivity. In Emmon Bach et al. (Eds.), *Quantification in Natural Languages, Volume 1*, 321–362. Dordrecht: Kluwer.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* PhD thesis, University of Massachussetts, Amherst.

Ioup, Georgette. 1977. Specificity and the interpretation of quantifiers. *Linguistics and Philosophy* 1(2):233–245.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Kaplan, David. 1979. On the logic of demonstratives. *Journal of Philosophical Logic* 8(1):81–98.

Kratzer, Angelika. 1979. Conditional necessity and possibility. In Rainer Bäuerle et al. (Eds.), *Semantics from Different Points of View*, 117–147. New York: Springer Verlag.

Kratzer, Angelika. 1980. The notional category of modality. In Hans-Jürgen Eikmeyer and Hannes Rieser (Eds.), *Words, Worlds, and Contexts*, 38–74. Berlin: Walter de Gruyter.

Liu, Feng-hsi. 1990. *Scope Dependency in English and Chinese.* PhD thesis, UCLA.

Ludlow, Peter, and Stephen Neale. 1991. Indefinite descriptions: in defense of Russell. *Linguistics and Philosophy* 14(2):171–202.

McCawley, James. 1977. Lexicographic notes on English quantifiers. In Woodford A. Beach et al. (Eds.), *Papers from the Thirteenth Regional Meeting, Chicago Linguistic Society*, 372–383.

McCawley, James. 1981. *Everything that Linguists have Always Wanted to Know about Logic but were Ashamed to Ask.* The University of Chicago Press.

Moltmann, Friederike. 1994. Intensional verbs and quantifiers. Unpublished manuscript.

Poesio, Massimo, and Alessandro Zucchi. 1992. On telescoping. In Chris Barker and David Dowty (Eds.), *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory (OSU WPL 40)*, 347–367. Columbus: The Ohio State University.

Roberts, Craige. 1989. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12(6):683–721.

Stalnaker, Robert. 1979. Assertion. In Peter Cole (Ed.), *Syntax and Semantics 9: Pragmatics*, 315–332. New York: Academic Press.

Szabolcsi, Anna. 1996a. Background notions in lattice theory and generalized quantifiers. In this volume.

Szabolcsi, Anna. 1996b. Strategies for scope taking. In this volume.

Vendler, Zeno. 1967. Each and every, any and all. In *Linguistics in Philosophy*, 70–97. Ithaca: Cornell University Press.

7

WEAK ISLANDS AND AN ALGEBRAIC SEMANTICS FOR SCOPE TAKING*

Anna Szabolcsi and Frans Zwarts[1]

Department of Linguistics
UCLA

[1]*Rijkuniversiteit Groningen*

1 INTRODUCTION†

This paper outlines a semantic approach to weak islands, a phenomenon that has traditionally been thought of as purely syntactic. Weak islands are environments that allow some, but not all, *wh*-phrases to extract:

(1) a. Which man didn't you invite?
b. * How didn't you behave?

(2) a. Which man do you regret that I invited?
b. * How do you regret that I behaved?

We propose that at least in a significant set of the cases the violation is semantic in nature. In agreement with É. Kiss (1992) and de Swart (1992), we informally characterize the role of interveners as follows:

(3) Weak island violations come about when an extracted phrase should take scope over some intervener but is unable to.

*Parts of this paper appeared in Szabolcsi (1992b). We had extremely helpful discussions with many colleagues, at UCLA and elsewhere, over the past two years. We wish to thank them all, especially F. Beghelli, D. Ben-Shalom, C. Dobrovie-Sorin, E. Engdahl, I. Heim, J. Higginbotham, J. Hoeksema, M. Krifka, L. Moritz, D. Oehrle, B. Partee, D. Pesetsky, B. Schein, D. Sportiche, E. Stabler, T. Stowell, and H. de Swart, and two anonymous NALS reviewers of an earlier version of this paper entitled 'Weak Islands and Algebraic Semantics.'

†This paper is reprinted from Natural Language Semantics 1 (1993, pp. 235–284). No attempt is made here to correct whatever global shortcomings it may have. However, a handful of footnotes are added by the alphabetically first author to enhance clarity. The new footnotes do not interfere with the original numbering; they are indicated by the traditional sequence of footnote symbols.

A. Szabolcsi (ed.), Ways of Scope Taking, 217–262.

(4) Harmless interveners are harmless only in that they can give rise to at least one reading of the sentence that presents no scopal conflict of the above sort: they can "get out of the way."

This paper's principal contribution consists in an algebraic semantic explication of scope taking. A scopal element SE as we will understand it is an item that can participate in scope ambiguities; of the cases in which a *wh*-phrase takes scope over some SE, we will restrict our attention to those which can be represented by letting the *wh*-phrase bind a variable within the scope of SE. The variable x can be of any logical type:[1]

(5) $\mathrm{WH}x[\ldots \mathrm{SE}(_{\alpha}\ldots x\ldots)]$

The specific semantics of SE does not tend to receive much attention in this context. But (5) is meaningful only if SE is given an appropriate argument. For instance, if SE is negation, the expression α must denote in a domain for which complements are defined. Whether α does so depends, to a great extent, on the semantics of WHx. This is the connection that we will explore.

(6) *Scope and operations*
Each scopal element SE is associated with certain operations (e.g., *not* with complements). For a *wh*-phrase to take scope over some SE means that the operations associated with SE need to be performed in the *wh*-phrase's denotation domain. If the *wh*-phrase denotes in a domain for which the requisite operation is not defined, it cannot scope over SE.

This approach requires at least a partial semantic analysis of extractees as well as interveners. For instance, to account for (1), it needs to be shown that *which man* ranges over a domain for which complements are defined, but *how* does not; further aspects of their meanings may remain obscure because they are not relevant. Or, to account for (2), at least one operation that is not defined for the *how*-domain needs to be identified in the factive context; further aspects of its meaning are not relevant. We will indeed propose such analyses for a few extractees and interveners, leaving others for further research.

Intuitively, we are assimilating the scopal failure in weak islands to the application of a numeral to a mass noun, which is unacceptable because counting cannot be defined for the mass domain:

(7) * six mists

[1]If F and G are (polymorphic) functions, we say that F can participate in scope ambiguities if for some G, $F \circ G$ and $G \circ F$ are not logically equivalent; see Keenan and Timberlake (1988). Generalized quantifiers, operators like negation, intensional verbs, etc. all fall under this definition. As regards scoping, the reference to binding a variable within the scope of SE merely serves the purpose of exposition and is not meant to commit us to any particular kind of representation.

This is to be distinguished from another type of scopal failure which is tied to particular syntactic configurations, for instance:

(8) a. Every man read few books.
   * 'There are few books such that every man read them'

   b. Few books were read by every man.
   'There are few books such that every man read them'

   c. Few books did every man read.
   'There are few books such that every man read them'

As Liu (1990) observes, part of the descriptive generalization is in semantic terms: downward monotonic quantifiers in object position do not take scope over the subject. But unlike the weak island cases, the reason here cannot be that the ensuing meaning would be incoherent, as it is in fact available in slightly different structures.

The present proposal builds on the results in Szabolcsi and Zwarts (1990, 1991), but also differs from it in fundamental ways. The earlier proposal was a rather direct semantic reinterpretation of the data underlying Rizzi's (1990) and Cinque's (1990) Relativized Minimality, and made the following main claims:

(9) *Island-escapers are individuals*
Wh-phrases that are sensitive to weak islands are the ones that range over partially ordered domains, rather than discrete individuals.

(10) *Weak islands and monotonicity*
Weak islands are environments in which the interveners between the *wh*-phrase and its trace cannot be composed into an upward monotonic function. The reason is that only upward monotonic functions preserve partial ordering.

It will be argued below that (9) is essentially correct, but interveners are to be characterized in terms of scope, rather than monotonicity properties, thus (10) is to be abandoned.

The discussion will be organized as follows. Section 2 reviews the core weak islands data, and outlines the accounts in Rizzi (1990) and Cinque (1990) on the one hand and in Szabolcsi and Zwarts (1990, 1991) on the other. Section 3 summarizes the monotonicity account and points out its problematic aspects, including some shared by Relativized Minimality. Section 4 proposes an alternative account in terms of scope. The present paper focuses on why non-individual *wh*-phrases do not take wide scope, cf. (3), with only a few remarks concerning (4). Section 5.1 outlines the connection between scope and algebraic operations, cf. (6). Section 5.2 presents detailed empirical justification for the individual vs ordered distinction, cf. (8), and Section 5.3 discusses its implementation in algebraic terms, and Section 5.4 predicative interveners. Section 6.1 introduces a novel set of data involving arguments of non-iterable predicates that support this account over ones in terms of discourse or thematic roles; Section 6.2 establishes a connection between event structure and whether the predicate denotes an ordered or an unordered set. Some formal details of scopal intervention are also spelled out. Section 7 is a brief conclusion.

2 WEAK ISLANDS: SOME FACTS AND TWO ACCOUNTS

2.1 Weak island facts and relativized minimality

Islands for extraction come in two varieties. *Strong islands* are absolute: they do not allow any *wh*-phrase to escape. Cinque (1990) argues that subject, complex NP, and adjunct islands belong here: the NP gap they may contain is an empty resumptive pronoun, not a trace. *Weak islands,* on the other hand, are selective: typically, phrases like *which man* can extract, but phrases like *why, how,* and *how many pounds* cannot. The cross-linguistically best known weak islands are infinitival/subjunctive/modal *whether*-clauses:

(11) a. Which man$_i$ are you wondering [whether to invite —$_i$]?

b. * How$_i$ are you wondering [whether to behave —$_i$]?

c. Welke man$_i$ heb jij je afgevraagd [of je —$_i$ moet
which man have you self wonder if you must
uitnodigen]?
invite
'Which man did you wonder whether you should invite?'

d. * Hoe$_i$ heb jij je afgevraagd [of je je —$_i$ moet
how have you self wondered if you self must
gedragen]?
behave
'How did you wonder whether you should behave?'

Extraction from embedded constituent questions is degraded or unacceptable for many speakers of English. In other languages these may either be strong islands (Dutch) or genuine weak islands (Hungarian):[2]

(12) ?/* Which man$_i$ are you wondering [who saw —$_i$]?

[2] Comorovski (1989) states, albeit without providing an explanation, that complements of *wonder*-type verbs constitute absolute islands. Her claim is at variance with both standard literature and our own judgments.

(13) * Welke man$_i$ heb jij je afgevraagd [wie —$_i$ gezien heeft]?
which man have you self wondered who seen has
'Which man did you wonder who saw?'

(14) Melyik embert$_i$ találgattad, [hogy ki látta —$_i$]
which man-Acc guessed-you that who saw
'Which man were you wondering who saw?'

Although the variation in (12) through (14) is not well-understood, we will follow standard practice both in assuming that the strong islandhood of certain *wh*-complements is syntactic in nature and in restricting our attention to examples that qualify as weak islands in the given dialect or language.

Drawing from work by H. Obenauer and J. Ross, Rizzi (1990) and Cinque (1990) observe that the same kind of selectivity is exhibited by many further environments: the presence of *beaucoup* 'a lot' in French, negation or negative quantifiers, *only*-phrases, adversative and factive predicates, and extraposition all create weak islands:

(15) a. Quel livre as-tu **beaucoup** consulté —?
what book have-you a lot consulted
'What book did you consult a lot?'

b. * Combien as-tu **beaucoup** consulté — de livres?
how-many have-you a lot consulted of books

(16) a. Which man did**n't** you/did **no one** think that I invited —?

b. * How did**n't** you/did **no one** think that I behaved —?

(17) a. Which man did **only John** think that I invited —?

b. * How did **only John** think that I behaved —?

(18) a. Which man did you **deny/regret** that I invited —?

b. * How did you **deny/regret** that I behaved —?

(19) a. Which man **was it a scandal** that I invited —?

b. * How **was it a scandal** that I behaved —?

Compare the following good *how*-extraction:

(20) How did **everyone/two men** think that I behaved —?

They propose the following uniform explanation for the contrasts in (10) through (20):

(21) *Referential wh*-phrases can be long-distance linked to their traces via referential indices; *non-referential wh*-phrases need to be linked to their traces via an (antecedent-) government chain.

(22) The government chain between a non-referential *wh*-phrase and its trace is broken

(i) by certain *interveners,* or

(ii) if the clause from which we extract is not sister of a theta-marking [+V] head.

(23) Referential *wh*-phrases are those that both bear a *thematic role* like Agent, Patient, etc. and are *Discourse-linked*; non-referential *wh*-phrases are those that bear a role like Reason, Manner, Measure, etc. or are not D-linked.

The majority of the weak island effects is attributed to (22i). What interveners break the government chain between the *how*-type phrase and its trace? Rizzi's answer is in terms of syntactic positions. Developing the theory of Relativized Minimality, he argues that since the extracted *wh*-phrase is in an A-bar specifier position, all and only intervening *A-bar specifiers* count as relevant interveners. Rizzi analyzes *whether, who, beaucoup, not, no one, only John* and *deny* as A-bar specifiers, at S-structure or at LF. In contrast, he points out that *everyone* or *two men* acquire their scope by adjunction according to May (1985), so they are predicted not to block non-referential extraction. Cinque adds that factive and extraposition islands are due to (22ii).

As regards referentiality, Rizzi draws the crucial line between those phrases that refer to *participants in the event* and those that do not; the latter are claimed never to be able to escape from weak islands. Drawing from Pesetsky's (1987), Comorovski's (1989), and Kroch's (1989) work, Cinque adds that even event participants have to be D-linked, i.e., "refer to specific members of a preestablished set," to be referential. Phrases differ in their ability to admit of a D-linked interpretation, so a scale is predicted:

(24) a. Which man do you regret that I saw —?

b. ? Who do you regret that I saw —?

c. ?? What do you regret that I saw —?

d. ?? How many books do you regret that I saw —?

e. * How much pain do you regret that I saw —?

f. * Who the hell do you regret that I saw —?

2.2 Recasting relativized minimality in semantic terms

Szabolcsi and Zwarts (1990, 1991) and Szabolcsi (1991)—henceforth Sz and Z—accept the above empirical generalizations and propose to reinterpret them in semantic terms. The main claims are as follows.

The distinction between good extractees and bad extractees can be characterized in denotational terms. Good extractees range over *individual* domains, bad ones over domains whose elements exhibit a *partial ordering* (a reflexive, transitive and antisymmetric relation; paradigmatically: inclusion). On their primary use, properties, amounts, manners, etc. belong to partially ordered domains. The term 'individual' is used to refer both to inherently discrete entities like John or Mary and to contextually individuated properties, amounts, etc.; individuation means that we expressly choose to ignore their overlaps.

The characterization of weak islands can be given in terms of the *monotonicity* properties of the items intervening between the extractee and its trace. Downward monotonic and nonmonotonic interveners block the extraction of non-individuals; upward monotonic ones are harmless. The connection lies in the fact that only upward monotonic environments *preserve* partial ordering. Since individuals are not ordered, they are not interested, so to speak, in whether ordering is preserved: they must be insensitive to weak islands. Non-individuals are ordered, so they can naturally require that the structure of their domain be preserved between the extraction site and the landing site.

These claims can be implemented in a grammar whether or not it has movement and traces. For instance, they can be expressed as a condition on wh-*trace* relations. Or, they can be implemented in a categorial grammar that handles extraction using *function composition*:[3]

(25)
$$\dfrac{\begin{array}{c}\text{How much milk}\\ S/(S/NP)_{\mathrm{MON}\uparrow}\end{array} \qquad \dfrac{\text{did(*n't) you drink}}{(S/NP)_{\mathrm{MON}\uparrow}}\ \text{compose}}{S}\ \text{apply}$$

Assume that *how much milk* is marked to apply to an expression of category S/NP only if it denotes an upward monotonic function. This assumption is methodologically analogous to (in fact, is inspired by) Zwarts's (1986) claim that negative polarity items must be arguments of downward monotonic functions. Categorial grammar assembles form and meaning simultaneously. Since monotonicity properties are inherited under composition, *did you drink* will be

[3]If the extracted phrase is an adjunct, a functor looking for it is created by lifting the category to be modified by the adjunct.

upward monotonic, whereas *didn't you drink* will inherit downward monotonicity from *n't*.[4]

In the following sections we discuss the empirical motivation for the monotonicity claims in some detail, and then go on to point out its problematic aspects. The individual vs. ordered distinction will be essentially maintained in the revised proposal; its empirical as well as algebraic elaboration is relegated to Sections 5.2 and 5.3.

3 WEAK ISLANDS AND MONOTONICITY

3.1 Summary of claims

Szabolcsi and Zwarts observe that the environments Rizzi and Cinque characterize as weak islands share some simple monotonicity properties: they are all either *downward monotonic* or *nonmonotonic.*

(26) a. A function f is upward monotonic iff for every A, B in its domain, if $A \subseteq B$, then $f(A) \subseteq f(B)$.

b. A function f is downward monotonic iff for every A, B in its domain, if $A \subseteq B$, then $f(A) \supseteq f(B)$.

c. A function f is nonmonotonic iff neither (a) nor (b).

Let us briefly review how the material in Section 2.1 fits these notions. *Not, no one,* and *deny* are clearly downward monotonic; by the same token, we predict that *few men* and *at most five men* also create weak islands. *Wh*-phrases, factives like *regret, only*-phrases, and *beaucoup* 'a lot' are analyzed as nonmonotonic. Since some of these items are focus-sensitive, we try to keep the focus structure of the examples constant:[5]

(27) a. [I know the answer to the question] who/whether he exercises $\not\leftarrow$ & $\not\rightarrow$

[4]The combinatory grammars in Steedman (1987) and Szabolcsi (1992a) have nothing to say about island constraints. To remedy this, Hepple (1990) introduces boundary modalities and what may be called a calculus of opacity. But he makes no empirical claims concerning what domains will be opaque for what relations, and why. The present paper tries to argue on empirical grounds that some of the island constraints are semantic in nature. It remains to be seen whether boundary modalities can encode the semantic generalizations or become, at least in this case, superfluous.

[5]Some comments on (28) and (30). (28) is clearly invalid in the b $\rightarrow$ a direction. The a $\rightarrow$ b direction may be tempting, but (b) has a more specific presupposition than (a), whence it cannot be entailed by (a). Some factives like *know* are upward monotonic if taken extensionally. See Ladusaw (1980) on both points. In (30), the non-monotonic analysis of *beaucoup, a lot,* etc. is inspired by Westerståhl (1985), who proposes four interpretations for *many,* two of which are non-monotonic due to context-dependence. Suppose John does nothing but push-ups for exercise. What he does may count as a lot of push-ups but not as a lot of exercise, if the norms associated with the two are different. De Swart (1992) points

b. [I know the answer to the question] who/whether he does push-ups

(28) a. John regrets that I exercise $\not\leftarrow$ & $\not\rightarrow$
b. John regrets that I do push-ups

(29) a. Only John exercises $\not\leftarrow$ & $\not\rightarrow$
b. Only John does push-ups

(30) a. John exercises a lot $\not\leftarrow$ & $\not\rightarrow$
b. John does push-ups a lot

By the same token, we predict that *exactly five men* and *often,* etc. also create weak islands. On the other hand, items like *think, John, everyone, two men,* etc., which do not create weak islands, are *upward monotonic.* (It is difficult to find a good sample of extraposition islands that are not also factive islands; no proposal is made for them in Szabolcsi and Zwarts.)

This descriptive characterization avoids some analytical problems that arise on Rizzi's and Cinque's analyses. They include the movement of *deny,* a head, into an A-bar specifier position at LF and the assumption that the complement of *regret* is not a sister to the verb. These have an alternative solution within Relativized Minimality, however: the adoption of Progovac's (1988) and Melvold's (1991) proposals to place empty operators in the [SPEC, CP] of the complements of *deny* and *regret,* which then serve as standard interveners. More important perhaps is the problem posed by the cross-linguistic variation in the syntax of negation. Recent work has attributed the variation to the fact that the negative particle may be a head, a specifier, or an adjunct. This would suggest that the island-creating effect of negation varies accordingly, but it does not: we are not aware of any language in which negation does not create a weak island. Rizzi (1992) proposes to solve this problem by assuming an empty A-bar specifier when NEG is a head, and vice versa. But this solution makes the original claim almost vacuous; it seems more natural to us to trace back the cross-linguistically uniform effect to the uniform semantics of negation.

The most important question is why downward monotonic and nonmonotonic environments constitute weak islands. The definitions in (26) make it clear that upward monotonicity means simply that the function *preserves partial ordering*; downward monotonic functions *reverse* it and nonmonotonic ones *obliterate* it. Now recall that in the previous section we claimed that island-sensitive phrases are characterized by the fact that they range over a partially ordered domain. It seems entirely natural for such a phrase to require that

out that on this view *seldom* would be nonmonotonic, too, which contradicts its ability to license negative polarity items. But this may be more of a problem for NPI-theories than for us: *only John* and *regret* are also NPI-licensers and nonmonotonic.

order be preserved by the path connecting it to its extraction site. On the other hand, *wh*-phrases that range over individuals do not have a partial order in their domain. Hence they cannot possibly be sensitive to the preservation of order and must be immune to weak islands—which they are.

3.2 Problems

The problems with the above proposal come in two varieties: descriptive and conceptual.

(31) There are downward monotonic and nonmonotonic interveners that for many speakers do not create weak islands.

(32) There are upward monotonic interveners which do create weak islands.

(33) Two downward monotonic items in the path do not (regularly) cancel out.

(34) The explanation of why downward monotonic and nonmonotonic paths are islands is not as strong as it should be.

Let us consider these in turn.

First, Szabolcsi and Zwarts predict that all non-upward monotonic interveners are equally bad. But many speakers report a contrast between (35a) and (35b,c):

(35)	a.	* How did **few people** think that you behaved?	MON$\downarrow$
	b.	How did exactly five people think that you behaved?	$-$MON
	c.	How did at most five people think that you behaved?	MON$\downarrow$

Second, Szabolcsi and Zwarts predict that all upward monotonic interveners are harmless.[6] De Swart (1992) examines *combien*-extraction and Dutch *wat voor*-split, and observes that clearly upward iterative adverbs like *twee keer* 'twice' create as bad islands as downward monotonic ones. She also reanalyzes *beaucoup, veel* 'a lot' as upward monotonic; this may be a matter of debate, cf. note 5, but 'twice' alone is sufficient to establish her case:

(36) a. Wat voor boeken heb je **twee keer** gelezen?
what for books have you twice read
'What (sort of) books have you read twice?'

b. * Wat heb je **twee keer** voor boeken gelezen?
what have you twice for books read

[6] Szabolcsi and Zwarts (1991) has a chapter on 'gradience,' but its data are not built into the theory. We will return to this below. See also Philip and de Villiers (1992).

Similarly, Hegarty (1992) argues that the class of matrix predicates that constitute weak islands is not that of *factives* but, rather, Cattell's (1978) *response stance* and *non-stance* verbs, in distinction to *volunteered stance* ones. Dukes (1992) notes that several of the new island creators in (37)-(38) are upward monotonic:

(37) Response stance: *deny, accept, agree, confirm, verify, admit*

(38) Non-stance: *know, regret, remember, surprise, realize, notice*

(39) Volunteered stance: *think, believe, suspect, allege, assume, claim*

Third, the most natural implementation of Szabolcsi and Zwarts's proposal, as was mentioned in Section 2.2, is to assume that interveners between the *wh*-phrase and its trace are composed into one big function, each contributing its own semantic properties to the result. This predicts that examples containing two downward monotonic interveners are grammatical, since the composition of two downward monotonic functions is upward monotonic. Now, there is at least one case, (40c), where this is borne out:

(40) a. * John is our hero, as you deny.
b. * John is our hero, as no one knows.
c. John is our hero, as no one denies.
d. John is our hero, as you know.

Many of our informants report that they sense an improvement with *wh*-extraction, too, but it does not prove significant under closer scrutiny:

(41) a. * How did he deny that you behaved?
b. ?? How did no one deny that you behaved?

In view of this last observation one may choose to abandon the path-minded formulation of the hypothesis, and use monotonicity properties to characterize bad interveners. This, however, makes the explanation somewhat stipulative.

Fourth, Szabolcsi and Zwarts point out that the link between the partially ordered nature of sensitive extractees and the non-upward monotonic nature of weak islands is not as strong as it should be. The theory explains clearly why individuals *cannot* be sensitive to weak islands, and why non-individuals *can* be. But it does not explain why they *are* sensitive, i.e., exactly what goes wrong when partial ordering is not preserved.

Individually, these descriptive and conceptual problems are not devastating; they might be seen as calling for further research. Together, however, they indicate that the explanation is on the wrong track.

To see an important source of the problems, let us recall a crucial assumption of the Relativized Minimality theory (RM). The theory of LF that RM relies on is that of May (1985). According to this theory, structure (usually) does not disambiguate scope. (42), for instance, is assigned a single structure in which *how* is higher than *everyone,* but they govern each other, whence they can be interpreted in either scope order or even independently. The adoption of this theory for the purposes of RM results in the assumption that it does not matter which reading of the sentence we are considering; all we have to know is that *everyone* is in an adjoined position, whence its intervention between *how* and its trace must be harmless. (43) is also assigned a single structure, but *no one* occupies an A-bar specifier position in it, whence it must block *how-extraction.*

(42) How did everyone behave?

(43) * How did no one behave?

Szabolcsi and Zwarts followed RM in this respect. The claim that certain interveners hurt because, being A-bar specifiers, they break a government chain, was replaced by the claim that they hurt because non-upward monotonic paths do not preserve partial order—but the assumption that upward monotonic interveners *qua interveners* are harmless became part and parcel of the theory.

Results by É. Kiss (1992) and de Swart (1992) indicate that this assumption is wrong. In addition to pointing out the island creating effect of iterative adverbs (cf. 36) de Swart notes that sentences like (44) are potentially ambiguous, and they are ungrammatical on the narrow scope universal reading.

(44) Combien ont-ils tous lu de livres?
how many have-they all read of books
'For each of them, tell me what number of books he read'
* 'For what number, they all read that number of books'

Similarly, É. Kiss points out that (42) has only readings (a) and (b), but not the narrow scope universal reading (c):

(42) How did everyone behave?

a. 'For every person, how did he behave?'

b. 'What was the uniform behavior exhibited by everyone?'

c. * 'For what manner, everyone behaved in that manner?'

In retrospect, the conclusion that upward quantifiers are not harmless when they expressly take narrow scope had been anticipated in Szabolcsi (1983, 1986) and in Szabolcsi and Zwarts's (1991) chapter on gradience in the strength of islands. Because of the conflict with RM, however, the pertinent data were excluded from the core set on which the Szabolcsi and Zwarts account was based.

We leave the question open whether Relativized Minimality can be restated to cope with these data. The restatement would have to involve a modified concept of LF and/or a modified definition of interveners. See Dobrovie-Sorin (1992) and Beghelli (1993) for work in this direction.

4 WEAK ISLANDS AND SCOPE

In what follows we will assume that weak islands are a scope phenomenon. That is, we adopt the following informal version of É. Kiss's (1992) and de Swart's (1992) proposals as a point of departure:

(45) The weak island effect comes about when the *wh*-phrase should take wide scope over some operator but it is unable to.

(46) Harmless interveners are only harmless in that they can give rise to at least one reading of the sentence that presents no scopal conflict of the above sort: they can "get out of the way."

É. Kiss and de Swart present their proposals in terms of filters:

(47) "Specificity Filter: If Opi is an operator which has scope over Opj and binds a variable in the scope of Opj, then Opi must be specific" (in the sense of Enç 1991) (É. Kiss 1992).

(48) "A quantifier Q1 can only separate a quantifier Q2 from its restrictive clause if Q1 has wide scope over Q2 (or is scopally independently from Q2)" (de Swart 1992, p. 402).

In developing a formal semantic explanation, at least two questions need to be asked:

(49) Why are certain *wh*-phrases restricted in their scope-taking abilities?

(50) What interveners are able to "get out of the way," and how?

In the following sections we will focus on (49). An answer to (50) is to be developed in Szabolcsi (1996) and Doetjes and Honcoop (1996, Section 5.3), within the context of how scope behavior is determined by the meanings of the specific quantifiers. Before turning to (49), however, we provide a brief overview of some results in the literature that pertain to (50), and indicate their relation to the monotonicity hypotheses in Szabolcsi and Zwarts.

An intervener[7] is harmless iff (i) it is scopeless, or (ii) it can take wide scope over the *wh*-phrase (family-of-questions reading),[8] or (iii) it can participate in a scope independent reading with the given *wh*-phrase (branching, cumulative, etc. readings). Recall the analysis of (42), *How did everyone behave?*: (42a) is a family-of-questions reading, (42b) is an independent reading, and (42c) is the ungrammatical, narrow scope universal reading. The reason why (Relativized Minimality and) Szabolcsi and Zwarts's proposal could be descriptively almost correct is that typically, though not all and only, upward monotonic items have options (i) through (iii).

Let us first restrict our attention to quantifiers. The case of (i) is rather straightforward: Zimmermann (1991) shows that principal ultrafilters (e.g. names) are *scopeless.* As regards (ii), both Groenendijk and Stokhof (1984) and Higginbotham (1991) claim that all quantifiers that are not downward monotonic can give rise to a *family-of-questions* reading. These quantifiers have non-empty *minimal elements* (i.e., smallest sets S of individuals such that $S \in GQ$); the question is to be answered for each individual in some minimal element. Definites and, in general, universals, denote filters: they have a unique, not necessarily empty minimal element (e.g., in the case of $[\![$*the men*$]\!]$ and $[\![$*every man*$]\!]$, the set of men). Here we get the classical pair-list answers. Indefinites have as a rule more than one minimal element (e.g., the minimal elements of $[\![$*two men*$]\!]$ are all two-member subsets of the set of men). In this case the answerer has to choose one minimal element and give a pair-list answer for the individuals in it. Groenendijk and Stokhof call this a choice reading.

(51) Who did every man see?
Man$_1$ saw Mary, man$_2$ saw Susan, ... , man$_n$ saw Lynn.

(52) Who did two men see?
For instance, John saw Susan, and Bill saw Jill.

It is remarkable that according to Groenendijk and Stokhof, both *exactly five men* and *at most five men,* which were found problematic in (35), give rise

[7]The notion of an intervener must be defined so as to cover the type **Who didn't destroy this city?,* in Section 6.1, which shows that any item that crucially enters into the computation of an answer counts as an 'intervener,' even if it syntactically does not intervene between the *wh*-phrase and its trace. We leave open the question whether the definition is to be syntactic or semantic.

[8]The identification of the family-of-questions reading with the quantifier scoping over the *wh*-word is theoretically not unproblematic (see Engdahl 1985, Chierchia 1992. Nothing much hinges on this analysis in the present paper. We tentatively adopt this analysis here in part because it makes it easier to express this section's generalizations, and in part because it is supported by Hungarian. Hungarian lacks both the family of questions and the choice readings. Since the language disambiguates the relative scopes of quantifier and *wh*/focus phrases in surface structure, the absence of these readings can be rather straightforwardly attributed to the fact that quantifiers never take scope over WH in Hungarian (whatever the explanation should be). See Szabolcsi (1983), É. Kiss (1991).

to the choice reading (the latter does because it is supposed to allow for an upward monotonic group reading). Downward monotonic quantifiers do not support the family of questions reading, since their minimal element is empty.

These generalizations need significant refinement; for instance, they do not explain the observation (de Swart's and our own) that adverbs like *twice, a lot,* and even *always,* and modified indefinites like *at least two men,* do not give rise to family-of-questions readings. Another salient fact to be explained is that the family-of-questions reading is not available in every language (cf. note 8). But even in this preliminary form they provide a partial explanation of why downward monotonic interveners were found to create weak islands. As they do not give rise to the family-of-questions readings, at least one option to "get out of the way" is unavailable to them.

As regards (iii), three kinds of scope independent readings have been noted in the literature: *branching* (Barwise 1979), *cumulative* (Scha 1981), and intermediate ones (Sher 1991).

(53) Three students read two books. (branching)
'There is a set S of three students, and there is a set B of two books, and every member of S read every member of B'

(54) Three students read two books. (cumulative)
'There is a set S of three students, and there is a set B of two books, and every member of S read at least one member of B, and every member of B was read by at least one member of S'

Liu (1990, 1992) conducted an empirical study of what noun phrases participate in branching readings in sentences like (53). She identifies a subset of noun phrases denoting upward monotonic quantifiers; she calls them *G(eneralized)-specific.* These include definites, universals, and indefinites not modified by *at least, at most,* or *exactly*; *wh*-phrases are also among them. A branching analysis is always available whenever both noun phrases are G-specific. Questions that may be analyzed as exhibiting these readings are as follows:[9]

(55) How many circles did everyone draw? (branching)
'Everyone drew the same number of circles—how many was it?'

(56) How many circles did these two people draw? (cumulative)
'Altogether how many circles did these two people draw?'

[9]In examples like (55) the fact that uniformity is presupposed, rather than asserted, is at least as relevant as branching itself. See (110)–(11) for an analysis. The constraints on cumulative readings have not been yet been studied in descriptive detail, but for a thorough theoretical discussion, see Schein (1993).

In a chapter on *gradience,* Szabolcsi and Zwarts (1991) observed that downward monotonic interveners create the most robust weak islands, while Liu's G-specific noun phrases are the most innocuous, even among the upward monotonic ones. These observations are immediately explained once we think about weak islands in terms of scope. Downward NPs have only a narrow scope reading, whereas G-specific NPs have the greatest number of non-narrow scope readings.

Going beyond quantifiers, note finally that intervening scopal particles (NEG) and verbs (*deny, regret*) have no chance to "get out of the way."* The same holds for intervening *wh*-phrases: if *who* in **How do you wonder who behaved?* took matrix scope, the subcategorization of *wonder* would be violated.

Although much more work is needed to clarify the semantic conditions of scope interaction between *wh*-phrases and quantifiers, with this we take it that the global plausibility of the scope account is established.

5 SCOPE, OPERATIONS, INDIVIDUALITY

The rest of the paper is concerned with the question why certain *wh*-phrases cannot take wide scope and are thus sensitive to weak islands. To be able to address this question, we will first propose a way of looking at scope that allows us to infer (certain) scope-taking abilities from the denotational semantic properties of the interacting expressions. The general idea is this:

(57) *Scope and operations*†

Each scopal element SE is associated with certain operations (e.g., *not* with complements). For a *wh*-phrase to take wide scope over some SE means that the operations associated with SE need to be performed in wh's denotation domain. If the *wh*-phrase denotes in a domain for which the requisite operation is not defined, it cannot scope over SE.

More specifically, we will adopt the claim, advanced in Szabolcsi and Zwarts, that the crucial property that island-sensitive *wh*-phrases have is that they do not range over individuals, but we interpret this very differently from Szabolcsi and Zwarts, as follows:

(58) *Individuality and wide scope taking*

When a *wh*-phrase ranges over discrete individuals, these can be collected

*It was noticed in Williams (1974) that stressed negatives do not create weak islands, e.g., *How DIDN'T you behave?* Neither we, nor the literature we are aware of has an account of this fact.

†The discussion in 5.1 will show that reference to "the *wh*-phrase's denotation domain" is simple but not quite precise. See the main text and the footnote there.

into unordered sets. All Boolean operations can be performed on such sets. When a *wh*-phrase does not range over discrete individuals, only a smaller set of operations (possibly none) are available in its denotation domain, hence answers cannot be defined in the general case.

The discussion will proceed in the following steps. Section 5.1 outlines how (57) and (58) work in principle. Section 5.2 presents detailed empirical justification for the individual vs. ordered distinction. Section 5.3 discusses its implementation in algebraic terms, and Section 5.4 analyses some further interveners. Section 6.1 presents further empirical support for the relevance of this distinction. It will be shown that when some argument of a verb necessarily denotes a sum, it is affected by weak islands, however "referential" it may be in thematic role or discourse terms. Section 6.2 argues that whether a domain consists of sums or unordered sets depends on whether the predicate is iterable and summative in the pertinent respect. 6.2 also lays out some formal details of how answers are defined.

5.1 Scope and operations

Let us begin by asking what "taking wide scope" means (for present purposes, at least). Consider the following questions, on the wide scope *who* reading:

(59) Who did Fido see?

(60) Who didn't Fido see?

(61) Who did every dog see?

(62) Who did at least two dogs see?

We assume that the interpretation of questions, whatever it should precisely be, ensures that an exhaustive list is determined by the answer. We will be concerned with how such a list can be defined or verified. The novelty of our approach consists in presenting standard procedures in such a way that a connection is established between the denotational semantic properties of the interacting expressions and their scope possibilities. The Boolean interpretations of the scopal interveners in (60)–(62) are as follows (see Keenan and Faltz 1985, pp. 84, 229 for precise definitions):

(63) a. *Negation* corresponds to taking *complements.*

b. *Universal* quantification corresponds to taking *intersections.*

c. *Existential* quantification corresponds to taking *unions.*

d. *Numerical* quantification corresponds to a combination of *intersections and unions.*

In the light of these, we can explicate (59)–(62) as follows.

To answer (59), we form the set of people that Fido saw, and list its members. For (60), we form the complement of this set. For (61), we take for each dog the set of individuals that it saw, intersect them, and list the members of the intersection. If (62) had *at least one dog,* we would simply take the sets of individuals that each dog saw and union them. The presence of *two* makes life more complicated: we have to take a lot of intersections in order to determine whether the same individual shows up in at least two sets, and then union the results. These cases contrast with the family-of-questions reading of (61), for instance: pair-list answers do not require Boolean operations.

The moral is that for a *wh*-phrase to take *wide scope* over some scopal element SE means that the definition/verification of the answer involves specific *operations* associated with SE.

This definition is rather general. First, it does not require for the narrow scope SE to become referentially dependent on the wide scope taker, hence SEs like negation are covered. Second, it would easily extend from wide scope *wh*-phrases to arbitrary wide scope quantifiers (at least as a necessary condition).

A simple consequence of the above is that a particular *wh*-phrase is able to take scope over some SE only if the requisite operations are available in the domain the *wh*-phrase ranges over. In (60)–(62) this is no problem. *Persons that Fido saw* denotes a set of individuals. Individuals can be collected into unordered sets. An unordered set is one that has no partial ordering defined on its elements; either because it would be impossible to define one or because we choose not to define one. Unions, intersections, and complements are defined for sets of individuals; unordered sets naturally form Boolean algebras. But are these operations available in the domains of all *wh*-phrases?

Szabolcsi and Zwarts argued that the distinctive descriptive property of island-sensitive *wh*-phrases is that they do not range over individuals but, rather, elements of *partially ordered* domains. We return to the empirical justification of this claim in Section 5.2. At this point, let us simply consider the following theoretical possibilities:[10]

(64) a. A partial ordering is a reflexive, transitive, anti-symmetric relation—paradigmatically: inclusion.

b. A Boolean algebra is a partially ordered set closed under unions, intersections, and complements.

[10]We use the qualification 'proper' to indicate that if the definition does not require the presence of an operation, it is indeed not present. A structure is 'closed under' an operation if the result of applying that operation to any element(s) of the structure is always an element of the structure.

c. A (proper) lattice is a partially ordered set closed under unions and intersections (but not complements).

d. A (proper) join semilattice is a partially ordered set closed under unions (but not complements or intersections).

e. A (proper) partial order is a partially ordered set (not closed under either complements, or intersections, or unions).

Proper lattices, join semilattices, and partial orders are thus increasingly poorer structures than Boolean algebras. From our present perspective this means that if a *wh*-phrase (on a certain interpretation) ranges over elements of such structures, some or all Boolean operations are unavailable in its domain. Consequently, it is predicted not to be able to take scope over SEs whose definition involves at least one of the missing operations. Thus, it is predicted to be sensitive to weak islands created by such SEs.

In Sections 5.3 and 6.1 we will argue that each of these cases is represented by *wh*-phrases. Specifically, number expressions denote in lattices, while collectives, manners, and amounts in join semi-lattices.*

*The explication of (59)–(62) and the subsequent reasoning in the main text show that the wording of (57) is not quite precise. To appreciate this, let us compare the derivations of *(I wonder) who you didn't see* and **(I wonder) how you didn't behave.* The denotations of the *you didn't see* and the *you didn't behave* segments need to be computed first. *You saw* denotes the set of individuals that you saw. Then, *you didn't see* is expected to denote the complement of this set. Since sets of individuals form Boolean algebras, complementation is fine. Once we got this far, combining the result with *who* cannot be a problem. On the other hand, we argue that the denotation of *you behaved* should not be conceived of as the set of manners that characterized your behavior but, rather, as the manner that characterized your behavior. Then, *you didn't behave* is expected to denote the complement of this manner. Since we argue that manners form join semi-lattices, in which complementation is not defined, the denotation of *you didn't behave* cannot be computed, and the derivation cannot proceed any further.

These derivations should make clear that the operation associated with the narrow scope SE *not* is actually performed in the course of computing the denotation of that segment of the sentence that the overt *wh*-phrase is to combine with. What role does the nature of the *wh*-phrase play, then? The segment we are looking at contains a gap associated with the *wh*-phrase. What this segment denotes is determined by what this gap is and, by transitivity, by what the *wh*-phrase is. It is for this reason that in (57) we used the oversimplified (perhaps misleading) formulation that the operation needs to be performed in "the *wh*-phrase's denotation domain."

It is clear, then, that while individuals themselves are not structured and manners are, what counts here (as claimed throughout the paper) is the Boolean structure of the sets of individuals and the semi-lattice structure of manners.

Let us compare, in this light, *(I wonder) who you didn't see* with **(I wonder) who you didn't get this letter from. Who* ranges over individuals (of type *e*) in both cases, but over individuals of different algebraic structures. In Section 6 we argue that *you got this letter from* does not denote the set of those who you got this letter from but, rather, the collective that you got this letter from. (This collective may be atomic or plural.) The standard assumption is that collectives are individuals that form join semi-lattices; hence the impossibility of

It may be important to point out a difference between the roles this proposal and Szabolcsi and Zwarts assign to partial ordering. Take the example of idiom chunks. According to Rizzi (1990), their extraction is sensitive to weak islands because they do not have a referential index. If idiom chunks do not have any reference at all, not even of an abstract kind, then Szabolcsi and Zwarts made the wrong prediction here because such things cannot exhibit a partial order, and hence cannot be interested in its preservation. In contrast, the present proposal makes the correct prediction: idiom chunks do not refer to things that can be collected into unordered sets, whence the Boolean operations are not available for defining an answer. Partial ordering here is not the defining characteristic of island-sensitive extractees but, rather, the most typical case of lack of individuality.

Anticipating the empirical results, consider the following problem. Is it correct to insist that answers be laboriously "computed"? Instead, we could just look at every individual in our universe and check whether it exhibits the property of being seen by Fido, not being seen by Fido, being seen by every dog, and being seen by at least two dogs. Let us call this the "look-up" procedure. For look-up, the properties in (60) through (62) are as simple as the property of being seen by Fido: look-up does not really take cognizance of the fact that *who* is taking scope over some scopal element. Look-up is viable because we assume that each individual is a "peg," from which all its properties are hanging (cf. Landman 1986).

But this procedure cannot be general. For one thing, we certainly do not want to exclude the possibility of being able to "compute" even those things that can be looked up. On the other hand, not everything that we can talk about is a "peg." For instance, it is natural to look at the Fido-peg and find that Fido is loud and weighs twenty pounds—but it is not natural to have a loudness peg with the information that Fido is loud, or a twenty-pounds peg with the information that Fido weighs twenty pounds. (Unless, of course, we are dealing with a contextual individuation of particular weights.) This means that a question like *How much do at least two dogs weigh?* cannot be answered by looking at every weight peg and finding out whether it exhibits the property that at least two dogs have it. The answer has to be "computed" by manipulating information obtained by looking at dogs—and then the question whether the requisite operations are available is crucial.

We are convinced that "look-up" plays an important role in a pragmatic/procedural model (which it will be necessary to develop to account for further aspects of the weak islands phenomenon). But it does not eliminate the need for

complementation is expected. Notice that in this minimal pair predictions cannot be made by simply looking at the *wh*-phrase; we need to know what the predicate (and thus the gap) is. We thank P. Jacobson and D. Cresti for pointing out the need for these clarifying remarks.

"computation," and hence it does not eliminate the vulnerability of *wh*-phrases that denote in an impoverished domain.

5.2 Individuation: semantics versus pragmatics

Consider a sample of *wh*-phrases: (i) *which person(s),* (ii) *who,* (iii) *what, how many men,* (iv) *who/what the hell,* (v) *how many pounds, how much attention, how tall, how, why.* Although the majority of scholars working on the subject do not examine the full sample, there is agreement that the phrases in (i) and (ii) extract most easily, and those in (iv) and (v) least easily, from weak islands. Furthermore, there is agreement that various degrees of contextualization enable practically any *wh*-phrase, save for *why,* to extract. The question is what distinguishes good and bad extractees and, in particular, what role contextualization plays. The arguments to be put forth in this section are consonant with Szabolcsi and Zwarts but are significantly more elaborate.

We argue that the crucial distinction is between *wh*-phrases that range over individuals and those that do not. We use the term *individual* to refer both to entities like John and Mary that are inherently discrete and to those, typically higher order, objects whose natural overlaps and complements we expressly choose to ignore. It follows that individuals can naturally be collected into unordered sets (cf. Section 5.1); in fact, this is what we take to be their defining property. Non-individuals are then characterized by the fact that they exhibit a partial ordering and this ordering is indeed taken into account; or else they are strictly non-referential, e.g., idiom chunks.

In our view, contextualization (Discourse-linking) comes into play in two main ways: a salient checklist or relevance criterion (i) may individuate a naturally ordered domain, and/or (ii) may speed up the manipulation of an already individual domain by making "look-up" available. For instance, (i) is the case in (65a):

(65) a. What don't we have good supplies of? Just bread and juice.

Contextualization is necessary not only to allow us to exclude, say, fire engines and phlogiston from consideration, but also to free us from listing supercategories and subcategories of bread and juice that we do not have good supplies of. Here contextualization saves a potentially unanswerable question. Similarly in (65b), which is acceptable if we have a list of potential scores and receivers' names on the blackboard:

(65) b. How many scores did no one receive? (Answer: 22 and 27.)
'Which of the figures on the blackboard has no name next to it?'

On the other hand, (ii) is the case in (66) when *who* ranges over persons under discussion:

(66) Who did everybody support? The candidate from Ohio.

When possible, it is indeed much faster to check a finite set of "candidate pegs" and see which of them have the property of being supported by everybody than to construct the intersection of everybody's supportees, as was described in Section 5.1. Here contextualization merely makes a question more felicitous.

These may be regarded as classical cases of D-linking: *what, how many,* and *who* now range over members of some salient set. What we wish to stress here is that *what* on its property reading can only do this if we make the strictly semantic move of collecting properties into an unordered set, i.e., if we expressly ignore the partial ordering that is otherwise inherent to them. Similarly for *how many scores.* Our explanation of the weak island phenomenon rests on this semantic aspect of individuation.[11,12]

Perhaps the clearest evidence that ranging over individuals, rather than ranging over contextually salient items, is the critical factor in extraction is provided by Dobrovie-Sorin (1992), whose views on this matter are very similar to ours. She discusses three distinct interpretations of *how many*-phrases: amount (67a), non-D-linked individual (67b), and D-linked individual (68). D-linked human direct objects in Romanian are clitic-doubled, which is extremely helpful in distinguishing readings (67b) and (68). The contrast in (67a, b) shows that 'how many women' on the amount interpretation cannot extract from a factive island, but on the individual interpretation it can extract even if it is not D-linked, i.e., not clitic doubled:[13]

(67)	Cîte	femei	regreţi	câ	ai	iubit?
	how-many	women	regret-you	that	have	loved

[11]The present notion of individuals is the same as in Szabolcsi (1983), a discussion of the focusing of Hungarian bare singulars in Montague Grammar. Our references to answerability are intuitively very compatible with Comorovski (1989). But she makes use of it technically in a very different way than we do. Restricting her discussion of weak islands to extraction from embedded constituent questions, she claims that a sentence like *Who do you know who invited?* presupposes that everybody was invited by someone. The question is not answerable unless this presupposition can be checked; and it is not checkable unless *who* ranges over a set of known membership. Thus our cases (i) and (ii) are on a par for Comorovski, even though *who,* as opposed to property-*what,* is independently capable of ranging over individuals. Furthermore, even if the presuppositional analysis of questions is correct, it is not clear how Comorovski's theory would extend to all the weak island cases that we intend to generalize over. For a discussion of referentiality, see also Chung (1992).

[12]The existence of individual correlates of properties (cf. Chierchia 1984) does not seem to automatically immunize properties against weak islands, as was pointed out to us by Alessandro Zucchi.

[13]Dobrovie-Sorin (1992) makes the crucial distinction in terms of restricted versus non-restricted quantification. Caveat: Dobrovie-Sorin paraphrases (67) on the (a) reading using the phrase 'for what number.' We changed this because in Section 5.3 we will argue that numeral expressions have a 'numbers' reading, distinct from the 'amount' reading.

a. * 'For what amount of women, you regret having loved that amount of women?' (Answer: Three.)

b. 'How many women are there such that you regret having loved them?' (Answer: There are three such women.)

(68) Pe cîte femei regreţi câ le ai iubit?
prep how-many women regret-you that cl have loved
'How many [=which] of the women do you regret having loved?'
(Answer: Three of them, namely, A, B, and C.)

Other authors who identify individualhood as the crucial factor (although for somewhat different reasons) are Aoun (1986) and Frampton (1990).

In the rest of this section we will provide informal empirical support for the claim that the core examples of island-sensitive extractees can be described as non-individual (partially ordered), and that the behavior of wh-*the-hell* expressions is also accountable for without making crucial reference to D-linking.

Wh-phrases like *which person* can easily be taken to range over individuals (as can plural *which persons,* as long as the predicate is distributive; certain nondistributive cases will be taken up in Section 6.1). Both *who* and *what* can range over individuals. But *what* (and marginally even *who*) also ranges over properties, which are ordered; see above. *Why* requires a propositional answer, and propositions are ordered by entailment, a special case of inclusion.[14]

How many N phrases have an individual interpretation but also, like *how many pounds* and *how much attention,* an amount interpretation (cf. 67a). Amounts can only be made sense of in terms of an ordering. The individual vs. amount ambiguity of numeral phrases is highlighted by the presence or absence of copula agreement in Italian clefts (an observation we owe to Filippo

[14]In the eighties *why* was the paradigmatic example of island-sensitivity, but it seems to us that it is in fact rather atypical. Its extraction is blocked by a wider range of interveners than that of any other *wh*-phrase. For instance,

i. Why did at least three men leave?
'Why did three, rather than only two, men leave?'
* 'What reason did at least three men have for leaving?'

ii. Why did you want me to quit?
'What reason did you have for wanting me to quit?'
* 'What reason did you want me to have for quitting?'

Informally, we may say that why is "captured" by the closest "interesting" thing in its own clause. This seems true even of German *warum,* which differs from *why* in being able to remain in situ (T. Kiss 1991; H. van Riemsdijk, p. c.). For this reason we will avoid *why*-examples. We have no account of its peculiar behavior for the time being.

On the formal side, note that the Boolean algebra associated with the propositional calculus consists of equivalence classes of propositions (usually referred to as the Lindenbaum algebra). The calculus itself is not Boolean in nature.

Beghelli). The agreeing version (a) is insensitive to weak islands, while the non-agreeing version (b) is sensitive:

(69) a. Sono cinque donne che non ho invitato.
are five women that not have-I invited
'There are five women who I didn't invite'

b. * È cinque donne che non ho invitato.
is five women that not have-I invited
'The amount such that I didn't invite that many women is five'

In French, *combien*-extraction unambiguously invokes the amount interpretation, although it is not a necessary condition for it:

(70) a. **Combien de livres** as-tu beaucoup consulté?
how-many of books have-you a lot consulted
'How many books are there that you have consulted a lot' or
'How many of the books have you consulted a lot?'

b. * **Combien** as-tu beaucoup consulté de livres?
how-many have-you a lot consulted of books
'For what amount, you consulted that many books a lot?'

c. * **Combien de cercles** as-tu beaucoup dessiné?
how-many of circles have-you a lot drawn
'How many circles did you draw a lot? [OK if circle-types]'

We argue that manners, the domain of *how,* are also ordered; in particular, the components of the manner characterizing each event do not form a set but a sum.[15] This intuition can be corroborated by a test involving *only. Only* has two interpretations: 'exclusively' and 'merely'. The first applies to elements of unordered sets, the second to elements of ordered ones. They may differ in their syntax (see Harada and Noguchi 1992); some languages even have different words for them.[16] See (71) for German and (72) for Dutch:

(71) a. Johann war **nur** 1953 und 1958 in London.
John was only 1953 and 1958 in London
'John was in London only (= exclusively) in 1953 and 1958'

b. Johanns Sohn wurde **erst** 1990 geboren.
John's son was only 1990 born
'John's son was born only (= as recently as) 1990'

[15] A sum is a nonminimal element of a join semilattice; see the next subsection.

[16] Some claim that *alleen* in (72a) is an independent adverb (whereas *slechts* is part of the subject). But *Alleen drie mannen woonden de vergadering bij* 'Only three men were at the meeting' is fine, and *alleen* is part of an XP in first position.

(72) a. Er zijn **alleen** drie stoelen in de kamer.
expl is only three chairs in the room
'There are only three chairs (and nothing else) in the room'

b. Er zijn **slechts** drie stoelen in de kamer.
expl is only three chairs in the room
'There are only three chairs (and no more) in the room'

We observe that Dutch *alleen* means 'exclusively' and *slechts* 'merely'. They can thus serve to diagnose adverbs:

(73) a. * Hij heeft het probleem om 2:00 **alleen** elegant opgelost.
he has the problem at 2:00 only elegantly solved
'He solved the problem at 2:00 only [= exclusively] elegantly'

b. ? Hij heeft het probleem om 2:00 **slechts** elegant opgelost.
he has the problem at 2:00 only elegantly solved
'He solved the problem at 2:00 only [= merely] elegantly'

c. Hij heeft het probleem om 2:00 **slechts** met tegenzin
he has the problem at 2:00 only with reluctance
opgelost.
solved
'He solved the problem at 2:00 only [= merely] reluctantly'

d. Zijn hele leven, heeft hij problemen **alleen/*slechts**
his whole life has he problems only
elegant opgelost.
elegantly solved
'In all his life, he solved problems only [= exclusively] elegantly'

(73a) with *alleen elegant* is unacceptable because the components of the manner in which the problem was solved on a particular occasion do not form a set; *alleen elegant* cannot mean 'of all manners, only elegantly'. (73b) with *slechts elegant* is somewhat strange, since elegance is towards the high end of the scale; (73c) with *slechts met tegenzin* is fine, since reluctance is towards the low end. (73d) switches to a bare plural object, whence we have a plurality of problem-solving events. Each has a manner of its own, and these manners as wholes can be collected into a set. Here *alleen elegant* can be used: it means that the manner of every problem-solving was elegant. The judgments are the same for the English counterparts. There is a corresponding improvement in extractability:

(74) a. * In what way didn't you solve the problem at 2:00?

b. In what way did you never solve problems?

(74a) may be acceptable, too, if the manner domain is turned into an unordered set by the brute force of D-linking, i.e., by providing an explicit list of manners to check and to report on in the answer.

The next question to ask is whether there remain cases that make invoking D-linking truly indispensable. *Wh-the-hell* expressions are a good candidate. Since Pesetsky (1987) it has been assumed that they form minimal pairs with their plain counterparts in that they are "aggressively non-D-linked," whereas plain *wh*-phrases are D-linkable. They seem to make a strong case for D-linking since they extract markedly less well than their counterparts, even when they contain individual expressions like *who*:

(75) a. Who are you wondering whether to invite?

b. ?? Who the hell are you wondering whether to invite?

We wish to argue that D-linkability is not a minimal difference between *wh-the-hell* expressions and their plain counterparts. Consider the following pair:

(76) a. Who saw John on the way home?

b. Who the hell saw John on the way home?

Let us ignore the rhetorical or cursing uses of (76b). Even so, the contexts in which the two questions are usable are not the same. The existential presupposition *wh*-questions carry does not prevent (76a) from being an open question, readily answerable by *Nobody.* (76b) on the other hand can only be asked if we have unquestionable evidence that someone saw John, and merely wish to identify the person(s). The strength of this requirement is illustrated by a context we owe to Bruce Hayes. When asked what a felicitous use of *Who the hell saw his mother?* would be, he answered, "If we know that whenever someone sees his mother, God sends purple rain, then upon seeing purple rain, I can ask: Who the hell saw his mother?" Now, lacking institutions like purple rain, we typically do not have unquestionable evidence about the rather complex situations that weak island violations tend to describe, e.g., that you are wondering whether to invite a particular person (cf. 75b). This provides an explanation of why such questions are notoriously bad. On the other hand, in those special situations where we do have such evidence, the *wh-the-hell* expressions become acceptable; for example, seeing someone madly searching through the dictionary, we may ask (77); or, one thief, seeing another trying to smuggle an item back to a house just robbed, may ask (78):

(77) What the hell do you still not know how to spell?

(78) What the hell are you upset that you took?

Tetsuya Sano (p. c.) informs us that these intuitions are paralleled by the interpretation and behavior of *ittai*-phrases in Japanese. We interpret these data as indicating that D-linking is not the critical factor in the behavior of *wh-the-hell* expressions; they are bad extractees for independent reasons.

These remarks have been intended to support the claim that the crucial feature of island-escapers is semantic. It appears that discourse context never makes a minimal difference for extractability. D-linking plays an important role when it forces, and facilitates, the individuation of a domain that is originally not individuated; but it is the ensuing semantic change, the creation of an unordered set, that matters for extractability.[17]

5.3 Structures for manners, amounts, and numbers

We assume that on the individual interpretation of *who, what, how many men,* etc. these expressions range over elements of unordered sets, whether or not they are D-linked: they invite us to list, or count, the members of such sets. Their immunity to weak islands is accounted for with reference to the fact that all Boolean operations are defined for unordered sets. (79) below illustrates the structure of a tiny Boolean algebra for sets of individuals:

(79)

$$\begin{tikzcd}[arrows=dash]
 & \{\text{John, Mary}\} \arrow[dl] \arrow[dr] & \\
\{\text{John}\} \arrow[dr] & & \{\text{Mary}\} \arrow[dl] \\
 & \emptyset &
\end{tikzcd}$$

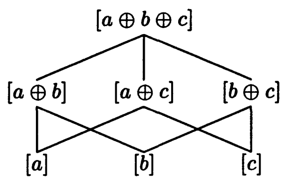
In this section we propose specific denotation domains for some island-sensitive phrases and show that they lack some or all of the Boolean operations. (The "domain" of idiom chunks trivially lacks the Boolean operations as they have no mentionable denotation at all).

The following structures will be considered; each is annotated with the kinds of expressions/readings we propose to assign to such a domain. The qualifi-

[17]There are significant cross-linguistic differences in the behavior of *wh*-phrases, which cautions against the careless use of "dictionary equivalents." Just two examples. English *which* is rather strictly D-linked; Dutch *welk(e)* and Hungarian *melyik* are much less so: they only require unicity. E.g., *In what year were you born?* is neutrally put as *Melyik évben születtél?* 'In which year were you born?' It is interesting to observe that *welk(e)/melyik*-phrases are just as good extractees as more D-linked *which*-phrases. Or, English *when* does not seem to be able to range over individuals even in D-linking contexts, whereas Hungarian *mikor* and Korean *encey* happily do so. Thus *Mikor$_i$ nem tudod, hogy kit kell meglátogatnod t_i?* and *ne-nun nwukwu-lul encey pangmwunha-eyaha-nunci al-ko sip-ni* 'When [= on what holidays] don't you know who you have to pay a visit?' are acceptable, in contrast to their English counterpart.

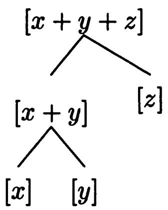
cation 'proper' is understood throughout; $\oplus$ stands for sum-formation, viz., semilattice unions:[18]

(80) Free Join Semilattice



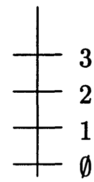
masses, collectives, manners

(81) Join Semilattice



amounts

(82) Lattice



numbers

The structure (80) has been proposed as the denotation domain of mass terms and plurals-as-collectives (see Landman 1991, pp. 254–267; 317–324 for a summary). We propose to add manners. Masses do not concern us in this paper; collectives will be explored in the next section. (81) and (82) have not received much attention in the literature. We argue below that numeral expressions on the amount reading denote elements of non-free join semilattices, whereas on the number reading (not yet discussed) they denote elements of chains, a special case of lattices.

The assumption of these structures leads us to predict that different extractees are sensitive to different interveners.[19] (80) and (81) have no bottom element (one that is smaller than any other in the structure). Hence they cannot be closed under complements (the complement of the top element $[a \oplus b \oplus c]$ or $[x + y + z]$ should be $\emptyset$) and under intersections (the intersection of two disjoint elements should be $\emptyset$). Both (80) and (81) are closed under unions, but only (80) is a fully articulated free structure. (Freedom means that whenever two pairs of elements are distinct, their unions are distinct, whereas in (81) $[y]$ and $[z]$ have no union distinct from $[x + y + z]$, for instance.) Thus *collectives, manners,* and *amounts* are predicted to be sensitive to weak islands created by SEs whose definition involves *complements* or *intersections,* but not to SEs involving just unions. (That is, if they turn out to be sensitive to the latter type, too, this must have an additional reason.) Finally, the chain in (82) is closed under unions (the least upper bound of 1 and 2 is 2) and, since it has a bottom element, it is closed under intersections as well (the greatest lower

[18] To make the discussion accessible to readers who are only familiar with the elements of set theory, we will talk about "union" and "intersection" even when technically, we should say 'join' and 'meet.'

[19] One might call this a truly semantic relativized minimality effect: the meaning of each extractee determines what interveners it is sensitive to.

bound of 1 and ∅ is ∅); it qualifies as a lattice. Note that it has no top element and cannot be closed under complements. Thus *numbers* are predicted to be sensitive to weak islands involving *complements only.*

Let us consider these cases one by one.

The fact that the join semilattice (80) is closed under unions expresses the cumulative reference property of masses and collectives. Section 6.1 will demonstrate that *collectives* are indeed sensitive to islands requiring complements or intersections. Do *manners* also denote elements of a join semilattice? We argued in Section 5.2 that they exhibit some partial ordering. It may be added that the ordering must be by unions, not intersections. If John behaved nicely but stupidly, his behavior is not one that has just the features common to niceness and stupidity but, rather, a behavior that subsumes both.[20] We have seen that *how*-extraction is sensitive to the standard weak islands, negative islands among them; moreover, that *how* does not take scope over a universal. To recapitulate the relevant observation in (42c), imagine that Bill behaved rudely and stupidly, Mary loudly and stupidly, and John nicely and stupidly; that is, everyone's complex behavior had stupidity to it. Under such circumstances, the question *How did everyone behave?* can nevertheless not be answered by *Stupidly.* This indicates that manners denote in (80) or in a mere partial order, the choice depending on whether we assume closure under unions or merely the existence of some unions. Intuitively, the sum of any two behaviors seems like a good candidate for being a more complex behavior, possibly including contradictory cases like *kindly and unkindly.* In accordance with this, questions in which *how* needs to take scope over a plain existential sound acceptable:[21]

(83) How did at least one person behave?

We will therefore assume that how denotes in a join semilattice (cf. 80).

Next, consider *how many (N).* Its individual reading (D-linked or not) has been discussed in Section 5.2; now we are concerned with its non-individual

[20]It is another matter whether the people who behave nicely but stupidly are in the intersection of those who behaved nicely and those who behaved stupidly. This depends on whether *behave* is taken to be distributive. If not, only a weaker relation will hold: that of having niceness to one's behavior. This weaker relation is comparable to a situation in which, if John and Bill lifted the table together, John did not lift the table but he participated in lifting the table.

[21]What is the answer to (83) in the situation sketched in connection with (42c)? ***Rudely and stupidly, loudly and stupidly, and nicely but stupidly*** may not sound acceptable. This may have an independent reason; in Section 6.2 we argue that the *behave*-relation is not summative. This problem can be avoided by answering, ***At least one person behaved rudely and stupidly, at least one loudly and stupidly, and at least one nicely but stupidly.*** Note that this is not a pair-list answer since it does not name the subjects (the first members of the pairs). If this reasoning is acceptable, manners indeed denote in a join semi-lattice. If not, we may assume, for instance, that some "contradictory" behaviors are impossible, whence we do not have closure under unions, and manners will be assigned some proper partial order. Nothing much seems to hinge on whether manners end up in (80).

readings. They come in two varieties: the well-known amount reading, as was given in (67a) for Romanian and (70b,c) for French, and what we will call the number reading.

(67) a. * Cîte femei regreţi câ ai iubit?
how-many women regret-you that have loved
'For what amount of women, you regret having loved that amount of women?' (Answer: Three.)

(70) b. * Combien as-tu beaucoup consulté de livres?
how-many have-you a lot consulted of books
'For what amount, you consulted that many books a lot?'

c. * Combien de cercles as-tu beaucoup dessiné?
how-many of circles have-you a lot drawn
'How many circles did you draw a lot? [OK if circle-types]'

Although the amount reading is well-known, there is no standard algebraic structure for it in the literature. The join semilattice discussed above does not seem to offer a way to capture the measuring aspect of amounts. The simplest alternative might be to turn to the chain of natural numbers. But that has too rich a structure: being a lattice, it lacks only complements. On the other hand, amounts seem sensitive to the intersections as well, cf. de Swart (1992):

(84) a. * How many circles did no kid draw?

b. * How many circles did every kid draw?

This shows that amounts denote in a poorer structure, possibly (81). Our argument now will proceed in two steps. We first argue that although (82) is not appropriate for the amount reading, it does correspond to another non-individual reading of numeral expressions. With this reading out of the way, we go on to justify the adoption of (81) for amounts.

It appears that there are contexts in which a non-individual *how many N* is able to take scope over universals and numerals. We may informally characterize these contexts as "counting-conscious." Suppose that we are evaluating how appropriate the midterm test was in comparison with the level of the class. We may then ask questions like,

(85) a. How many problems did every student solve?
'For what number, every student solved at least that number of problems?'

b. How many problems did at least 50% of the students solve?
'For what number, at least 50% of the students solved that number of problems?'

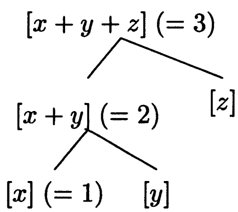
Here *how many problems* is not D-linked: it is not intended as 'which of these problems' or 'which of the numbers that we have listed'; nor is it meant as 'how many problems are there such that ... .' It asks for a purely numerical value; we will call it the *number reading.* Or, imagine a situation in which individuation is rather inconceivable: we agreed that the swimming team can take a break when everybody covers at least 50 laps. Feeling that the break is drawing closer, we may ask,

(86) [At least] How many laps has every swimmer covered by now?

One interesting aspect of these examples is that the narrow scope universal does not make them unacceptable. Another is that they require maximal answers. If every student solved 23 problems (but not everyone solved more) or every swimmer has covered 46 laps (but not everyone has covered more), the answerer cannot play it safe by answering *One/Ten*; the answers have to be *Twenty-three* and *Forty-six,* respectively. The question arises whether this is a semantic effect or a Gricean one. The adoption of the lattice structure (82) predicts that it is semantic. The narrow scope universal requires that we take intersections, which just gives the greatest lower bounds 23 and 46 in these cases.

In sum, the chain in (82) has linguistic relevance but, exhibiting the rich structure numbers have, is not appropriate as a denotation domain for island-sensitive *amounts.* What we need is a structure that resembles (82) in that it allows for an interpretation of measuring but is nevertheless not a chain. We argue that the structure in (81) may do the job. Below the nodes are annotated with $(= n)$ to highlight the intended interpretation:

(81′)



The "backbone" of (81) is a chain like (82). Formally, we may look upon (81) as a *witnessed* version of (82): if p is a proper part of q, there is some part of q (the witness) that does not overlap with p (Landman 1991, p. 314). The branching that the witness property guarantees is sufficient to eliminate closure under intersections, which is what we are aiming at.

But what is the intuitive content of (81)? We propose that (81) is an abstraction of (80). The elements $[a]$, $[b]$, $[c]$, etc., in (80) represent real stuff, therefore the sum of $[a]$ and $[b]$ needs to be distinguished from the sum of $[a]$ and $[c]$: even if they happen to have the same size, they have their own identity.

What we do in (81) is take away the identity of bits of stuff (we might say "individuality" in the everyday sense, were 'individual' not a technical term with a different meaning in this paper). Here $[x]$, $[y]$, $[z]$, etc., are all unit-sized, though they are not unit-sized bits of concrete stuff, but arbitrary—and therefore abstract—unit-sized bits. Fixing an arbitrary unit-sized $[x]$ to start with, $[y]$ stands for the equivalence class of all unit-sized bits of stuff whose addition to $[x]$ yields a two-unit-sized bit, and $[z]$ stands for the equivalence class of all unit-sized bits of stuff whose addition to $[x+y]$ yields a three-unit-sized bit. Thus amounts are construed as abstract bits of stuff. Being abstract and allowing for the definition of a scale, (81) resembles (82) more than it does (80). On the other hand, the witness property seems to capture what distinguishes amounts from numbers. (81) reflects the intuition that three cups of milk (or three men) is obtained by adding another cup of milk to two cups of milk (or another man to two men), rather than just moving higher on a scale.[22]

5.4 Operations for further interveners

So far we have primarily restricted our attention to scopal interveners that are straightforwardly Boolean (compounds). Increasing the descriptive coverage significantly would go beyond the scope of this paper; for instance, we do not present an analysis of the most famous of weak islands, i.e., *wh*-islands, although we believe that they belong here.* Some discussion of two specific cases may be methodologically interesting, however.

According to Cinque (1990), complements of factives are one paradigmatic type of weak islands. Recall, however, Hegarty's (1992) observation that the empirically correct class of predicates is, rather, one that comprises Cattell's (1978) response stance and non-stance verbs, in distinction to volunteered stance ones (on the intended readings):

(37) Response stance: *deny, accept, agree, confirm, verify, admit*

(38) Non-stance: *know, regret, remember, surprise, realize, notice*

(39) Volunteered stance: *think, believe, suspect, allege, assume, claim*

[22]We follow Krifka (1990) in taking *man* as a measure for men. We leave open the question exactly how amounts without canonical measures should be treated, e.g., *(how) much attention*.

*We might adopt Groenendijk and Stokhof's (1984) semantics for interrogatives, according to which in a world where John and Mary walk, *Who walks?* denotes $\lambda x[\text{walk}(x)] = \{\text{john}, \text{mary}\}$ (see Szabolcsi 1996 for a review, and Gutiérrez Rexach 1996 for an alternative way to get the same effect). Although Groenendijk and Stokhof do not analyze *who* as a quantifier, they point out that this analysis of the interrogative is equivalent to assigning universal force to the *wh*-phrase, cf. $\forall x[\text{walk}(x) \leftrightarrow x = \text{john} \lor x = \text{mary}]$. Thus *wh*-islands can be expected to be at least as bad as islands created by *everyone*.

Does this mean that under the present approach the accommodation of these facts requires an in-depth analysis of the meanings of these verbs? It does not: it is sufficient to identify one Boolean operation in their meanings that is not defined for the domain of sensitive *wh*-phrases; a circumstance that makes the present enterprise globally feasible.

Dukes (1992) presents a preliminary analysis in this spirit. He observes that a sentence with a factive matrix predicate can be paraphrased as follows:

(87) I regret that John left.
regret(I)(that John left) & fact(that John left)

According to this analysis, the proposition *that John left* is an argument of both the matrix verb and the sentential predicate *fact.* Following Molnár (1982) and Sántha (1980), this approach naturally extends to non-factive examples in (37) and (38). For instance, *fact* in (87) can be replaced by some predicate like *assumption* in (88):

(88) I confirm that John left.
confirm(I)(that John left) & assumption(that John left)

The relevant point here is that the paraphrase involves conjunction, viz. intersection. This may be identified as the Boolean operation that creates a weak island. On the other hand, there are no natural sentential predicates for complements of volunteered stance verbs; at best a tautological cognate can be found, in which case the conjunction is semantically irrelevant:

(89) a. I thought that John left.
think(I)(that John left) & thought(that John left)

b. I suspected that John left.
suspect(I)(that John left) & ???(that John left)

Therefore, the analysis of volunteered stance verbs does not necessitate this kind of conjunction, wherefore these contexts are predicted not to create weak islands.*

Another case that deserves mentioning is that of intensional verbs like *want* and *seek.* The standard assumption is that they are scopal elements. Nevertheless, they obviously do not create weak islands:

(90) a. How many circles do you want to draw?

b. How many unicorns are you seeking?

*D. Dowty (p.c.) points out that if all presuppositions are represented as conjuncts, we make a host of incorrect predictions. Moltmann's (1994) event-based analysis of attitude reports provides a framework within which the proposal in the main text can be naturally implemented and avoid this problem.

This is predicted by the current theory if we assume that the scopal properties of these verbs are not Boolean in nature—which seems correct. (Note that no theory that treats scope as a primitive can make the correct distinction here.)

6 ISLAND-SENSITIVE COLLECTIVES AND THE CONDITIONS FOR SET FORMATION

6.1 Unique arguments and weak islands

In this section we will discuss a set of extractees which have not been considered in previous literature and which, as far as we can see, share nothing else but the lack of Boolean structure with the standard items discussed so far, and are nevertheless systematically subject to weak islands. The distinction between iterable and 'one time only' predicates is familiar from the aspectual literature. For instance, *show this letter to Mary* and *get a letter from Mary* are iterable: it is possible to show the same letter (token) to Mary, or to get a letter from Mary, more than once. *Get this letter from Mary, burn this letter,* and *win the Rimet Cup in 1978* are 'one time only' predicates: it is not possible to get the same letter (token), or to burn the same letter (token), more than once; similarly for winning the Rimet Cup, a unique object, in a given year. But *get one's favorite letter from Mary* is again not a 'one time only' predicate, due to the bound variable.

Here we will be concerned with a specific consequence of the 'one time only' property, namely, that it imposes a unicity requirement on the arguments and the adjuncts of the predicate. This can be demonstrated as follows. In the iterable (91) examples the distributive answer *John did and Bill did* is as acceptable as *John and Bill did.* In 'one time only' (92), the former is unacceptable: John and Bill must form a collective recipient . Similarly, in (91) the short (exhaustive) answer *Bill* can be modified by *only.* In (92) it cannot or, more precisely, if *only* is acceptable, it must mean 'alone' and not 'exclusively'. The effect disappears in (93).

(91) a. Who showed this letter$_{\text{token}}$ to Mary?
John and Bill did / John did and Bill did.
Bill did / Only Bill did.

b. Who got a letter from Mary?
John and Bill did / John did and Bill did.
Bill did / Only Bill did.

(92) a. Who got this letter$_{\text{token}}$ from Mary?
John and Bill did / *John did and Bill did.
Bill did / (*)Only Bill did.

b. Who burned this letter$_{\text{token}}$?
John and Bill did / *John did and Bill did.
Bill did / (*)Only Bill did.

c. Who won the Rimet Cup in 1978?
Argentina did / *Only Argentina did.

(93) Who got his favorite letter from Mary?
John and Bill did / John did and Bill did.
Bill did / Only Bill did.

The same observations apply to other arguments and adjuncts, e.g.,

(94) From whom did you get this letter$_{\text{token}}$?
From Mary / (*)Only from Mary.

(95) When did you get this letter$_{\text{token}}$?
Yesterday / Only yesterday [= not earlier].

This phenomenon, together with its consequences for scope, was observed in Szabolcsi (1986, pp. 334–7). In what follows we will somewhat enlarge the set of data and spell out the explanation in terms of the present proposal.

(96) and (97) indicate that the *who* subject or experiencer of an iterable predicate can take scope over negation or a universal, while the *who* subject or source of a 'one time only' predicate cannot. (An existential would eliminate the 'one time only' property in the latter case, so it cannot be tested.) (98) and (99) show a similar contrast with a factive and a *wh*-island; a PP argument is extracted in order to eliminate irrelevant difficulties with subject extraction.

(96) a. Who didn't show this letter$_{\text{token}}$ to Mary?
To whom didn't you show this letter$_{\text{token}}$?

b. * Who didn't get this letter$_{\text{token}}$ from Mary?
* From whom didn't you get this letter$_{\text{token}}$?

(97) a. Who showed every letter$_{\text{narrow scope}}$ to Mary?
To whom did you show every letter$_{\text{narrow scope}}$?

b. * Who got every letter$_{\text{narrow scope}}$ from Mary?
* From whom did you get every letter$_{\text{narrow scope}}$?

(98) a. To whom do you regret having shown this letter$_{\text{token}}$?

b. * From whom do you regret having gotten this $\text{letter}_{\text{token}}$?

(99) a. To whom do you wonder whether I showed this $\text{letter}_{\text{token}}$?

b. * From whom do you wonder whether I got this $\text{letter}_{\text{token}}$?

The sensitivity of these arguments to weak islands cannot be explained with reference to thematic roles or discourse factors. The thematic roles are equally "referential" in all cases, and there can hardly be a coherent notion of D-linking or specificity that would distinguish the 'one time' arguments from the others. On the other hand, the absence of the unicity requirement means that *show this $letter_{token}$ to Mary* denotes a set of individuals of whom the predicate holds independently, whereas the presence of the unicity requirement means that *get this $letter_{token}$ from Mary* denotes a sum of whose parts the predicate does not hold independently:

(100) ιx[get this letter from Mary(x)] = [John $\oplus$ Bill]

Since sums form a semilattice, the explanation in the previous section carries over.

A last interesting point to note here is that exactly the same effect is observed no matter whether the sum-term is a subject or a source, although in the former case negation and the object universal do not syntactically intervene between the *wh*-phrase and its trace. This supports the definition of wide scope taking given in the previous section, which refers to the necessity of performing certain operations in the definition/verification of the answer, rather than to the wide scope taker's binding a variable within the syntactic scope of the other operator.

6.2 Event structure and set formation

In this section we propose a connection between certain properties of predicates and the question whether the denotation of a particular parameter is an element of an ordered or of an unordered set. 'Parameter' serves as a cover term for both arguments and adjuncts in the grammatical sense. Details of how question interpretation is defined will also be made more precise, although we are not offering a full formalization here.*

The basic idea derives from Carlson's (1984, p. 274) suggestion that bearers of thematic roles are unique per event. "If there is a proposed event with, say, two themes, then there are (at least) two events and not one." Informal though his proposal is, Carlson is careful to note that on the group reading of *John and Bob threw the chest into the ocean* we have a single event with the collective of

*In this section, a few changes have been made in the formalization to enhance its readability. They do not affect the content of the claims.

John and Bill as its unique Agent, and in *Bob washed the car,* the car is the Theme, and its parts are not.

We dub events characterized by thematic uniqueness *minimal events* ($e_{m/i}$):

(101) a. visit([Rome])([John])([$e_{m/i}$]) entails (b), (c)

b. ιx[visit([Rome])(x)([$e_{m/i}$])] = [John]

c. ιx[visit((x)[John])([$e_{m/i}$])] = [Rome]

Enclosed in square brackets are objects coming from "overpopulated" Linkean domains (join semilattices) of various sorts. In adherence to Carlson's intuition, [John $\oplus$ Bob], i.e., the sum of John and Bob, is used only if the predicate does not distribute over the parts of the plural object. We will call semilattice objects "slobjects" and usually suppress the square brackets. How do we come to think of the denotations of *visited Rome* and *John visited* as sets of slobjects? We submit that the reason is that these predicates allow us to lump several minimal events together and, at the same time, to collect the unique slobjects corresponding to the pertinent parameter into an unordered set. This requires that the relation between events and objects be summative:[23]

(102) A relation R [between events and objects] is *summative* iff

$$R(e,x) \wedge R(e,' x') \ \rightarrow \ R(e \cup e,' x \cup x')$$

Visited Rome is summative: If John visited Rome and Bill visited Rome, then John and Bill visited Rome—according to the present intuition, the last clause describes a non-minimal event. Similarly for *John visited.* We assume that summativity has to be non-vacuous: it presupposes that it is possible for there to be two distinct events that we can lump together. If the description of the predicate itself involves a parameter, then this means the relation has to be iterable with respect to that parameter. It must be possible for there to be two distinct events involving the same object:

(103) A relation R [between events and objects] is *iterable* iff

$$\Diamond \exists e \exists e' \exists e'' \exists y [e' \subseteq e \ \wedge \ e'' \subseteq e \ \wedge \ e' \neq e'' \ \wedge \ R(e,' y) \ \wedge \ R(e',' y)]$$

The x *visited* relation between a minimal event and Rome is iterable. On the other hand, the x *destroyed* relation between a minimal event and Rome is not iterable (in the token sense to which we adhere): the same city cannot be destroyed more than once.

[23]This definition as well as (103) and (116) are borrowed from Krifka (1990).

Non-iterability means that the predicate describes a biunique relation between slobjects and minimal events. We encode this by writing the event parameter as a *function* of that other parameter with respect to which the event is not iterative:

(104) destroy(Rome)(Bob)(f_e(Rome))

(The agent may be so written, too, but it does not seem necessary.)

Prior to proceeding to events involving manners and amounts, let us see how the above assumptions are utilized in set formation. We will use 'set' to mean unordered set, unless otherwise specified.

We stipulate that *set formation* takes place only if the predicate is both summative and iterable. On the basis of (101) we can form the standard denotation of the predicate *visit Rome,* the set of those who visit Rome, as follows:

(105) $\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ of visiting Rome within some fixed event range I & x is the unique agent of some $e_{m/i} \subseteq e]\}$

The empirical claim that is being made here is that non-iterable and/or nonsummative relations do not feed set formation. For instance, the linguistic fact that there can be at most one slobject that destroyed Rome might be expressed by saying that it is an element of the singleton set denoted by *destroy Rome*—but we will *not* do so. Instead, the denotation of a non-iterable predicate remains a slobject. The intuition behind this is that a predicate denotes a set only if it can in principle hold of more than one thing independently. Empirical support for this intuition comes from the data reviewed in Section 6.1, i.e., the fact that the questioning of a unique parameter is sensitive to weak islands.

The definition of an answer to *Who visited Rome?* now involves (105), but that of an answer to *Who destroyed Rome?* can involve only (106):

(106) $\iota x[\text{destroy(Rome)}(x)(f_e(\text{Rome}))] = ?$

As regards *Who didn't visit Rome?, Who visited every city?,* and *Who visited a(ny) city?,* the reasoning in 6.1 can be reproduced as follows. If we have sets, as in (105), we can form their complements, or we can intersect and union them with others. The outputs also feed the Boolean operations.

(107) $-\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ of visiting Rome within some fixed event range I & x is the unique agent of some $e_{m/i} \subset e]\} = ?$

(108) $\bigcap_{n \in N}\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ of visiting city$_n$ within some fixed event range I & x is the unique agent of some $e_{m/i} \subset e]\} = ?$

(109) $\bigcup_{n \in N}\{x$: $\exists e[e$ is the sum of minimal events $e_{m/i}$ of visiting city$_n$ within some fixed event range I & x is the unique agent of some $e_{m/i} \subset e]\} = ?$

But since *destroy Rome* does not denote a set, no complement can be formed, and *Who didn't destroy Rome?* is correctly predicted to be ungrammatical.[24] Similarly, *Who destroyed every city?* cannot have a reading parallel to (108). The same sentence is grammatical on the family-of-questions reading (which does not concern us here) and on the reading which presupposes that the same agent (slobject) destroyed every city, cf. (42b). This latter will be expressed roughly as follows:

(110) $\iota x \forall z[\text{destroy}(\text{city}_z)(x)(f_e(\text{city}_z))] = ?$

It might be tempting to revise the set formation assumptions to allow for an alternative representation of this reading. The intersection of singletons is non-empty iff the singletons are identical:

(111) $\bigcap_{i \in I}\{x : [\text{destroy}(\text{city}_i)(x)(f_e(\text{city}_i))]\} = ?$

However, this interpretation asserts, rather than presupposes, that the same agent destroyed every city, which seems counterintuitive. Furthermore, it would predict that as a next step, a complement can be formed: *Who didn't destroy every city?* This is wrong, so (110) must be the correct representation.

The grammatical *Who destroyed a(ny) city?* may be puzzling: the destruction of each city is non-iterable, but that of an arbitrary city is iterable. Due to the first fact we cannot use (109). But we can capitalize on the fact that precisely in this case the event parameter is a function of the theme, whence they share an index:

(112) $\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ that are destructions of some city$_i$ within some fixed event range I & x is the unique agent of some $e_{m/i} \subseteq e]\} = ?$

With these considerations in mind, we can turn to the classical cases of manners and amounts.

First, the slobject denoted by the manner parameter is typically a sum:

(113) $\text{behave}([\text{kindly} \oplus \text{stupidly}])([\text{John}])([e_{m/i}])$

Second, while both the *behaved kindly but stupidly* and the *John behaved* relations are iterable, summativity fails (we never get cumulative readings):

[24]Alternatively, if *Who didn't destroy Rome?* is interpreted as '$\iota x[\text{destroy}(\text{Rome})(x)] \neq ?$,' then an exhaustive answer like *Hannibal* leads to absurd consequences (e.g., Hannibal is the unique slobject not identical to Rome's destroyer, ergo every other slobject is identical).

(114) John behaved kindly at event e and John behaved stupidly at event $e' \not\rightarrow$ John behaved kindly and stupidly at $e \cup e'$

(115) John behaved kindly at event e and Bob behaved stupidly at event $e \not\rightarrow$ John and Bob behaved kindly and stupidly at event e

As a consequence, set formation does not take place. *How didn't you behave?* and *How did everyone behave?* are both out on the wide scope *how* reading. The latter sentence has a family-of-questions reading and one analogous to (110).

Amounts may arise in two different ways, cf. *John weighs ninety pounds* and *John visited two cities.* Both require an additive measure: the value assigned to the sum of two non-overlapping slobjects z and z' is the sum of the values assigned to z and to z' ($\circ$ stands for 'overlap'):

(116) The function μ is an additive measure iff

$$(\neg(z \circ z') \ \wedge \ \mu(z) = n \ \wedge \ \mu(z') = n') \ \rightarrow \ \mu(z \cup z') = n + n'$$

For the sake of simplicity, we will only examine the *two cities* type. Following Krifka (1990), we take *city* to be the measure function. As long as the measured objects do not overlap, the summativity tests that failed above will work here, and we get cumulative readings:

(117) John visited six cities at e and John visited five cities at e' $\rightarrow$ John visited eleven cities at $e \cup e'$

(118) John visited six cities at e and Bob visited five cities at e $\rightarrow$ John and Bob visited eleven cities at e

These measures are not part of the characterization of the minimal event: measuring is an operation performed on sets or slobjects assembled on the basis of minimal events. In *How many cities did John visit?*, for instance, the set of cities that John visited is constructed and μ is applied to that set:

(119) $\mu(\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ of John visiting a city within some fixed event range I & x is the unique theme of some $e_{m/i} \subseteq e]\}) = ?$

Similarly, a good reading can be constructed for *How many cities didn't you visit?*, etc. by measuring the complement of the set of cities visited:[25]

[25] This option is not available for ***How many circles didn't John draw?*** if drawing is understood as creation, and ***John*** is not contrastive. This question is equivalent to ****How many circles aren't there?***; there is no complement that could be formed. We suggest to capture this by measuring non-iterable events directly. The elaboration of this suggestion goes beyond the scope of this paper. (See Doetjes and Honcoop 1996 for related ideas, however.)

(120) $\mu(-\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ of John visiting a city within some fixed event range I & x is the unique theme of some $_{em/i} \subseteq e]\}) = ?$

For the cumulative reading of *John and Bob visited eleven cities,* the two sets of cities are unioned before measuring (we do not provide a general algorithm here):

(121) $\mu(\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ of John visiting a city within some fixed event range I & x is the unique theme of some $e_{m/i} \subseteq e]\}$
$\mu\{x : \exists e[e$ is the sum of minimal events $e_{m/i}$ of John visiting a city within some fixed event range I & x is the unique theme of some $e_{m/i} \subseteq e]\}) = ?$

Measuring differs from the Boolean operations in two respects: its input does not have to be a set, and its output is certainly not a set. For the latter reason μ cannot be followed by the Boolean operations. *How many cities didn't you visit?* is ungrammatical on the reading that asks for the complement of the number of cities visited, and so on.

In other words, there are two reasons why Boolean operations may be unavailable: one is that we were never able to form sets in the first place, and the other is that our sets were subjected to an operation whose value is itself not a set.

7 WEAK ISLANDS—SYNTAX OR SEMANTICS?*

The traditional analysis of weak islands is purely syntactic: it relies on argument/adjunct asymmetries and escape hatches. Recent literature indicates that the generalizations holding for a wider natural class of weak islands have a semantic flavor: D-linking and intervening operators have been shown to play a role. Nevertheless, the theoretical terms in which Relativized Minimality is formulated are syntactic. In this paper we have argued that at least a significant subset of the data can be explained in semantic terms. It may be interesting to ask what the scope of the proposal is.

The present paper has made two independent claims. One is that many weak island violations are due to the failure of the *wh*-phrase to take scope over some intervening operator; see also Dobrovie-Sorin (1992), É. Kiss (1992), and de Swart (1992). Neither these works, nor the present paper has demonstrated, however, that *all* weak islands are scopal. The other claim, entirely

*This section originally contained tentative analyses of two further phenomena. These have been eliminated in the interest of brevity.

our own, concerns the semantic explanation of one type of scopal failure. We have argued that a *wh*-phrase (or quantifier) can take scope over another scopal element SE only if the operations associated with SE are defined for its denotation domain. If the requisite operations are not defined, the intended reading is simply incoherent. We have offered an analysis in this spirit of a suggestive set of examples, many of which do not seem to have an independent syntactic account: consider the claim that different *wh*-phrases are sensitive to different weak islands, and the claim that arguments of non-iterable predicates are sensitive to weak islands.

If our semantic claim concerning scope-taking is logically correct, then it captures an absolute limitation on what meanings are expressible. It is not a matter of elegance whether one invokes it in the explanation of certain phenomena: it will be in effect even if the readings it excludes can be excluded in syntactic terms as well. In this sense it is truly not a rival of syntactic accounts. We expect that the syntactic and semantic explanations of weak island facts will eventually properly overlap. We expect many of the semantic constraints to have syntactic correlates: ones that have independent syntactic motivation, or ones that are semantically motivated but are compatible with independent syntactic considerations. There may remain important cases that are excluded only semantically or only syntactically. Dobrovie-Sorin's (1992) and Beghelli's (1993) work appears to point to this conclusion.

REFERENCES

Aoun, Joseph. 1986. *Generalized Binding.* Dordrecht: Foris.

Barwise, Jon. 1979. On branching quantifiers in English. *Journal of Philosophical Logic* 8(1):47–80.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In Amy J. Schafer (Ed.), *Proceedings of the North East Linguistic Society 23, Volume One*, 65–80, Amherst. GLSA Publications, University of Massachussetts.

Carlson, Gregory. 1984. Thematic roles and their role in semantic interpretation. *Linguistics* 22(3):259–279.

Cattell, Ray. 1978. On the source of interrogative adverbs. *Language* 54(1):61–77.

Chierchia, Gennaro. 1984. *Topics in the Syntax and Semantics of Infinitives and Gerunds.* PhD thesis, University of Massachussetts, Amherst.

Chierchia, Gennaro. 1992. Functional WH and weak crossover. In Dawn Bates (Ed.), *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 75–90, Stanford. CSLI.

Chung, Sandra. 1992. Wh-agreement and 'referentiality' in chamorro. Unpublished manuscript, UCSC.

Cinque, Guglielmo. 1990. *Types of A'-Dependencies.* Cambridge: MIT Press.

Comorovski, Ileana. 1989. *Discourse and the Syntax of Multiple Constituent Questions.* PhD thesis, Cornell University.

Dobrovie-Sorin, Carmen. 1992. LF representations, weak islands, and the ECP. *GLOW Newsletter* 27:14–15.

Doetjes, Jenny, and Martin Honcoop. 1996. On event-relatedness, symmetry and weak islands. In this volume.

Dukes, Michael. 1992. Factives, 'stance' predicates, and weak islandhood. Unpublished manuscript, UCLA.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry* 22(1):1–25.

Engdahl, Elisabet. 1985. *Constituent Questions.* Dordrecht: Reidel.

Frampton, John. 1990. The fine structure of WH movement and the proper formulation of the ECP. Unpublished manuscript.

Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers.* PhD thesis, University of Amsterdam.

Gutiérrez Rexach, Javier. 1996. Questions and generalized quantifiers. In this volume.

Harada, Yasunari, and Naohiko Noguchi. 1992. On the semantics and pragmatics of dake (and only). In Chris Barker and David Dowty (Eds.), *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory (OSU WPL 40)*, 125–145, Columbus. The Ohio State University.

Hegarty, Michael. 1992. Adjunct extraction without traces. In Dawn Bates (Ed.), *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 209–222, Stanford. CSLI.

Hepple, Mark. 1990. *The Grammar and Processing of Order and Dependency: A Categorial Approach.* PhD thesis, Edinburgh University.

Higginbotham, James. 1991. Interrogatives I. In Lisa Cheng and Hamida Demirdash (Eds.), *More Papers on WH-Movement (MIT WPL 15)*, 47–76, Cambridge. MIT.

Keenan, Edward, and Leonard Faltz. 1985. *Boolean Semantics for Natural Language.* Dordrecht: Reidel.

Keenan, Edward, and Alan Timberlake. 1988. Natural language motivation for extending categorial grammar. In Richard Oehrle et al. (Eds.), *Categorial Grammars and Natural Language Structures*, 265–296. Dordrecht: Kluwer.

Kiss, Tibor. 1991. Some properties of negation in German. Stuttgart, Unpublished manuscript.

É. Kiss, Katalin. 1991. Logical structure in linguistic structure: The case of Hungarian. In James Huang and Robert May (Eds.), *Logical Structure and Linguistic Structure*, 387–426. Dordrecht: Kluwer.

É. Kiss, Katalin. 1992. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11(1):85–120.

Krifka, Manfred. 1990. Four thousand ships passed through the lock. *Linguistics and Philosophy* 13(5):487–520.

Kroch, Anthony. 1989. Amount quantification, referentiality and long WH-movement. Unpublished manuscript.

Ladusaw, William. 1980. Affective OR, factive verbs, and negative polarity items. In Jody Kreiman and Almerindo Ojeda (Eds.), *Papers from the Sixteenth Regional Meeting: Chicago Linguistics Society*, 170–184.

Landman, Fred. 1986. *Towards a Theory of Information.* Dordrecht: Foris.

Landman, Fred. 1991. *Structures for Semantics.* Dordrecht: Kluwer.

Liu, Feng-hsi. 1990. *Scope Dependency in English and Chinese.* PhD thesis, UCLA.

Liu, Feng-hsi. 1992. Branching quantification and scope independence. In Jaap van der Does and Jan van Eijck (Eds.), *Generalized Quantifier Theory and Applications*, 297–331. Chicago: CSLI/University of Chicago Press.

May, Robert. 1985. *Logical Form: Its Structure and Derivation.* Cambridge: MIT Press.

Melvold, Janis. 1991. Factivity and definiteness. In Lisa Cheng and Hamida Demirdash (Eds.), *More Papers on WH-Movement (MIT WPL 15)*, 97–117. Cambridge: MIT.

Molnár, Ilona. 1982. Existential relations in hogy 'that'-sentences. In Ferenc Kiefer (Ed.), *Hungarian Linguistics*, 111–147. Amsterdam: John Benjamins.

Moltmann, Friederike. 1994. Attitude reports, events and partial models. Unpublished manuscript.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric Reuland and Alice ter Meulen (Eds.), *The Representation of (In)definiteness*, 98–130. Cambridge: MIT Press.

Philip, William, and Jill de Villiers. 1992. Monotonicity and the acquisition of weak islands. Talk Presented at the Stanford Conference on Child Language.

Progovac, Ljiljana. 1988. *A Binding Approach to Polarity Sensitivity.* PhD thesis, University of Southern California.

Rizzi, Luigi. 1990. *Relativized Minimality.* Cambridge: MIT Press.

Rizzi, Luigi. 1992. Argument/adjunct (a)symmetries. In *NELS 22: Proceedings of the North East Linguistic Society*, 365–381. GLSA Publications, University of Massachussetts, Amherst.

Sántha, Miklós. 1980. Some problems of the semantic treatment of that-clauses in Montague Grammar. *Computational Linguistics and Computer Languages* 14(1):67–79.

Scha, Remko. 1981. Distributive, collective, and cumulative quantification. In Jeroen Groenendijk et al. (Eds.), *Formal Methods in the Study of Language*, 483–512. Amsterdam: Mathematisch Centrum.

Schein, Barry. 1993. *Plurals and Events.* Cambridge: MIT Press.

Sher, Gila. 1991. *The Bounds of Logic. A Generalized Viewpoint.* Cambridge: MIT Press.

Steedman, Mark. 1987. Combinatory Grammar and parasitic gaps. *Natural Language and Linguistic Theory* 5(3):403–440.

de Swart, Henriëtte. 1992. Intervention effects, monotonicity and scope. In Chris Barker and David Dowty (Eds.), *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory (OSU WPL 40)*, 387–406, Columbus. The Ohio State University.

Szabolcsi, Anna. 1983. Focussing properties, or the trap of first order. *Theoretical Linguistics* 10(2/3):125–145.

Szabolcsi, Anna. 1986. From the definiteness effect to lexical integrity. In Werner Abraham and Sjaak de Meij (Eds.), *Topic, Focus and Configurationality*, 321–348. Amsterdam: John Benjamins.

Szabolcsi, Anna. 1991. Islands, monotonicity, composition, and heads. *GLOW Newsletter* 26:58–59.

Szabolcsi, Anna. 1992a. Combinatory Grammar and projection from the lexicon. In Ivan Sag and Anna Szabolcsi (Eds.), *Lexical Matters*, 241–268. Stanford: CSLI.

Szabolcsi, Anna. 1992b. Weak islands, individuals, and scope. In Chris Barker and David Dowty (Eds.), *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory (OSU WPL 40)*, 407–438, Columbus. The Ohio State University.

Szabolcsi, Anna. 1996. Quantifiers in pair-list readings. In this volume.

Szabolcsi, Anna, and Frans Zwarts. 1990. Semantic properties of composed functions and the distribution of Wh-phrases. In *Proceedings of the Seventh Amsterdam Colloquium*, 529–555. Amsterdam: Institute for Language, Logic, and Information.

Szabolcsi, Anna, and Frans Zwarts. 1991. Unbounded dependencies and algebraic semantics. Lecture Notes from the 3rd European Summer School in Language, Logic and Information.

Szabolcsi, Anna, and Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1(3):235–284. Reprinted in this volume.

Westerståhl, Dag. 1985. Logical constants in quantifier languages. *Linguistics and Philosophy* 8(4):387–413.

Williams, Edwin. 1974. *Rule Ordering in Syntax.* PhD thesis, MIT.

Zimmermann, Ede. 1991. Scopeless quantifiers and operators. In Jaap van der Does and Jan van Eijck (Eds.), *Generalized Quantifier Theory and Applications*, 207–223. Amsterdam: Dutch Network for Language, Logic and Information.

Zwarts, Frans. 1986. *Categoriale Grammatika en Algebraïsche Semantiek.* PhD thesis, University of Groningen.

8

THE SEMANTICS OF EVENT-RELATED READINGS: A CASE FOR PAIR-QUANTIFICATION*

Jenny Doetjes and Martin Honcoop

Department of Linguistics
University of Leiden

1 INTRODUCTION: THE PROBLEM AND THE MAIN CLAIMS

In this paper we will be concerned primarily with ambiguities of the type exhibited in (1), where OR stands for *O*bject-related *R*eading and ER for *E*vent-related *R*eading. These readings have been extensively discussed and thoroughly analyzed in Krifka (1990).

(1) Last year, 4,000 ships passed through the lock (OR/ER)

*This joint paper is a natural extension of our individual efforts to clarify the status of ORs and ERs in grammatical theory, as reported on in Honcoop (1992) and, especially in connection to quantification at a distance in French, Doetjes (1994). We would like to thank the following persons for insightful discussion on the subject matter of this paper, as well as related issues: Filippo Beghelli, Dorit Ben-Shalom, Rose-Marie Déchaine, Astrid Ferdinand, Jacqueline Guéron, Javier Gutiérrez-Rexach, Elena Herburger, Helen de Hoop, Hans-Georg Obenauer, Tetsuya Sano, Barry Schein, Tim Stowell, Henriëtte de Swart, and Elisabeth Villalta. Thanks are also due to the people who attended any one of the three different talks on related matters in Paris in February and December of 1994, the audience attending the TIN-dag in Utrecht in 1994, as well as to the audience at the syntax/semantics colloquium which was held in the Spring of 1995 at UCLA, and the audience at WCCFL XIV, which was held in March 1995 at USC. Crit Cremers, Teun Hoekstra and Anna Szabolcsi deserve special mention for their ongoing support and constructive criticism. Many thanks furthermore to Frits Beukema for correcting our English. But most of all, we would like to express our gratitude to Carmen Dobrovie-Sorin, who encouraged us to collaborate on this paper, and who at various stages actively participated in the project herself. Needless to say, if it wasn't because of her, this paper could not have been written in the first place. We owe an obvious debt to her. During his stay at the linguistics department at UCLA in Winter and Spring of 1995, the research of the second author was partially supported by a grant of the Netherlands Organization of Scientific Research (NWO), which is hereby gratefully acknowledged. Finally, the second author would like to thank the people at the Linguistics department at UCLA for their warm hospitality. For any remaining errors, we are entirely to blame.

A. Szabolcsi (ed.), Ways of Scope Taking, 263–310.

OR: 4,000 ships are such that each of them passed through the lock last year

ER: there were 4,000 events in which a ship passed through the lock last year

It is the second reading of (1) which will be of special interest to us. The distinctive feature of this type of reading is that it allows for what we might call the recycling of individuals. That is, the sentence in (1) may be true in a situation where there are only 2,000 ships, each of which passed through the lock on two occasions last year. Note that this scenario falsifies (1) on its OR.

Our main interests in embarking on this particular enterprise are twofold. First of all, there is the obvious, purely representational, problem as to how to capture the distinct truth-conditions that are involved in the two readings. This issue will be dealt with in Section 3. We will advance and defend the claim that ER can be derived by quantifying over ordered pairs of events and objects, as illustrated for (1) in (2) (the denotations of expressions will henceforth be represented by a prime).

(2) $4000\langle e, x\rangle : \text{ship}'(x) \wedge \text{passed-through-the-lock}'(e, x) \wedge \text{last-year}'(e)$

The attraction of the pair-quantificational approach to ER resides in the fact that it immediately accounts for the phenomenon of recycling: whenever a particular ship passed through the lock last year on more than one occasion, we will have just as many ⟨event,ship⟩ pairs that satisfy the restrictive description in (2).

Although it is true that a pair-quantificational analysis of ER yields the right results for cases such as (1), it is also true that it stands in need for some further qualification, in the light of the fact that mass noun constructions such as *4,000 tons of radioactive waste* may also give rise to such readings. More generally, we will formally demonstrate why the class of those constructions for which ER may be derived by quantifying over ⟨event, object⟩ pairs, in the way indicated in (2), exactly coincides with the class of those constructions in which the relevant QP (*Q*uantifier *P*hrase) argument is headed by a count noun and in which the verbal predicate is fully distributive with respect to that argument. Both of these points will be taken up in Section 3.1. Having thus established the lack of generality of the pair-quantificational approach on its most straightforward implementation, we may then proceed to its modification. In doing so, we draw heavily on the essentials of Krifka's (1990) approach to ER, the most thoroughly worked out analysis of this phenomenon to date. By implementing his proposal in a pair-quantificational setting, we can deal with the problems noted above in connection to pair-quantificational approaches to ER in an elegant and straightforward way. This constitutes the subject matter of Section 3.2.

Secondly, our interest in ER derives from the fact that the ambiguity between OR and ER is not just some accidental quirk of nature. Instead, we will show in Sections 4 and 5 that the two readings are deeply anchored in natural language grammar. It is here that the pair-quantificational approach to ER proves to be fruitful and explanatory.

In Section 4, we shall first see that ER is sensitive to whether the determiner is symmetric (weak) or non-symmetric (strong). Specifically, we will observe that only QPs headed by weak determiners allow for both OR and ER in the default case. To illustrate the point, although the sentence in (1) can have an OR as well as an ER, the sentence in (3) only admits of an OR.

(3) Last year, most ships passed through the lock (OR/*ER)

OR: most ships are such that each of them passed through the lock last year

ER: * most events in which a ship passed through the lock (last year) occurred (last year)

The pair-quantificational approach to ER outlined in Section 3.2 allows us to derive the contrast between (1) and (3) as a simple corollary. A crucial prerequisite for the well-formedness of pair-quantifiers is that both the event-variable and the object-variable occur in their restriction. We will argue that the difference between symmetric and non-symmetric determiners is due to the fact that the former allow the eventive verbal predicate to join the nominal predicate in their restriction by semantic inference.

Pursuing the implications of this analysis, we will furthermore observe that strong QPs may also support ER if the sentence contains a focused constituent or if the relevant QP is modified by a relative clause which embeds an eventive predicate. Again, given plausible assumptions with respect to the semantics of focus, this immediately falls out under a pair-quantificational approach to ER. Both focus and modification by a relative clause yield a structure in which an eventive predicate is mapped into the restrictive clause of the strong quantifier, together with the head noun. Our account of these data will in part be modeled on Chierchia's (1992) treatment of donkey anaphora in terms of dynamic semantics. The dynamic aspect of our pair-quantificational approach to these facts will be substantially motivated by the striking similarities that exist between ER and so-called *symmetric* (or pair-quantificational) readings of donkey sentences. We will carefully establish and discuss these similarities in Section 4.5.

In Section 5, we show that ER is sensitive to weak islands (WI). That is, QPs which block *how*-extraction when occupying the subject position also block ER when occupying the object position in simple transitive clauses. Furthermore, we will observe that sentence negation, which is known to impair extraction of *how,* also blocks ER for the subject, as shown in (4).

(4) Last year, 4,000 ships didn't pass through the lock (OR/*ER)

OR: 4,000 ships are such that each of them didn't pass through the lock last year

ER: * there were 4,000 events in which a ship did not pass through the lock last year

These facts can be neatly accounted for if we adopt Szabolcsi and Zwarts's (1993) algebraic semantic approach to WIs. Szabolcsi and Zwarts argue that the critical property that many weak island sensitive expressions share is that they denote elements of impoverished algebraic structures, e.g., mere join semi-lattices. The weak island sensitivity of ER automatically follows then on our analysis. Events are standardly thought of as having a join semi-lattice structure, and in Section 3 we will argue that ⟨event, object⟩ pairs inherit this structure from their event component.

Prior to presenting the analysis, in Section 2 we discuss some issues that pertain to the (in)felicity of ER. Although we take the position in this paper that natural language grammar should characterize the distribution of ER, a position that is in fact forced upon us in the light of the observations that we discuss in Sections 4 and 5, this is not to deny that these readings are heavily constrained by various pragmatic factors. In Section 2 then, we will briefly discuss how some of these factors may conspire to make one reading more felicitous than the other, given a specific context of utterance.

2 THE (IN)FELICITY OF ER

Without doubt, ER is severely constrained by pragmatic factors that may make these readings implausible under certain discourse conditions. In brief, we suggest that a necessary condition for the felicity of ER is that the identity of the individuals (or of the particular portions of matter) that we are talking about be both irrelevant and easy to ignore.

By way of example, it is easy to construct situations in which Krifka's example is felicitous on both counts:

(5) Last year, 4,000 ships passed through the lock

First, imagine that the cost-effectiveness of the maintenance of the lock or the workload of its personnel is under discussion. Under these circumstances the only relevant thing to know is how many lock passages occurred; whether some of these passages were by the same individual ships makes no difference. Second, the fact that 4,000 is a large figure makes it easy to ignore the identity of the ships. For the role of this latter factor, compare (6a) with (6b):

(6) a. Last year, 500 ships passed through the lock
b. Last year, only The Flying Dutchman and The Merry Ploughman passed through the lock; and they did so 500 times in total

Even though the toll collected and the amount of work involved depend only on the number of passages, (6b) may be preferred to (6a) when the identity of the ships is known and/or salient.

In other contexts, the irrelevance of identities may suffice to justify ER even if the figures are small. Consider:

(7) You get out a fresh tablecloth after every fifth customer, even if the old one looks clean

(8) Your toy fountain spouted up 10 liters of water yesterday; we'll need a new battery

(9) If she wants me to examine 20 students, she must block out more than one day for me

It may be that the same customers return frequently to a table, or that the fountain recycles the water it uses, or that the students to be examined come from three classes with overlapping enrollment, and the discourse partners may even be aware of this. However, the wear of the table cloth and the battery and the time needed to conduct the exams will not be affected, which is why ER is entirely felicitous.

We will not attempt to analyze the pragmatic conditions under which ER is felicitous any further. We assume that it is in principle possible to specify those conditions, however complex they might be. The question relevant for the rest of this paper is as follows: is ER the product of those pragmatic factors? If yes, then it does not belong within the scope of grammar (semantics) proper, but may be assimilated to conversational implicatures, for instance. Or is it the case that grammar (semantics) routinely supplies ER, but they are often discarded on pragmatic grounds? We argue that the latter is the case. As was indicated in the introduction, the availability of ER is constrained, not only by more or less hazy pragmatic factors, but also by quite clearcut grammatical ones: the logical properties of the determiner, the "bracketing" of the sentence, and the presence of standard weak island inducers. This is possible only if ER is a proper grammatical (semantic) phenomenon.

3 THE REPRESENTATION OF ER

What would a semantics for ER look like? In 3.1, we shall see that although the pair-quantificational approach goes a long way towards capturing the truth conditions of ER, it is simply not general enough on its most straightforward implementation: it fails to extend to those cases where the ER is effected by a mass noun. We will demonstrate formally why we can compute ER by quantifying over pairs of events and objects when a distributive, verbal predicate combines with a count noun. In Section 3.2, we briefly discuss the essentials of Krifka's (1990) approach to ER, which generalizes over both the count noun and the problematic mass noun constructions. We will adopt the tools developed by Krifka in an essentially pair-quantificational setting, so that the problems noted in Section 3.1 in connection with the pair-quantificational approaches to ER can be solved in an elegant and straightforward way. This minor modification of Krifka's framework will receive its justification in Sections 4 and 5, where we discuss the grammatical restrictions on ER.

3.1 Quantification over ⟨event, object⟩ pairs: its attractions and problems, and their sources

When looking at the ER of sentence (1) above, repeated below as (10), one might have the impression that ER in general can be obtained by simply counting the number of minimal events with respect to which a given verb may be truthfully predicated. That is, if we take verbs to denote relations between events and objects that participate in these events, as originally proposed in Davidson (1967), we might wish to paraphrase ER as in (11) and to represent it formally as in (12).

(10) 4,000 ships passed through the lock (last year)

OR: 4,000 ships are such that each of them passed through the lock (last year)

ER: there were 4,000 events in which a ship passed through the lock (last year) ($\approx$ 11)

(11) 4,000 times a ship passed through the lock (last year)

(12) $|\{e \in \text{EVENTS} | \exists x : \langle e, x\rangle \in \text{passed-through-the-lock}' \wedge x \in \text{ship}'\}| = 4000$

Alternatively, we could have represented the truth conditions of the ER of (10) in set-theoretic terms as in (13a), where we count the number of ⟨event, object⟩ pairs that fall under the extension of passed through the lock. (13a) provides the interpretation of the pair-quantification structure in (13b).

(13) a. $|\{\langle e,x\rangle \mid \langle e,x\rangle \in \text{passed-through-the-lock}' \wedge x \in \text{ship}'\}| = 4000$
b. $4000\langle e,x\rangle : \text{ship}'(x) \wedge \text{passed-through-the-lock}'(e,x)$

The reason why (12) and (13) are equivalent lies in a widely shared assumption concerning the thematic properties of eventive predicates (cf. for instance Carlson 1984, Schein 1993, among others), namely, that one and the same event cannot have two distinct objects as agent, patient, or whatever. Carlson calls this property thematic uniqueness. It may be defined as in (14).[1]

(14) DEFINITION (THEMATIC UNIQUENESS) *For any event-relation R:*

$$\forall e \forall x \forall x'[\langle e,x\rangle \in R \wedge \langle e,x'\rangle \in R \rightarrow x = x']$$

Given thematic uniqueness, it is easy to show that the following proposition holds.[2]

(15) FACT $e \neq e' \leftrightarrow \langle e,x\rangle \neq \langle e',x'\rangle$

Therefore, on the basis of (14), we may conclude that, in general, the number of events in the domain of an event-relation R is the same as the number of ⟨event, object⟩ pairs that are members of R. And this in turn means that to the extent that we are justified in deriving ER by quantifying over events, we are justified in deriving these readings using pair-quantification as well, at least in the form as we presented the latter mechanism here. When it comes to capturing the property of recycling that sets ER apart from OR, both representations fit the bill equally well: whenever a particular ship is recycled in events of passing through the lock, we will have just as many events or ⟨event, ship⟩ pairs over which the quantifier 4,000 may quantify.

It turns out, however, that these analyses fail to extend to those cases where ER involves a mass noun, as first noted by Krifka (1990). For in these cases, trying to paraphrase ER by means of a QP adverb, as done in (17) for instance, is inevitably doomed to failure: (16) does not give the slightest reason to believe that the amount of waste involved in each passage was a ton; nor does it say anything else about the sizes of the portions.

[1] The definition in (14) is intended to apply to plural, non-minimal events as well on the assumption that plural events involve plural objects (collectives) as their unique, thematic participants. Plural events and objects, and their algebraic properties, will be discussed below.

[2] *Proof*: By definition,

$$\langle x,y\rangle \neq \langle x',y'\rangle \leftrightarrow x \neq x' \vee y \neq y'$$

L $\Rightarrow$ R follows from the definition above.
R $\Rightarrow$ L: Suppose $\langle e,x\rangle \neq \langle e',x'\rangle$. Assume $e = e'$ for contradiction. By thematic uniqueness, it follows that $x = x'$, contradicting our assumption that $\langle e,x\rangle \neq \langle e',x'\rangle$. Therefore, $e \neq e'$. □

(16) Last year, 4,000 tons of radioactive waste passed through the lock

OR: 4,000 tons of waste are such that they passed through the lock last year

ER: the total weight of waste that passed through the lock last year (in one or more different lock passages) was 4,000 tons ($\not\approx$ 17)

(17) 4,000 times a ton of radioactive waste passed through the lock last year

Thus, it is reasonably clear that we cannot represent ER in general by having the "responsible" determiner-element directly quantify over events, and given our result in (15), pair-quantification would not do much good either.

In a subset of the cases counting minimal events or ⟨event, object⟩ pairs would give a correct interpretation for ER. To briefly show what cases these are, we need to make the assumption that the universe of a model has the structure of a *join semi-lattice.* This assumption merely boils down to saying that the domain of entities forms a set which is partially ordered by the $\subseteq$ relation (part of), and which is furthermore closed under the $\oplus$ operation (sum, or join), so that for any two individuals a and b which are members of E, the join of a and b ($a \oplus b$) is also a member of E (cf. Chapter 1 for some exposition of lattices). This assumption concerning the structure of the universe of models has proved to be fruitful in the study of plurality (cf. Link 1987) and, when extended to the domain of EVENTS, as we will assume from now on,[3] in the study of the aspectual structure of verbal predicates (cf. Krifka 1989), among various other things.

The cases in which event-counting would yield correct results for ER have two crucial properties. One, the noun phrase whose event-related reading is at issue contains a count noun; in other words, the denotation of the noun presupposes a unique articulation into minimal parts. This property of count nouns is captured by saying that they are atomic predicates:

(18) DEFINITION (ATOMICITY) *A predicate P may be called* atomic *just in case*

$$\forall x[P(x) \rightarrow \exists y[y \subseteq x \wedge ATOM(y,P)]]$$

In words, (18) says that any object of which an atomic predicate P holds should have parts that are atomic in P: atomic parts have no proper subparts of which P holds. Mass nouns, on the other hand, are non-atomic (divisive) predicates, which means that not all entities of which P holds have P-atoms as parts:

[3]Note that this extension requires postulating a so-called *sorted* universe E, which consists of the mutually exclusive but jointly exhaustive subdomains O (for objects) and S (for events, or situations).

(19) DEFINITION (DIVISIVITY) *A predicate P may be called* divisive *just in case*

$$\neg\forall x[P(x) \rightarrow \exists y[y \subseteq x \wedge ATOM(y,P)]]$$

The second crucial property is what we will refer to as Distributivity:

(20) DEFINITION (DISTRIBUTIVITY) *An event relation R satisfies* Distributivity *just in case for any nominal predicate N*

$$\forall e \forall x \forall x' \in N[R(e,x) \wedge x' \subseteq x \rightarrow \exists e'[e' \subseteq e \wedge R(e',x')]]$$

In words, whenever a complex object $x \in N$ is involved in a global event as specified by the verbal predicate, each and every part $x' \in N$ of this object has a subevent corresponding to it in the extension of the same predicate.

The event relation expressed by the verbal predicate ***pass through the lock*** satisfies Distributivity. Suppose for instance that there is an event e in which the composite object *Candida* $\oplus$ *Eleonore* performed one lock passage. Then, obviously, it must be the case that there are two subevents e' and e'' in which *Candida* and *Eleonore* performed one lock passage, respectively. If the verbal predicate is then combined with a count noun construction such as ***4,000 ships,*** the resulting event predicate must be atomic since the partitioning of some global event into its constituent subevents, induced by the constraint in (20), will stop as soon as the atoms that generate the extension of the count noun are reached. On the other hand, if the same verbal predicate is combined with a mass noun construction such as ***4,000 tons of radioactive waste,*** the resulting event predicate must be divisive as there is in principle no limit now on the partitioning of the global event into still smaller and smaller subevents, as triggered by Distributivity. Thus such a predicate does not tell us how to construct minimal events and consequently, simple event counting cannot do justice to ER.

3.2 Pair-quantification revisited

In this section, we will proceed with the necessary modification of the pair-quantificational approach to ER, so as to cover the mass noun constructions in a satisfactory, uniform manner. To this end, we will briefly discuss the bare essentials of Krifka's (1990) analysis of ER, which generalizes over both count and mass nouns, and then implement his proposal in a pair-quantificational setting. Our overall aim is to show that the grammatical restrictions on ER, which we will discuss in depth in the following two sections, fall out naturally from a pair-quantificational analysis of these readings, and that these restrictions can be accounted for while preserving a fully uniform treatment of both count and mass noun constructions.

Intuitively, the problem that the pair-quantificational approach to ER faces when confronted with mass nouns consists in the fact that the measure functions with which these predicates can be semantically combined, like *tons (of), meters (of)* etc., cannot be directly applied to ⟨event, object⟩ pairs. For instance, it is hard to make sense of 4,000 tons of ⟨event, radioactive-waste⟩ pairs. To put it differently, if we were to apply these measure functions to ordered pairs of events and objects, we had better make sure that they refer, in some way or other, to the relevant, quantitative properties of the nominal co-argument.

Using Krifka's (1990) terminology, the latter requirement means that any measure function μ' on ⟨event, object⟩ pairs should be *standardized* with respect to its corresponding measure function μ on the domain of objects. Now, we may observe that under very special circumstances μ' can be straightforwardly standardized with respect to its corresponding measure function μ, in that μ' should simply yield the same value as μ. Evidently, we want the measure function μ' on ordered pairs of events and objects to yield the same value as the corresponding measure function μ on objects whenever the OR and ER of a sentence truth-conditionally coincide. Krifka notes that the latter situation obtains just in case the eventive predicate denotes a *non-iterative* event, like *be burned* in (21) below. Obviously, the OR of (21) could not possibly differ from its ER, as this would imply that the same paper can be burned more than once, contrary to fact.

(21) 4,000 tons of paper will be burned tomorrow (OR = ER)

The aspectual property of *iterativity* may be formally defined as follows (cf. Krifka 1989, p. 93):

(22) DEFINITION (ITERATIVITY [ITER]) *For any event e, object x, and event relation R,*

$$ITER(e,x,R) \leftrightarrow R(e,x) \wedge \exists e'\exists e''\exists x'[e' \subseteq e \wedge e'' \subseteq e \wedge e' \neq e'' \wedge x' \subseteq x \wedge R(e',x') \wedge R(e'',x')]$$

In words, (22) says that an event relation R is iterative with respect to an event e and an object x just in case there is a part of x which is involved in two different parts of e, as specified by R.

As Krifka (1990) observes, this particular procedure for standardizing one measure function with respect to another can be generalized on account of the fact that every iterative event can be partitioned into non-iterative subevents. On the basis of that observation, we may briefly summarize Krifka's analysis as follows. His proposal essentially boils down to the claim that all varieties of ER can be computed by a *measure function on events* that simply refers to the quantitative properties of the relevant object directly in case we are dealing with a non-iterative event (*Standardization*), or indirectly by first partitioning the iterative event into non-iterative subevents. After we have counted or measured the relevant objects in each cell of the partition (Standardization), we can generalize over all the partial results by claiming additivity for the pertinent measure function on events. In keeping with Krifka's terminology, we will call the last step *Generalization.* The property of *additivity* for measure functions can be defined as in (23) (Krifka 1990, p. 494).[4]

(23) DEFINITION (ADDITIVITY FOR MEASURE FUNCTIONS)

$$\neg x \circ y \ \wedge \ \mu(x) = n \ \wedge \ \mu(y) = n' \ \rightarrow \ \mu(x \oplus y) = n + n'$$

Our modification of Krifka's proposal consists in the fact that we will use the two steps identified above to define a measure function μ' on ⟨event, object⟩ pairs in terms of the measure function μ on objects, whereas Krifka defines a measure function μ' on events in terms of μ. Sections 4 and 5 will be devoted to a substantial motivation of this modification.

In the remainder of this subsection, we proceed as follows. First, for a measure function μ' to be well-defined on ordered pairs of events and objects, we first need to ensure that these pairs exhibit a partial ordering, given the fact that measure functions in general are only defined on partially ordered domains (cf. Krifka 1989, 1990). Then, we identify the algebraic structure of the domain of ⟨event, object⟩ pairs as a join semi-lattice on the basis of the join semi-lattice structure of the domain of events. Having clarified the algebraic structure of the domain of ⟨event, object⟩ pairs, we may then offer our formalization of both Standardization and Generalization in a pair-quantificational setting. We will conclude this section by showing that our modified pair-quantification approach yields a uniform treatment of both count nouns and mass nouns, as desired.

To tackle the first point above, we may simply define a partial ordering on ordered ⟨event, object⟩ pairs in terms of the partial ordering on events and the partial ordering on objects, both of which we have already assumed to be well-defined.[5]

(24) a. $\langle e, x\rangle \subseteq \langle e', x'\rangle \ \leftrightarrow \ \langle e, x\rangle \oplus \langle e', x'\rangle = \langle e', x'\rangle$

[4] $\circ$ stands for *overlap.* We say that an object x overlaps an object y just in case there is an object z which is part of both x and y.

$$x \circ y \ \leftrightarrow \ \exists z[z \subseteq x \ \wedge \ z \subseteq y]$$

[5] Since, by definition, $x \subseteq y \ \leftrightarrow \ x \oplus y = y$, (24a) and (24b) together entail that

$$\langle e, x\rangle \subseteq \langle e', x'\rangle \ \leftrightarrow \ e \subseteq e' \ \wedge \ x \subseteq x'$$

b. $\langle e,x\rangle \oplus \langle e',x'\rangle = \langle e'',x''\rangle \leftrightarrow e \oplus e' = e'' \wedge x \oplus x' = x''$

Given the partial ordering on ⟨event, object⟩ pairs as defined in (24), we can easily determine the lattice-algebraic structure of the domain of ⟨event, object⟩ pairs on the basis of the lattice-algebraic structure of the domain of events. We have already assumed, following standard literature, that the domain of events is structured in a join semi-lattice. We simply note here that this assumption follows from the hypothesis that events are cumulative but the domain of events has no bottom element (cf. Chapter 1 for why this implication holds). That is, the lattice sort for events is assumed to be constrained by the following axiom:

(25) AXIOM (THE LATTICE SORT FOR EVENTS HAS NO BOTTOM ELEMENT)
$\neg\exists e \forall e' : e \subseteq e'$

This axiom expresses the intuition that no event can be claimed to be included in all other events. On the basis of (24) and (25), it is relatively easy to show that the domain of ⟨event, object⟩ pairs cannot have a bottom element either. That is, we can prove the following fact:[6]

(26) FACT *If the lattice sort for events has no bottom element, then the lattice sort for ⟨event, object⟩ pairs has no bottom element either.*

$$[\neg\exists e \forall e' : e \subseteq e'] \rightarrow [\neg\exists e,x\, \forall e',x' : \langle e,x\rangle \subseteq \langle e',x'\rangle]$$

Given that we have defined the $\oplus$ operation on ⟨event, object⟩ pairs in terms of the $\oplus$ operation on events and the $\oplus$ operation on objects, it also follows that the domain of ⟨event, object⟩ pairs is closed under sums, as long as the domains of events and objects are. This observation then entails, together with the fact in (26), that the domain of ⟨event, object⟩ pairs is structured in a join semi-lattice as well. Since this will turn out to be of crucial importance when we turn to the Weak Island effects on ER in Section 5, we state this result in (27) for ease of reference.

[6] *Proof*: Suppose $\neg\exists e \forall e' : e \subseteq e'$, and assume there is a $\langle e,x\rangle$ such that for all $\langle e',x'\rangle$, $\langle e,x\rangle \subseteq \langle e',x'\rangle$, aiming for a contradiction. From the latter assumption it follows by definition that for all $\langle e',x'\rangle$:

$$\langle e,x\rangle \oplus \langle e',x'\rangle = \langle e',x'\rangle$$

That is, for all e',x': $e \oplus e' = e'$ and $x \oplus x' = x'$. But then there is an e such that for all $e' : e \subseteq e'$, contradicting our initial assumption that $\neg\exists e \forall e' : e \subseteq e'$. Therefore:

$$\neg\exists e,x \forall e',x' : \langle e,x\rangle \subseteq \langle e',x'\rangle \qquad \Box$$

(27) The domain of ⟨event, object⟩ pairs constitutes a (proper) join semilattice.

Having thus determined the algebraic structure of the domain of ⟨event, object⟩ pairs, we can now turn to the implementation of Krifka's proposal in pair-quantificational terms. Recall from the preceding discussion that the core of Krifka's analysis could be stated in two separate clauses: the Standardization clause, in which he standardized ER with respect to OR, and the Generalization clause, in which he generalized over all the partial results, each obtained by Standardization as applied to the respective non-iterative subevent. Instead of using the two clauses to define a measure function on events in terms of its corresponding measure function on objects, as Krifka did, we will make use of them to define a measure function μ' on ⟨event, object⟩ pairs in terms of its corresponding measure function μ on objects. (28) below then formalizes the Standardization clause and the Generalization clause in a pair-quantificational setting. Note that Generalization simply claims additivity (cf. 23 above) for measure functions on ⟨event, object⟩ pairs.

(28) i. **Standardization**

For any event e, object x, and for any event relation R and nominal predicate N,

$$[\neg\mathrm{ITER}(e,x,R) \wedge N(x) \wedge R(e,x)] \rightarrow [\mu'(e,x) = n \leftrightarrow \mu(x) = n]$$

ii. **Generalization**

For any events e and e', and for any objects x and x',

$$[\neg\langle e,x\rangle \circ \langle e',x'\rangle \wedge \mu'(e,x) = n \wedge \mu'(e',x') = n'] \rightarrow [\mu'(\langle e,x\rangle \oplus \langle e',x'\rangle) = n + n']$$

To illustrate the workings of (28), let us have a look again at the ER of our model sentence *4,000 ships passed through the lock,* repeated here as (29a). According to the present account, the truth conditions for both its OR and ER can be captured as in (29b) and (29c), respectively.[7]

(29) a. 4,000 ships passed through the lock

[7] As a sidenote, we may observe that, in a sense that can be made precise, a measure function such as COUNT determines the quantificational force in representations of the type exemplified in (29). To see this, it may be instructive to think of the x's in (29) in terms of the notion *witness set,* as employed in Generalized Quantifier theory (cf. Barwise and Cooper 1981, and Chapter 1). In terms of this notion, we could represent the truth conditions of the OR of (29a), as in (ia) or (ib).

b. $\exists e \exists x[\text{ships}'(x) \wedge \text{passed-through-the-lock}'(e,x) \wedge \textsc{count}(x) = 4000]$

c. $\exists e, x[\text{ships}'(x) \wedge \text{passed-through-the-lock}'(e,x) \wedge \textsc{count}'(e,x) = 4000]$

Leaving the unproblematic OR as represented in (29b) aside, we may recall from the preceding discussion that the unqualified pair-quantification approach was perfectly well equipped to handle the ER of (29a) (cf. 13 above), in virtue of the fact that the eventive predicate *pass through the lock* is fully distributive with respect to its external argument. Now, it is fairly easy to see that the revised pair-quantificational analysis, as stated in (28), does not upset this result. Suppose we evaluate the representation in (29c) with respect to an iterative event e. (28) would then require that we partition e first into non-iterative subevents $e_1, \dots, e_m$. With respect to these non-iterative subevents, Standardization in (28i) allows us to infer the facts in (30).

(30) a. $\text{ships}'(x_1) \wedge \text{passed-through-the-lock}'(e_1, x_1)$
$\wedge\ \textsc{count}'(e_1, x_1) = \textsc{count}(x_1) = n_1$

$\vdots$

m. $\text{ships}'(x_m) \wedge \text{passed-through-the-lock}'(e_m, x_m)$
$\wedge\ \textsc{count}'(e_m, x_m) = \textsc{count}(x_m) = n_m$

That is, for each non-iterative subevent e_i, $\textsc{count}'$ as applied to the pertinent pair $\langle e_i, x_i \rangle$ simply gives the number of ships n_i that passed through the lock in e_i, in accordance with Standardization. This recalls Krifka's earlier observation that non-iterative events blur the distinction between OR and ER. Note now that this observation is also captured by the unqualified pair-quantificational

(i) a. $4000x : \text{ship}'(x)(\exists e[\text{passed-through-the-lock}'(e,x)])$

b. $\exists e \exists W[W \subseteq \text{ships}' \wedge W \subseteq \{x \mid \text{passed-through-the-lock}'(e,x)\} \wedge |W| = 4000]$

Thus, the similarity between (29b) and (ib) becomes apparent in that COUNT counts the atomic members in the plural object x, whereas $|\dots|$ in (ib) counts the members in the witness set W. By extending the analogy then, we may refer to the representation in (29c) as a pair-quantificational structure, where COUNT now counts the atoms in the "plural" ⟨event, object⟩ pair.

Let us not forget, however, that the equivalence in (i) only holds for monotone increasing quantifiers. For non-monotonic and decreasing quantifiers we crucially need to refer to maximality conditions. A similar need for maximality conditions when confronted with non-monotonic and decreasing quantifiers is also recognized by those authors who are concerned with a lattice-algebraic characterization of natural language quantification (cf. for instance Link 1987, Krifka 1989). For concreteness, we will simply adopt Krifka's (1989, 1990) solution to the maximality problem in the latter framework, which crucially refers to the notion of a maximal event (at a reference time). Cf. especially Krifka (1989) for detailed proposals along these lines with respect to the treatment of quantification in event semantics.

approach to ER. We only need to observe that on the basis of the definition in (22), the following fact must hold.[8]

(31) FACT *For any event relation R:*

$$\begin{array}{ll} \textit{if} & \forall e \forall e' \forall x[\langle e,x\rangle \in R \ \wedge \ \langle e',x\rangle \in R \ \rightarrow \ e = e'] \\ \textit{then} & x \neq x' \ \leftrightarrow \ \langle e,x\rangle \neq \langle e',x'\rangle \end{array}$$

That is, for all non-iterative events e, whether we count the number of ⟨event, ship⟩ pairs that were involved in a passing through the lock in e, or simply the number of actual ships themselves that passed through the lock in e, the result should be the same, in the light of the fact in (31).

By Generalization (28ii) then, if we fuse all the non-overlapping $\langle e_i, x_i\rangle$ pairs into the global, ordered pair $\langle e, x\rangle$,[9] we can add up the values for the non-iterative cells of the partition in (30) (i.e. $n_1 + \ldots + n_m$), the total sum of which should equal 4,000 if (29c) is to come out true in e.[10] Evidently, this should be identical to the total number of ordered ⟨event, ship⟩ pairs, as any one of these pairs will not show in more than one cell of the partition. This means that for constructions in which a count noun is combined with a distributive predicate, the unqualified pair-quantification approach yields the same, correct result as its modified variant in (28), as desired.

However, (28) is significantly more general than the unqualified pair-quantification approach in that it readily extends to ER effected by mass nouns. Recall from the preceding section that the latter approach ascribed truth conditions to these readings that are satisfied by only a proper subset of the admissible models (cf. 17 above). According to our present account, the ER of the sentence in (16) for example, repeated here as (32a), has the truth conditions stated in (32b).

---

[8]The proof of this fact is similar to the proof of Fact 15.

[9]In general, if $\neg e \circ e'$, then $\neg\langle e,x\rangle \circ \langle e',x'\rangle$ for any x and x', since $\langle e,x\rangle \circ \langle e',x'\rangle$ just in case $e \circ e'$ and $x \circ x'$, as the reader may check for him/herself on the basis of the definition of overlap (cf. footnote 4) and (24) above. In fact, we could strengthen the above implication into a bi-conditional, by showing that for any x and x',

$$(e \circ e') \ \rightarrow \ (\langle e,x\rangle \circ \langle e',x'\rangle)$$

holds as well. Informally, the latter claim straightforwardly follows from the intuitively plausible assumption that if two events e and e' overlap, then the participants in e and e' must overlap as well. For reasons of space, we will refrain here from presenting the formal proof.

[10]Notice that this step assumes *summativity* for eventive predicates. It may be defined as follows (cf. Krifka 1989, 1990).

(i) DEFINITION (SUMMATIVITY)

$$\forall e \forall e' \forall x \forall x'[R(e,x) \ \wedge \ R(e',x') \ \rightarrow \ R(e \oplus e', x \oplus x')]$$

Along with Krifka (1989, 1990), we will assume that this property holds of all eventive predicates throughout the rest of this section.

(32) a. 4,000 tons of radioactive waste passed through the lock (last year)

b. $\exists e, x[\text{radioactive-waste}'(x) \wedge \text{passed-through-the-lock}'(e, x) \wedge$ $\text{TONS-OF}'(e, x) = 4000]$

To demonstrate the full generality of (28), suppose we are given the toy model in (33), which verifies (32a) on its ER.

(33) a. $\langle e_1, x_1 \rangle \in$ passed-through-the-lock$'$; where TONS-OF$(x1)$ =2,500.00

b. $\langle e_2, x_2 \rangle \in$ passed-through-the-lock$'$; where TONS-OF$(x2)$ = 389.12

c. $\langle e_3, x_2 \rangle \in$ passed-through-the-lock$'$; where TONS-OF$(x2)$ = 389.12

d. $\langle e_4, x_3 \rangle \in$ passed-through-the-lock$'$; where TONS-OF$(x3)$ = 212.64

e. $\langle e_5, x_4 \rangle \in$ passed-through-the-lock$'$; where TONS-OF$(x4)$ = 509.12

4,000.00

Since the same bit of radioactive waste participates in e_2 and e_3, we cannot apply the Standardization clause in (28i) directly, $(e_1 \oplus e_2 \oplus e_3 \oplus e_4 \oplus e_5)$ being an iterative event in the sense of Definition (22). However, we may partition the global event into the non-iterative subevents $(e_1 \oplus e_2)$ and $(e_3 \oplus e_4 \oplus e_5)$, for which Standardization will yield the values 2,889.12 and 1,110.88 respectively, since 2,889.12 tons of radioactive waste passed through the lock in the first non-iterative cell, whereas in the last non-iterative cell 1,110.88 tons of radioactive waste passed through the lock. That is, Standardization allows us to infer the facts in (34).

(34) a. $\text{passed-through-the-lock}'(e_1 \oplus e_2, x_1 \oplus x_2) \wedge \text{radioactive-waste}'(x_1 \oplus x_2) \wedge \text{TONS-OF}'(e_1 \oplus e_2, x_1 \oplus x_2) = \text{TONS-OF}(x_1 \oplus x_2) = 2889.12$

b. $\text{passed-through-the-lock}'(e_3 \oplus e_4 \oplus e_5, x_2 \oplus x_3 \oplus x_4) \wedge \text{radioactive-waste}'(x_2 \oplus x_3 \oplus x_4) \wedge \text{TONS-OF}'(e_3 \oplus e_4 \oplus e_5, x_2 \oplus x_3 \oplus x_4) = \text{TONS-OF}(x_2 \oplus x_3 \oplus x_4) = 1110.88$

Given the fact that $(e_1 \oplus e_2)$ does not overlap $(e_3 \oplus e_4 \oplus e_5)$, $\langle (e_1 \oplus e_2), (x_1 \oplus x_2) \rangle$ does not overlap $\langle (e_3 \oplus e_4 \oplus e_5), (x_2 \oplus x_3 \oplus x_4) \rangle$ either. By Generalization (28ii), we may therefore infer that

(35) $\text{TONS-OF}'(\langle e_1 \oplus e_2, x_1 \oplus x_2 \rangle \oplus \langle e_3 \oplus e_4 \oplus e_5, x_2 \oplus x_3 \oplus x_4 \rangle) =$
$\text{TONS-OF}'(e_1 \oplus e_2 \oplus e_3 \oplus e_4 \oplus e_5, x_1 \oplus x_2 \oplus x_3 \oplus x_4) =$
2889.12 (cf. 34a) + 1110.88 (cf. 34b) = 4000

which is the desired result.[11]

[11]Krifka's (1990) analysis has recently been challenged by Moore (1994), who observes that the requirement on partitioning into non-iterative subevents, crucial in Krifka's framework,

Summarizing the discussion in this section then, we have seen that the naive pair-quantificational analysis of ER should at least be modified in light of the fact that constructions with mass nouns may exhibit ER as well. The last construction type points to the need of defining a measure function on ⟨event, object⟩ pairs in terms of its corresponding measure function on objects: the measure functions with which mass nouns combine, like *tons (of),* or *meters (of)* etc., simply cannot be applied directly to ordered pairs of events and objects. Technically, this means that measure functions on ⟨event, object⟩ pairs should be standardized with respect to their corresponding measure function on objects. Krifka's (1990) approach to ER offers the tools to resolve this dilemma. Implementing his basic insight that ER may be standardized with respect to OR, either directly in case we are dealing with a non-iterative event, or indirectly by partitioning the iterative event into non-iterative cells, we could define the "measure" with which our quantifier over ⟨event, object⟩ pairs is associated in terms of the measure with which the pertinent nominal argument is associated. Given the fact that our modified pair-quantification approach in (28) yields the same, correct results for distributive, verbal predicates as the naive pair-quantification approach, essentially on account of the fact in (31) above and the results of Section 3.1, we will henceforth restrict our attention to the ER of fully distributive, verbal predicates and their representation in terms of naive pair-quantification, to keep things maximally simple. In the next sections, we will discuss the effects of strong versus weak quantifiers on ER, and the sensitivity of this reading to weak islands. The pair-quantificational approach to ER that we developed in this section allows us to capture these facts in an elegant and relatively straightforward way.

---

becomes highly problematic and counterintuitive when we consider the ER of a sentence such as *57,000,760 liters of water passed through the pump last month,* imagining a closed system with one million liters of water circulating through a pump (cf. also Doetjes 1994 for similar observations). The problem here is that it seems hard, if not impossible, to impose a non-iterative partitioning on the global event referred to in this sentence. Taking it to its extreme, it would require us to identify every bit of water, down to the molecular level, to check whether it occurs twice in the same cell of the partition. Needless to say, Moore's criticism carries over to our analysis, as developed in this section, as well. Although we fully agree with the points raised by Moore (1994), it still remains to be seen whether an alternative approach to ERs, based on Moore's as yet informal notion of *eventual objects,* can account for the grammatical restrictions on ER that we will discuss in the next two sections. There we will argue that a pair-quantificational analysis of ERs, which incorporates some of the tools developed by Krifka, yields these restrictions in a natural and uniform fashion.

4 SYMMETRIC VERSUS NON-SYMMETRIC DETERMINERS AND THE AVAILABILITY OF ER

In this section we will see that ER is sensitive to the weak/strong distinction in the class of determiners. Specifically, we will observe that, whereas noun phrases with symmetric (weak) determiners may freely give rise to both OR and ER, noun phrases with non-symmetric (strong) determiners support ER only under special circumstances. We argue that this contrast falls out immediately from the pair-quantificational approach to ER. Pair-quantifiers require both the event-variable and the object-variable to be introduced in their restriction. This entails that ER is available only if the predicate that supplies the event variable occurs in the determiner's restriction—something that does not happen across the board. But there are at least three distinct possibilities for it to happen. First, symmetric determiners allow the eventive predicate to be mapped into their restriction together with the head noun by semantic inference. Second, the presence of focus may give rise to the necessary rebracketing. Third, when a relative clause modifies the determiner's restriction, its eventive predicate may serve the purpose. Of these three options, the second and the third are available to noun phrases with non-symmetric (strong) determiners as well. Part of the discussion in this section will be cast in a *dynamic* setting (cf. Dekker 1990, 1993, Groenendijk and Stokhof 1991, Chierchia 1992, Kanazawa 1994b, 1994a), which allows us to address some striking similarities between donkey-anaphora and ER. These similarities concern the ways in which anaphoric dependencies between (implicit event) arguments interact with either ER or so-called symmetric (pair-quantificational) readings for donkey sentences.

4.1 Symmetry, or How to get an eventive predicate in the restriction of the ER pair-quantifier by semantic inference

Up to this point, one might have the impression that ER is simply available across the board. However, such an impression would be mistaken. There are severe restrictions on the availability of ER. Consider the sentences in (36), which allow for both OR and ER, and the sentences in (37), which only allow for OR:

(36) a. Last night, (at least) 4,000 ships passed through the lock(OR/ER)

OR: there are (at least) 4,000 ships each of which passed through the lock last night

ER: there were (at least) 4,000 events in which a ship passed through the lock last night

b. Last night, exactly 4,000 ships passed through the lock (OR/ER)

c. Last night, many ships passed through the lock (OR/ER)

d. Last night, some ships passed through the lock (OR/ER)

(37) a. Last night, most ships passed through the lock (OR/*ER)

OR: most ships are such that each passed through the lock last night

ER: * most events in which a ship passed through the lock (last night) occurred (last night)

b. Last night, every ship passed through the lock (OR/*ER)

c. Last night, 60% of the ships passed through the lock (OR/*ER)

d. Last night, the 4,000 ships passed through the lock (OR/*ER)

The following question then naturally suggests itself: What accounts for the stark contrast between (36) and (37)? One important clue can be found in the fact that, correlating with the distinction in available readings, there is a sharp distinction in denotational properties between the determiners that head the subject QPs: those in (36) are symmetric on at least one of their readings, whereas those in (37) are non-symmetric on every reading. The notion of *symmetry* is of course a familiar one in Generalized Quantifier (GQ) theory:

(38) DEFINITION (SYMMETRY) *A quantifier Q is* symmetric *just in case*

$$Q_E(A)(B) \quad \textit{iff} \quad Q_E(B)(A)$$

Thus, on its so-called weak (or non-specific) reading, the quantifier AT LEAST 4,000 allows its NP-argument and its VP-argument to swap places without any difference in meaning, as illustrated in (39) below. However, this is not the case for the determiners in (37). To see this, it suffices to note that the left-hand side of the biconditional in (40b) is not equivalent in truth conditions to the right-hand side on any possible construal for *most.*

(39) a. At least 4,000 doctors are activists

b. $\models$ at least 4,000 doctors are activists $\leftrightarrow$ at least 4,000 activists are doctors

(40) a. Most doctors are activists

b. $\not\models$ most doctors are activists $\leftrightarrow$ most activists are doctors

Assuming this to be the right characterization of the semantic parameter that differentiates between those QPs that allow for ER and those that do not, we might ponder the question why symmetry would make a difference for ER. Recall now that we assumed ER to be represented by means of pair-quantification, an assumption that would ascribe the logical form in (41c) to the ER of example (36a), repeated here as (41a). We will simply take it that the OR of (41a) is to be represented as in (41b).

(41) a. Last night, 4,000 ships passed through the lock

b. OR: $4000x : \text{ship}'(x)(\exists e[\text{passed-through-the-lock}'(e,x)])$

c. ER: $4000\langle e,x\rangle : \text{ship}'(x) \ \wedge \ \text{passed-through-the-lock}'(e,x)$

Now, the thing to note about (41c) is that, as part of a more general, syntactic requirement on the format of quantification, both the event-variable and the object-variable need to be introduced in the restriction of the pair-quantifier, as crucially both variables participate in determining the restricted domain over which the pair-quantifier ranges.[12] This means then that if we want to interpret the numeral *4,000* as a quantifier over ordered pairs of events and objects, we at least need to be able to "transfer" the eventive predicate, which contains the Davidsonian (event) argument, from the Nuclear Scope (the second argument) of the quantifier 4,000 into its Restrictive Clause (the first argument). (We will also need to dispose of the existential quantifier that binds the event-variable, an issue that is to be taken up in what follows.) This is precisely the point where symmetry becomes crucial. It is fairly easy to show that, given a conservative quantifier Q (cf. definition 42), all and only those Q's that are symmetric may be translated, while preserving their truth conditions, into predicative (or 1-placed) quantifiers in which the original Restrictive clause and Nuclear Scope now jointly exhaust their only argument. That is, on the

[12]Technically, the term ***restriction*** cannot be applied to quantificational structures of the type exhibited in (41c), the reason being that this term strictly applies to the first argument of a two-placed quantifier Q,

$$Q : \mathcal{P}(E^n) \times \mathcal{P}(E^n) \ \rightarrow \ \{0,1\}$$

However, we may note that for all and only the symmetric Q's, we have

$$Q_{E^n}(P^n \cap Q^n) \quad \text{iff} \quad Q_{E^n}(P^n \cap Q^n, E^n)$$

for arbitrary n-placed predicates P and Q. In light of this, we could have expanded (41c) into (i).

(i) $4000\langle e,x\rangle : \text{ships}'(x) \ \wedge \ \text{passed-through-the-lock}'(e,x)(\text{EVENTS} \times \text{OBJECTS})$

assuming a sorted universe $E = \text{EVENTS} \cup \text{OBJECTS}$. With respect to (i) then, we can indeed say that the pair-quantifier has a restriction. However, to avoid unnecessary complications, we will continue talking about the restriction of a quantifier, even in cases where the relevant quantifier has only one argument.

basis of (38) and (42); we may derive the fact in (43) (cf. Chapter 1 of this volume).[13,14]

(42) DEFINITION (CONSERVATIVITY) *A quantifier Q is conservative just in case*

$$Q_E(A)(B) \quad \textit{iff} \quad Q_E(A)(A \cap B)$$

(43) FACT *For all conservative and symmetric determiners Q,*

$$Q_E(A)(B) \quad \textit{iff} \quad Q_E(A \cap B)(A \cap B)$$
$$\textit{iff} \quad Q_E(E)(A \cap B)(\stackrel{\text{def}}{=} Q_E(A \cap B))$$

By semantic inference, (43) allows us to map an eventive predicate in the Nuclear Scope (henceforth, NS) into the restriction of a symmetric quantifier. That is, on the basis of (43), we can infer that (41b) is truth-conditionally equivalent to (44), where the eventive predicate is now contained in the restriction of the symmetric quantifier 4,000.

(44) $4000x : \text{ship}'(x) \;\wedge\; \exists e[\text{passed-through-the-lock}'(e, x)]$

4.2 Pair-quantification and existential disclosure

We now need to address the issue of how to get from (44) to the pair-quantification structure in (41c). Putting it crudely, the problem is how to get rid of the existential quantifier in (44) so that the quantifier 4,000 can "unselectively" bind its event-variable. We suggest that this problem is similar to the task of having the QP adverb *always* in (45a) quantify over the indefinite *a cat* so as to obtain a reading that could be represented as in (45b) (cf. Pelletier and Schubert 1989).

(45) a. Always, when a cat has blue eyes, it is intelligent.

b. $\forall x[\text{cat}'(x) \;\wedge\; \text{has-blue-eyes}'(x) \;\rightarrow\; \text{intelligent}'(x)]$

We will adopt Chierchia's (1992) proposal for the latter problem, and extend it to our case. Chierchia's (1992) solution to the problem presented by (45) is couched in a dynamic framework, which is essentially a variant of Groenendijk

[13] It is standardly assumed that every natural language quantifier is conservative.

[14] Note that we have made use in (43) of the following notational convention:

$$Q_E(A)(B) \stackrel{\text{def}}{=} Q_A(B)$$

and Stokhof's (1991) system of Dynamic Predicate Logic (DPL), enriched with Generalized Quantifiers. For our purposes, it suffices to look at DPL as an extension of ordinary Predicate Logic (PL) in which the notion of the syntactic scope of a quantifier has been strengthened. Note that although in ordinary PL the equivalence in (46) below does not hold, in DPL it does, thus demonstrating the dynamic principle according to which the scope of an existential quantifier can be extended to the right.[15] The equivalence in (46) is of key importance in Chierchia's solution to the problem case in (45), and will figure prominently in the remainder of this section as well.

(46) $\exists x[Px] \wedge Qx \equiv \exists x[Px \wedge Qx]$

As Chierchia points out, (46) suggests a general procedure for abstracting over a variable which was originally bound by an existential quantifier. If we replace Qx in (46) by the equation $x = x'$, where x' is a free variable, we get the desired result, as shown in (47) below.

(47) $\lambda x'[\exists x[Px] \wedge x = x'] \equiv \lambda x'[\exists x[Px \wedge x = x']]$ $(\equiv \lambda x'[Px'])$

This procedure of abstracting over a variable which was originally quantified over by an existential quantifier is called *Existential Disclosure* and was first proposed and developed in Dekker (1990, 1993). It can be defined as in (48), where ϕ contains the existential quantifier which is in need of disclosure.

(48) EXISTENTIAL DISCLOSURE (ED)

$$\lambda x[\phi] \stackrel{\text{def}}{=} \lambda x'[\phi \wedge x = x']$$

Chierchia then observes that if Existential Disclosure is applied to the indefinite *a cat* in (45), the QP adverb *always* may subsequently quantify over its "disclosed" variable. Assuming that (45) and the ER of (36a) should be treated on a par, we simply note that application of Existential Disclosure to (44), and subsequent "unselective" binding of the disclosed event argument by the quantifier 4,000 will yield the desired pair-quantification structure in (41c), as shown below.[16]

(49) a. $4000 : \lambda e'\lambda x[\text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e', x)]]$

[15] The equivalence in (46) allows DPL to treat cross-sentential anaphora successfully, as in *A man came in. He whistled.* For more details on DPL and its applications to anaphor binding, the reader should consult Groenendijk and Stokhof (1991).

[16] We have included λ-expressions in (49) to highlight the fact that Existential Disclosure functions as a type-shifting rule, converting an n-ary predicate into an $n+1$-ary predicate. In the remainder of this section, we omit the λ-expressions from representations like (49) for the sake of readability.

b. $4000 : \lambda e \lambda x[\text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e', x)] \wedge e' = e]$ ED

c. $4000 : \lambda e \lambda x[\text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e', x) \wedge e' = e]]$ (46)

d. $4000 : \lambda e \lambda x[\text{ship}'(x) \wedge \text{passed-thru-the-lock}'(e, x)]$ (47)

A few notes are in order here. First of all, claiming that a determiner-function may "unselectively" bind an eventive variable which is abstracted over by Existential Disclosure is tantamount to claiming that, corresponding to every ordinary determiner-function Q that relates sets of individuals, there is a quantifier Q' that quantifies over ordered pairs of events and objects, and which is furthermore expressed by the same lexical element. Thus, the pair-quantifier $4000\langle e, x\rangle$ in (49) is by no means derived from the quantifier $4000x$: it is simply another interpretation which is available to the numeral *4,000,* requiring both the object and event-variable to be properly introduced in its restriction, as was already observed above. Secondly, we need to ensure that the process of abstracting over existentially quantified variables by Existential Disclosure is sufficiently constrained. Unless we do so, we predict, given the assumptions thus far, that a sentence like *Four boys saw a dog* has a reading which is true in a situation where a single boy saw four dogs. We will stipulate that a determiner-function may "unselectively" bind an existentially quantified variable after Existential Disclosure has applied to it just in case that variable is an event-variable.[17]

4.3 Non-symmetric determiners

Having taken care of the availability of ER for symmetric determiners, we are now in the position to offer an explanation for why the same readings are not available for non-symmetric (strong) determiners. Note first that (43) entails that strong determiners will not allow an eventive predicate in the NS to be mapped into their restriction by semantic inference. And indeed, as (43) would lead us to expect, we observe that the left-hand side of the biconditional in (50a), the natural language paraphrase of which is given in (50b), is not truth-conditionally equivalent to the right-hand side.

(50) a. $\not\models \textsc{Most}x : \text{ship}'(x)(\exists e[\text{passed-through-the-lock}'(e, x)]) \leftrightarrow$
$\textsc{Most}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e, x)]$

[17] We might speculate that this constraint derives from the fact that only ⟨event, object⟩ pairs can be treated as nominal entities over which determiner-functions may quantify. That is, we might look on our ⟨event, object⟩ pairs as an alternative implementation of Carlsonian stages (i.e. spatio-temporal slices of individuals; cf. Carlson 1977, Carlson 1982), if any sense can be made of these entities independently of the phenomenon we are considering here. This possibility merits further investigation.

b. $\not\models$ Most ships passed through the lock (OR) $\leftrightarrow$ Most ships that passed through the lock exist (OR)

Since the pair-quantifier of ER needs both the event and the object variables to be introduced in its restriction, as we already observed above, (43) straightforwardly derives that, other things being equal, non-symmetric determiners cannot give rise to ER. This accounts for the patterns of available readings observed in (36) and (37).

Summarizing the discussion up to this point, we have seen that only symmetric (weak) determiners may give rise to both OR and ER in the default case. This observation essentially follows from the pair-quantification approach to ER. Since the ER pair-quantifier requires both the event-variable and the object-variable in its restriction, we would expect there to be a difference between symmetric and non-symmetric quantifiers in the first place, for only symmetric quantifiers allow an eventive predicate in the NS to be mapped into their restriction by semantic inference, as illustrated in (44). We then stipulated that Existential Disclosure applies to the resulting structure, enabling the symmetric quantifier to "unselectively" bind the "disclosed" event-variable.

4.4 Focus

In the following two subsections we discuss two constructions (focus and modification by a relative clause respectively) in which strong quantifiers can give rise to ER. We argue that these constructions may be taken to create the required bracketing in which an eventive predicate is contained in the Restrictive Clause of the strong quantifier, thus allowing the quantifier to "unselectively" bind the existentially disclosed, eventive variable. It is here that we will make full use of the dynamic property in (46) in order to capture some striking similarities between ER and donkey-anaphora. These similarities in turn suggest that our previous call on (46) is well-motivated.

Krifka (1990) observes that strong quantifiers suddenly allow for ER when another constituent of the sentence receives focal accent.[18] (51) below illustrates this effect for the sentences in (37), where capital letters indicate focal accent.

(51) a. Most ships passed through the lock AT NIGHT (OR/ER)

OR: most ships that passed through the lock are such that each passed through the lock at night (cf. Partee 1991)

ER: most events in which a ship passed through the lock occurred at night

[18] Note that Krifka does not correlate this focus effect to the distinction between weak and strong quantifiers. This connection was first observed in Honcoop (1992), who referred to the latter distinction as a Definiteness Effect.

b. Every ship passed through the lock AT NIGHT (OR/ER)

c. 60% of the ships passed through the lock AT NIGHT (OR/ER)

d. The 4,000 ships passed through the lock AT NIGHT (OR/ER)

Now, it is a generally acknowledged descriptive fact that focal accent on a predicate expression correlates with a specific semantic effect. Thinking in terms of a relational quantification structure $Q(A, B)$, where A is the restrictive term and B is the NS, Partee (1991) describes the semantic effect of focus as follows: focus semantically maps a predicate expression to the NS B, whereas the focus neutral material (or focus frame) is forced into the restrictive term. The following pair of sentences, from Rooth (1985), illustrates this effect, where the meaning of the sentence in (52a) can be paraphrased as in (52b), while the meaning of the example in (53a) allows for the paraphrase in (53b).[19]

(52) a. Officers always escorted BALLERINAS

b. Always, when officers escorted someone, officers escorted ballerinas

(53) a. OFFICERS always escorted ballerinas

b. Always, when someone escorted ballerinas, officers escorted ballerinas

According to the above generalization then, we would expect the focused adverbial phrase in (51a) to be mapped into the NS of the quantifier MOST, subsequently forcing the focus frame into its restrictive term. That is, analogous to the cases discussed in (52) and (53), we would expect focus here to effect the rebracketed structure in (54a), whose meaning can be spelled out as in (54b).

(54) a. Most ships that passed through the lock (at some event), passed through the lock at night

b. $\textsc{Most}x : \text{ship}'(x) \land \exists e[\text{passed-through-the-lock}'(e, x)]$
$(\text{passed-through-the-lock}'(e, x) \land \text{at-night}'(e))$

Although (54b) has the appearance of an illicit binding configuration, on account of the fact that the existential quantifier, being trapped in the restrictive term of MOST, is unable to bind the corresponding event-variables in the NS, this is not so under the dynamic assumptions we adopted. Assuming dynamic generalized quantifiers to be conservative just as much as their static counterparts are (cf. Chierchia 1992, Kanazawa 1994b, 1994a), we may note that

[19] Herburger's (1994) focus-affected readings illustrate much the same point. She observes for instance that a sentence like *Many SCANDINAVIANS won the Nobel prize* (based on Westerståhl 1985b) can have an interpretation according to which it is claimed that, relative to the number of Nobel prize winners, there were many that were Scandinavian. This semantic effect points to the same descriptive generalization as the facts in (52) and (53).

(54b) is equivalent to (55a) below under Conservativity. Thanks to the dynamic property in (46), (55a) in turn reduces to (55b), which is equivalent under Conservativity to (55c).[20]

(55) a. $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)] \wedge$
$\text{passed-through-the-lock}'(e,x) \wedge \text{at-night}'(e))$ CONS

b. $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x) \wedge$
$\text{at-night}'(e)])$ (46)

c. $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\exists e[\text{passed-through-the-lock}'(e,x) \wedge \text{at-night}'(e)])$
CONS

All in all, then, if no special assumptions are made with respect to the focus-induced structure in (54b), we simply obtain the (focus-affected) OR for the sentence in (51a), since (54b) = (55c) comes out true just in case most of the ships that passed through the lock, passed through the lock at night. That is, whether or not most ships that passed through the lock were responsible for less than half of the events in which a ship passed through the lock at night is irrelevant to determining the truth of (55c).[21] The other cases in (51) can be treated in a similar fashion.

[20] With respect to the last step in the derivation, observe that

$$Q(A \cap B, A \cap B \cap C) \text{ iff } Q(A \cap B, B \cap C)$$

is just a particular instance of (42).

[21] There is a small caveat here, in that (55c) does not express that *all* the lock passages in which most of the ships were involved occurred at night, as the focus-affected OR of (51a) would require. To remedy this, we need to be more specific about the contents of dynamic Conservativity, without dwelling too much on all the formal details. Kanazawa (1994b) shows that the following two definitions of a dynamic generalized quantifier Q^D in terms of its static counterpart Q satisfy his version of dynamic Conservativity.

(i) $Q^D_W x(\phi,\psi) \leftrightarrow Qx(\phi, \phi \wedge \psi)$ (the *weak* definition)

(ii) $Q^D_S x(\phi,\psi) \leftrightarrow Qx(\phi, \phi \rightarrow \psi)$ (the *strong* definition)

Note that our use of dynamic Conservativity in the main text appeals to the definition in (i). But since (ii) is just as much consistent with Kanazawa's formulation of dynamic Conservativity, we could have appealed to this definition instead in (55). This would have led to the representation in (iii), which does capture the truth-conditions of the focus affected OR of (51a). To see this, we simply note that in DPL the equivalence in (iv) is valid (cf. Groenendijk and Stokhof 1991).

(iii) $\text{MOST}x : \text{ship}'(x) \wedge$
$\exists e[\text{passed-through-the-lock}'(e,x)](\forall e[\text{passed-through-the-lock}'(e,x) \rightarrow$
$\text{passed-through-the-lock}'(e,x) \wedge \text{at-night}'(e)])$

(iv) $\exists x[\phi] \rightarrow \psi \equiv \forall x[\phi \rightarrow \psi]$

As Kanazawa (1994b) argues, the choice between the two definitions in (i) and (ii) is not random. Rather, it depends on the left and right monotonicity of the static quantifier Q: the dynamic counterpart of both left and right monotone increasing quantifiers ($\uparrow$ MON $\uparrow$; cf.

But note now that, since focus has forced the eventive predicate *passed through the lock* into the restrictive term of MOST, we could have abstracted over the eventive variable by Existential Disclosure in (54b), allowing MOST to pair-quantify over it. That is, since (54b) is equivalent in all relevant respects to the symmetric case discussed in (44) above, we would expect Existential Disclosure to be able to apply to it in much the same way as it did for (44) in (49). (56) below plots the relevant derivational steps.

(56) a. $\text{MOST}\langle e,x\rangle : \text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e',x)] \wedge e' = e$
$(\text{passed-thru-the-lock}'(e',x) \wedge \text{at-night}'(e'))$ ED

b. $\text{MOST}\langle e,x\rangle : \text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e',x)] \wedge e' = e$
$(\text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e',x)] \wedge e' = e \wedge \text{passed-thru-the-lock}'(e',x) \wedge \text{at-night}'(e'))$
CONS

c. $\text{MOST}\langle e,x\rangle : \text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e',x) \wedge e' = e]$
$(\text{ship}'(x) \wedge \exists e'[\text{passed-thru-the-lock}'(e',x) \wedge e' = e \wedge \text{at-night}'(e')])$ (46)

d. $\text{MOST}\langle e,x\rangle : \text{ship}'(x) \wedge \text{passed-thru-the-lock}'(e,x)$
$(\text{passed-thru-the-lock}'(e,x) \wedge \text{at-night}'(e))$
CONS & (47)

We may observe that (56d) adequately expresses the (focus-affected) ER of (51a). Its truth conditions are satisfied just in case most events in which a ship passed through the lock occurred last night. We have seen then that focus provides one means of creating the required bracketing in which an eventive predicate together with the head noun determines the restriction of a strong quantifier, as shown in (54), given some plausible assumptions with respect to the semantic effects of focus placement. Since structures such as the one

---

AT LEAST TWO, for example) and left and right monotone decreasing quantifiers ($\downarrow$ MON $\downarrow$; cf. NO, for example) is defined as in (i), whereas the dynamic counterpart of the remaining monotonic quantifiers (i.e. $\uparrow$ MON $\downarrow$ and $\downarrow$ MON $\uparrow$; cf. NOT EVERY and EVERY respectively) is defined as in (ii). Since dynamic Conservativity is at the heart of dynamic treatments of donkey-anaphora, as we will see in the following subsection, Kanazawa correctly predicts that the donkey-sentence *No farmer who owns a donkey beats it* only has the *weak* reading "no farmer who owns a donkey beats a donkey he owns," given that NO^D is defined as in (i), whereas the corresponding donkey-sentence with *every* strongly prefers the strong reading "every farmer who owns a donkey beats every donkey he owns," given that EVERY^D is defined as in (ii).

All in all, if our call on dynamic Conservativity for the cases in (51) is justified, we would predict that the sentence *No ship passed through the lock AT NIGHT* has the weak reading "no ship that passed through the lock at some event, passed through the lock at night at some event," but not the strong reading "No ship that passed through the lock at some event, passed through the lock at night at every event in which it passed through the lock," which is correct. The same remarks concerning the exact contents of dynamic Conservativity will carry over to our discussion of the relative clause cases as well.

in (54) satisfy the requirement on pair-quantification that both variables be introduced in the restriction of the pair-quantifier, subsequently abstracting over the event-variable by Existential Disclosure allows the strong quantifier to pair-quantify over it, thus giving rise to (focus-affected) ER. More generally, we may take the focus effects in (51) as further support for a pair-quantificational analysis of ER.

4.5 Relative clauses

Although we did not explicitly state this, we are in fact right in the business of treating ER on a par with donkey-anaphora, given our previous use of dynamic Conservativity to extend the scope of an existential quantifier in the restriction of a superordinate strong determiner into the latter's NS. Dynamic Conservativity as well as the principle in (46) are exploited by Chierchia (1992) to handle donkey-type anaphoric dependencies. Chierchia's approach may be exemplified by briefly considering how his analysis allows *a donkey* in (57a) below to extend its scope to the matrix clause. The reader may check that we have made use of the same procedure as in (57b-e) in our previous discussion of focus-affected OR (cf. for instance 56).

(57) a. Usually, if a farmer owns a donkey, he beats it

b. $\textsc{Most}x : \text{farmer}'(x) \wedge \exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y)](\text{beats}'(x,y))$

c. $\textsc{Most}x : \text{farmer}'(x) \wedge \exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y)]$
$(\text{farmer}'(x) \wedge \exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y)] \wedge \text{beats}'(x,y))$ CONS

d. $\textsc{Most}x : \text{farmer}'(x) \wedge \exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y)]$
$(\text{farmer}'(x) \wedge \exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y) \wedge \text{beats}'(x,y)])$ (46)

e. $\textsc{Most}x : \text{farmer}'(x) \wedge \exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y)]$
$(\exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y) \wedge \text{beats}'(x,y)])$ CONS

To pursue the analogy between ER and donkey-anaphora further, consider the following examples, where a strong determiner embeds a relative clause that contains an eventive predicate.

(58) a. Most ships that passed through the lock transported radioactive waste (OR/ER)

OR: most ships that passed through the lock are such that each transported radioactive waste

ER: most events in which a ship passed through the lock were events in which a ship transported radioactive waste

b. 60% of the people who visited the art exhibition bought some postcards (OR/ER)

OR: 60% of the people who visited the art exhibition are such that each bought some postcards

ER: 60% of the events in which a person visited the art exhibition were events in which a person bought some postcards

Interestingly, having an OR or ER for these sentences correlates with what we might call *sequencing of events.* More specifically, the OR for sentence (58a) for example is compatible with a situation in which each and every passing through the lock coincides exactly with an event of transporting radioactive waste, as well as with a situation in which each and every passing through the lock is truly distinct from an event of transporting radioactive waste. Let us call the first type of sequencing of events *dependent events,* and the second type *independent events.* In contradistinction to OR, ER for (58a) is only compatible with a dependent events reading: each and every event of passing through the lock coincides exactly with an event of transporting radioactive waste.

Note, however, that our notion of dependent events should be sufficiently general, so as to cover not only those cases where two events exactly coincide, but also those cases where the event referred to by the predicate in the restrictive clause is part of the stative, quasi-permanent event denoted by the matrix predicate. The latter case obtains in the sentences in (59) below, both of which, interestingly enough, may easily give rise to ER. For instance, the example in (59b) on its ER claims that most events in which a ship passed through the lock were events in which a ship had picked up its cargo in the port of Rotterdam. This interpretation therefore requires the matrix predicate to denote a property that persists through time, thereby allowing events to be included as parts that temporally follow the actual picking up of the cargo in the port of Rotterdam.[22]

(59) a. Most ships that passed through the lock had a red mast (OR/ER)

[22] Observe that, while an ER for the sentence in (59a) is perfectly possible, the same reading is simply absent in (i).

(i) 4,000 ships had a red mast

On the assumption that predicates such as *had a red mast* (individual-level predicates in the sense of Kratzer 1995 and Diesing 1992) are non-iterative (cf. de Hoop and de Swart 1989), this is immediately accounted for. To be more precise, on the assumption that the eventive predicate in (i) and (59a) is non-iterative, the ER is simply indistinguishable from the OR in (i). To account for the "true" ER of (59a) then, we may assume that individual-level predicates *IP* are constrained by the meaning postulate in (ii).

(ii) $\forall e \forall x[IP(e,x) \rightarrow \exists e'[e \subseteq e' \wedge IP(e',x)]]$

This postulate captures the inclusion relation under a dependent events reading that holds between an event referred to by predicate in the restrictive clause and the event denoted by the matrix individual-level predicate, as noted in the main text. But it also allows the

b. Most ships that passed through the lock had picked up their cargo in the port of Rotterdam (OR/ER)

Furthermore, the dependency involved in dependent events readings is indirect in that for cases such as (60) below, only the reference time of the matrix predicate is required to exactly coincide with the event of passing through the lock, due to the presence of the temporal adverbial phrase *three hours later.* More generally, (60) suggests that in a dependent events reading, the events referred to by the predicate in the restriction coincide with the events denoted by the matrix predicate just in case the reference time of the matrix predicate coincides with its event time, as is the case with the sentences in (58).

(60) Most ships that passed through the lock arrived in the port of Rotterdam three hours later (OR/ER)

Now, the observation that ER is incompatible with an independent events reading leads us to a sharp prediction. If we can force the events denoted by the eventive predicate in the relative clause to be distinct from the events denoted by the matrix predicate, say by adding in some appropriate time adverbial, we would predict that only an OR will survive. This prediction is indeed borne out, as shown by (61).

(61) Most ships that passed through the lock are now in the port(OR/*ER)

What is it that accounts for the correlation between OR and ER and the two types of sequencing of events distinguished above? We suggest that thinking of sequencing of events in terms of anaphoric relations between different event arguments will allow us to derive the above correlation from a more general, dynamic theory of donkey-type anaphora of the sort exemplified in (57) above.[23]

We will now proceed to work out this suggestion. Let us return to the OR of (58a), and concentrate first on its independent events reading. We can simply model independent events by having the event argument in the matrix predicate bound by its "own" existential quantifier, as indicated in (62a) below. For present purposes, the representation in (62a) is the same as (62b), which is intended to capture the truth conditions of (57a) above, on the reading where the pronoun simply receives a deictic interpretation.

(62) a. $\textsc{Most}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e, x)]$
$(\exists e'[\text{transported-radioactive-waste}'(e', x)])$

iterative events that satisfy the restrictive clause in (59a) to be included in the non-iterative events denoted by the individual-level predicate in the NS.

[23]This suggestion fits in nicely with Partee's (1989) claim that implicit arguments may participate in the same donkey-type anaphoric dependencies as ordinary pronouns. Cf. also Chierchia (1992) for a dynamic treatment of the anaphoric dependency that holds between the event-arguments in the restrictive and matrix clauses in sentences such as *If John is in the bathtub, he usually sings.*

b. $\text{MOST}x : \text{farmer}'(x) \wedge \exists y[\text{donkey}'(y) \wedge \text{owns}'(x,y)](\text{beats}'(x,z))$

Now, the OR of (58a) on its dependent events reading is more interesting. It seems natural to analyze dependent events in terms of an anaphoric dependency between an antecedent event argument, which in this case is contained in a relative clause adjoined to the subject QP, and an anaphoric event argument, which is located in the matrix predicate, not c-commanded by its antecedent. However, making this assumption is tantamount to treating the dependent events reading of (58a) as a donkey-type anaphoric dependency. To see this, it suffices to note that (63a) below is structurally equivalent to (57b). Given this structural equivalence then, we can extend the scope of the existential quantifier over events into the NS, so as to capture the dependent events reading, in much the same way as we extended the scope of the indefinite *a donkey* into the matrix clause, viz. by appealing to dynamic Conservativity and the principle in (46). (63b-d) trace the relevant derivational steps for the OR of (58a) on its dependent events reading. They are the same as those plotted in (57c-e) above.[24]

(63) a. $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\text{transported-radioactive-waste}'(e,x))$

b. $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$\wedge\ \text{transported-radioactive-waste}'(e,x))$ CONS

c. $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)$
$\wedge\ \text{transported-radioactive-waste}'(e,x)])$ (46)

d. $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\exists e[\text{passed-through-the-lock}'(e,x)$
$\wedge\ \text{transported-radioactive-waste}'(e,x)])$ CONS

Finally, we need to address the issue why the ER of the sentences in (58) is only compatible with a dependent events reading. Observe first that the relative

[24] In representations such as (63), to enhance readability we have omitted making reference to the fact that in dependent events the anaphoric dependency between event arguments is mediated through the reference time of the matrix predicate. To accomodate the mediating role of reference times, we would have to assume that their semantics can be modeled by a covert predicate that expresses the intended relation between the reference and event time, as illustrated in (i) for the OR of (58a) on its dependent events reading, analogous to the treatment of overt time adverbials such as ***three hours later***, which express a more complex relation between reference and event times.

(i) $\text{MOST}x : \text{ship}'(x) \wedge \exists e[\text{passed-through-the-lock}'(e,x)]$
$(\exists e \exists e'[\text{passed-through-the-lock}'(e,x)$
$\wedge\ \text{transported-radioactive-waste}'(e',x) \wedge \text{at}'(e,e')])$

clauses in these examples syntactically provide for the necessary eventive predicate in the restrictive term of the strong determiner. As we have already observed in various places, the requirement on pair-quantification is now satisfied, in that both the object-variable and the event-variable are properly introduced in the restriction of the strong determiner. This means, then, that if we subsequently abstract over the eventive variable by Existential Disclosure, the strong determiner may happily pair-quantify over it. But for a pair-quantification structure to be well-formed, the arity of the restrictive clause should match the arity of the NS, just as much as this is true for any quantification structure. In this case, the first and the second argument of the pair-quantifier should be two-placed predicates. Note first that this requirement constrains the availability of so-called *symmetric* and *asymmetric* readings for donkey-sentences. That is, although the sentence in (64a) below has a pair-quantificational, or symmetric reading, represented in (64b), which allows for a situation in which a single farmer beats all of his donkeys, the total sum of which outnumbers the donkeys that are owned by all the other farmers, the sentences in (65), which do not contain an anaphor in the consequent that is coreferent with *a donkey,* obviously do not. The latter only allow for the so-called (subject) asymmetric reading, in which the QP adverb *usually* only quantifies over the indefinite subject *a farmer* in the antecedent clause (cf. Bäuerle and Egli 1985).[25]

(64) a. Usually, if a farmer owns a donkey, he beats it

b. $\text{MOST}\langle x, y\rangle : \text{farmer}'(x) \wedge \text{donkey}'(y) \wedge \text{owns}'(x,y)(\text{beats}'(x,y))$

(65) a. Usually, if a farmer owns a donkey, he is rich

b. Usually, if a farmer owns a donkey, he beats his wife

In dynamic terms, the requirement on matching arities means that as soon as we abstract over an indefinite in the restrictive term by Existential Disclosure, we need to see to it that the corresponding pronoun in the NS is made anaphorically dependent on it, so that this position will be abstracted over as well, in virtue of dynamic Conservativity and the principle in (46). The derivation in (66) traces these steps as they apply to example (64a).

[25]Thanks to Anna Szabolcsi for pointing out the relevance of these examples. Note that we cannot have quantification over farmer-donkey pairs in (i) either, as predicted given the absence of an anaphor corresponding to *a donkey* in the NS:

(i) If a farmer sees a donkey, he screams.

At first sight there seems to be quantification over farmer-donkey pairs in (i), contrary to our expectations, because there is a screaming event corresponding to every farmer-donkey pair. However, we are not quantifying over farmer-donkey pairs, but over events in which a farmer sees a donkey. If a farmer sees a donkey several times, he screams at each of these occasions, and not at one of them. This shows that the number of screaming events corresponds to the number of events in which a farmer sees a donkey, and not to the number of farmer-donkey pairs.

(66) a. $\textsc{Most}\langle x,y\rangle$: farmer$'(x)$ $\wedge$ $\exists y'[$donkey$'(y')$ $\wedge$ owns$'(x,y')]$ $\wedge$ $y' = y$
(beats$'(x,y')$) ED

b. $\textsc{Most}\langle x,y\rangle$: farmer$'(x)$ $\wedge$ $\exists y'[$donkey$'(y')$ $\wedge$ owns$'(x,y')]$ $\wedge$ $y' = y$
(farmer$'(x)$ $\wedge$ $\exists y'[$donkey$'(y')$ $\wedge$ owns$'(x,y')]$ $\wedge$ $y' = y$ $\wedge$ beats$'(x,y')$) CONS

c. $\textsc{Most}\langle x,y\rangle$: farmer$'(x)$ $\wedge$ $\exists y'[$donkey$'(y')$ $\wedge$ owns$'(x,y')$ $\wedge$ $y' = y]$
(farmer$'(x)$ $\wedge$ $\exists y'[$donkey$'(y')$ $\wedge$ owns$'(x,y')$ $\wedge$ $y' = y$ $\wedge$ beats$'(x,y')]$) (46)

d. $\textsc{Most}\langle x,y\rangle$: farmer$'(x)$ $\wedge$ donkey$'(y)$ $\wedge$ owns$'(x,y)$(beats$'(x,y)$)
(cf. 65b; CONS & 47)

Conversely, if there is no pronoun in the matrix clause anaphorically dependent on the abstracted-over indefinite in the restrictive clause, as is the case in the sentences in (65), the requirement on matching arities will inevitably be violated. This takes care of our earlier observation that these sentences only allow for (subject) asymmetric readings.

By the same token, if we were to abstract over the event argument in the relative clauses of (58) by Existential Disclosure, so as to obtain ER, we again need to ensure that the event argument of the matrix predicate is anaphorically dependent on it, so that this position will be abstracted over as well. Otherwise, the constraint on matching arities is violated. But note now that the obligatory anaphoric dependency of the two event arguments will force a dependent events reading, thus capturing our earlier observation with respect to (58) that ER is only compatible with this type of sequencing of events. (67) shows the relevant derivational steps for the ER of (58a); they are equivalent to the inferential steps displayed in (66).

(67) a. $\textsc{Most}\langle e,x\rangle$: ship$'(x)$ $\wedge$ $\exists e'[$passed-thru-the-lock$'(e',x)]$ $\wedge$ $e' = e$
(transported-radioactive-waste$'(e',x)$) ED

b. $\textsc{Most}\langle e,x\rangle$: ship$'(x)$ $\wedge$ $\exists e'[$passed-thru-the-lock$'(e',x)]$ $\wedge$ $e' = e$
(ship$'(x)$ $\wedge$ $\exists e'[$passed-thru-the-lock$'(e',x)]$ $\wedge$ $e' = e$ $\wedge$ transported-radioactive-waste$'(e',x)$) CONS

c. $\textsc{Most}\langle e,x\rangle$: ship$'(x)$ $\wedge$ $\exists e'[$passed-thru-the-lock$'(e',x)$ $\wedge$ $e' = e]$
(ship$'(x)$ $\wedge$ $\exists e'[$passed-thru-the-lock$'(e',x)$ $\wedge$ $e' = e$ $\wedge$ transported-radioactive-waste$'(e',x)]$) (46)

d. $\textsc{Most}\langle e,x\rangle$: ship$'(x)$ $\wedge$ passed-thru-the-lock$'(e,x)$
(transported-radioactive-waste$'(e,x)$) CONS & (47)

Conversely, if there is no event argument in the matrix predicate that can be made anaphorically dependent on the abstracted-over event argument in the restrictive clause, as is the case in (61) above due to the time adverbial *now,* we again run into a conflict with the requirement on matching arities. This accounts for the fact that sentences such as (61) do not admit of ER.

Concluding our discussion of the cases involving relative clauses, we have seen that symmetric and asymmetric readings for donkey-sentences correlate with the presence versus absence of anaphoric dependencies. Furthermore, we observed that ER and OR correlate with dependent versus independent events. Interestingly, the two correlations show very similar properties. That is, the observation that only the symmetric pair-quantificational reading of (64a) requires the pronoun *it* to be anaphorically dependent on *a donkey* is replicated by the observation that only the ER of sentence (58a) forces the event argument of the matrix predicate *transported radioactive waste* to be anaphorically dependent on the event argument of the antecedent predicate *passed through the lock,* resulting in a dependent events reading. This is a compelling argument in favour of the pair-quantification approach to ER, in that a small set of independently motivated dynamic assumptions yields the above correspondence between ER and symmetric readings for donkey-sentences simply as a matter of principle. It is hard to see for instance how any other theory of ER could treat the incompatibility between ER and independent events readings, as evidenced in (61), and the absence of a symmetric pair-quantificational construal for the donkey-sentences in (65) completely on a par.

4.6 Conclusions

In this section, we first observed that only symmetric (weak) determiners can freely give rise to both OR and ER. This contrast receives a natural interpretation on the pair-quantificational approach to ER. Since the ER pair-quantifier requires that the event-variable as well as the object-variable be introduced in its restriction, we would expect there to be a difference between strong and weak determiners in the first place, as only symmetric (weak) determiners allow an eventive predicate in their NS to be mapped into their restrictive term by semantic inference. Subsequently abstracting over the event-variable by Existential Disclosure enables the symmetric determiner to "unselectively" bind it. We then argued that both focus and relative clause formation may effect the required bracketing in which an eventive predicate, together with the head noun, determines the restriction of a strong determiner. The fact that in these constructions strong determiners can give rise to ER will thus not be surprising.[26] In the process, we have exploited some elementary tools of

[26] Helen de Hoop (p.c.) observes that context may facilitate ER for strong determiners even in the absence of focus or an overt relative clause. For instance, imagining a situation in

Dynamic Predicate Logic, enriched with Generalized Quantifiers. These tools provide us with the means to treat a number of striking similarities between ER and donkey-type anaphora on a par. These similarities in turn furnish further support for the pair-quantificational analysis of ER.

5 WEAK ISLAND EFFECTS ON ER

In this section, we discuss weak island effects on ER. We will first observe that the class of those QPs that block *how*-extraction when occupying the subject position exactly coincides with the class of those QPs that block ER for simple transitive clauses in which they occupy the object position. Furthermore, we observe that sentence negation, which is known to induce weak island effects, also blocks ER that targets the subject QP in simple transitive clauses. Assuming Szabolcsi and Zwarts's (1993) algebraic approach to this phenomenon, we argue that these effects are naturally accounted for by the pair-quantificational analysis of ER, as outlined in Section 3. We argued there that the domain of ⟨event, object⟩ pairs is structured in a join semi-lattice (cf. 27). Thus we predict that the ER pair-quantifier is unable to have scope over those scopal expressions (i.e. QPs, sentence negation, QP-adverbs, etc.) that are semantically associated with the Boolean operations *meet* or *complementation,* operations that are simply not defined on a join semi-lattice. The Weak Island effects on ER then follow in virtue of two auxiliary observations: (i) those scopal expressions that block ER need to take narrow scope with respect to the subject; and (ii) the same scopal expressions require either meet or complementation to be performed.

---

which a police officer reports his findings on the number of cars that crossed some intersection and their speed to his superiors, a sentence such as *Most cars exceeded the speed limit* can naturally have an ER, which might be paraphrased as "Most events in which a car crossed the intersection were events in which a car exceeded the speed limit." Consistent with the present analysis, we could analyze this effect in terms of Westerståhl's (1985a) *context set,* a property X that intersects with the "live on" set A of all determiners Q_E (i.e. $Q_E(A \cap X)$) and whose exact content is fixed by the context of utterance. Context sets not only enable us to account for the fact that, in normal discourse, the domain of quantification is restricted to some contextually salient subset of the "live on" set, they also provide us with the means to deal with de Hoop's observation in essentially the same way as we accounted for the focus and relative clause facts in the main text. To wit, if we fix X to be the property $\lambda x \exists e[\text{crossed-the-intersection}'(e, x)]$, we observe that both the object and event variable are now properly introduced in the restriction of the strong determiner MOST. This means that MOST may happily pair-quantify over both variables if we choose to abstract over the event variable by Existential Disclosure.

5.1 Demonstration of the Weak Island effect

Object QPs as well as sentence negation can influence the availability of ER. The sentences in (68) contrast markedly with those in (69)–(71) in that only the former display both OR and ER. The ER for the sentences in (69)–(71) is missing.

(68) a. 4,000 people visited the Rijksmuseum last year (OR/ER)
OR: 4,000 people are such that each of them visited the Rijksmuseum last year
ER: there were 4,000 events in which a person visited the Rijksmuseum last year
b. 4,000 people visited (the) three museums last year (OR/ER)
c. 4,000 people visited every museum last year (OR/ER)

(69) 4,000 people didn't visit the Rijksmuseum last year (OR/*ER)
OR: 4,000 people are such that each of them did not visit the Rijksmuseum last year
ER: * there were 4,000 events in which a person did not visit the Rijksmuseum last year

(70) a. 4,000 people visited no museum last year (OR/*ER)
OR: 4,000 people are such that each of them visited no museum last year
ER: * there were 4,000 events in which a person visited no museum last year
b. 4,000 people visited few museums last year (OR/*ER)

(71) a. 4,000 people visited at most three museums last year (OR/*ER)
OR: 4,000 people are such that each of them visited at most three museums last year
ER: * there were 4,000 events in which a person visited at most three museums last year
b. 4,000 people visited exactly three museums last year (OR/*ER)

Taking these facts in turn, the contrast between (68) on the one hand and (69) and (70) on the other furnishes initial support for the idea that ER is sensitive to weak islands. We note that the above distinction is similar to a Negative Island effect, a standard subcase of the more familiar weak island (WI) effect, which also subsumes *wh*-islands. Negative islands are known to block extraction of *how,* as illustrated by the contrast of (72) against (73) and (74).

(72) a. How did the student behave?
b. How did (the) three students behave?
c. How did every student behave?

(73) * How didn't the student behave?

(74) a. * How did no students behave?
b. * How did few students behave?

In recent years a number of authors, including for instance Comorovski (1989), Kroch (1989) and Cinque (1990), have observed that extracting *wh*-phrases over WIs leads to well-formed results just in case the extracted *wh*-phrase receives a Discourse-linked (or presuppositional) interpretation in the sense of Pesetsky (1987). If true, we could use this descriptive generalization as an interesting heuristic. That is, given our assumption that in ER the pertinent quantifier ranges over ⟨event, object⟩ pairs, we may now expect ER to be sensitive to WIs as well, on the intuitively plausible assumption that ordered pairs of events and objects are not the sorts of objects that can be made contextually salient by discourse. For one thing, ⟨event, object⟩ pairs do not persist through time and space. Temporarily ignoring the issue as to whether this is the right characterization of those QPs that are insensitive to WIs, we suggest that the similarity in contrasts in (68) versus (69) and (70) and (72) versus (73) and (74) is consistent with such a description. The pattern carries over to further quantificational expressions that have been discussed in the literature on WIs:

(75) a. ?? How did at most three students behave?
b. ?? How did exactly three students behave?

The contrast between (72) and (75) corresponds with the contrast between (68) and (71). This shows that the "opacity" effects on ER in (69)–(71) can be properly described as WI effects.

5.2 Weak islands and scope

In their 1993 paper, Szabolcsi and Zwarts argue that the specificity effects that have been observed with respect to WIs may be reduced to semantic, algebraic constraints on scope-taking, thereby further elaborating on suggestions to this effect by Dobrovie-Sorin (1992, 1993), Kiss (1992) and de Swart (1992). Leaving aside how this might be of use in explaining the intervention effects on *how*-extraction, we will show in the remainder of this subsection that their insight that it is scope which is at the very core of the WI phenomenon proves to be crucial in coming to terms with the WI effects on ER as well.

As a first indication that we should be looking for a scopal theory of WIs if we want to come to grips with the ER facts, we observe that as soon as we move the harmful object QPs in (70) and (71) into a position where they can take surface scope over the relevant subject QPs, the resulting structures readily allow for ER, as the sentences in (76) and (77), respectively, testify.[27]

(76) a. No museum was visited by 4,000 people last year (OR/ER)

OR: no museum is such that 4,000 people visited it last year

ER: no museum is such that there were 4,000 events in which a person visited it last year

b. Few museums were visited by 4,000 people last year (OR/ER)

(77) a. At most three museums were visited by 4,000 people last year (OR/ER)

OR: at most three museums are such that 4,000 people visited them last year

ER: at most three museums are such that there were 4,000 events in which a person visited them last year

b. Exactly three museums were visited by 4,000 people last year (OR/ER)

Likewise, if we apply passivization so that the negative particle *n't* can take surface scope over the demoted subject QP *4,000 people,* we again observe that the resulting structure admits of an ER, as evidenced in (78).

(78) The Rijksmuseum wasn't visited by 4,000 people last year (OR/ER)

OR: it is not the case that there are 4,000 people such that they visited the Rijksmuseum last year

ER: it not the case that there were 4,000 events in which a person visited the Rijksmuseum last year

Clearly, this argument would be without any force if it were the case that the object QPs in (70) (*no museum* etc.) and (71) (*at most three museums* etc.), as well as sentence negation in (69), could freely take (inverse) scope over a c-commanding subject. That is, if these object QPs were indistinguishable from the object QPs in (68) in terms of their ability to take inverse scope, it is hard to see why scope would make a difference. However, the two classes of object QPs can in fact be distinguished from each other in terms of their ability to support inverse scope readings, as has been firmly established by a

[27] We owe the observations in (76)–(78) to Jacqueline Guéron.

number of researchers (cf. Liu 1990, Beghelli 1993, Ben-Shalom 1993, Beghelli 1995), thus reinforcing the idea that it is indeed scope which is at the heart of the WI effects on ER.

Let us first introduce some terminology. The object QPs in (68) and (70)-(71) fall into two distinct natural classes: QPs such as *three museums/every museum* (may) denote principal filters, whereas QPs such as *at most three museums/exactly three museums* do not.[28] (79) provides the paradigm case, adapted from Ben-Shalom (1993), illustrating that for non-principal filter QPs such as *no museum* it is almost impossible to take inverse scope.

(79) **Only principal filter QPs can take inverse scope over another c-commanding QP** (cf. Ben-Shalom 1993)

a. Two referees read every abstract (S > O, O > S)

S > O: there is a set of two referees each of whom read every abstract

O > S: there is a set containing all abstracts each of which was read by two (possibly different) referees

b. Two referees read three abstracts (S > O, ?O > S)

c. Two referees read at least three abstracts (S > O, ??O > S)

d. Two referees read exactly three abstracts (S > O, *O > S)

e. Two referees read more than three abstracts (S > O, *O > S)

f. Two referees read fewer than three abstracts (S > O, *O > S)

g. Two referees read no abstracts (S > O, *O > S)

To these observations we may add that sentence negation cannot scope over a c-comanding expression either.

What is it that makes ER incompatible with the obligatory narrow scope construal for either a non-principal filter QP or sentence negation? Intuitively, the problem here is that under the narrow scope construal for any of these scopal expressions, we need to perform operations that are implausible at best. Consider for instance (69), *4,000 people didn't visit the Rijksmuseum last year,* on the obligatory narrow scope construal for the sentence negation marker. The sentence requires us to construct ⟨event, person⟩ pairs that participated in negative events of visiting the Rijksmuseum. Or, consider (71b), *4,000 people*

[28] Cf. Chapter 1 for some discussion of ***principal filters***. Note that the availability of ER for the sentence with *three museums* hinges on the specific interpretation of the indefinite object QP, whereby it comes to mean something like *the three museums*. We will adopt Ben-Shalom's (1993) proposal according to which the specific interpretation of bare numeral QPs can be modeled by principal filters. The same remarks also apply to the conditions under which the bare numeral object QP in (79b) can take inverse scope.

visited exactly three museums last year. In order to find out whether the same ⟨event, person⟩ pair participated in exactly three events of visiting some museum, we would need to be able to hold ⟨event, object⟩ pairs constant across different visiting events. It would seem that both of these requirements are impossible to meet. We cannot fulfill the first requirement because by its very nature a negative event resists recycling of objects (negative events in which someone does something being hard to make sense of to begin with), and we cannot fulfill the second requirement because an ⟨event, object⟩ pair is a spatiotemporally bounded entity that by definition cannot be held constant across events.[29] On the other hand, the OR for the same sentences differs from the ER in that, on the basis of the same intuition, it is perfectly possible to construct the set of persons who were not involved in an event of visiting a museum, or to keep a person constant across several events of visiting a museum.

5.3 A formal account of the island sensitivity of ER

It is precisely this connection between scope and the (im)possibility of performing certain operations that Szabolcsi and Zwarts explicate in lattice-algebraic terms. On the basis of that formal restatement, they are able to derive WI effects in general in a natural, uniform manner. The core of their analysis can be stated as in (80).

(80) Scope and operations (cf. Szabolcsi and Zwarts 1993, p. 6)

> Each scopal element SE is associated with certain Boolean operations. For a *wh*-phrase [or any QP, for that matter; D&H] to take scope over some SE means that the operations associated with SE need to be performed in the *wh*-phrase's denotation domain. If the *wh*-phrase denotes in a domain for which the requisite operation is not defined, it cannot scope over SE.

Briefly recapitulating, when we say that each SE is associated with certain Boolean operations, we simply mean to say that each and every SE in conjunction with a verbal predicate can be interpreted as a Boolean combination of singular predications. The examples in (81) provide a quick illustration of what this entails for QPs when in construction with an intransitive verb such as

[29] The impossible ER of (71b) with *exactly three museums* in object position should be distinguished from the reading on which there were 4,000 package tours that all involved visiting exactly three museums; this seems perfectly possible. Here we are not holding ⟨event, person⟩ pairs constant across different events of visiting some individual museum. Rather, the event member of each pair is composed of smaller events of visiting an individual museum. The whole force of the "package tour" scenario is that it allows us to ignore the atomic events of visiting some museum. In other words, on the "package tour"-interpretation, each visiting event affects a collection of exactly three museums as though it was a single, irreducible unit.

walks, assuming a model in which John, Bill and Mary are the only students.[30] We will find it useful to have some visual illustration of what particular Boolean operations are associated with what SE.

(81)	a.	*John walks*	=	$W(j)$
	b.	*every student walks*	=	$W(j) \wedge W(b) \wedge W(m)$
	c.	*a student walks*	=	$W(j) \vee W(b) \vee W(m)$
	d.	*no student walks*	=	$\neg[W(j) \vee W(b) \vee W(m)]$
	e.	*at most one student walks*	=	$\neg[(W(j) \wedge W(b)) \vee (W(b) \wedge W(m)) \vee (W(j) \wedge W(m))]$

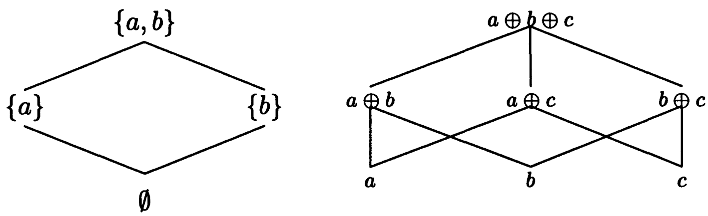
To return to our earlier, informally stated, connection between scope and the (im)possibility of performing certain operations, we may now take the subject QP *4,000 people* to range over ordinary (atomic) individuals on the OR, as seems natural. In the spirit of (80), individuals may be collected into unordered sets, and one may perform all Boolean operations on sets, since the set of all sets $\mathcal{P}(E)$ over some unordered domain E forms a Boolean algebra, as is illustrated in (82c) below, assuming a two-membered domain E. Note now that this line of reasoning captures in algebraic terms our earlier intuition that it is perfectly possible either to construct the set of persons who were not involved in some event, or to keep a person constant across several events. To construct the set of persons who were not engaged in some particular event simply means that we form the complement of the set of persons who were engaged in that event, and to keep a person constant across several events merely boils down to intersecting the sets of persons who were involved in those events. Ultimately, since both complementation and intersection (meet) are defined on a Boolean algebra, (80) automatically derives our earlier observation that OR is insensitive to WI effects.

(82) a. A *Boolean algebra* is a partially ordered set A which is closed under all Boolean operations (i.e. *meet, join,* and *complementation*).

b. A *(proper) join semi-lattice* is partially ordered set A which is closed only under *join.*

[30] We will assume, as is standard, that negation is directly associated with Boolean complementation. Cf. Chapter 1, Section 1.1.2., for more discussion on the connection between scopal expressions and Boolean operations.

c. **Boolean algebra** **join semi-lattice**



Recall now that we informally attributed the WI sensitivity of ER to the problem that, on the obligatory narrow scope construal of sentence negation and the non-principal filter object QPs in (69) as well as (70) and (71), respectively, we were forced either to construct ⟨event, person⟩ pairs that participated in negative events of visiting some museum, or to keep an ⟨event, person⟩ pair constant across several visiting events. Intuitively, it is impossible to perform any of these tasks.

This intuition can be captured in formal, algebraic terms as well. To see this, note first that the principle in (80) entails that the obligatory narrow scope construal for those object QPs that do not denote principal filters and sentence negation requires their associated Boolean operations to be performed in the denotation domain of the subject QP *4,000 people.* Since all of these scopal expressions have at least Boolean complementation associated with them, as partly illustrated in (81d, e), and since Boolean complementation is not defined on a join semi-lattice, we would immediate derive the unavailability of ER if we could show that the numeral *4,000* ranges over entities that are structured in a join semi-lattice.

Now, on the pair-quantificational approach to ER that we have developed and defended in Sections 3 and 4 above, *4,000* does in fact range over entities that are structured in a join semi-lattice. We argued in Section 3 that the domain of ⟨event, object⟩ pairs constitutes a (proper) join semi-lattice (cf. 27 above). This essentially followed from the fact that the domain of events has no bottom element, from which it could be concluded that the domain of ⟨event, object⟩ pairs does not have a bottom element either (cf. Fact 26). Thus, this fact not only accounts for the WI effects on ER observed above, it also preserves our earlier intuitive explanation of what goes wrong in these cases. In accordance with the principle in (80) and our result in (27), the impossibility of constructing ⟨event, person⟩ pairs participating in negative events of museum visiting is formally explicated in terms of the impossibility of forming the complement of the relevant ⟨event, person⟩ pairs, whereas the impossibility of keeping an ⟨event, person⟩ pair constant across several events of museum visiting follows from the impossibility of performing meet on the relevant ⟨event, person⟩ pairs.

Note that we get this result essentially for free, as we already argued in Section 3.2. for completely independent reasons that the domain of ⟨event, object⟩ pairs is structured in a join semi-lattice, thus lending further support to our pair-quantificational approach to ER.

What about the "good guys," then? As the paradigm in (79) makes clear, principal filter QPs, such as the ones in (68), can get out of the denotation domain of the subject QP by taking inverse scope over it. This means that their associated Boolean operations need not be performed in the domain over which the determiner of the subject QP ranges. This possibility for saving ER is illustrated in (83). Thus, (83) shows that the Boolean meets associated with the universal object QP *every museum* are not performed in the domain of ⟨event, object⟩ pairs, but instead serve to conjoin propositions, assuming some model where m_1 and m_2 are the only museums.

(83) a. 4,000 people visited every museum (last year) (O > S)

ER: every museum is such that there were 4,000 events in which a person visited it (last year)

b. $\text{E}\textsc{very}y : \text{museum}'(y)(4,000\langle e,x\rangle : \text{people}'(x) \ \wedge \ \text{visited}'(e,x,y))$

c. $(|\{\langle e,x\rangle \mid \langle e,x,m_1\rangle \in \text{visited}' \ \wedge \ x \in \text{people}'\}| = 4000) \ \wedge$
$(|\{\langle e,x\rangle \mid \langle e,x,m_2\rangle \in \text{visited}' \ \wedge \ x \in \text{people}'\}| = 4000)$

Even though we argued above that the Boolean operations associated with the principal filter QPs do not have to be performed in the denotation domain of the subject, this is of course no longer the case on their narrow scope, distributive construal. But then, Szabolcsi and Zwarts's scope principle in (80), together with the fact that the domain of ⟨event, object⟩ pairs constitutes a (proper) join semi-lattice (cf. 29), will predict that ER for a sentence such as (68c) is incompatible with a narrow scope, distributive construal of *every museum,* as principal filters generally are defined in terms of Boolean meet (cf. for instance 82b), and this operation is not defined on a join semi-lattice.

Again, the prediction neatly matches the facts. For example, the ER of (83a) simply does not admit a narrow scope, distributive construal for the universal object QP. Although the sentence does admit of a "package tour"-interpretation (cf. fn. 29), it does not allow for an ER in which each ⟨event, person⟩ pair participates in an event of visiting some individual museum. This is so because meet is not defined on the join semi-lattice in which ⟨event, object⟩ pairs are structured.[31]

[31] Note that this particular "opacity" effect on ERs has its analogue in the *how*-extraction cases as well, as observed originally by Williams and taken up by de Swart. If we force a universal subject QP to have distributive, narrow scope with respect to *wh*-extracted how, the resulting structure is ill-formed, as attested by (i).

Another possibility for principal filter QPs to get out of the denotation domain of the subject QP would be to directly denote a singular or plural individual. Typically, on the latter, collective construal, a QP such as *the (three) museums* does not support scopal interactions with other QPs, and can therefore be said to be scopeless. Descriptively, the singular or plural individual reading is only available to the expressions in (84).

(84) **Scopeless expressions**

i. singular definite descriptions, specifically interpreted singular indefinites, and proper names;

ii. plural definite descriptions and bare numeral QPs that are interpreted collectively.

Thus, on their singular or plural individual reading, these expressions essentially behave like names. Since names do not have have any Boolean operation associated with them, we predict the expressions in (84) to be completely transparent with respect to ER. This prediction squares with what we observe. For instance, on its collective construal the object QP *the three museums* allows a (cumulative) ER that might be paraphrased as in (85).

(85) 4,000 people visited the three museums (last year)

ER: there were (altogether) 4,000 events in which a person visited any one of the three museums (last year)

Whichever of the two options the object QPs choose for avoiding discord (taking inverse scope over the subject or supporting a singular or plural individual reading), it is clear that we do not have to perform Boolean operations that may be undefined on the denotation domain of the determiner of the subject QP. Thus, this analysis covers the pattern of OR and ER evidenced above.

In this section we have observed that ER is sensitive to WI effects. That is, the very same QPs that block *how*-extraction when occupying the subject position also block ER when occupying the object position. Furthermore, we have noticed that sentence negation, which is known to induce WI effects, also blocks ER that targets the subject QP in simple transitive clauses. Under Szabolcsi and Zwarts's (1993) algebraic-semantic approach to WIs, the observation that ER is sensitive to WI effects provides a very strong argument in favour of the pair-quantificational approach to ER, as we developed it in Section 3. Given that we argued there that the domain of ⟨event, object⟩ pairs forms a join semi-lattice, i.e. a partially ordered set which is neither closed

(i) * How did the students each behave?

under meet nor complementation, the WI effects on ER follow automatically from Szabolcsi and Zwarts's account in virtue of two auxiliary observations: (i) those scopal expressions that impair ER need to take narrow scope with respect to the subject; and (ii) the very same scopal expressions require either meet or complementation to be performed.

REFERENCES

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2):159–219.

Bäuerle, Rainer, and Urs Egli. 1985. Anapher, Nominalphrase und Eselsätze. *Arbeitspapiere des Sonderforschungsbereichs* 99.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In Amy J. Schafer (Ed.), *Proceedings of the North East Linguistic Society 23, Volume One*, 65–80, Amherst. GLSA Publications, University of Massachussetts.

Beghelli, Filippo. 1995. Asymmetries with decreasing QPs. In Raul Aranovich et al. (Eds.), *The Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, 467–482, Stanford. CSLI.

Ben-Shalom, Dorit. 1993. Object wide scope and semantic trees. In Lahiri (Ed.), *Proceedings of SALT III.*

Carlson, Gregory. 1977. *Reference to Kinds in English.* PhD thesis, University of Massachussetts, Amherst.

Carlson, Gregory. 1982. Generic terms and generic sentences. *Journal of Philosophical Logic* 11(2):145–181.

Carlson, Gregory. 1984. Thematic roles and their role in semantic interpretation. *Linguistics* 22(3):259–279.

Chierchia, Gennaro. 1992. Anaphora and dynamic binding. *Linguistics and Philosophy* 15(2):111–183.

Cinque, Guglielmo. 1990. *Types of A'-Dependencies.* Cambridge: MIT Press.

Comorovski, Ileana. 1989. *Discourse and the Syntax of Multiple Constituent Questions.* PhD thesis, Cornell University.

Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (Ed.), *The Logic of Decision and Action*, 81–95. University of Pittsburgh Press.

Dekker, Paul. 1990. Existential disclosure. Presented at the Third Symposium on Logic and Language, Révfülöp.

Dekker, Paul. 1993. *Transsentential Meditations.* PhD thesis, University of Amsterdam.

Diesing, Molly. 1992. *Indefinites.* Cambridge: MIT Press.

Dobrovie-Sorin, Carmen. 1992. LF representations, weak islands, and the ECP. *GLOW Newsletter* 27:14–15.

Dobrovie-Sorin, Carmen. 1993. *The Syntax of Romanian: Comparative Studies in Romance.* New York: Mouton de Gruyter.

Doetjes, Jenny. 1994. Quantification at a distance and event relatedness. In Crit Cremers (Ed.), *Linguistics in the Netherlands 1994*, 13–24. Amsterdam: John Benjamins.

Groenendijk, Jeroen, and Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14(1):39–100.

Herburger, Elena. 1994. Focus and NP quantification. In Erin Duncan et al. (Eds.), *The Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, 517–533, Stanford. CSLI.

Honcoop, Martin. 1992. ER or OR: Quantification, intervention and fission. Master's thesis, University of Leiden.

de Hoop, Helen, and Henriette de Swart. 1989. Over indefiniete objecten en de relatie tussen syntaxis en semantiek. *GLOT* 12:19–35.

Kanazawa, Makoto. 1994a. Dynamic generalized quantifiers and monotonicity. In Makoto Kanazawa and Christopher Piñon (Eds.), *Dynamics, Polarity, and Quantification*, 213–249. Stanford: CSLI.

Kanazawa, Makoto. 1994b. Weak versus strong readings of donkey sentences and monotonicity inference in a dynamic setting. *Linguistics and Philosophy* 17(2):109–158.

É. Kiss, Katalin. 1992. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11(1):85–120.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. In Gregory Carlson and Jeff Pelletier (Eds.), *The Generic Book*, 125–175. The University of Chicago Press.

Krifka, Manfred. 1989. Nominal reference, temporal constitution, and quantification in event semantics. In Renate Bartsch et al. (Eds.), *Semantics and Contextual Expression*, 75–115. Dordrecht: Foris.

Krifka, Manfred. 1990. Four thousand ships passed through the lock. *Linguistics and Philosophy* 13(5):487–520.

Kroch, Anthony. 1989. Amount quantification, referentiality and long WH-movement. Unpublished manuscript.

Link, Godehard. 1987. Generalized quantifiers and plurals. In Peter Gärdenfors (Ed.), *Generalized Quantifiers*, 151–180. Dordrecht: Reidel.

Liu, Feng-hsi. 1990. *Scope Dependency in English and Chinese.* PhD thesis, UCLA.

Moore, Steven. 1994. Events, objects, and eventual objects. In Erin Duncan et al. (Eds.), *The Proceedings of the Twelfth West Coast Conference on Formal Linguistics*, 577–590, Stanford. CSLI.

Partee, Barbara. 1989. Binding implicit variables in quantified contexts. In Caroline Wiltshire et al. (Eds.), *Papers from the 25th Annual Regional Meeting of the Chicago Linguistics Society, Part One: The General Session*, 342–365. University of Chicago.

Partee, Barbara. 1991. Topic, focus, and quantification. In Steven Moore and Adam Zachary Wyner (Eds.), *Proceedings of SALT I*, 159–187. Cornell University, Working Papers in Linguistics #10.

Pelletier, Jeff, and Lennart Schubert. 1989. Generically speaking. In Gennaro Chierchia et al. (Eds.), *Properties, Types and Meaning, Volume 2*, 193–268. Dordrecht: Kluwer.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In Eric Reuland and Alice ter Meulen (Eds.), *The Representation of (In)definiteness*, 98–130. Cambridge: MIT Press.

Rooth, Mats. 1985. *Association with Focus.* PhD thesis, University of Massachussetts, Amherst.

Schein, Barry. 1993. *Plurals and Events.* Cambridge: MIT Press.

de Swart, Henriëtte. 1992. Intervention effects, monotonicity and scope. In Chris Barker and David Dowty (Eds.), *SALT II: Proceedings of the Second Conference on Semantics and Linguistic Theory (OSU WPL 40)*, 387–406, Columbus. The Ohio State University.

Szabolcsi, Anna, and Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1(3):235–284. Reprinted in this volume.

Westerståhl, Dag. 1985a. Determiners and context sets. In Johan van Benthem and Alice ter Meulen (Eds.), *Generalized Quantifiers in Natural Language*, 45–71. Dordrecht: Foris.

Westerståhl, Dag. 1985b. Logical constants in quantifier languages. *Linguistics and Philosophy* 8(4):387–413.

9

QUANTIFIERS IN PAIR-LIST READINGS*

Anna Szabolcsi

Department of Linguistics
UCLA

In this paper the term pair-list reading will be applied to both types (1) and (2):

(1) Who did every dog bite?
'For every dog, who did it bite?'

(2) Who did six dogs bite?
'For six dogs of your choice, who did each bite?'

Type (1) will be referred to as a fixed domain reading and type (2) as a choice reading, when the distinction is necessary.

Pair-list readings arise when the interrogative contains a quantifier; the issue to be addressed is what role this quantifier plays. The standard view is that the quantifier here does not have the same kind of quantificational force as in other, "normal" contexts; instead, it contributes a restriction on the domain of the question. Furthermore, it is assumed that interrogatives on the pair-list reading are lifted, i.e., denote generalized quantifiers over individual questions. Abstracting away from certain differences between authors (Groenendijk and Stokhof 1984, Higginbotham 1991, Chierchia 1993), matrix as well as complement pair-list readings are assigned the following kind of interpretation:

(3) $\lambda P \exists X[X$ a set determined by the quantifier & P(which $x \in X$ bit whom)]

where P is a variable ranging over properties like being a secret, being known by John or being wondered about by John.

In this paper I argue that (3) should be traded for two distinct interpretations. The arguments for the revisions are empirical. They come from observing

*This paper is a revision of sections 1 through 3 of my paper in the proceedings of the Ninth Amsterdam Colloquium; section 4 of that paper is developed in 'Strategies for scope taking' (this volume). This research was partially supported by NSF grant #SBR 9222501.

A. Szabolcsi (ed.), Ways of Scope Taking, 311–347.

exactly what quantifiers support pair-list readings in matrix and in complement contexts and analyzing what logical apparatus is necessary for accommodating these possibilities. In other words, the argument is guided by the heuristic formulated in Szabolcsi (1996):

(4) What range of quantifiers participates in a given phenomenon is suggestive of exactly what that phenomenon consists in.

The two proposed interpretations are as follows. In distinction to (3), (5) has domain restriction but no lifting, and (6) has lifting but the quantifier operates in its own usual manner.

(5) Matrix questions and complements of *wonder*-type verbs:
which $x \in A$, which y [x bit y]
where A is the unique set determined by the quantifier

(6) Complements of *find out*-type verbs:
$\lambda P[Q(\lambda x[P(\text{which}y[x\text{bit}y])])]$
where Q is the usual interpretation of the given quantifier

Since various papers in this volume argue that bare indefinites, universals and modified numerals contribute differently to the interpretation of the sentence, (6) cannot mean that in extensional complements, all types of noun phrases are "quantified in" in the sense of Montague, for instance. Rather, "the usual interpretation of the quantifier" needs to be read as a cover term: each type of noun phrase induces a pair-list reading in the same syntactico-semantic fashion that is characteristic of it in other scopal contexts. This is supported by the fact that the different types of noun phrases induce pair-list readings in somewhat different syntactic configurations. This issue is discussed in great detail in the next chapter (Beghelli 1996), with specific reference to the syntax and semantics of distributivity.

The organization of the paper is as follows:

Section 1 provides a brief summary of the pair-list literature, singling out some points that are particularly relevant for the coming discussion.

Section 2 shows that the dilemma of quantification versus domain restriction arises only in extensional complement interrogatives. In matrix questions and in intensional complements, only universals support pair-list readings, whence the simplest domain restriction treatment suffices. Related data, including coordination and cumulative readings, are discussed.

Section 3 argues that in the case of extensional complements, the domain restriction treatment is inadequate for at least two independent reasons. One has to do with the fact that not only upward monotonic quantifiers support pair-list readings, and the other with the derivation of "apparent scope out" readings. The reasoning is supplemented with some discussion of the semantic properties of "layered quantifiers."

The above will establish the need for quantification, so the question arises how the objections explicitly enlisted in the literature against quantification can be answered. Section 4 considers the de dicto reading of the quantifier's restriction, quantificational variability, and the absence of pair-list readings with *whether*-questions, and argues that they need not militate against the quantificational analysis. Section 5 summarizes the emergent proposal.

Finally, section 6 discusses the significance of the above findings for the behavior of weak islands. It has been claimed that one way to evade a weak island violation is for the potentially offending quantifier to support a pair-list reading. The present paper predicts, then, that the quantifiers that give rise to weak island violations in matrix questions and intensional complements are not the same as those that do in extensional complements. The data will be shown to bear out this prediction, which in turn provides additional support for the scopal approaches to weak islands and to pair-list readings.

1 SOME PROPOSALS IN THE LITERATURE

As was mentioned in the introduction, it is currently assumed that quantifiers do not behave in their usual manner when supporting pair-list readings; rather, they uniformly provide a domain restriction for the question. Why is quantification into interrogatives a problematic issue? Detailed discussions of the problems and how they are handled in the literature can be found in Groenendijk and Stokhof (1984, 1989), Higginbotham (1991), Lahiri (1991), and Chierchia (1993). The present section merely singles out a few points that will be relevant in the coming discussion.

The crux of the matter is that quantification is defined for domains of type t (expressions that can be true or false), and interrogatives are not such. Now essentially two strategies can be followed. One is to find a suitable subexpression or superexpression of type t, and quantify into that. Another is to let the quantifier contribute to the interpretation of the interrogative in some non-quantificational way which, however, gives the same semantic result.

In the discussion below, when a question contains a quantifier, I will be concerned only with the pair-list reading. No mention will be made of the other (primary) reading.

1.1 Karttunen

To begin with, Karttunen (1977) interprets the *wh*-question (7a) as the set of true propositions which have the semantic format 'Fido bit a.' E.g., if Fido bit Mary and Judy and no one else, then (7b) denotes the set of propositions $\{{}^\wedge$Fido bit Mary, ${}^\wedge$Fido bit Judy$\}$.

(7) a. Who did Fido bite? *or* who Fido bit

b. $\lambda p \exists x[{}^\vee p \;\&\; p = {}^\wedge(\text{Fido bit } x)]$

Trying the "subexpression trick," quantification into (7b) would give the following result:

(8) a. Who did every dog bite? *or* who every dog bit

b. * $\lambda p \forall y[\text{dog}(y) \rightarrow \exists x[{}^\vee p \;\&\; p = {}^\wedge(y \text{ bit } x)]]$

As Karttunen points out, (8b) is not what we want: the set of propositions in (8b) is empty whenever there is more than one dog. Thus (in the 1977 paper) he effectively invokes the "superexpression trick." He proposes that pair-list readings are obtained by quantifying into a superordinate clause; for matrix questions he assumes embedding under a silent performative verb. Using Hendriks' (1993) technique to generate extraclausal scope, we may restate this solution by postulating a uniformly lifted representation for pair-list readings. (9b) is the set of properties (like being known to John) such that for every dog y, the set of true answers to the question who y bit has those properties:

(9) a. Who did every dog bite? *or* who every dog bit

b. $\lambda P \forall y[\text{dog}(y) \rightarrow P(\lambda p \exists x[{}^\vee p \;\&\; p = {}^\wedge(y \text{ bit } x)])]$

1.2 Groenendijk and Stokhof

Compare this with Groenendijk and Stokhof's (1984) proposal. These authors interpret the basic interrogative as a single proposition: the set of those worlds i in which the things Fido bit are the same as in the real world j.[1] E.g., if Fido bit Mary and Judy and no one else, then (10b) denotes the proposition ${}^\wedge(\{x \mid \text{Fido bit } x\} = \{\text{Mary, Judy}\})$.[2]

(10) a. Who did Fido bite? *or* who Fido bit

[1] Groenendijk and Stokhof appeal to explicit quantification over possible worlds. Take, for instance, $\lambda x[\text{bit}(i)(x)(\text{Fido})]$. Here *bit* is understood as denoting the intension of the verb *bit*. Then *bit(i)* is its extension in world i. The whole lambda expression denotes the set of those who Fido bit in world i. $\lambda i[\text{bit}(i)(\text{King})(\text{Fido})]$ is the set of worlds in which Fido bit King, i.e., the proposition that Fido bit King. Montague (1974) would have written this as ${}^\wedge(\text{bit(King)(Fido)})$. Groenendijk and Stokhof cannot use this simpler notation because it would not enable making reference to worlds, which they need in (10b). These notational complications are quite independent of our main concern.

[2] A proposal to treat questions as generalized quantifiers is presented in Gutiérrez Rexach (1996). This can be regarded as an extensional version of Groenendijk and Stokhof in view of the fact that *Who did Fido bite?* is interpreted as $\lambda P[\lambda x[\text{person}(x) \;\&\; \text{Fido bit}(x)] = P]$, where P is an answer set.

b. $\lambda i[\lambda x[\mathrm{bit}(i)(x)(\mathrm{Fido})] = \lambda x[\mathrm{bit}(j)(x)(\mathrm{Fido})]]$

In this case the "subexpression trick" does work for some examples:

(11) a. Who did every dog bite? *or* who every dog bit

b. $\lambda i \forall y[\mathrm{dog}(j)(y) \rightarrow (\lambda x[\mathrm{bit}(i)(x)(y)] = \lambda x[\mathrm{bit}(j)(x)(y)])]$

However, as Groenendijk and Stokhof point out, this does not give the desired result when the quantifier is an indefinite:

(12) a. Who did six dogs bite? *or* who six dogs bit

b. * $\lambda i \exists_6 y[\mathrm{dog}(j)(y)\ \&\ (\lambda x[\mathrm{bit}(i)(x)(y)] = \lambda x[\mathrm{bit}(j)(x)(y)])]$

The crucial difference is that in (11a) we have a universal and, consequently, the question has a unique true and complete answer. The question in (12a) on the other hand does not have a unique true and complete answer, i.e., it does not denote a unique proposition. On the intended interpretation, one first has to choose some set of six dogs (hence the name choice question); only after this is accomplished can the real question be asked and receive a true and complete answer. To accommodate the existential quantifier that captures choice, lifting is necessary. The format of the simplest amendment of (12b) would come quite close to (my expression of) Karttunen's (9b):

(13) a. Who did six dogs bite? *or* who six dogs bit

b. $\lambda P \exists_6 y[\mathrm{dog}(j)(y)\ \&\ P(j)(\lambda i[\lambda x[\mathrm{bit}(i)(x)(y)] = \lambda x[\mathrm{bit}(j)(x)(y)]])]$

The properties P that a lifted question takes as argument are like being a secret or being known by John.

Groenendijk and Stokhof (1984) discuss Karttunen's quantificational approach in detail and reject it. The ultimate reason is that (9b) as well as (13b) interpret the predicate *dog* de re. I return to this in section 4.1 below. They propose a quite different approach, namely, that the quantifier, whether it be a universal or an existential, does not act in pair-list readings the same way it does elsewhere. Instead, it determines a set that serves to restrict the domain of the question. The crucial insight is that the universal in *Who did every dog bite?* functions in the same way as the *wh*-in-situ in *Who did which dog bite?*. Similarly, *Who did six dogs bite?* may be interpreted as *Who did which of the six dogs that you chose bite?*.

The set that serves to restrict the domain of the question is a witness set of the quantifier. (For some background, see Chapter 1.)

(14) A set W is a witness of the generalized quantifier GQ iff W is an element of GQ and is also a subset of the smallest set GQ lives on.

Universals have a unique witness. The witness of $[\![every\ dog]\!]$ is the set DOG. An indefinite containing the numeral n has as many witnesses as there are distinct n-tuples in the relevant part of the universe; e.g. any set that contains at least six dogs and no non-dogs is a witness of $[\![six\ dogs]\!]$.[3] Thus the general format of pair-list readings in Groenendijk and Stokhof (1984) is as follows:

(15) a. Who did every/six dog(s) bite? *or* who every/six dog(s) bit

b. $\lambda P \exists W[W$ a witness of $[\![every/six\ dog(s)]\!]$ &
$P(j)(\lambda i[\lambda x_{x \in W} \lambda y[\mathrm{bit}(i)(y)(x)] = \lambda x_{x \in W} \lambda y[\mathrm{bit}(j)(y)(x)]]])]$

Since $[\![every\ dog]\!]$ has a unique witness, the set DOG, lifting in this case might be dispensed with and we could have (16):

(16) $\lambda i[\lambda x_{x \in \mathrm{DOG}} \lambda y[\mathrm{bit}(i)(y)(x)] = \lambda x_{x \in \mathrm{DOG}} \lambda y[\mathrm{bit}(j)(y)(x)]]$

Many complicated-looking details of (15b) are irrelevant for the coming discussion, so from now on I will abbreviate it as follows:

(17) Schematic representation of the pair-list reading using domain restriction:

$$\lambda P \exists W[\mathrm{witness}(W, [\![QP]\!]) \ \&\ P(\text{which } x \in W \text{ bit whom})]$$

Technically, both (17) and (9b) contain a property variable P that applies to a question denotation. The main difference is that in (9b), the quantifier occurs outside this question denotation, whereas in (17), only the choice of W does. The rest of the action associated with the quantifier (cf. *which* $x \in W$) occurs inside the question denotation. In 3.2 we shall see that this is what eventually qualifies (9b) "quantificational" and (17) "non-quantificational."

This is the background that I assume below. A few further comments are in order.

As Chierchia (1993) explains in detail, there is a to some extent terminological debate concerning whether pair-list readings involve scope and quantification into interrogatives. Groenendijk and Stokhof (1984) maintain that the quantifier is assigned scope over the *wh*-phrase, but the phenomenon is not quantification. Higginbotham and May (1981), May (1985) and Higginbotham (1991, 1993) on the other hand argue that we are dealing with both scope and quantification; however, their explication of what quantification amounts to in this context is, in the pertinent respects, logically equivalent to Groenendijk

[3] Groenendijk and Stokhof actually use minimal elements, and Chierchia, minimal witnesses, in the definition of domain restriction. Minimality causes a problem because it collapses $[\![at\ least\ six\ dogs]\!]$, $[\![more\ than\ five\ dogs]\!]$ and $[\![exactly\ six\ dogs]\!]$ on the one hand, and all decreasing quantifiers on the other. Plain witness gives the correct results. I presuppose this improvement in the main text.

and Stokhof's (1984) allegedly non-quantificational solution. For this reason, I do not discuss this theory separately; unless otherwise indicated, whatever I say about Groenendijk and Stokhof is assumed to hold of Higginbotham and May, too.

1.3 Chierchia

Engdahl's (1985) approach, which inspired Chierchia's (1992, 1993) and Jacobson's (1992), constitutes a genuine alternative. Engdahl takes functional questions (18) to be paradigmatic and proposes that individual questions (19) as well as pair-list questions (20) are but special cases:

(18) a. Who did every dog/no dog bite?
which f, every/no dog x bit $f(x)$

b. Its (own) master.
f = master-of

(19) a. Who did Fido bite?
which f, Fido bit f(Fido)

b. Mary.
$f = \{\langle\text{Fido, Mary}\rangle\}$

(20) a. Who did every dog bite?
which f, for every dog x, x bit $f(x)$

b. Fido bit Mary, Spot Fido, and King my cat.
$f = \{\langle\text{Fido, Mary}\rangle, \langle\text{Spot, Fido}\rangle, \langle\text{King, my cat}\rangle\}$

As Chierchia explains, the parallelism between the classical functional reading and the so-called pair-list reading is grounded in the fact that a function may be defined "intensionally," pointing out a generalization, or "extensionally," simply specifying a set of ordered pairs. The classical functional reading obtains when a generalization is available, cf. (18b). The pair-list reading obtains when we are content with an extensional definition, cf. (20b).

This approach differs from all the above in that it does not assume a QP>WH scope relation in pair-list readings: the *wh*-phrase has widest scope. Chierchia enlists a novel empirical reason for adopting this analysis: he proposes to explain the well-known subject/object asymmetry in pair-list licensing with reference to a Weak Cross-over effect induced by the "layered trace" fx. It seems, however, that a wider range of data exhibits intricate patterns that can by no means be reduced to WCO. There are differences between matrix and complement and between *every* and *each* that WCO cannot predict, and even the behavior of VP-internal arguments seems to diverge from the WCO pattern. See Beghelli (1996) for a detailed discussion of the relevant data. In the light of these, I do not find Chierchia's syntactic argument compelling.

Exactly how interrogatives are interpreted and how the quantifier contributes to the pair-list reading are matters that are independent of the above choice. Chierchia (1993) combines Engdahl's functional approach with Karttunen's interpretation of interrogatives and Groenendijk and Stokhof's innovation of letting the quantifier contribute a domain restriction. The result is as follows:

(21) a. Who did every/six dog(s) bite? *or* who every/six dog(s) bit

b. $\lambda P \exists W [W$ a witness of $[\![$*every/six dog(s)*$]\!]$ &
$P(\lambda p \exists f_{f \in [W \to \text{ANIMATE}]} \exists x_{x \in W} [{}^{\vee}p \;\&\; p = {}^{\wedge}\text{bit}(x, f(x))])]$

1.4 Summary

Singling out a few points that are particularly relevant for the coming discussion, let me conclude this section with the following observation:

(22) Groenendijk and Stokhof's, Higginbotham's and Chierchia's approaches to pair-list readings share the following properties (overtly or in view of logical equivalence):

a. No descriptive or theoretical distinction is made between matrix and complement cases;

b. All pair-list readings are (allowed to be) interpreted as generalized quantifiers over individual questions;

c. The quantifier is assumed to contribute a set that serves to restrict the domain of the question.

I will challenge these assumptions on the basis of data concerning what quantifiers support pair-list readings.

2 THE MATRIX VERSUS EXTENSIONAL COMPLEMENT ASYMMETRY

This section will demonstrate that the ranges of quantifiers that support genuine pair-list readings in matrix and in complement contexts are quite different. In brief, only universals do so in the matrix and in intensional (*wonder*) complements, whereas almost all quantifiers do so in extensional (*find out*) complements.[4] Anticipating the detailed data to be presented in the sections below, let us see what the significance of these observations might be.

The traditional ideal of formal elegance requires that the accounts of quantificational phenomena be designed to be as general as possible. The results presented in this volume as well as elsewhere indicate, however, that very often only particular subsets of quantifiers participate in a given process. One way to deal with this is to supplement the general accounts with filters. Another is to go for specialized accounts from the very beginning. A specialized account is one that builds on the distinctive properties of that subset of quantifiers that actually participate in the phenomenon to be accounted for and does not try to be applicable to others. This is the strategy I am following. Therefore, the accounts of matrix and complement pair-list readings must match the respective participating quantifiers.

Consider matrix questions first. As Section 1 made clear, interpreting pair-list readings as generalized quantifiers over individual questions ("lifting") is necessitated only by choice questions, which do not have a unique true and complete answer. If only universals need to be taken care of, then, using a Groenendijk and Stokhof-style interpretation of interrogatives, either the simplest form of quantification (12b) or the simplest form of domain restriction (17) will do.

(12b) $\lambda i[\forall y[\mathrm{dog}(j)(y) \rightarrow (\lambda x[\mathrm{bit}(i)(x)(y)] = \lambda x[\mathrm{bit}(j)(x)(y)])]$

(17) $\lambda i[\lambda x_{x\in\mathrm{DOG}}\lambda y[\mathrm{bit}(i)(y)(x)] = \lambda x_{x\in\mathrm{DOG}}\lambda y[\mathrm{bit}(j)(y)(x)]]$

More precisely, only (12b) is really contingent on adopting Groenendijk and Stokhof's particular interpretation of interrogatives. Recall that the gist of (17) is that it assimilates *Who did every dog bite?* to *Who did which dog bite?*. Thus an analog of (17) should be possible to devise in any theory that handles multiple interrogation.

Let us assume that domain restriction is the adequate account of matrix pair-list questions. Does some form of it extend to complement interrogatives in general, as suggested in the literature? It will be pointed out in 3.1 that the answer depends on the monotonicity properties of the participating quantifiers. Given that the quantifiers that support complement pair-list readings are not all filters and, furthermore, some of them are not even upward monotonic, we shall see that the data lead the conclusion that the answer is No.

It is clear, then, that on my analysis pair-list readings do not constitute a unitary phenomenon. It may be a little unsettling to assign divergent semantic analyses to matrix and intensional/extensional complement cases but, in fact, Beghelli (1996) points out that they must diverge in syntax, too.

[4]The intensional/extensional qualification of these complements comes from Groenendijk and Stokhof.

2.1 Universals versus modified numeral indefinites

The basic observation that indefinites only support pair-list readings in complements was made in the course of joint work with Frans Zwarts in 1992. A study by St. John (1993) confirmed this and revealed the significance of cumulative readings; Doetjes (1993) independently made consonant suggestions. I thank S. Spellmire for assistance with the field work from which come the more detailed data which the present paper rests on.

Consider first (23) versus (24):

(23)	Who	did *every dog* bite?	OK Fido bit X, Spot bit Y, ...
	Which boys		OK Fido bit X, Spot bit Y, ...
	Which boy		% Fido bit X, Spot bit Y, ...
	What boy		% Fido bit X, Spot bit Y, ...
(24)	Who	did *more than two dogs* bite?	* Fido bit X, Spot bit Y, ...
	Which boys		* Fido bit X, Spot bit Y, ...
	Which boy		* Fido bit X, Spot bit Y, ...
	What boy		* Fido bit X, Spot bit Y, ...

There is a clear contrast between the two sets. *Every dog* is a basically good inducer of pair-list readings (although not as good as is assumed in the literature: many speakers reject the examples that contain an overtly singular *wh*-phrase, see the %'s). On the other hand, no speaker is tempted to answer the question containing *more than two dogs* with a list of pairs. Similar to *more than two dogs* are all "modified numerals," e.g. *two or more dogs, exactly two dogs, fewer than two dogs, many/few dogs.* As to *Who did at least two dogs bite?,* some speakers are willing to answer it with a pair-list, but this is probably a pragmatic "mention some" effect induced by a non-logical use of *at least.* The reason to believe this is that (i) logically equivalent *two or more dogs* never elicits pair-list answers, and (ii) speakers who answer the *at least two dogs* question with a pair-list tend pick just two dogs, rather than three or eleven.

But the contrast between universals and "modified numeral indefinites" vanishes in complements, together with the mysterious (to me) marginality of singular *wh*-phrases. For instance:

(25) a. John found out who/which boys every dog bit. cf. (23)
OK 'John found out about every dog who/which boys it bit'

b. John found out which boy every dog bit.
OK 'John found out about every dog which boy it bit'

(26) John found out which boy more than two dogs bit. cf. (24)
OK 'John found out about more than two dogs which boy each bit'

Similar to *more than two dogs* are practically all the modified numerals listed above; the data will be discussed more closely in 3.1.

To be more precise, the matrix effects disappear only in a subset of the complement cases. The paradigm below indicates that the complement of *wonder* behaves exactly like matrix questions:

(27) a. John wonders who every dog bit. cf. (23), (25)
OK pair-list

b. John wonders which boy every dog bit.
% pair-list

(28) John wonders who more than two dogs bit. cf. (24), (26)
?? pair-list

The demarcation line is between matrix verbs that Groenendijk and Stokhof call extensional versus the ones that they call intensional. The names are due to the fact that on their approach, the extension of a question is its answer. The sentence *John found out who came* means 'John found out the answer to the question *Who came?*.' On the other hand, *John wonders who came* means 'John stands in the wonder-relation with the question *Who came?*.' Apparently, one cannot stand in the wonder-relation to a question which, not being a possible matrix question, cannot be asked in its own right.[5]

What the data show, then, is that modified numeral indefinites support a pair-list reading only in (extensional) complements. One possibility is that the asymmetry between matrix (24) and complement (26) hinges on the very notion of choice. Intuitively, the desired reading seems to require more than the existence of a witness set about whose elements the question may be asked. Rather, it seems to require that the indefinite be able to "offer up sets for choice." Modified numeral indefinites are apparently unable to do so, and this is not surprising: as Szabolcsi (1996) shows, they are essentially counters, not set denoters. In contrast, pair-list readings in the complement do not involve any "choice." As the paraphrases indicate, they involve counting (here: counting the dogs about whom John found out which boy they bit). For the modified numeral, this is business as usual.

2.2 The natural habitat of lifted questions

In this section, I wish to take another look at the explanation for the matrix versus subordinate asymmetry offered at the end of 2.1.

[5]There are other respects in which complements of *wonder* behave like matrix questions. Munsat (1986) notes a variety of such points, including the licensing of negative polarity items. Berman (1990) draws a parallelism in the context of quantificational variability. G. Carlson (p.c.) points out that in some American English dialects, *wonder*-complements exhibit inversion, together with sequence of tenses.

I suggested that the reason why modified indefinites do not support matrix pair-list questions is that, being counters, they cannot offer up sets for choice. If this is true, then bare numeral indefinites, which are known to introduce set (group) referents, are expected to be great inducers of choice readings. But, while certainly there is improvement, they are just not good enough:

(29) Which man/who did more than two dogs bite?
* Fido bit X, King bit Y, and Spot bit Z.

(30) Which man/who did two dogs bite?
? Fido bit X, and King bit Y.

Moreover, while English allows this use of bare numerals, Dutch does so to a much lesser extent (even Groenendijk and Stokhof themselves (1984, pp. 555–6) express serious doubts about the availability of these readings in Dutch):

(31) Welk boek lazen twee jongens?
what book read two boys
?* Jaap read War and Peace, and Henk read Magic Mountain.

Thus the possibility arises that no matrix interrogative ever involves choice and (30), to the extent it is acceptable, is an instance of something else.[6] But what can be wrong with choice?

[6]In this note I offer an analysis of what this "something else" might be. I admit, however, that I do not yet have a fully satisfactory pretheoretical grasp of these particular data, whence the analysis may need to be revised in the future. I expect that this will not affect the rest of the proposal in this paper.

Bare numeral indefinites in English appear to be able to support matrix pair-list (choice) readings. I will first claim that these are not really pair-list cases.

Krifka (1991) and Srivastav (1992) discuss questions with definites, and argue that they support not pair-list but cumulative readings. Consider:

(i)	Who		OK Fido bit X and Spot bit Y.
	Which boys	did *the dogs* bite?	OK Fido bit X and Spot bit Y.
	Which/what boy		* Fido bit X and Spot bit Y.

They argue that the "real answer" here would be ***The dogs, Fido and Spot, bit X and Y (between them),*** and the apparent pair-list answers are just more cooperative ways of spelling out how exactly the bitings were distributed. (The same basic observation had been made in Szabolcsi 1983, p. 128, in response to Haïk 1984.)

I suggest that the indefinites data in (30) is to be interpreted in the same way. Namely, an answer of the pair-list format is acceptable only insofar as it merely disambiguates an acceptable answer of the cumulative format:

(ii)	Who		OK Fido bit X and King bit Y = They bit X and Y
	Which boys	did *two dogs* bite?	OK Fido bit X and King bit Y = They bit X and Y
	Which/what boy		?? Fido bit X and King bit Y $\neq$ They bit X and Y

What is the evidence for this analysis?

Prior to proposing an answer, it is in order to note that the claim that only universals support pair-list readings is in some sense not new. At earlier stages of their work, both Groenendijk and Stokhof (1982) and Chierchia (1992) had made such a claim and offered their own explanations for the restriction. The critical difference between these theories and mine is that they assumed that all matrix and complement cases work identically, whereas I am observing a descriptive contrast and am therefore seeking an explanation that holds for matrix questions but not, for instance, the *wh*-complement of ***find out.***

As was pointed out in 1.2, the fact that an interrogative does not have a unique true and complete answer requires lifting the question and interpreting it as a generalized quantifier, viz. as a set of properties like being a secret,

---

Two dogs questions resemble ***the dogs*** questions in that they typically require *wh*-phrases that are not overtly singular. This is unfortunately of less diagnostic value than Krifka and Srivastav think, since many speakers of English reject the singular even with fixed domain readings. However, I have found reliable informants who do accept the singular in the case of ***every dog*** (hence the % in (23)) and nevertheless reject it in the case of ***the dogs*** and ***two dogs,*** which squares with the analysis.

Another relevant fact is that plain X ***and*** Y is itself an acceptable answer to ***Who/ Which boys***[plural!] ***did two dogs bite?*** in the cumulative situation where one dog bit X and the other bit Y, i.e., when the dogs bit one person each.

Why is the cumulative option unavailable to modified numeral indefinites, cf. (29)? The term "cumulative" may be a little misleading here, since Scha (1981) introduced it in connection with cardinalities, and indeed, ***More than two dogs bit fewer than six boys between them*** is fine. The readings in (30)/(ii) should rather be called "distributed group" readings. I suggest that ***more than two dogs*** and its brothers do not participate in such readings because, unlike ***the/two dogs,*** they are not potential group denoters in the relevant sense. This accords with Kamp and Reyle's (1993) observations; for further discussion, see Szabolcsi (1996) and Beghelli (1996).

The claim that bare numerals support cumulative (distributed group) readings, not pair-list readings does not directly solve our basic problem, however. Since groups consisting of two dogs can be many, choice is involved here, too. Also, there is a type of data that has not been mentioned yet: disjunctive questions.

(iii) Who did Fido or King bite? OK King bit John.

On Groenendijk and Stokhof's analysis, these are choice questions, too (and, according to their judgment, the intuitively best case).

What I am going to suggest is that (iii) is not an instance of the choice reading. Rather, the sole interpretation of the question is one where the *wh*-phrase has widest scope, i.e., 'Who is such that either Fido or King bit him?' The answer given above is a partial answer (presented in a co-operatively explicit format à la Srivastav and Krifka), which is elicited under particular pragmatic circumstances that Groenendijk and Stokhof (1984) call "mention-some" contexts. In the same vein, I assume that the pertinent distributed group reading of (ii) ***Who/which boys did two dogs bite?*** is also a "mention-some" example, rather than a choice reading. Thus its representation is as in (iv), with B a variable over groups of boys and D over groups of dogs:

(iv) $\lambda i[\lambda B \exists D[|\mathrm{ATOMS}(D)| = 2 \,\&\, \forall b \leq B[\exists d \leq D[d \,\mathrm{bit}(i)\, b]] \,\&$
$\forall d \leq D[\exists b \leq B[d \,\mathrm{bit}(i)\, b]] = \lambda B \exists D[|\mathrm{ATOMS}(D)| = 2 \,\&$
$\forall b \leq B[\exists d \leq D[d \,\mathrm{bit}(j)\, b]] \,\&\, \forall d \leq D[\exists b \leq B[d \,\mathrm{bit}(j)\, b]]]$

being known by John, etc. Now, such an interpretation is entirely natural for complement interrogatives, which are indeed syntactic arguments of expressions denoting such properties. But matrix interrogatives never combine with expressions of this sort (unless we literally adopt Ross's silent performative hypothesis). Instead, they are genuine questions. Thus it seems natural to interpret them in a way that directly links them to possible answers, which is what the unlifted interpretations do; and it is not natural to interpret them as lifted questions. The natural habitat of lifted questions is the argument position.[7]

In general, I assume that it is justified to interpret an expression E as a function over properties P only if E actually combines with denoters of such properties. This constraint properly distinguishes lifted questions from questions-as-generalized quantifiers in the sense of Gutiérrez Rexach (1996). In the former case, the question takes properties denotable by matrix clauses as argument; in the latter case, the question takes properties denotable by elliptical answers as argument. The latter is fully justified for a matrix question.

If this reasoning is correct, then only interrogatives which in virtue of their semantic form have a unique answer are possible in the matrix. This claim excludes choice questions with indefinites, whether modified or bare. Interestingly, it also makes predictions for data in a related domain: conjunctions and disjunctions of questions.

2.3 Conjunctions and disjunctions of interrogatives

Groenendijk and Stokhof point out that both fixed domain questions and choice questions come in (at least) two varieties:

(32) a. What did every girl read?

b. What did Mary read? And, what did Judy read?

(33) a. What did some girl read?

b. What did Mary read? Or, what did Judy read?

The parallelism between the (a) and the (b) cases is of course based on the fact that universal quantification reduces to conjunction (intersection), and existential quantification to disjunction (union). What is relevant to us here is that the two-question sequence in (32b) has a unique complete and true answer, exactly as (32a) does, while the two-question sequence in (33b) lacks one (involves a choice), exactly as (33a) does. This entails that the representation of (32b) can go unlifted, but that of (33b) cannot:

---

[7] I thank G. Chierchia for discussion on this point.

(34) What did Mary read? And, what did Judy read?

$$\begin{aligned}
&\lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{mary})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{mary})]] \cap \\
&\lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{judy})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{judy})]] \\
&= \lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{mary})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{mary})]\ \& \\
&\quad \lambda x[\mathrm{read}(j)(x)(\mathrm{judy})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{judy})]]
\end{aligned}$$

(35) What did Mary read? Or, what did Judy read?

$$\begin{aligned}
* \quad &\lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{mary})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{mary})]] \cup \\
&\lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{judy})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{judy})]] \\
&= \lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{mary})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{mary})] \vee \\
&\quad \lambda x[\mathrm{read}(j)(x)(\mathrm{mary})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{judy})]]
\end{aligned}$$

$$\begin{aligned}
&\lambda P[P(j)(\lambda j \lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{mary})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{mary}]]) \vee \\
&\quad P(j)(\lambda j \lambda i[\lambda x[\mathrm{read}(j)(x)(\mathrm{judy})] = \lambda x[\mathrm{read}(i)(x)(\mathrm{judy})]])]]
\end{aligned}$$

We are now predicting that question disjunctions are unavailable wherever pair-list readings with indefinites are. Let us see how this prediction works out.

Question disjunctions that illustrate the choice reading in the literature invariably come in an inter-sentential format, as in (33b). If this were an irrelevant detail, the *or* connecting the two sentences could easily be moved into intra-sentential position. But it cannot:

(36) a. Who did you marry? *Or,* where do you live?

b. ?? Who did you marry *or* where do you live?

This suggests that the *or* in (33b) and (36a) does not really offer a choice but, instead, is an idiomatic device that allows one to cancel the first question and replace it with the second. This idiomatic character is corroborated by the fact that the Hungarian equivalents are entirely unacceptable unless *inkább* 'rather, instead' is added; something that we do not expect if the connective acts as a standard Boolean operator. The marginality of (36b) indicates, then, that questions cannot be directly disjoined.

Just as pair-list readings with indefinites are perfect in extensional complements, disjunction becomes impeccable, too. But the claim that questions cannot be directly disjoined is confirmed by the fact that (37) only has a wide scope *or* (distributive) interpretation obtained by lifting both disjuncts:

(37) John found out who you married *or* where you live.

i. 'John found out who you married or found out where you live'

ii. * 'John found out (who you married or where you live)'

Naturally, for this distinction to make sense, the two readings must be distinct. According to Groenendijk and Stokhof (1989), *know wh-ϕ and/or wh-ψ* is logically equivalent to *know wh-ϕ and/or know wh-ψ.* I disagree with this in the case of *or.* Take (38):

(38) a. Bill knows where John lives or knows who Sue married.

b. Bill knows (where John lives or who Sue married).

If Bill never heard of Sue, (38a) may be true but (38b), if grammatical at all, seems implausible. Intensional verbs, as above, retain the matrix effect:

(39) ?? John wonders where you live or who you married.
'John would be happy to know either'

The claim that interrogatives must be lifted first to become disjoinable is corroborated by syntactic data from Hungarian and Korean (Seungho Nam, p.c.). In these languages, even *wh*-complements are introduced by a subordinator morpheme. The Hungarian subordinator is *hogy,* and the counterpart of (37) is unacceptable unless both disjuncts contain a *hogy*:

(40) János megtudta, hogy kit vettél feleségül vagy *(hogy) hol
John found-out *that* whom you married or *(*that*) where
laksz.
you-live

In Korean, *ci* is the subordinator:

(41) a. * na-nun Mary-ka etiey sal-**kena** Kathy-ka etiey
I-top Mary-nom where live-**or** Kathy-nom where
sal-nun-**ci** al-ayo
live-pres-**comp** know

b. na-nun Mary-ka etiey sal-nun-**ci** **hokun** etiey
I-top Mary-nom where live-pres-**comp** **or** where
sal-nun-**ci** al-ayo Kathy-ka
live-pres-**comp** know Kathy-nom
'I know where Mary lives or where Kathy lives'

That our predictions are borne out for the right reason (that is, for a semantic, not a logico-syntactic one) is corroborated by the fact that conjunctions pattern like universals. *And* can be moved into intra-sentential position, and the repetition of the subordinator (*hogy*/*ci*) is optional:

(36)′ a. Who did you marry? *And,* where do you live?

b. Who did you marry *and* where do you live?

(40)′ János megtudta, hogy kit vettél feleségül és (hogy) hol laksz
John found-out *that* whom you married *and* (*that*) where you-live

(41a)′ na-nun Mary-ka etiey sal-**ko** Kathy-ka etiey sal-nun-**ci** al-ayo
I-top Mary-nom where live-**and** Kathy-nom where live-pres-**comp** know
'I know where Mary lives and where Kathy lives'

To conclude, it seems plausible that the reason why matrix choice questions (whether they involve modified numeral indefinites, bare numeral indefinites, or disjunction) do not exist is that matrix clauses cannot denote generalized quantifiers of the pertinent kind. (For a preliminary account of some residual cases, see note 6.)

It is worth noting that my findings refute the letter, but not the spirit, of Groenendijk and Stokhof's theory of choice questions. It is true that the data turn out to be different than they assumed. But what their theory says really is that *if* choice questions exist, they *have to* be lifted. The fact that choice questions do not exist in a context where it is reasonable to assume that denoting lifted questions is impossible is perfectly consistent with this theory.

3 THE NECESSITY OF QUANTIFICATION INTO (EXTENSIONAL) COMPLEMENT INTERROGATIVES

3.1 Domain restriction and monotonicity

Let us from now on focus solely on (extensional) complement interrogatives.

In what follows I will assume that all complement interrogatives denote generalized quantifiers. The question, then, is whether the domain restriction schema in (17) is an adequate general representation of complement pair-list readings:

(42) a. ... who QP bit

b. $\lambda P \exists W[\text{witness}(W, [\![QP]\!])\ \&\ P(\text{which } x \in W \text{ bit whom})]$

I argue that it is not adequate, for at least two independent reasons. The first has to do with monotonicity. The second has to do with "apparent scope out" readings, to be discussed in 3.2.

The simple point to be made in this subsection is that domain restriction requires upward monotonicity. Why? "Domain restriction" means that we pick a set and restrict our attention to its members, ignoring whatever happens outside. But we can only safely do so if that set is determined by an increasing quantifier. To illustrate with non-interrogative examples,

(43) a. *(At least) two men walk* = There is a set of (at least) two men who walk
(it does not matter if men outside this set also walk)

b. *Exactly two men walk* $\neq$ There is a set of exactly two men who walk
(we must guarantee that all walking men are in this set)

c. *Less than two men walk* $\neq$ There is a set of less than two men who walk
(we must guarantee that all walking men are in this set)

The schema in (42) faces exactly the same problem as the paraphrases in (43). For instance, if P is replaced by *John knows,* we get that there is a witness W of QP about whose members John knows who they bit, ignoring whatever else John knows. (42) misinterprets any sentence in which the QP inducing the pair-list reading is not upward monotonic.

At this point the empirical question of exactly what quantifiers support pair-list readings becomes crucial. It is sometimes claimed in the literature that only upward monotonic cases work. The data justifying this claim tend to involve only matrix questions with *no* N, however. That is, neither other decreasing quantifiers, nor non-monotonic quantifiers (which pose exactly the same logical problem) are investigated.

In 2.1 I have anticipated that, in distinction to matrix questions, almost all quantifiers support pair-list readings in extensional complements. Let us now take a closer look at the data.

One type of context I used to elicit the relevant judgments is as follows. We are in the business of finding out how dangerous each neighborhood dog is and get together to compare notes. This context simply makes the competing non-pair-list reading of the complement irrelevant, without being either pragmatically or syntactically too special to produce representative judgments. A sample of the results is as follows:

(44) a. I found out who *three dogs* bit.

b. I did a lot better! I found out who *more than five dogs* bit.

c. John is not here but I have glanced at his list, and I estimate that he found out who *more than five but certainly fewer than ten dogs* bit.

d. And I know that Judy found out who *exactly four dogs* bit.

e. ? Bill is very lazy: he only found out who *at most three dogs* bit.

f. * Mary is even worse: she found out who *no dog* bit.

g. Don't worry; I think we now know who *every dog* bit.

What we see is that the only type of quantifier that is clearly excluded in this context is *no dog.* With this one exception, increasing (44a, b, g), non-monotonic (44c, d), and decreasing (44e) quantifiers are found to support a pair-list reading. It is true that decreasing examples seem to require the presence of *only* in the matrix and even so, they may be somewhat worse than the rest. The crucial fact is, however, that upward monotonicity is not a sine qua non for the pair-list reading.

The conclusion is that the domain restriction schema (42) needs amending. Let us consider three alternatives.

The first, (45) just adds an ad hoc maximality condition to (42), so that it will not go wrong if QP is not upward monotonic.

(45) $\lambda P \exists W[\text{witness}(W, [\![QP]\!]) \; \& \; P(\text{which } x \in W \text{ bit whom})$
$\& \; \forall x[x \notin W \;\rightarrow\; \neg P(\text{whom } x \text{ bit})]]$

The second version, (46) departs from this most radically: it is standard quantification into a lifted interrogative, assigning wide scope to *Q dogs* over the *wh*-phrase.

(46) $\lambda P Q x[\text{dog}(x), P(\text{who } y[x \text{ bit } y])]$

The third version, (47) is an interesting intermediate case. If we read the original (42) as a noble, though empirically incorrect, attempt to express that only increasing quantifiers support pair-list readings, then (47) just expresses, in the same spirit, what seems to emerge from (44) as the correct empirical generalization, namely, that all quantifiers except for the type *no dog* do so. This is how (47) works. QP is required to have a non-empty witness A. "Negative" quantifiers like *no dog* are distinguished by having the empty set as their unique witness, so this formulation lets all others in (given a universe that is not trivially too small).[8] The rest ensures that all and only the members of A count:

[8] If data involving other decreasing quantifiers are not judged to be quite good enough, (47) can be reformulated as follows:

$\lambda P \exists A \exists B[\text{non-}\emptyset \text{ minimal witness}(B, [\![QP]\!])$
$\& \text{ witness}(A, [\![QP]\!]) \; \& \; \forall x[P(\text{whom } x \text{ bit}) \text{ iff } x \in A]]$

(47) $\lambda P \exists A[\text{non-}\emptyset \text{ witness}(A, [\![QP]\!])\ \&\ \forall x[P(\text{whom } x \text{ bit}) \leftrightarrow x \in A]]$

At first sight (47), too, seems like an innocent improvement over (42): the maximality condition is no longer added like an afterthought. But the new formulation makes a crucial difference. In (42), both reference to the relevant witness and universal quantification over its members took place inside the argument of the property variable P: cf. P(**which** $\boldsymbol{x} \in \boldsymbol{W}$ bit whom). In (47), both take place outside P: cf. $\forall \boldsymbol{x}[P(\text{whom } x \text{ bit})$**iff** $\boldsymbol{x} \in \boldsymbol{A}$. This has the consequence that (47) is every bit as "quantificational" as (46) is.

How shall we choose between these formulations?

(45), let's face it, is quite ugly. But notice that there is a certain similarity between it and a schema discussed in Beghelli, Ben-Shalom, and Szabolcsi (1996): the schema for branching quantification proposed in Sher (1991). Informally, Sher's definition of branching goes as follows: There are two sets A and B such that their cross-product $A \times B$ is in the relation R, and $A \times B$ is the largest cross-product in R. Both schemata start out with a formulation that makes sense only when increasing quantifiers are involved, namely, a formulation involving existential quantification over elements/witnesses of the quantifier. Then both schemata are supplemented with an independent maximality condition to take care of the non-monotonic and decreasing cases. So, if Sher's schema is acceptable (independently of what natural language examples correspond to it), (45) should be acceptable, too. Or should it? It seems to me that there is a difference. Namely, in the case of branching there is extremely good motivation for appealing to existential quantification over sets. This is what captures a core ingredient of our intuition about branching, namely, that it involves two sets that are chosen independently. In other words, our intuition about branching is heavily based on the increasing case, whence this "modular" approach seems justified. On the other hand, I do not believe we have a comparable core intuition about increasing cases in complement pair-list readings. (The matrix case is different!) Therefore, it seems to me that (45) can be ruled out on purely aesthetic grounds.

Aesthetics notwithstanding, it remains to be seen whether there is hard empirical evidence in favor of any of these alternatives. In 3.2 I argue that there is.

Here we require QP to have a non-empty minimal witness B. This excludes all decreasing quantifiers (and also non-continuous quantifiers with a decreasing component, e.g. *fewer than two or more than six dogs,* which does not seem problematic). But we cannot stick with B: the minimal witnesses of *exactly three dogs* are the same as those of *three or more dogs* and *more than two dogs,* but sentences containing these QPs are not synonymous. We must be allowed to pick an appropriately big enlargement A of B to do the real work. This is what my formulation exploits. I thank D. Ben-Shalom for discussion on these matters.

3.2 "Apparent scope out" phenomena

3.2.1 Evidence for quantification into lifted interrogatives

It is generally agreed that whatever rule assigns scope to QPs like *every student,* it operates within the boundaries of one clause. A typical example is (48):

(48) Some librarian or other found out that every student needed help.
* 'every > some'

It is striking, then, that a comparable reading of (49) is entirely natural. Notice that on this reading not only the existence of students can be inferred in the matrix, but also the matrix subject becomes referentially dependent: the librarians vary with the boys:

(49) Some librarian or other found out which book every student needed.
OK 'every > some'

Should we allow *every student* to distributively scope out of its own clause? The qualification "distributively" is of utmost importance here. It is observed in Farkas (1996) and Beghelli and Stowell (1994) that both universals and bare numeral indefinites can take unbounded scope. This, however, pertains only to (some subset of) their restrictor; they do not make extraclausal quantifiers referentially dependent. Thus it would be quite exceptional for (49) to rely on such a possibility.

Moltmann and Szabolcsi (1994) argue that distributive scoping out is not necessary. It is proposed that the critical reading arises when the complement clause (i) has a pair-list reading and (ii) is assigned scope over the matrix subject. This latter is of course a clause-internal step. That is, the derivation is not (50) but (51):

(50) * [every student]$_i$ [some librarian found out which book x_i needed]

(51) [$_{\text{pair-list}}$ which book every student needed]$_i$ [some librarian found out v_i]

Apart from saving the clause-boundedness of *every* N's distributive scope, there are specific reasons for assuming (51). I will come back to these in 3.2.3, but first let us consider how the issue at hand helps evaluate the alternatives introduced in the previous section.

The question is what formal interpretation the pair-list reading must have for (51) to yield the "apparent scope out" effect. Let's see. In (52) through (54), I quantify (45) through (47) into *some librarian found out* p:

(52) $\lambda P \exists W[\text{witness}(W, [\![\text{every student}]\!]) \;\&$
$P(\text{which } x \in W \text{ needs which book}) \;\& \;\textsc{maximality}]$

$(\lambda p[\exists z[\text{librarian}(z) \mathbin{\&} \text{found-out}(z,p)]]) =$
$\exists W[\text{witness}(W, [\![\text{every student}]\!]) \mathbin{\&}$
$\exists z[\text{librarian}(z) \mathbin{\&} \text{found-out}(z, \text{which } x \in W \text{ needs which book})] \mathbin{\&}$
$\forall x[x \notin W \rightarrow \neg\exists z[\text{librarian}(z) \mathbin{\&}$
$\text{found-out}(z, \text{which book } x \text{ needs})]]]]$

(53) $\lambda P \forall x[\text{student}(x) \rightarrow P(\text{which book } y[x \text{ needs } y])]$
$(\lambda p[\exists z[\text{librarian}(z) \mathbin{\&} \text{found-out}(z,p)]]) =$
$\forall x[\text{student}(x) \rightarrow \exists z[\text{librarian}(z) \mathbin{\&}$
$\text{found-out}(z, \text{which book } y[x \text{ needs } y])]]$

(54) $\lambda P \exists A[\text{non-}\emptyset \text{ witness}(A, [\![\text{every student}]\!]) \mathbin{\&}$
$\forall x[P(\text{which book } x \text{ needs}) \text{ iff } x \in A]]$
$(\lambda p[\exists z[\text{librarian}(z) \mathbin{\&} \text{found-out}(z,p)]]) =$
$\exists A[\text{non-}\emptyset \text{ witness}(A, [\![\text{every student}]\!]) \mathbin{\&} \forall x[\exists z[\text{librarian}(z) \mathbin{\&}$
$\text{found-out}(z, \text{which book } x \text{ needs})] \text{ iff } x \in A]]$

Recall that (45) is Groenendijk and Stokhof's original domain restriction interpretation of the pair-list reading, supplemented by an ad hoc maximality condition to take care of not upward monotonic QPs. (52) shows that quantifying (45) into the matrix clause does not make the librarians vary with the students. It is easy to see why: as was mentioned above, in (45) all quantificational action takes place inside the argument of P that matrix material will replace. Thus matrix and complement quantifiers cannot interact scopally.

On the other hand, both (46) and (47) give the desired result: the librarians vary with the students. This confirms that they are variations on the same quantificational theme.

To summarize, first we have seen that not only upward monotonic quantifiers support pair-list readings. Restricting the domain of the question to a witness of a non-upward quantifier is logically incorrect unless a maximality condition is supplied. Two ways of stating the maximality condition plus a purely quantificational alternative were offered. Second, we have seen that of the two ways of handling maximality, only one can also cope with apparent scope out readings. This, however, is in every pertinent respect equivalent to the quantificational alternative.

The conclusion is, then, that the interpretation of complement pair-list readings must involve quantification into lifted questions. This, however, may be formulated in slightly different ways, e.g. (46) or (47).

3.2.2 Decreasing quantifiers

A minor issue, the choice between (46) and (47), is still left open. As they stand, both presuppose that the failure of (some or all) decreasing QPs to support pair-list readings has an independent explanation; technically, they differ in that (47) stipulates this restriction explicitly, while (46) requires some additional device.

Groenendijk and Stokhof, as well as Higginbotham (1991) offer an independent explanation in pragmatic terms: a question that asks you to remain silent is not felicitous:

(55) Who did no dog bite?
 * 'For no dog, tell me who it bit (=don't tell me anything)'

This explanation, however, does not extend to complement cases like (44f). It would make perfect pragmatic sense for ***Mary found out who no dog bit*** to mean that Mary did not find out about any dog who it bit; nevertheless, speakers do not accept this reading. Likewise, the pragmatic explanation, being question-specific, does not account for Moltmann's (1992) and Schein's (1993) observation that parallel readings are absent from other *wh*-constructions:

(56) a. John is taller than [how tall] no other student is.
 * 'John isn't taller than any other student'

 b. John read what no student wrote.
 * 'John didn't read any student's writing'

Moltmann (1992) proposes that the reason is that decreasing quantifiers do not take inverse scope. Matters may not be that simple, though. As we have seen, only the type *no N* is entirely unable to support a pair-list reading, while the range of quantifiers that practically do not take inverse scope is much larger (see 3.2.3 and Szabolcsi 1996). As of date, I am not aware of an enlightening syntactic or semantic explanation for the exceptional behavior of *no N*.

3.2.3 "Layered quantifiers"

The "apparent scope out" phenomenon bears a great burden in ruling out (45), the domain restriction schema (amended by an ad hoc maximality condition). Now, the use of quantification into a lifted interrogative yields results that are logically equivalent to quantification into a superordinate clause (see Hendriks 1993 for a general theory that bears this out). Thus it is worth making an excursus and show that the proposed analysis, called the "layered quantifier" analysis in Moltmann and Szabolcsi (1994), is empirically justified. Below I will review two types of supporting evidence. First, consider (57):

(57) More than one librarian found out which book every boy stole from her.

Here the complement contains a pronoun to be bound by the matrix subject. The matrix subject is chosen so that it can exhibit variation and can bind a singular pronoun, but not coreler with it, cf.:

(58) a. Some librarian lost *her* hat. *She* was sad.

b. More than one librarian lost *her* hat. **She* was sad.

Let us examine the "librarians vary with boys" reading and ask whether *more than one librarian* can bind *her* on that reading. The derivation in (50) would predict that it can, since only *every boy* is quantified into the matrix: the rest of the complement, including *her,* is within the scope of *more than one librarian*:

(59) [every boy$_2$ [more than one librarian$_1$ found out which book t_2 stole from her$_1$]]

On the other hand, (51) predicts that binding is not possible, because the whole complement is quantified in and is thus outside the scope of *more than one librarian:*

(60) * [which book every boy stole from her$_1$]$_3$[more than one librarian$_1$ found out t_3]

Speakers judge that the critical reading is in fact unavailable, i.e., (60) is the correct representation.

The second type of evidence has to do with some restrictions on when the apparent scope-out reading is available. Consider, for instance, (61). It does have a pair-list reading ' ... found out about more/fewer than six boys which book they needed,' but we have a fixed librarian: librarians cannot vary with boys.

(61) Some librarian or other found out which book more/fewer than six boys needed.

The analysis in (50) would require a new stipulation to the effect that *every boy,* but not *more/fewer than six boys,* can scope out of its clause. In contrast, Moltmann and Szabolcsi (1994) correlate the differential interpretations with the fact that *every boy,* but not *more/fewer than six boys,* is a good inverse scope taker in itself, and show that the analysis in (51) automatically predicts that the generalized quantifier representing the pair-list reading inherits its scopal abilities from its internal wide scope quantifier.

Since we are dealing with a property of all "layered quantifiers" that has some interest of its own, let us examine the general case first. A "layered quantifier" is any generalized quantifier that has another one quantified into it. For instance, in noun phrases this other quantifier may be a genitive or prepositional phrase. Notice now that the examples in (62) can be paraphrased so that the determiner of the internal wide scope quantifier becomes the determiner of the whole layered quantifier (and an existential appears):

(62) a. every girl's fingerprint = every fingerprint that belongs to some girl

b. more/fewer than three girls' fingerprints = more/fewer than three fingerprints that each belong to some girl

Why is this interesting? Most semantic properties of a noun phrase can be predicted from what its determiner is. Thus when equivalences like in (62) obtain, it is likely that the whole quantifier's behavior will match that of its internal wide scope quantifier. Scope behavior is one relevant semantic property, and witness:

(63) a. Someone saw every girl.
OK 'every girl > someone'

b. Someone saw more/fewer than three girls.
?* 'more/fewer than three girls > someone'

(64) a. Someone saw every girl's fingerprint.
OK 'every girl's fingerprint > someone'

b. Someone saw more/fewer than three girls' fingerprints.
?* 'more/fewer than three girls' fingerprints > someone'

In what cases does the above equivalence obtain? Makoto Kanazawa (p.c.) drew our attention to the following simple rule:

(65) The following equivalence, in which D is the internal wide scope quantifier's determiner,

$$\lambda P[Dx[R(x)][P(fx)]] = \lambda P[Dy\exists x[R(x) \ \& \ (y = fx)][P(y)]]$$

holds for any D when f is a one-to-one function. It holds even without f being one-to-one iff D is $\exists$, $\forall$, or their negations, or D is simply decreasing in its VP-argument.

It is worth emphasizing that the lefthand side of the equivalence is any faithful interpretation of the noun phrase, not necessarily its "standard logical form." Consider, for instance, *every girl's fingerprint.* All we are interested in now is that its meaning can be faithfully expressed as (66), where the *fingerprint of* relation is one-to-one; we are not asking whether exactly (66) should be the format in which the grammar produces its logical form:

(66) $\lambda P\forall x[\mathrm{girl}(x), P(\iota y[\text{fingerprint-of}(x)(y)])]$

Note also that f need not map individuals to individuals, it may operate on sets/groups. Thus, for instance, *fewer than six girls' books* is not problematic, because we can construct a one-to-one function that maps each girl to the set of all her books:

(67) a. fewer than six girls' books $\neq$ fewer than six books that belong to some girl

b. fewer than six girls' books = fewer than six maximal book-sets that each belong to some girl

On the other hand, even if by definition every girl has a unique favorite movie, whence *favorite movie of* is a function, many girls may share a favorite, whence this function is not one-to-one. It is easy to check, however, that with the determiners in (68) the equivalence still holds:

(68) a. every girl's favorite movie = every movie that is some girl's favorite

b. a girl's favorite movie = a movie that is some girl's favorite

c. no/not every girl's favorite movie = no/not every movie that is some girl's favorite

d. fewer than three girls' favorite movies = fewer than three movies that are each some girl's favorite

When does the equivalence fail? One type is where the function is not one-to-one and D is a non-decreasing numerical determiner. Observe that in (69a,b) there is no guarantee that there are at least three distinct movies that are each some girl's favorite; it may be that every girl's favorite is either "Aladdin" or "Jurassic Park." Another type is where there is no function at all, as in (69c): the *a-poem-by* relation is not a function.

(69) a. three girls' favorite movies $\neq$ three movies that are each some girl's favorite

b. exactly three girls' favorite movies $\neq$ exactly three movies that are each some girl's favorite

c. a poem by every poet $\neq$ every poem that is by a poet

In fact, examples in which the "head noun" of the layered quantifier has its own overt determiner typically pattern with (69c) in failing to exhibit the interesting equivalence.

Having considered the general case, let us return to pair-list readings. Recall that we are interested in deriving the fact that the complement interrogative on its pair-list reading inherits its semantic properties from its internal wide scope quantifier. Consider:

(70) (I found out) which book every boy/more than six boys needed.

Here we always have a one-to-one function from boys to questions: for each boy x, we have a unique question of the form *which book x needed.* Therefore, pair-list readings exhibit the equivalence in (65):

(71) which book D boy(s) needed =
D question(s) such that for some boy x, the question is which book x needed

Consequently, D indeed determines the scopal abilities of the pair-list quantifier. *Which book every boy needed* is predicted to be able to make the matrix subject referentially dependent, *which book more than six boys needed* is predicted not to.

4 EMPIRICAL OBJECTIONS TO THE QUANTIFICATIONAL APPROACH

Observe that the output of my analysis of quantifiers in complement pair-list readings is (semantically) the same as that of Karttunen (1977). The difference is that while Karttunen quantifies directly into a superordinate clause, I quantify into a lifted interrogative. We have seen that in isolation, these two are logically equivalent, although the present choice turns out to be preferable when more complex data are considered.

Recall now that Groenendijk and Stokhof as well as Chierchia do not merely propose another, domain restriction analysis; they also argue explicitly against quantification.[9] The present section briefly comments on specific empirical issues that arise in connection with the de dicto reading of the quantifier's restriction (4.1), quantificational variability (4.2), and the absence of pair-list readings with *whether*-questions (4.3). I wish to thank U. Lahiri and F. Moltmann for discussions on these matters.

4.1 The "de dicto" reading of the restrictor

One important reason why Groenendijk and Stokhof object to Karttunen's (1977) treatment of pair-list readings in terms of quantification into interrogatives is that this does not account for the fact that the common noun part of the QP is interpreted "de dicto." Consider (72). Karttunen's analysis says that, for every individual who is a criminal, John knows what candy that individual craves—but John himself need not know that the individual is a criminal. The restrictor *criminal* is outside the scope of *know,* i.e., it is read "de re." Groenendijk and Stokhof claim that this is not sufficient for the truth of (72): John

[9]Karttunen and Peters (1980) also propose a pair-list analysis different from Karttunen's (1977), which however has ad hoc features and has not been pursued further.

himself must also know that those individuals are criminals, i.e., the restrictor must occur inside the scope of *know* and be read "de dicto."[10]

(72) John knows what candy every criminal craves.

This objection carries over to my (46) and (47) in the following sense. If the complement clause is interpreted as an extensional object of *know, know* is part of what replaces the variable P. Thus reference to the common noun or witness set of QP is made only outside the scope of *know.* It is of course also possible to interpret the whole generalized quantifier that stands for the complement as an intensional object, in which case the problem does not arise.[11]

Now, it appears to me that Groenendijk and Stokhof's own stronger claim is in fact too strong, in two respects. First, compare (72) with (73):

(73) John has just discovered what candy every criminal craves.

This sentence need not mean that John has just discovered that the guys are criminals, although it may be natural to require that he be independently aware of them being criminals. That is, it seems that we are dealing with presupposed awareness and not with an entailment expressible strictly in terms of whatever the matrix verb happens to be (here: *discover*). The fact that Groenendijk and Stokhof consistently use *know* in their examples masks this difference.

Second, even the presupposition of awareness is restricted to cases where the matrix subject is an intelligent being acting knowingly. In (74), the experiment will neither reveal that the guys are criminals, nor does it have any awareness of this.

(74) This experiment will reveal what candy every criminal craves.

The same holds of John in (75), in case he informs us inadvertently, in an indirect way:

(75) If we trick him into rambling about his customers, John will tell us what candy every criminal craves.

Third, it seems that on the "varying librarians" reading (which I argued involves quantifying the whole complement, not merely its QP, into the matrix clause) librarians need not be aware that the person whose book needs they found out about is a student:

[10] More precisely, in addition to the domain restriction derivation, Groenendijk and Stokhof allow for quantifying into the matrix, too. Naturally, the "de dicto" claim does not apply to this latter case. This coexistence of two alternative derivations does not make the empirical predictions easy to check.

[11] My understanding is that the verb *know* takes an intensional complement in a different sense than *wonder* does. The argument of *know* is the intension of a lifted interrogative; the complement of *wonder* is that of an unlifted one. I assume that the complement of *know,* like that of *seek,* may be either extensional or intensional in the pertinent sense.

(76) Some librarian or other found out which book every student needed.

All in all, it appears that the data do not compel us to adopt Groenendijk and Stokhof's specific formulation. It is not my aim in this paper to develop an alternative proposal. Let me assume that some theory of presuppositions and intensionality is able to handle the facts that are undoubtedly there. (The datum in (76) may indeed suggest that the phenomenon Groenendijk and Stokhof observe is contingent on the whole complement being interpreted as an intensional object. Intensional interpretation is excluded when the complement is quantified into the matrix to make the subject referentially dependent.)

4.2 Quantificational variability

Another objection may be derived from a point made in Chierchia (1993). Chierchia mentions that one important advantage of his treatment of pair-list readings, which is in many respects like Groenendijk and Stokhof's, is that Lahiri's (1991) proposal for the treatment of the "quantificational variability effect" straightforwardly extends to it. To recap, the QVE is the phenomenon that, in the presence of a quantificational adverb like *usually* or *for the most part, which students* may wind up meaning 'most students.' The pioneering analysis of these data is Berman's (1990), who appeals to unselective binding. Lahiri's alternative does not involve unselective binding but reproduces the same intuitive result. He interprets (77) roughly as follows:

(77) Mary knows, for the most part, which students came.
'Mary knows most parts of the complete answer to the question which students came = For most students, Mary knows whether they came'

Chierchia (1993, p. 218) comments on the extension of this analysis to pair-list readings, "In a situation with three people a, b, and c, where a loves b, b loves c, and c loves a, if Mary knows that a loves b and b loves c, sentence [78] would be true."

(78) Mary knows, for the most part, who everyone loves.

He notes that the QVE obtains only with universals and not with indefinites, e.g.:

(79) Mary knows, for the most part, who six students love.

The absence of a QVE is predicted on the domain restriction analysis. The complement interrogative in (79) has no unique complete answer, so Lahiri's algorithm—correctly—cannot apply.

How can the QVE data be possibly accounted for if the pair-list reading is derived using quantification? Although the problem initially looks staggering, Groenendijk and Stokhof (1993) offer a trick that does the job.[12] In this paper, the authors propose a general account of the QVE that relies crucially on both fundamental assumptions and indepedently motivated particular techniques of dynamic semantics. I review the pertinent aspects of their proposal without trying to justify the underlying theory here.

In standard first order logic, the equivalence in (80) holds only if x is not free in ψ:

(80) $\forall x[\phi \to \psi] = \exists x\phi \to \psi$

It is a defining property of dynamic semantics that the equivalence holds without such a restriction. Thus we can trade the original universal of the sentence for an existential. $\exists x\phi \to \psi$ can then be subjected to existential disclosure, which removes the existential quantifier and makes x available for further quantification. Thus *most* can effectively quantify over the variable originally bound by the universal. So, (78) is interpreted as (81):

(81) 'For most persons, Mary knows (completely) who that person loves'

Fortunately, these equivalences do not hold if we replace *every* with an indefinite.

With the main job thus done, let us ask whether this result is exactly the same as Lahiri's and Chierchia's. This question is not easy to answer because they do not spell out what count as parts of a pair-list answer, but it seems they would quantify over pairs, as in (82), not over loving persons, as (81) does:

(82) 'For most person$_1$/person$_2$ pairs, Mary knows whether p_1 loves p_2'

In the model that Chierchia considers for (78) *love* is a one-to-one function, so the two readings cannot be distinguished; but this need not be so. Consider two models that make a distinction. *R*'s are lovers and *d*'s are loved ones. Bold face indicates that Mary knows that the relevant r loves that particular d:

(83) a.
$$\begin{array}{llll} r_1 & d_1 & \mathbf{d_2} & \mathbf{d_3} \\ r_2 & d_4 & \mathbf{d_5} & \mathbf{d_6} \\ r_3 & d_7 & \mathbf{d_8} & \mathbf{d_9} \end{array}$$

b.
$$\begin{array}{ll} r_1 & d_1 \\ r_2 & d_2 \\ r_3 & \mathbf{d_3}, \ldots, \mathbf{d_{500}} \end{array}$$

In (83a), Mary knows most of the pairs but her knowledge of each individual lover is partial. (83b) is the by now classical test case in which one lover is a member of overwhelmingly many pairs, and while Mary does not have any knowledge about any majority of the lovers, she does about this person. My judgment is that (78) is false in both models, thus in fact (81) is correct.

This means that (if the assumptions of dynamic semantics are generally tenable) the quantificational approach to pair-list readings can be married with a fully satisfactory treatment of quantificational variability.

[12] I. Heim (p.c.) points out that Groenendijk and Stokhof's proposal is preliminary in that it does not spell out a compositional treatment.

4.3 Complements with *whether*

Finally, as is noted already in Karttunen and Peters (1980), quantifying into interrogatives incorrectly predicts that interrogatives with *whether* have pair-list readings:

(84) John found out whether everyone left.
* 'John found out about everyone whether he left'

It is argued in Moltmann and Szabolcsi (1994) that this is really part of a bigger problem of why quantification into clauses lacking a variable binding operator is not attested:

(85) a. whether every girl walks * $\lambda P \forall x[\text{girl}(x) \rightarrow P(\text{whether } x \text{ walks})]$
b. that every girl walks * $\lambda P \forall x[\text{girl}(x) \rightarrow P(\text{that } x \text{ walks})]$

How do we know that (85b) is not available? If it were, then, assuming that complement clauses can be quantified into the matrix, as is suggested in Section 3, quantifiers in the complement would systematically appear to scope over quantifiers in the matrix. But this is not the case, cf. (48). Moltmann and Szabolcsi offer preliminary speculations, but this particular problem remains largely open for the time being.

5 "QUANTIFICATION:" A DIVERSE PHENOMENON

In sum, I have defended the view that the variety of quantifiers that support pair-list readings in extensional complements necessitates a treatment that can be called "quantificational" in truth-conditional terms. Now, various papers in this volume argue that bare indefinites, universals and modified numerals contribute differently to the interpretation of the sentence (where the differences may be representational/procedural, rather than truth-conditional). In the light of this, the claim concerning quantification cannot mean that in extensional complements, all types of noun phrases are simply "quantified in" in the sense of Montague, for instance. Rather, "quantification" needs to be read as a cover term. The intended interpretation is that each type of noun phrase induces a pair-list reading in the same syntactico-semantic fashion that is characteristic of it in other scopal contexts. This is what contrasts with the claim that the uniform contribution of QPs to pair-list readings is in terms of domain restriction.

According to the typology in Szabolcsi (1996), quantifiers fall into two main categories. Universals and bare numeral indefinites are argued to introduce discourse referents. In the case of universals, the referent is the unique witness of the quantifier; in the case of indefinites, it is a plural individual whose atoms are the elements of a minimal witness. In both cases, the referent is associated with a separate distributive operator. Hence, the interpretation of (86) will be roughly as in (87). (87) is like (47), simplified by removing the "non-empty witness" qualification and the biconditional that ensures maximality. These simplications are possible, because the quantifiers at issue are all monotonically increasing.

(86) ... who every dog/two dogs bit

(87) $\lambda P \exists A[\text{witness}(A, [\![\textit{every/two dog(s)}]\!]) \ \& \ \forall x[x \in A \rightarrow P(\text{whom } x \text{ bit})]]$

The other category of quantifiers comprises modified numerals and other decreasing items; these are argued to perform a counting operation on a predicate denotation, in the manner of generalized quantifiers. Hence, (88) can be represented straightforwardly in the manner of (46):

(88) ... who more/fewer than six dogs bit

(89) $\lambda P \text{more/fewer-than-six}\, x[\text{dog}(x), P(\text{who } y[x \ \text{ bit } \ y])]$

As was noted in 3.2.2, this latter formula presupposes an independent account of why the type of *no dog* cannot appear here.

The claim that whereas universals in matrix questions and intensional complements behave in an unusual way that can be assimilated to multiple interrogation, the various quantifiers that support pair-list in extensional complements do so in their own usual manner, is corroborated, quite spectacularly, by the syntactic analysis in Beghelli (1996). Since those facts are quite complex, I do not attempt to summarize them here; the reader is referred to Beghelli's work in the next chapter.

6 CONSEQUENCES FOR WEAK ISLANDS

Finally, let me explore the consequences of the above observations for the phenomenon that originally prompted me to investigate pair-list readings: weak islands. Szabolcsi and Zwarts (1993) propose that weak island violations are in fact a scope phenomenon:

(90) Weak island violations come about when an extracted phrase should take scope over some intervener but is unable to. Harmless interveners are harmless only in that they can give rise to at least one reading of the sentence that presents no scopal conflict of the above sort: they can "get out of the way."

Consider the following contrast:

(91) a. How much milk did every kid drink?

b. * How much milk did fewer than five kids drink?

The claim is that neither example has a reading on which *how much milk* is scoping over the subject quantifier (the reason why this is so is discussed in detail in that paper). For (91a), suppose that Billy drank a pint of milk, Johnny drank a quart, and Pete drank a tiny cup. On the plain WH $>$ $\forall$ reading, (91a) should be answered as "A tiny cup," i.e. the smallest amount that a kid drank. But this is not a good answer. The reason why (91a) is nevertheless grammatical is that *every kid* can "get out of the way" by supporting two other readings. One is where we presuppose that every kid drank the same amount of milk and want to identify this amount. One might say that *every kid* is scopeless, or scope independent of WH, on this reading. The other good reading is the pair-list reading, which may be dubbed the $\forall$ $>$ WH reading. In contrast to (91a), (91b) is ungrammatical because *fewer than five kids* can only take narrow scope; it doesn't have a single chance to "get out of the way."

We focus on pair-list readings now. Szabolcsi and Zwarts (1993, section 4) did not present novel observations but merely stated, with reference to then-current literature, that indefinites and universals, in distinction to decreasing quantifiers, are expected not to create weak islands, because they can support choice readings and fixed domain readings, respectively.

The present paper has made novel claims concerning the distribution of pair-list readings. Let us see what the consequences are for weak islands.

The most important descriptive claim made above is that different quantifiers support pair-list readings in the matrix or intensional complements and in extensional complements (universals versus almost all quantifiers). This predicts that a much wider range of quantifiers creates weak islands in the first type of context (providing, of course, that supporting a pair-list reading is the only option for the quantifiers in question to "get out of the way").

Examples with decreasing quantifiers bear this prediction out quite spectacularly. They create a weak island in the matrix and in intensional complements, but not in extensional complements:

(92) a. * How did fewer than five kids behave?

b. * I wonder how fewer than five kids behaved.

c. (He didn't do well in his survey.) He only found out how fewer than five kids behaved.

Modified numerals also present the same kind of contrast, although some speakers feel that the matrix examples are not entirely out, either:

(93) a. ?/?? How did more than five kids behave?
?/?? How did between ten and twenty kids behave?

b. ?/?? I wonder how more than five kids behaved.
?/?? I wonder how between ten and twenty kids behaved.

c. I found out how more than five kids behaved.
I found out how between ten and twenty kids behaved.

How can (93a, b) be relatively acceptable? It seems to me that they are acceptable to the extent these sentences presuppose that more than five / between ten and twenty kids behaved uniformly, and ask to identify this uniform behavior. The extensional complement examples on the other hand are impeccable and do not need such a presupposition: they have a pair-list reading. (For some reason, decreasing quantifiers do not lend themselves to a uniformity presupposition.)[13]

Definites and bare numeral indefinites have been claimed not to induce pair-list readings. Nevertheless, matrix questions/intensional complements involving these are also acceptable:

(94) a. How did the boys behave?
How did three boys behave?

b. I wonder how the boys behaved.
I wonder how three boys behaved.

c. I found out how the boys behaved.
I found out how three boys behaved.

Here we have a variety of salvaging options. Definites, and possibly indefinites, can support distributed group readings that are superficially quite similar to pair-list (see note 6). Furthermore, both *the boys* and *three boys* can denote groups and, as Doetjes and Honcoop (1996) point out, plural individuals being scopeless, they are as innocuous as proper names.

[13]Szabolcsi and Zwarts report that many speakers find even *at most five people* an acceptable intervener, as opposed to *few people,* for instance (their (35c)). They refer to Groenendijk and Stokhof's claim that these quantifiers may support an increasing group reading. Although at present I do not know which of the normally decreasing quantifiers have such an alter ego, Groenendijk and Stokhof's specific claim indeed seems to be confirmed. E.g. S. Spellmire points out to me the following contrast:

i. At most/fewer than five men ever went to the beach.

ii. At most/fewer than five men each went to the beach.

iii. * At most/fewer than five men each ever went to the beach.

I should add, though, that the reason why the group version eludes the weak island effect is presumably that it supports a uniformity presupposition and not, as was conjectured in Szabolcsi and Zwarts, that it supports a choice reading.

In sum, it appears that the current account of pair-list readings, in conjunction with the scope account of weak islands, correctly predicts a complex set of data that no other proposal in the literature does.

The observation that matrix choice questions do not exist necessitates some revision of Szabolcsi and Zwarts' preliminary account of examples involving indefinites; given, however, that these items have other options to "get out of the way," the general coverage of the account is not diminished.

Further subtle predictions come from considering the syntactic positions that quantifiers need to occupy to support pair-list readings. This topic is discussed in detail by Beghelli (1996) in the next chapter.

REFERENCES

Barker, Chris, and David Dowty (Eds.). 1992. *Antecedent Contained Deletion in a Variable Free Semantics*, Columbus. The Ohio State University.

Beghelli, Filippo. 1996. The syntax of distributivity and pair-list questions. In this volume.

Beghelli, Filippo, Dorit Ben-Shalom, and Anna Szabolcsi. 1996. Variation, distributivity, and the illusion of branching. In this volume.

Beghelli, Filippo, and Tim Stowell. 1994. The direction of quantifier movement. *GLOW Newsletter* 32:56–57.

Berman, Stephan. 1990. Towards the semantics of open sentences: Wh-phrases and indefinites. In Martin Stokhof and Leen Torenvliet (Eds.), *Proceedings of the 7th Amsterdam Colloquium*, 53–79. ILLC/Department of Philosophy, University of Amsterdam.

Chierchia, Gennaro. 1992. Functional WH and weak crossover. In Dawn Bates (Ed.), *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 75–90, Stanford. CSLI.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1(2):181–234.

Doetjes, Jenny. 1993. Wide and independent scope in Wh-questions. Columbus, Unpublished manuscript.

Doetjes, Jenny, and Martin Honcoop. 1996. On event-relatedness, symmetry and weak islands. In this volume.

Engdahl, Elisabet. 1985. *Constituent Questions.* Dordrecht: Reidel.

Farkas, Donka. 1996. Evaluation indices and scope. In this volume.

Groenendijk, Jeroen, and Martin Stokhof. 1982. Semantic analysis of wh-complements. *Linguistics and Philosophy* 5(2):175–233.

Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers.* PhD thesis, University of Amsterdam.

Groenendijk, Jeroen, and Martin Stokhof. 1989. Type-shifting rules and the semantics of interrogatives. In Gennaro Chierchia et al. (Eds.), *Properties, Types and Meaning, Volume 2: Semantic Issues*, 21–68. Dordrecht: Kluwer.

Groenendijk, Jeroen, and Martin Stokhof. 1993. Interrogatives and adverbs of quantification. In Katalin Bimbó and András Máté (Eds.), *Proceedings of the 4th Symposium on Logic and Language*, 1–29. Budapest: Akadémiai Kiadó.

Gutiérrez Rexach, Javier. 1996. Questions and generalized quantifiers. In this volume.

Haïk, Isabelle. 1984. Indirect binding. *Linguistic Inquiry* 15(2):185–223.

Hendriks, Herman. 1993. *Studied Flexibility.* PhD thesis, University of Amsterdam.

Higginbotham, James. 1991. Interrogatives I. In Lisa Cheng and Hamida Demirdash (Eds.), *More Papers on WH-Movement (MIT WPL 15)*, 47–76, Cambridge. MIT.

Higginbotham, James. 1993. Interrogatives. In Ken Hale and Samuel J. Keyser (Eds.), *The View from Building 20*, 195–228. Cambridge: MIT Press.

Higginbotham, James, and Robert May. 1981. Questions, quantifiers and crossing. *Linguistic Review* 1(1):41–80.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Karttunen, Lauri. 1977. The syntax and semantics of questions. *Linguistics and Philosophy* 1(1):3–44.

Karttunen, Lauri, and Stanley Peters. 1980. Interrogative quantifiers. In Christian Rohrer (Ed.), *Time, Tense and Quantifiers: Proceedings of the Stuttgart Conference on the Logic of Tense and Quantification*, 181–206. Tübingen: Niemeyer.

Krifka, Manfred. 1991. Definite Noun-Phrases are not quantifiers. *Linguistic Inquiry* 23(2):156–163.

Lahiri, Utpal. 1991. *Embedded Interrogatives and Predicates that Embed Them.* PhD thesis, MIT.

May, Robert. 1985. *Logical Form: Its Structure and Derivation.* Cambridge: MIT Press.

Moltmann, Friederike. 1992. *Coordination and Comparatives.* PhD thesis, MIT.

Moltmann, Friederike, and Anna Szabolcsi. 1994. Scope interactions with pair-list quantifiers. In Mercé Gonzâlez (Ed.), *NELS 24: Proceedings of the North East Linguistic Society, Volume Two*, 381–395. GLSA Publications, University of Massachussetts, Amherst.

Montague, Richard. 1974. The proper treatment of quantification in ordinary english. In Richmond H. Thomason (Ed.), *Formal Philosophy: Selected Papers of Richard Montague*, 247–270. New Haven: Yale University Press.

Munsat, Stanley. 1986. Wh-complementizers. *Linguistics and Philosophy* 9(2):191–217.

Scha, Remko. 1981. Distributive, collective, and cumulative quantification. In Jeroen Groenendijk et al. (Eds.), *Formal Methods in the Study of Language*, 483–512. Amsterdam: Mathematisch Centrum.

Schein, Barry. 1993. *Plurals and Events.* Cambridge: MIT Press.

Sher, Gila. 1991. *The Bounds of Logic. A Generalized Viewpoint.* Cambridge: MIT Press.

Srivastav, Veneeta. 1992. Two types of universal terms in questions. In *NELS 22: Proceedings of the North East Linguistic Society*, 443–457. GLSA Publications, University of Massachussetts, Amherst.

St. John, Julie. 1993. Quantifiers and Wh-phrases in matrix and embedded clauses. Columbus, Unpublished manuscript.

Szabolcsi, Anna. 1983. Focussing properties, or the trap of first order. *Theoretical Linguistics* 10(2/3):125–145.

Szabolcsi, Anna. 1996. Strategies for scope taking. In this volume.

Szabolcsi, Anna, and Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1(3):235–284. Reprinted in this volume.

10

THE SYNTAX OF DISTRIBUTIVITY AND PAIR-LIST READINGS*

Filippo Beghelli

Department of Linguistics
UCLA

1 INTRODUCTION

The phenomenon of pair-list (PL), or family of questions, readings is exemplified by (1).

(1) What did every student read?
'for every student x, what did x read?'

It is well-known that such readings obtain only under restricted syntactic circumstances. Compare (1) with (2), for example; similar contrasts have been observed among VP-internal arguments:

(2) Who read every book?
* 'for every book x, who read x?'

All syntactic theories of PL attempt to explain these contrasts. May (1985) attributes them to the Empty Category Principle, implemented using Pesetsky's Path Containment Condition. Aoun and Li (1993) attribute them to the Minimal Binding Requirement. Chierchia (1993) subsumes them under Weak Cross-Over. Hornstein (1995) develops Chierchia's proposal.

In this paper I observe that the availability of PL exhibits a much more intricate pattern than is predicted by any of these theories. Simply, the distribution of PL does not quite match either the ECP or the MBR or the WCO

*Much of the research presented in this paper results, directly or indirectly, from discussions with Tim Stowell and Anna Szabolcsi. Many thanks to both for generously sharing their insights with me. I am grateful to Susan Spellmire for insights and discussion of much crucial data. The following people in the Linguistics Department at UCLA have provided judgements and shared ideas: Paola Crisma, Katherine Crosswhite, Ed Garrett, Chai-Shune Hsu, Ed Keenan, Robert Kirchner, Rachel Lagunoff, Felicia Lee, Bill Leonard, Pino Longobardi, Matt Pearson, Brian Potter, Susan Spellmire, Dominique Sportiche, Ed Stabler, Richard Wright. Thanks to all of them. This work was supported in part by NSF grant #SBR 9222501.

A. Szabolcsi (ed.), Ways of Scope Taking, 349–408.

pattern. I will argue that, instead, it matches the patterns of distributivity that are laid out in Beghelli (1995), Beghelli and Stowell (1994), and present a theory that explains why that should be so.

To whet the reader's appetite, let me present here a small sample of the data that are not predicted by the standard theories. While binding and WCO order direct objects (DO) above indirect objects (IO), PL is more readily available when the universal is an IO:

(3) a. What did you show to every man? [? PL]
b. To whom did you show every picture? [?? PL]

But this may be thought of as a relatively minor point. More importantly, the contrasts in (1)–(2)–(3) disappear in certain *wh*-complements:

(4) a. I know what every student read. [ok PL]
b. I know who read every book. [ok PL]
c. I know what you showed to every man. [ok PL]
d. I know who you showed every picture to. [ok PL]

Further, when *what* and *who* are replaced by a *which*-phrase in the matrix, only subject universals support a solid PL:

(5) a. Which book did every student read? [ok PL]
b. Which student read every book? [* PL]
c. Which book did you show to every man? [* PL]
d. To which man did show every picture? [* PL]

But if *every* is replaced by *each,* all the examples in (5) become acceptable, as has been observed by Williams (1986), e.g.,

(6) Who read each book? [ok PL]

Finally, as was observed in Szabolcsi (1996a), the QPs that support PL in the matrix and in *wonder*-complements are not the same as those that support PL in *find out*-complements. For example:

(7) a. What did more than five boys read? [* PL]
b. I found out what more than five boys read. [ok PL]

(8) a. What did fewer than five boys read? [* PL]
b. I only found out what fewer than five boys read. [? PL]

The view that I take in this paper is that the inadequacies of current approaches to PL to deal with the above diversity stem from two sorts of standard assumptions. The first (and more general) is the assumption that "all quantifiers are created equal," i.e. that they all participate in the same scope assignment mechanism. Given that standard theories have taken quantifiers like *every* and *some* as basic representatives of the species, this has led to the expectation that scope interactions are symmetric. The second, and related, assumption is that distributivity is a free component of scope: i.e. that all QPs support distributive readings in the same manner.

I will argue below that both of these hypotheses are simplistic, and should be abandoned in favor of a more differentiated account of scopal interactions. This in turn will allow us to capture the empirical generalizations that are overlooked by current approaches.

The structure of this paper is as follows: Section 2 presents a brief overview of current approaches to the syntactic (LF) derivation of PL. Section 3 introduces some basic distinctions between different types of QPs, and presents data to show that these types use distinct mechanism of scope assignment. This section includes, in particular, a presentation of two distinct distributivity patterns that are associated with two types of QPs. In Section 4 I outline a theory of scope in LF in which scopal diversity can be expressed. The following Section, 5, focuses on the quantifiers *every* and *each,* studying their behavior when they are in the scope of negation; the purpose of this section, which is based in part on Beghelli and Stowell (1996), is to provide independent evidence for the claims to be made Section 6, which is devoted to an analysis of PL in matrix questions with the quantifiers *every* and *each.* Section 7 presents an account of the interactions between *every/each* and *wh*-elements in embedded questions, focusing on the cases where *wh* does not carry interrogative force. Finally, Section 8 extends the analysis to types of QPs other than *every/each.*

2 APPROACHES TO PL

In this section I present an overview of current LF approaches to PL. I limit the presentation to just a few proposals: May (1985), Aoun and Li (1993), Chierchia (1993), and (briefly) Hornstein (1995). One of the focuses of this overview will be the issue of "diversity" presented in the Introduction. It is not my purpose to offer a complete overview of this research field, but rather to provide a context for the discussion in the rest of the paper. For a summary of the semantic approaches to PL, see Szabolcsi (1996a).

Both May's (1985) and Aoun and Li's (1993) accounts treat *wh*/QP interactions on par with QP/QP interactions, at least in terms of the representation that these are assigned in LF.

As is well known, QP/QP interactions are derived by freely applying the scope assignment rule of QR, which routinely adjoins QPs to VP, IP, or PP. Extending the account to QP/*wh* interactions requires a more complex set of principles. May derives the distribution of PL via the combined application of a general well-formedness contraint on LF, Pesetsky's (1982) Path Containment Condition (PCC), and an LF interpretive principle, his Scope Principle.

QR applies freely to the S-Structure representations of (1) and (2), as it does in sentences with non-*wh* QPs. As a result of the application of QR, non-*wh* QPs in argument positions are adjoined either to VP or to IP. After some of the resulting LFs are filtered out as ill-formed by the PCC, the Scope Principle assigns scopal readings to the surviving ones. Whenever two QPs end up, after QR, in positions which govern each other, they are free to take on any relative scope. Otherwise, scope is determined by c-command.[1]

In the case of (2), the PCC rules out the LF which is obtained by QR-ing *every book* to IP, a position from where, by May's own definition of government, the QP and *wh* would govern each other (and thus yield PL). Since in (1) no

[1]The PCC requires that if A′-categorial paths have a (non-empty) intersection, they must embed, not overlap. (A path is defined as the sequence of nodes $n_1, \ldots, n_k$ connecting the binder n_1 to its bindee n_k; in a path, every n_j immediately dominates $n_j + 1$.) Its initial motivation lies in deriving contrasts between sentences like *What do you wonder who saw?* vs. **Who do you wonder what saw?*.

Technically, the derivation proceeds as follows. After QR applies freely to the S-Structure representations of sentence (2), determining the LFs in (2a′, b′), the PCC rules out (2b′).

(2) Who read every book (on the reading list)?

a′. $[_{CP}$ who$_1$ $[_{IP}$ t_1 $[_{VP1}$ every book$_2$ $[_{VP2}$ read $t_2]]]]$?

b′. $[_{CP}$ who$_1$ $[_{IP1}$ every book$_2$ $[_{IP2}$ t_1 $[_{VP}$ read $t_2]]]]$?

This is because in (2b′) the paths overlap. The path of *who*$_1$ = ⟨IP2, IP1, C′, CP⟩; the path of *every book*$_2$ = ⟨VP, I′, IP2, IP1⟩; IP2 and IP1 are the overlapping portion. In (2a′) the paths are disjoint, thus the PCC (vacuously) allows this LF. (In (2a′), the path of *who*$_1$ = ⟨IP, C′, CP⟩, and the path of *every book*$_2$ = ⟨VP2, VP1⟩).

On the other hand, the following LF is available for (1):

(1) What did every student read?

a′. $[_{CP}$ what$_2$ did $[_{IP1}$ every student$_1$ $[_{IP2}$ t_1 $[_{VP}$ read t_2 $]]]]$?

This is a well-formed LF by the PCC since the paths embed. The path of *what*$_2$ = ⟨VP, I′, IP2, IP1, C′, CP⟩ inludes as its central portion the path of *every student*$_1$, which is ⟨IP2, IP1⟩.

Once the output of QR passes the PCC filter, the Scope Principle determines which readings are associated to each LF. The Scope Principle assigns symmetric scope to two QPs when these govern each other at LF. Since by May's definition, this condition is satisfied by *what* and *every student* in (1a′), both the scoping *what* > *every student* and *every student* > *what* are assigned to it. The latter derives the PL interpretation of (1). But the QPs *who* and *every book* do not govern each other in (2a′), and consequently only the scoping *who* > *every book* is assigned to it, corresponding to the individual answer reading.

such PCC violation takes place when the quantifier is raised to IP, PL is a possible interpretation of the sentence.

Some of the consequences of this approach are questionable. First, the adoption of the Scope Principle forces LFs to be scopally ambiguous. The same LF can be the vehicle for two (truth conditionally distinct) scope assignments. Whether this is desirable or not (and it would seem not to be), it should be pointed out that it is not necessitated by the account of non-*wh* QP/QP interactions.

Second, to extend his treatment to questions where the quantifier is embedded in a complement clause, May needs to assume that QPs built with *every* are not upwardly bound in their scope by *that*-complements.

To derive the PL reading speakers normally assign to (9), May raises the universal quantifier, by successive-cyclical application, to the matrix IP, so that mutual government with the *wh*-element obtains.

(9) a. Who do you think everyone saw at the rally?

b. $[_{CP}$ who$_2$ do $[_{IP1}$ everyone$_1$ $[_{IP1}$ you think $[_{IP2}$ x_1 $[_{VP}$ saw x_2 $]]]]$

This assumption is problematic. Although there seem to be cases where the scope of *every/each* does extend outside the boundaries of their clause, *that*-complements invariably appear to block their upward scope.

Most importantly, as observed above, May's theory is inherently unable to handle the "diversity" issues. The different ability of distinct quantifiers (*each, every, most, few,* etc.) to support PL in various syntactic configurations cannot be stated in his theory since none of the principles it invokes makes any reference to quantifier-types, with the possible exception of the focus reading of *each.*[2]

The same remark applies to Aoun and Li (1993). Their treatment of *wh*/QP interactions is based largely on the approach in May (1985). The LF representations for the relevant sentences are in fact similar. The difference between the two approaches is in the principles that are invoked to derive the correct set of LF representations.[3]

---

[2]T. Stowell (p.c.) points out to me that the assumption that QR may raise a universal out of its clause, cf. (9), in fact correctly predicts that it is possible to void any asymmetries in complement PL. May himself does not seem to be aware of either this prediction or of the data that may support it. In any case, this same device could not account for the whole array of the "diversity" issues mentioned in Section 1.

[3]Aoun and Li object to resorting to the PCC as the relevant well-formedness condition. They note that there are languages, like Chinese and Spanish, where the PCC does not seem applicable. As observed by Huang (1982), in Chinese sentences like (i.a) can receive either interpretation (i.b) or (i.c):

(i) a. He wondered who$_1$ bought what$_2$

b. 'for which x_1 did he wonder which x_2 is such that x_1 bought x_2?'

Aoun and Li also reject May's Scope Principle, pointing to empirical shortcomings in the account of Double Object constructions, and to its inability to handle cross-linguistic variation in the availability of scope construals.[4]

The PCC and May's Scope Principle are replaced by the Minimal Binding Requirement (MBR), which requires variables to be bound by the most local potential A′-binder, and by a new Scope Principle, which states that a quantifier Q1 has scope over a quantifier Q2 if Q1 c-commands a member of the chain containing Q2. Aoun and Li's proposal makes use of well-formedness principles provided by the Generalized Binding framework.[5]

---

c. 'for which x_2 did he wonder which x_1 is such that x_1 bought x_2?'

In other words, these sentences are ambiguous between subject or object *wh* being paired with the *wh* in matrix Comp. The interpretation in (iib) should be excluded (by the PCC) on a par with the ungrammatical **Who did you wonder what saw?*.

Yet questions with quantifiers show the same type of ambiguity as their English counterparts: Chinese questions corresponding to English *Who bought everything (for Zhangsan)?* are unambiguous (only individual answer), whereas questions like *What did everyone buy (for Zhangsan)?* are again ambiguous between an individual answer and a pair-list answer. This is a contradiction for the PCC. Aoun and Li point to similar inadequacies for Spanish.

[4]May's Scope Principle appears also to have empirical problems, in English as well as in Chinese. As pointed out by Larson (1990), VP-internal arguments in Double-Object constructions display a curious lack of scope ambiguity, which is found both in English and Chinese. The second object is scopally "frozen" in its position. There is no inverse scope reading *every problem* > *one student* in (i.a).

(i) a. John assigned one student every problem [unambiguous: * *every* > *one*]

b. John assigned one problem to every student [ambiguous]

This contrasts with (i.b), exemplifying a Prepositional Indirect Object construction (PIO), where scope ambiguity resurfaces.

May's scope principle would assign either scoping to the quantifiers in (i.a), since there would be well formed representations (by the PCC) where both quantifiers would govern each other. For example, we could adjoin *one student* to IP, and then also adjoin *every problem* to IP. Since no maximal projections intervene between them, the two QPs govern each other.

[5]The LF-structure given by May (1985) to represent the PL interpretation of (1) is the same structure that Aoun and Li use to derive PL. What needs to be demonstrated is that this interpretation doesn't violate the MBR.

The structure in question is given below in (i), corresponding to Aoun and Li's (1993) ex. [41], ch. 2.

(i) $[_{CP}$ what$_2$ $[_{IP}$ every student$_1$ $[_{IP}$ vbl$_1$ $[_{Agr}$ $[_{VP}$ t_2 $[_{VP}$ t_1 read vbl$_2$]]]]]]

This structure does not violate the MBR because *every student*$_1$ is not a potential A′-binder for vbl$_2$: co-indexing vbl$_2$ with *every student*$_1$ implies co-indexing vbl$_2$ to vbl$_1$ as well; but this creates a Principle C violation (vbl$_2$ ends up being bound by either t_1 or vbl$_1$), as variables have status comparable to R-expressions for the purposes of the Binding theory. The MBR, however, is formulated so that *every student*$_1$ is only a potential binder for variable x if coindexing x and *every student*$_1$ does not incur a grammatical violation. Thus *every student*$_1$ does not count as a potential binder for vbl$_2$. The MBR is accordingly not violated.

The structure in (i) is scopally ambiguous. For Aoun and Li, only A′-positions are relevant to the application of the Scope Principle; neither of the theta positions occupied by t_1 and

Aoun and Li's proposal does not differ much from May's with respect to diversity issues, but represents a refinement of that theory, both empirically and theoretically. Consider for example their Scope Principle. By stating that scope relations are computed in terms of whole chains (rather than just their heads), they effectively have the ability to reconstruct A′-moved items. Thus they do not incur the difficulty with (9), where the quantifier is contained in an embedded *that*-complement.

A genuinely alternative view of PL is given in Chierchia (1992, 1993). He analyzes PL as a subcase of the functional reading (which, in turn, can be seen as a type of individual reading, trading functions for individuals). Chierchia's insight is that in a PL answer, the function is given "extensionally" (by listing a set of pairs), rather than in its normal "intensional" form (as a procedure to compute pairs). At LF, a PL question like (1), repeated below:

(1) What did every student read?

is accordingly analyzed as 'what is the function f such that for every student x, x read $f(x)$?' The answerer proceeds to define the function extensionally, by enumerating, for each student x, what x read.

Chierchia capitalizes on this observation to provide a syntactic and semantic account of PL. He proposes that the LF representations of questions that receive either a functional or a PL interpretation differ from that of individual questions in that the *wh*-trace has complex (internal) structure. Such a functional trace contains both an empty category with a f(unction)-index, functioning as a standard *wh*-trace, and an empty element with an a(rgument)-index, which acts like a bound pronominal. The a-index corresponds to the bound pronoun in the functional answer (e.g. *What did every student read?* Her *assigned book*). Crucially, the pronominal a-index is bound by the QP, whereas the f-index is the same as that of the *wh*-element.

There are, thus, three possible LF representations for questions with quantifiers. These are given below (cf. Chierchia 1993, p. 211):

(10) a. $[_{CP}$ Who$_j$ $[_{IP}$ every Italian man$_i$ $[_{IP}$ t_i fears t_j]]]
[individual interpretation]

---

vbl$_2$ matters. The scoping *every student* > *what* is obtained by *every student* c-commanding the intermediate trace left, in VP-adjoined position, by *what*; whereas the other scoping, *what* > *every student,* follows from *what* c-commanding *every student.* Note that this effectively amounts to applying reconstruction to the *wh*-phrase.

Consider now (2), which does not support PL. The LF is given below in (ii). This LF would be ruled out by the PCC, but is not ruled out by the MBR.

(ii) $[_{CP}$ who$_1$ $[_{IP}$ vbl$_1$ $[_{VP}$ every book$_2$ $[_{VP}$ t_1 read vbl$_2$]]]

As before, given that t_1 (and also vbl$_2$) don't count, (ii) can only be read with the scoping *who* > *every book,* as desired.

b. [$_{CP}$ Who$_j$ [$_{IP}$ every Italian man$_i$ [$_{IP}$ t_i fears t_j^i]]]
[functional interpretation]

c. [$_{CP}$ [Who$_j$ every Italian man$_i$] [$_{IP}$ t_i fears t_j^i]]
[PL interpretation]

Chierchia assumes May's QR applies to the subject QP. The individual interpretation (10a) has an LF with a standard *wh*-trace. The functional (10b) and the pair-list (10c) interpretations require a functional trace. In addition, PL requires the QP to absorb with the *wh*-in-Comp. Absorption is the syntactic process that Higginbotham and May (1981) use to deal with multiple *wh*-questions (as well as other phenomena). An absorbed *wh*-QP yields an interpretation where the question operator ranges over both the domain of the quantifier and that of the *wh*-element.

This analysis makes a direct prediction regarding the distribution of PL: this reading will be excluded whenever the functional trace c-commands (in LF) the trace of the QP. If the functional trace c-commands the thematic position of the QP, we have a Weak-Crossover (WCO) violation, given that the QP is co-indexed with the argument-trace, which is a pronominal element. The configuration in (11) is structurally analogous to the classic WCO pattern in (12):

(11) a. Who read every book? [=(2), *PL]

b. [$_{CP}$ Who$_j$ [$_{IP}$ every book$_i$ [$_{IP}$ $t_j^{[+pro]_i}$ read t_i]]] [WCO]

(12) a. ?? His$_i$ mother loves everyone$_i$

b. [$_{IP}$ everyone$_i$ [$_{IP}$ his$_i$ mother loves t_i]] [WCO]

c. ?? Who$_i$ does his$_i$ boss like?

d. [$_{CP}$ Who$_i$ [$_{IP}$ his$_i$ boss like t_i]] [WCO]

Chierchia therefore captures May's basic generalization on the appearance of PL without the assumption that PL arises, in LF, by the QP taking scope over the *wh*. The problematic part of Chierchia's account is that the application of absorption is somewhat stipulative. We are left wanting to know (i) why absorption should be blocked, in matrix interrogatives, when quantifiers like *more than n, fewer than n,* etc. interact with *wh*-phrases (cf. 7, 8); (ii) why absorption goes through in complement interrogatives (cf. 4, and again 7, 8); and (iii) why—and whether—absorption should never be available in languages like Hungarian given that these lack PL entirely.

Finally, note that Chierchia's account cannot discriminate between the behavior of *every* and *each* (cf. 6), since these two quantifiers show identical WCO effects.

Hornstein (1995) adopts Chierchia's proposal in his novel approach to QP-scope. He presents an elegant theory of scope that dispenses with A′-movement in LF (and thus, with QR). Scope assignment is derived as the by-product of movement to Case/agreement positions (A-movement). Hornstein assumes a principle of Chain-Pruning, whereby only a single link in a chain is interpreted at the Conceptual-Intentional interface. Accordingly, all but one member of a multi-membered chain must be deleted. Applied to chains headed by QPs, this principle yields the same results as reconstruction. The classic scope ambiguity in (13a) is for example derived as follows (elements in parentheses have been pruned):

(13) a. some student read every book

b. $[_{\text{AgrSP}}$ some student $[_{\text{TP}}$ $[_{\text{AgrOP}}$ every book $[_{\text{VP}}$ (some student) $[_{\text{V}'}$ read (every book)]]]]] [*some* > *every*]

c. $[_{\text{AgrSP}}$ (some student) $[_{\text{TP}}$ $[_{\text{AgrOP}}$ every book $[_{\text{VP}}$ some student $[_{\text{V}'}$ read (every book)]]]]] [*every* > *some*]

Given Hornstein's rejection of A′-movement in LF, which he motivates on Minimalist grounds, his analysis of PL is especially attractive, since it does not rely on moving the QP to a position where it takes scope over the *wh*-element. Hornstein adopts Chierchia's basic insights without, however, the part of Chierchia's account that involves the use of absorption.

For Hornstein, then, individual readings correlate with deleting all but the copy of the *wh*-element in Comp; functional and PL readings are possible (subject to WCO) when the copy in [Spec, CP] is deleted. To illustrate:

(14) a. Who does everyone love?

b. $[_{\text{CP}}$ Who$_j$ $[_{\text{AgrSP}}$ everyone$_i$ $[_{\text{VP}}$ t_i love (who)]]]

c. $[_{\text{CP}}$ (Who) $[_{\text{AgrSP}}$ everyone$_i$ $[_{\text{VP}}$ t_i love who^i_j]]]

By adopting Chierchia's analysis, Hornstein can provide an account of *wh*/QP interactions that meets two basic desiderata: (i) it does not involve ECP-like principles (which could not be used in a Minimalist setting without considerable reformulation); (ii) it does not involve applying QR-like processes to the QP (in order to move it to a position where it can quantify over the *wh*-element in [Spec, CP]). Instead, it assumes something analogous to reconstruction of the *wh*-element.

However, since Hornstein assumes that absorption (which is applied with A′-movement) does not exist, it is not clear how he proposes to reformulate Chierchia's distinction between PL and functional readings. As is well known, functional readings do not display the distributional asymmetries that PL does (for example, functional readings are supported by negative quantifiers: *Who* *does no Italian married man love? His mother-in-law*; but negative quantifiers do not support pair-list. The same holds for a number of other quantifiers. As we will see in Section 8, absorption gives Chierchia a way to account for the lack of PL with quantifiers like those featured in (7): he assumes that absorption, which is necessary for PL, simply does not apply to these kinds of quantifiers. The fact that Hornstein rejects absorption while adopting the rest of Chierchia's proposal makes his account ill-equipped to handle diversity issues. Plus, Chierchia locates the potential source of cross-linguistic variation as to the existence of PL in whether the language has absorption (of the requisite sort). Hornstein's account presumably loses this cross-linguistic prediction, too.

After this review of some of the main syntactic approaches to PL, it is time to introduce the basic assumptions of the approach that will be presented in this paper.

3 PATTERNS OF DISTRIBUTIVITY

One of the basic hypotheses of this paper is that the notion of distributivity plays an important role in the phenomenon of PL. The particular implementation to be proposed will allow us to derive the contrast between the distribution of PL with *every*-QPs and *each*-QPs, as observed in (2) and (6). Our discussion of distributivity will eventually not only make sense of this contrast, but also give us some clues toward the other diversity issues mentioned in the introduction. Futher assumptions about the different behavior of QP types in declarative (including negative) contexts will provide the tools to tackle all of the tasks that we have set before us concerning the distribution of PL.

In this section, I outline the premises of a theory of distributivity that is based on Beghelli (1995), developing the approach pursued in Beghelli and Stowell (1996).

Let's begin by introducing a definition of distributivity. Given a sentence S and a QP α in it, α has a distributive interpretation relative to a reading R of S iff α is able to induce co-variation between individuals in its domain (=restrictor set) and another (overt or silent) quantificational element in S under reading R. For example, in *Three kids climbed some tree,* the QP *three kids* is distributive insofar as we can construe the sentence as being true in a model where for at least two different kids there are two different trees and events of climbing.[6]

[6]This notion of distributivity differs from the standard notion in logical semantics. According to the latter, a quantifier is distributive iff the predicate that interprets its nuclear scope applies individually to two or more members of the quantifier's domain. Essentially, our definition requires co-variation, whereas the standard notion doesn't. One area in which the two notions make different predictions is plural numerical QPs. Take a sentence like:

On the basis of this definition, the theory that I will present makes the following claims: (a) distinct QP-types do not support distributivity in the same way: the syntactic configurations in which distributivity is supported vary with QP-type; (b) distributive readings emerge via the agency of distinct mechanisms, which should be represented at the syntax-semantics interface (LF).

Before we can present the relevant generalizations, it is necessary to discuss QP types.

3.1 Types of quantifier phrases

For reference in the discussion, I introduce below an informal classification of QP-types (this typology is largely based on Szabolcsi 1996b and previous work). After commenting briefly on some of the types, I return in the next subsection to distributivity.

It is essential to the account to be developed that we distinguish, in addition to WhQPs (the familiar *wh*-words like *who, what, which man, which of the men, ...*) and negative QPs (NQPs) such as *no one, none of the men, ...* , the following three classes of QPs:

(15) **QP Types** (cf. Szabolcsi 1996b, also Beghelli 1993, 1995)

DISTRIBUTIVE-UNIVERSAL QPs (DQPs). Universal distributive with singular agreement: *every, each.*

GROUP-DENOTING QPs (GQPs). "Plain indefinites": *some, several*; "bare-numeral": *one student, two students, ...* ; definites: *the stu-*

(i) Four students lifted five chairs

Let's focus on the object wide scope readings of (i). We can distinguish three sets of interpretations that can *in principle* be associated with (i)—leaving aside for the moment which ones are *actual* readings: (a) 'there are five chairs such that each was lifted by a possibly different set of four students (where they acted individually or collectively); (b) 'there is a stack of five chairs such that it was lifted by four students (where they acted individually or collectively, but they surely lifted the whole stack on every lifting event); (c) 'there are five chairs each of which was lifted by the same set of four students (who acted individually or collectively).'

Under our definition, only (a) corresponds to a distributive use, since distinct chairs may be related to distinct sets of four students. In the standard notion, (c) is also a distributive reading of *five chairs,* since each chair was lifted by four students, but there is no co-variation between individual chairs and sets of four students. One interpretation of (c) is the branching reading, where each chair was lifted by each student; the other corresponds to a distributive′ (in the standard sense) interpretation of *five chairs* and a collective reading of *four students.* We would not call either interpretation of (c) distributive. ((b) of course would not be called a distributive interpretation of *five chairs* under either definition of distributivity.)

Based on our definition of distributivity, we will claim later on in this section that examples like (25), *Two of the students read three books,* which are similar to (i), do not support an object wide-scope distributive reading.

dents, these students, ... ; partitives: *one of the students, two of the students, ...*

COUNTING QPs (CQPs). Cardinality expressions typically built with "modified numberal" quantifiers: *few men, fewer than five men, at most six men, ...* ; *more than five men, at least six men, ...* ; *between six and nine students, more (students) than (teachers), ...*

It is well known that the QP-types in (15) do not support collective vs. distributive predication in the same way. DQPs are incompatible with collective predicates, whereas both GQPs and CQPs are compatible with either. The examples in (16), (17) and (18) bear this out: recall that *sneeze* does not support collective readings, whereas *surround* is only collective.

(16) a. ?? Every/each soldier surrounded the fort

b. Every/each soldier sneezed

(17) a. Twenty/the soldiers surrounded the fort

b. Twenty/the soldiers sneezed

(18) a. More/fewer than twenty soldiers surrounded the fort

b. More/fewer than twenty soldiers sneezed

Even though both support collective readings (although in different ways, see Szabolcsi 1996b), CQPs differ from GQPs in a number of respects: the most striking are that whereas GQPs are relatively free in their upward scope and support non-local anaphora, CQPs do neither (cf. Beghelli 1995). I will not discuss anaphora properties extensively (the reader is referred to Szabolcsi 1996b).[7]

[7] As regards anaphora, CQPs do not support coreference across clausal boundaries, witness the contrast:

(i) a. A student was happy. But he had no reason to be.

b. * More than one student was happy. But he had no reason to be.

Especially relevant to the characterization of the anaphora properties of these QPs is the data presented in Kamp and Reyle (1993). These data point to the inability of CQPs to introduce a discourse referent:

(ii) a. [Five men]$_i$ hired [a secretary that they$_i$ liked]
ok *they* = the five men as a collective

b. [More than five men]$_i$ hired [a secretary that they$_i$ liked]
* *they* = the collective of the more than five men

Kamp and Reyle observe that *they* in (ii.a) can refer to the group of five men (the men hired a secretary that they as a group liked), whereas in (ii.b), the pronoun *they* must be construed as a bound variable. Accordingly, the predicate in the relative clause must

Differences among the types listed in (15) are especially striking when we turn to scopal behavior (cf. Beghelli 1995 for extended discussion). Consider simplex SVO sentences like those in (19) where a GQP occupies the subject position. It can be seen in (19a, b) that when the object QP is built with *every/each,* both S > O (19′c) and O > S (19′a) readings are available. But when the object is built with a counting Q (*more than five, more ... than ...*, *few(er than five)*), neither the scopal construal in (19′a) nor (19′b) is normally available to speakers (when (19c, d, e) are uttered with neutral intonation). Only the S > O scoping (in 19′c) is.

(19) a. Some student/one of the students read every book.
b. Some student/one of the students read each book.
c. Some student/one of the students read more than five books
d. Some student/one of the students read more books than magazines
e. Some student/one of the students read few(er than three) books

(19′) a. 'for each one of Qx, x a book, there is a (possibly different) student who read x'
b. 'there is a set X, containing Q books, such that some student read X'
c. 'there is some student y such that for Qx, x a book, y read x'

Let's refer to construals where a QP α scopes over another QP β even though β c-commands α at Spell-Out as inverse scope. We can conclude, then, that CQPs do not support inverse scope over (this type of) a subject. In this respect, they differ markedly from DQPs.

Consideration of a further test environment allows us to refine the above generalization about the scope behavior of CQPs. When a CQP occurs as the complement of an intensional predicate, as in (20), only *de dicto* interpretations are possible. *De re* interpretations, as in (20′), obtained by scoping the CQP above the intensional predicate, are generally not available.

(20) a. John is looking for more than five unicorns (for the stables of his new villa)

---

be construed distributively. These examples provide a significant test environment because when a pronoun occurs inside a VP, it cannot use the same VP to build an antecedent for itself. Pronouns cannot corefer to a quantified set that they help to define. The data strongly support the conclusion that CQPs do not introduce discourse referents that pronouns can use; but GQPs do.

Though in English DQPs show a behavior comparable to CQPs in this respect, Szabolcsi (1996b) argues that there are empirical reasons to make a three-way distinction as outlined in (15).

b. John is looking for more unicorns than hippogryphs (for the stables of his new villa)

c. John is looking for few unicorns (for the stables of his new villa)

(20′) * 'there is a set/group X of Q unicorns, such that John is looking for X'

These data, plus other types of evidence reviewed in Beghelli (1995), support the following conclusion (cf. also Beghelli 1993):

(21) CQPs take scope *in situ*; they do not avail themselves of scope movement in LF.

3.2 Distributivity with GQPs and DQPs

Leaving aside CQPs for the time being, let's consider the scopal properties of the other two types. The observations to be made in this respect have direct bearing on our theme, distributivity. Of particular importance in this paper will be a comparison of the distributive properties of DQPs and GQPs. To begin with, consider the scope properties of GQPs. It is well known that GQPs have fairly free upward scope. For example, they can take scope over a subject DQP by scoping out of intensional contexts, freely supporting *de re* interpretations:

(22) Every officer wanted to go out with two ballerinas.
OK 'there are two ballerinas (say, Jill and Sue) such that every officer wanted to go out with them'

It is also standardly assumed that GQPs support distributive readings. Given an appropriate predicate, these readings are supposed to be in free variation with group readings.

There are data that point to the opposite conclusion, however, and indicate that GQPs do not support distributive readings in exactly the same sense as DQPs do. This evidence is provided by the distribution of QPs like *a different book* in one of their readings. The reading in question can be characterized as the "distinct share" interpretation. In (23a) this corresponds to the construal where distinct students are associated with distinct books. (The other interpretation of *a different book* in (23a) is the "anaphoric" reading, which is irrelevant to us here. This consists in construing a book as wide scope, and asserting that this particular book is distinct from a previously mentioned book).

The distribution of *a different N,* in the "distinct share" reading, shows that only DQPs are licensors; no other type of QP, including GQPs and CQPs, is able to license it.

(23) a. Every student read a different book
b. Each student read a different book
c. * Five students read a different book
d. * The students read a different book
e. * More than five students read a different book
f. * No students read a different book

The examples below further show that licensing of *a different N* by a DQP can take place via LF movement:

(24) a. A different student read every book
b. A different student read each book
c. * A different student read five books
d. * A different student read the books
e. * A different student read more than five books
f. * A different student read no books

Under the standard view that GQPs freely support distributive readings, their inability to license *a different book* is surprising, given that (i) the contribution of the modifier *different* is merely to enforce a one-to-one distributive dependency, and given that (ii) DQPs are good licensers from any position. These data suggest that distributivity with GQPs and DQPs may be the result of different underlying mechanisms, i.e., that the distributivity of GQPs is in some way "weaker" than that of DQPs.

A second type of data confirms this impression: the existence of scopal asymmetries in the availability of distributive readings with GQPs. The relevant observation can be found in Verkuyl (1988), van der Putten (1989), Ruys (1993), Abusch (1994), and Beghelli (1995). GQPs appear to function as distributors when they take scope over a QP that they c-command (at Spell-Out), as in (25a), but they do not support distributivity when they take inverse scope, as shown by the unnaturalness of the reading in (25b). Only (25c) represents an available inverse scope reading of (25) (cf. note 7 for a more detailed discussion of the readings of 25.)

(25) Two (of the) students read three books

a. OK 'for each of two students, there is a possibly different set of three books that they read'

b. * 'for each of three books, there is a possibly different set of two students who read that book'

c. OK 'there is a set Y of three books such that (each of) two students read Y.'

There is a sharp constrast between DQPs and GQPs in this respect. Unlike GQPs, the (inverse) scope of DQPs is always accompanied by the ability to distribute:

(25) d. Two (of the) students read every/each book
OK 'for every book x, there is a possibly different set of two students who read x'

The provisional conclusion that can be drawn from these data is that whereas DQPs are always (and exclusively, as seen above) distributive, GQPs do not support distributive construals with *a different N,* and display distributive readings only in certain scopal configurations.

On the basis of the data in (23)–(25), and of other scope data to be considered later on, let's assume that only DQPs are properly 'distributive,' and that distributivity is not an inherent property of GQPs, but arises via some additional mechanism available only in certain configurations. (In the account that I will propose at the end of this section, distributivity with GQPs arises via the agency of an operator that is not part of, nor is necessitated by, GQPs themselves.)

As a convenient descriptive label, I will refer to the pattern of distributivity displayed by DQPs—in (non-negative) declarative sentences (this qualification will be discussed at length below)—as STRONG DISTRIBUTIVITY (henceforth, SD).

The distributive dependency holding between a QP built with *every* or *each* and a clausemate indefinite is not constrained by their relative position, as has long been recognized. It does not only hold when the quantifier, acting as the distributor, c-commands the indefinite, which takes the role of distributee (26a), but also when the indefinite c-commands the quantifier (26b).

(26) a. Every student read some book

b. Some student read every book

The following examples give further illustrations of the availability of strong distributivity for various choices of indefinites and relative positions of distributor and distributee. They show that strong distributivity equally holds when distributor and distributee are respectively in the positions of direct object-indirect object (27a); indirect object-subject (27b); and indirect object-direct object (27c).

(27) a. John gave every book to a different student/to two of these students/to fewer than five students.

b. A different student/two of the students/more than one student introduced John to every professor

c. John showed a different book/two of the books/fewer than six books to every student

We have seen that as distributors, GQPs do not behave at all like DQPs. Let's thus refer to the emergence of distributive readings with GQPs as PSEUDO-DISTRIBUTIVITY (PD). PD patterns unlike SD in the following two respects: the availability of distributive readings depends (i) on the type and interpretation of the distributee; and (ii) on the relative argument positions of distributor and distributee. This is more thoroughly laid out in (28) and (29) below, where I state the main descriptive generalization governing the distribution of SD and PD readings. The rest of this subsection is devoted to illustrating the pattern.

It will be helpful, in the presentation of the data, to refer to subtypes of GQPs. I introduce the following distinctions. BARE GQPs are the plain type, built with determiner plus head noun (*some student, two students,* ...); PARTITIVE GQPs are those built with the addition of *of the*: *some of the students, two of the students,* etc. The distinction is relevant to interpretation: as pointed out by Diesing (1992), bare GQPs can be interpreted either as presuppositional or as cardinal; partitive GQPs are generally interpreted as presuppositional.

(28) **The pattern of Strong Distributivity (SD)**

a. *Type of the distributee.* Distributivity is supported over any type of distributee: both cardinal (=non-specific) and presuppositional indefinites.

b. *Position of the distributor.* Distributivity is supported from any argument or adjunct position: both direct and inverse distributive scope are possible.

(29) **The pattern of Pseudo-Distributivity (PD)**

a. *Type of the distributee:*

(i) PD is generally possible between two co-arguments when the distributor is presuppositional, unless the distributee is a subject.

(ii) PD is generally impossible when the distributee is presuppositional, unless the distributor is a subject.

b. *Position of the distributor.* When both distributor and distributee are bare GQPs, PD is possible when the syntactic position of distributor and distributee observes the following argument hierarchy: subject > indirect object/adjunct > direct object.

None of the restrictions listed in (29) hold for SD. The ability of DQPs to support distributive readings is affected neither by the type of the distributee nor by syntactic position. The differences between DQPs and GQPs thus go well beyond the observation that DQPs only possess a distributive reading, whereas GQPs also display group readings.

GQPs and DQPs are somewhat comparable only when the distributor is interpreted presuppositionally and the distributee is a non-subject interpreted cardinally (cf. 29a). This is illustrated by the okay readings assigned to the examples in (30) (when prefixed to readings, the mark 'ok' means that the reading is natural; '?' that it is possible; '??', '*' that it is very hard, or impossible):

(30) a. Five of the students read two books
OK 'five of the students each read a (possibly different) set of two books'
* 'there is a set of two books such that for each of them, there is a (possibly different) set of five of these students who read it'

b. John showed five of the books to two students
OK/? 'five of the books are such that for each of them, John showed it to a (possibly different) group of two students'
* 'there is a set of two students such that each was shown a (possibly different) set of five books'

c. John showed two books to five of the students
OK/? 'five of the students are such that to each of them, John showed a (possibly different) group of two books'
* 'there is a set of two books such that each was shown to a (possibly different) set of five students'

The similarity between GQPs and DQPs breaks down in most other cases. First, recall from our discussion of (25) that inverse distributive scope over a subject is generally not available: similarly, (30a) does not support a reading like 'there is a set of two books such that for each of them, there is a (possibly different) set of five of these students who read it.' Inverse (distributive) construals are not generally available with GQPs. Neither (30b) nor (30c) support distributive readings where *five of the books* and *five of the students* are the distributee, respectively.

This brings us to another parameter of PD. There is a marked difference in the ability of DQPs and GQPs to support distributivity when the distributee is chosen to be a presuppositional indefinite, like *two of these books*. This is not limited to inverse scope, but holds even when the GQP that acts as distributor c-commands the distributee. The claim is expressed in (29a(ii)).

DQPs are insensitive to whether the distributee is presuppositional or not, but this makes a difference with GQPs. The examples below contrast DQPs and GQPs in their ability to induce co-variation, as subjects, in a presuppositionally interpreted QP. These examples test the possibility of construing *two of these women* as the distributee, when the quantifier serving as distributor has as domain the set of men:

(31) a. Every/each man visited two of the(se) women [ok distribution]
b. Three men visited two of the(se) women [ok/? distribution]
c. John introduced two of the(se) women to three men [?? distribution]
d. John introduced three men to two of the(se) women [?? distribution]
e. Two of the(se) women visited three men [* distribution]

(31) shows that only subject GQPs can accomplish this task: whereas such subjects can somewhat marginally support distribution over presuppositional indefinites, GQP complements are typically unable to.[8]

In sum, GQPs freely distribute only over CQPs and cardinally-interpreted GQPs, but not over QPs that receive a presuppositional interpretation. Note also that the determiner *a* seems (at least with many speakers) to prefer a non-presuppositional interpretation, and is a good (pseudo-)distributee with GQPs.

Let's now consider the availability of distributive readings when the distributor and the distributee are both bare GQPs. Inverse distributive scope is still impossible, at least when the distributee is a subject; example (25) showed this. Consider next the availability of distributive readings between two bare GQPs in complement positions.

In the position of direct object (DO), GQPs like *five boys/three books* do not naturally support distributive readings, even when the distributee is a prepositional indirect object (P-IO), or an adjunct:

(32) a. I showed five books to a student [?? distributive scope *five books* > *a student*]
b. I introduced five boys to a girl [?(?) distributive scope *five boys* > *a girl*]

[8]Adjuncts seem to pattern with complements, as the following additional examples show:

(i) a. I left five of these books for two people [* distributive scope *two people* > *five of these books*]
b. I saw five of these students at two conferences [* distributive scope *two conferences* > *five of these students*]

(33) a. I left five books for a friend
[?(?) distr. scope *five books* > *a friend*]

b. I saw five students at a conference
[?(?) distr. scope *five students* > *a conference*]

So, GQPs are best as distributors when subjects; worst when direct objects. P-IOs (and certain adjuncts, it seems) are somewhere in between. Consider the configuration where the distributor is a P-IO or an adjunct, and the distributee is a DO. In this configuration distributive readings are more easily available than when the distributor is a DO.

(34) a. I showed two papers to three students
[ok/? distr. scope *three students* > *two papers*]

b. I introduced two women to three men
[ok/? distr. scope *three men* > *two women*]

This might be surprising, since in terms of phenomena such as pronominal binding, WCO, and NPI licensing, DOs appear to c-command P-IOs and adjuncts (cf. Barss and Lasnik 1986). The examples above show that scope is not computed using the same type of structural relations.

Having introduced the SD and PD patterns of distributivity, I turn now to outlining the theory of scope in LF on which this paper is based. After doing so, I will return to proposing an account for the patterns of distributivity reviewed in this section.

4 A "TARGET LANDING SITES" THEORY OF SCOPE

4.1 New functional projections

The discussion in the previous section has provided some basic evidence against scope being a uniform phenomenon, a point that is taken up in several of the contributions to this volume. In the context of theories of LF, the principle that "all QPs are created equal" can be expressed as follows:

(35) **Uniformity of scope assignment**

a. The scope assignment rule does not apply differently to different QPs. Each QP-type has access to the same scope positions as any other QP-type.

b. There are multiple scope positions accessible to each QP-type. The choice among these does not depend on QP-type.

The assumption in (35) is incorporated into May's and Aoun and Li's theories, and to a large extent, also in Hornstein's. The theory of scope proposed in Beghelli and Stowell (1994) (cf. Beghelli and Stowell 1996 and Beghelli 1995) rejects (35), in favor of the following principle:

(36) **Scopal Diversity** Distinct QP types have distinct scope positions and participate in distinct scope assignment processes.

Asymmetries of scope of the kind observed in the previous section are taken as an indication that the idiosyncratic scope assignment mechanisms which are available to different QP-types give them access to correspondingly different scope positions. This is the premise of the alternative theory of scope to be outlined in this section.

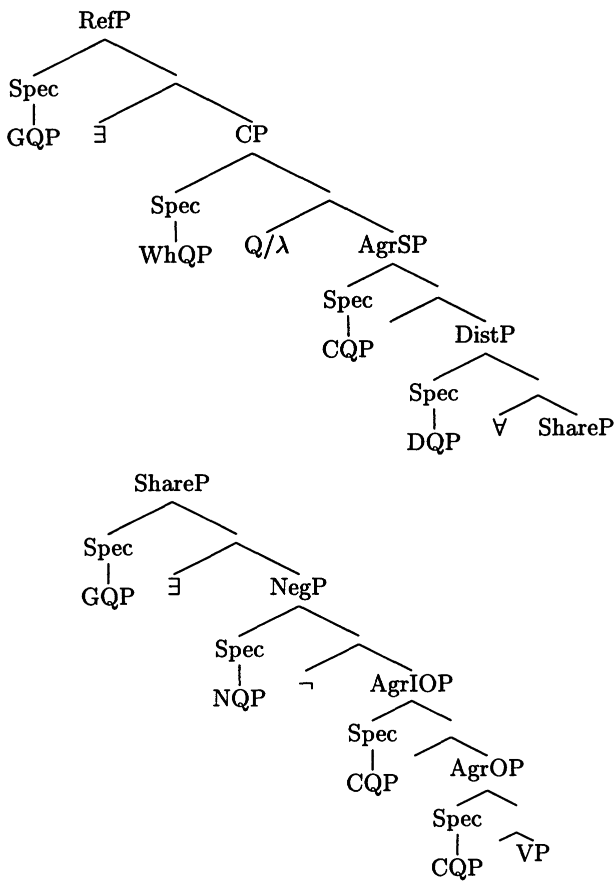
By allowing QP-types to have their own "target" scope positions, all the asymmetries of scope that we have observed in the previous section can be accounted for. First, to account for the very local scope of CQPs we should prevent them from undergoing scope movement in LF above and beyond movement to Case/Agreement positions. Conversely, the (relatively) unbounded scope of GQPs can be expressed by giving them access to a high scope position in the clause, from where they can gain wide(st) scope over (the positions) of other QP types, including DQPs. Furthermore, we can separate (strong) distributivity from scope by assuming the existence of a scope position associated with distributivity, which will be accessible to DQPs, but not to GQPs.

A theory of scope that derives the foregoing claims can be built as follows. I will refer to it as a "Target Landing Sites" theory of scope (TLS) (cf. Beghelli 1995).[9] The basic assumption of the approach is that the functional structure of the clause provides for a number of X^0 positions for Operators that serve logico-semantic functions. Current syntactic theories already provide for some of these positions: C^0 is the site of the Question operator that is associated with interrogative *wh*-words and direct question, or of the "lambda"(-like) operator invoked with non-interrogative *wh* and complementizers; Neg^0 hosts the Negative operator which is associated with clausal negation and NQPs.

In addition, the TLS theory assumes that there is a position for the Distributive operator associated with DQPs ($Dist^0$), and two distinct positions for Existential Operators: one for the wide scope construal of GQPs (Ref^0) and one for when they are narrow scope with respect to a higher quantifier ($Share^0$). These Operators are taken to project their own functional XPs, determining the following fixed hierarchy of projections in the clause. The three Agreement heads are not indicated in (37) in order to keep the diagram more transparent.

[9] In Beghelli and Stowell (1996), the TLS approach is referred to as a "checking theory of scope."

(37)



QPs are licensed in LF by moving into the specifier position of that functional category whose operator suits the semantic and/or morphological properties of their (sub-)type.

More precisely, QPs are assumed to be endowed with logico-semantic features which are type-specific. To realize ("check") these features, QPs are driven to move to that functional projection that hosts the appropriate Operator. In accordance with this logic, it is standardly assumed that *wh*-QPs move to [Spec, CP] where they can check their interrogative features by Spec-Head agreement with the Question operator (or, if they do not have interrogative force, they can check other kinds of *wh*-features with the appropriate *wh*-Operator). Similarly, a recent research tradition (cf. Zanuttini 1991, Haegeman 1994, Moritz and Valois 1994, etc.) has argued that NQPs are driven to move to NegP, the projection that hosts the Negative Operator, to check their negative feature.

Extending these directions, the TLS approach assumes the following types of movement in LF:

(38) **Target Landing Sites for QPs**

a. DQPs (*every man, each man*) are driven to move to [Spec, DistP] to check the features associated with their type, including a morphological feature of singular agreement and (logico-semantic) distributive features.

b. Definite GQPs (*the men*) move to [Spec, RefP] to check morphological and semantic features of definiteness against the wide scope Existential operator in Ref^0.

c. Indefinite (*some men*) and bare-numeral GQPs (*two men*) are also driven to [Spec, RefP] when they are associated with the topic-like feature of being a "subject of predication."

d. When narrow scope, these same GQPs are driven to move to [Spec, ShareP], where they check the feature of contributing a group referent to the interpretation of the clause; in this position, they can serve as distributive share to a DQP.

This approach finds some of its theoretical justification in the semantic theory of QP-types proposed in Szabolcsi (1996b). GQPs introduce DISCOURSE REFERENTS, i.e., variables, in the form of GROUPS. Groups are singular or plural *individuals*, corresponding to WITNESS SETS of the quantifier.[10] The group referent introduced by the GQP *two men* is, for example, any group that contains two men and no non-men.

Szabolcsi departs from standard DRT as it concerns DQPs. She proposes that DQPs also contribute a discourse referent to the interpretation of the sentence, in the form of a SET VARIABLE. Thus the QP *every man* introduces a variable X to be assigned the set containing all the men in the situation.

CQPs do not contribute discourse referents. They are interpreted as generalized quantifiers.

I adopt Szabolcsi's semantic theory of QP-types and incorporate it in the TLS proposal outlined above as follows.

The individual (group) variable contributed by GQPs must be bound by an existential operator. This takes place as GQPs move to [Spec, RefP] or [Spec, ShareP], where they enter into Spec-Head agreement with the existential operators available in those projections. GQPs are thus interpreted as individual terms existentially closed at their scope positions.

---

[10] A witness set of a quantifier is an element of the quantifier that is also a subset of its restrictor (the smallest set it lives on). See Chapter 1 for some background on this notion.

Things are more complex with DQPs; these both enter into a relation with a distributive operator (in Dist^0) and introduce a set variable. The distributive operator, by definition, binds individuals in the set contributed by a DQP. Thus it cannot bind the set itself. In declarative, non-negative contexts, the set variable is bound by the existential operator in Ref^0. (As we will see later on, other types of operators may be involved in the licensing of the variable in other contexts.) We distinguish therefore between the scope of the distributive operator, fixed in DistP, and the scope of the set, which is normally higher, in RefP.[11]

CQPs, being interpreted as generalized quantifiers, do not contribute variables, and cannot be bound by operators. They remain in the LF position where they are independently driven to move for Case/Agreement checking. We can assume they land in [Spec, AgrXP] (following Chomsky 1993), or remain *in situ*. When indefinite and numeral GQPs support only a cardinal interpretation (cf. Milsark 1974, Diesing 1992), they behave as CQPs do. Since they neither fulfill the role of subjects of predication nor introduce a group referent, they are not driven to move to RefP or to ShareP, and instead take scope in AgrXP, that is, *in situ*.

Theories of scope like May's and Aoun and Li's assume that the driving force of QR is the need of QPs to establish their scope, *qua* quantificational expressions. Since (by definition) this need is uniform across QPs, it yields a uniform mechanism of scope assignment, as in (35). Within a TLS approach, scope is instead the by-product of different morphological and logico-semantic needs, individual to (sub-)types of QPs, as detailed above. Thus, it results in a differentiated mechanism of scope assignment.

An important application of the theory outlined in this section to the syntax of Hungarian is provided in Szabolcsi (1996b). Her findings lend empirical support to the hypotheses we have entertained regarding the functional structure of the clause. Hungarian is a language that "wears LF on its sleeve." QPs are typically moved to clause initial positions and linearly ordered according to relative scope. Ordering of QPs is done on the basis of a hierarchy of positions endowed with invariant logico-semantic functions (cf. É. Kiss 1987). Szabolcsi

[11] Some justification for this assumption is given by the *de re* scope of DQPs. As pointed out by Farkas (1996), in a sentence like (i) below, the DQP *every man in this room* can be interpreted *de re* (hence wide scope) with respect to the intensional predicate, although it cannot be construed distributively with respect to *a witch*: the sentence is about a singular witch:

(i) A witch claimed that every man in this room had contact with her

This apparent paradox is accounted for in our proposal because the scope of the set of all men in the room is separated from the scope of the distributive operator. Whereas the former can be bound by an existential operator in the highest RefP, the latter remains local to the embedded DistP.

shows that the hierarchy of positions in Hungarian and their logico-semantic functions closely correspond to the hierarchy outlined in (36).

4.2 Deriving the distributivity patterns

4.2.1 The SD pattern

On the basis of the theory of QP licensing in LF of the previous section, we can provide a principled characterization of the patterns of distributivity sketched in Section 3.

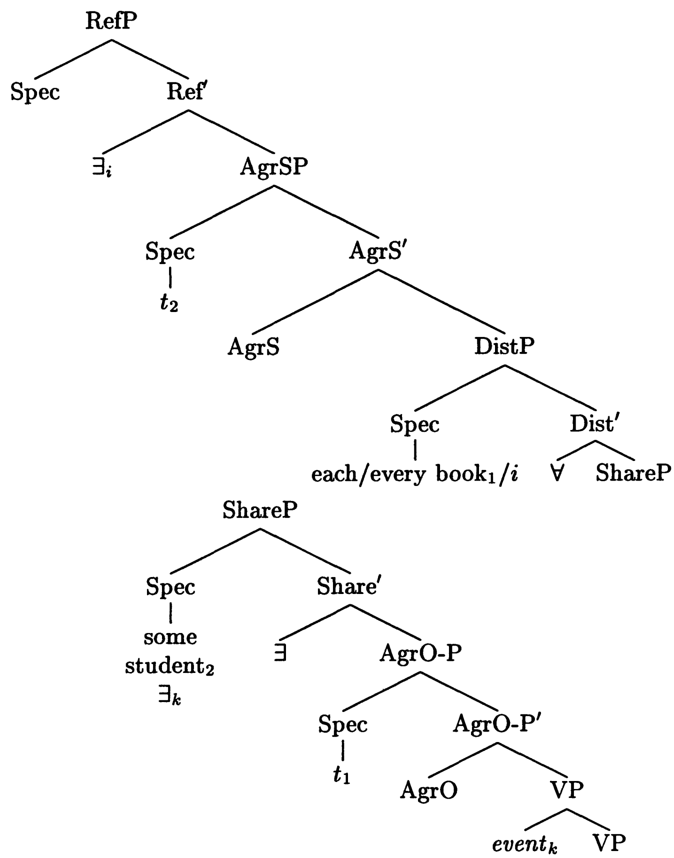
Let's begin with the SD pattern, which is uniformly displayed by DQPs in declarative, non-negative contexts. In the account given in Beghelli and Stowell (1994, 1996), SD is viewed, following an insight of Choe (1987), as a binary relation, requiring the simultaneous presence of a distributor and a distributee.

For the distributive relation to hold, the head of DistP (which remains unpronounced) must be "activated" by Spec-Head agreement with a DQP (the distributor). Dist selects ShareP: this projection hosts the share of distribution, or distributee. The share is required to be, semantically, a QP that can co-vary with the distributor, such as an existentially quantified term. An existential quantifier over the event argument and/or a GQP can occur in this position, since these are interpreted as existential quantifiers over groups. (It is assumed that scope positions can be filled multiply.) Thus both GQPs, which introduce group referents, and/or an existential quantifier over the event argument have access to Spec of ShareP.

To illustrate the proposal, I review below some derivations. The following LF represents the O > S reading of the sentence in (39a) below:

(39) a. Some student read each/every book
OK 'for every book x, there is some (possibly different) event e and student y such that y read x at e'

b.



The derivation involves the use of reconstruction. The subject indefinite *some student* is reconstructed to the Spec of ShareP position, through which it has moved on its way to its Case position in [Spec, AgrSP]. Reconstruction obeys the same principles that are imposed on movement to scope positions: a QP cannot reconstruct to positions that are incompatible with its semantic interpretation.

Note that both the Spec of DistP and the Spec of ShareP positions must be filled with the appropriate types of elements at LF for the strong distributivity relation to be satisfied.

Further illustrations of the LF scheme for SD are given by the examples below:

(40) a. Every student laughed
'for every student there is an event of laughing in which (s)he is the agent'.

b. $[_{\text{AgrSP}}\ t_1\ [_{\text{DistP}}$ every student$_1$ $[_{\text{ShareP}}\ \exists e_2\ [_{\text{VP}}$ laugh$_2$]]]]

(41) a. Every student read two books
'for every student x there is an event e of reading and a group Y of two books such that x is the agent and Y the theme at e'

b. $[_{\text{AgrSP}}\ t_1\ [_{\text{DistP}}$ every student$_1$ $[_{\text{ShareP}}\ \exists e_3$, two books$_2$ $[_{\text{VP}}$ read$_3$]]]]

The free availability of distributive readings with DQPs follows from this account. DQPs are driven to move to [Spec, DistP] to satisfy their features, which results in their supporting distributive construals whenever there is a suitable element to serve as share, be it a GQP or the (existentially quantified) event argument. The appearance of distributive readings with DQPs is therefore predicted whenever movement (to the relevant positions) is allowed.

4.2.2 The PD pattern

Let's come now to the PD pattern. The hypothesis that I present here is that differences in the SD and PD patterns reflect entirely different grammatical mechanisms at work in the two cases. SD comes about as the distributor moves to Spec of DistP, and the distributee to Spec of ShareP. In PD, DistP is not involved. Following Link (1983), Roberts (1987) and others, I propose that PD takes place through the agency of a covert distributive operator. Syntactically, I treat this distributor (which is distinct from the distributive operator in Dist^0) as an adverbial. In this way, I maintain that the group reading is the default reading of GQPs. My proposal goes beyond the current literature in that I observe empirical constraints on where distribution may occur and attempt to provide a syntactic account of them.

The distributive operator in PD can be compared to a "silent floated *each.*" More precisely, I assume that this element corresponds to the LF distributive adverbial that underlies what Safir and Stowell (1989) call "binominal *each,*" as in (42).

(42) Two students read a book each

In the spirit of Safir and Stowell's analysis, let's assume that the postnominal modifier *each* in (42) raises in LF to a position where it can be anteceded by the distributor *two students* and c-commands the distributee *a book.* "Silent *each*" is a covert counterpart of binominal *each,* which occurs in the position to which binominal *each* has raised. Thus, in LF, silent *each* is comparable to a floated quantifier.

It is natural to assume that binominal *each* is related to so-called "adverbial" *each,* i.e. the floated quantifier that occurs in examples like:

(43) Two students each read a book

I will not pursue the details of a unified analysis of adverbial and binominal *each*. What is most relevant for our purpose here is that binominal *each* has some of the properties that we need for our silent distributive element. First, it does not license *a different N*; second, as noted by Safir and Stowell, it cannot modify a subject (i.e., subjects cannot be marked as distributees by binominal *each*):

(44) a. ?? Five students read a different book each

b. * A student each visited five professors

Thus, in the analysis proposed here, silent *each* is an LF floated quantifier with the properties of binominal *each*. Cinque (1994a, 1994b) suggests that FQs are generated in AgrXP projections. Slightly extending this proposal, I assume that silent *each* can occur in AgrXP (X=S, IO, O) and in ShareP. I will leave it as an open question where exactly FQ are projected within these projections; this is not crucial to the analysis. Silent *each* will not be generated in RefP, in accordance with the fact that FQs do not occur in positions higher than the (canonical) subject position, and with the fact that binominal *each* cannot mark a subject as distributee.

Concerning the distribution of silent *each,* the null hypothesis is that it will match that of floated quantifiers. Data from Romance languages like French and Italian show that quantifiers can be floated off any argument position. However, there is a strict positional order, depending on which argument antecedes the floated Q (cf. Cinque 1994a, 1994b), as follows:

(45) $FQ_{S(ubject)} > FQ_{I(ndirect)\ O(bject)} > FQ_{D(irect)\ O(bject)}$

The data in (46)–(47), relative to a floated quantifier that corresponds to English *all,* illustrates the pattern (cf. also Cinque 1994b; the differences in gender marking help to identify the antecedents of the floated quantifiers).[12]

(46) a. Les cadeaux, ils les leur ont [French]
the gifts(masc) they(S) them(DO) to them(IO) have
toutes tous donnés
all(fem)(IO) all(masc)(DO) given(masc)

b. * Les cadeaux, ils les leur ont
the gifts(masc) they them to them have
tous toutes donnés
all(masc) all(fem) given(masc)

[12]I am grateful to Dominique Sportiche (p.c.) for providing the data in (46)–(47). Comparable data are presented by Cinque (1994a, 1994b).

(47) a.

Ils	les	ont	tous
they(masc)(S)	them(DO)	have	all(masc)(S)
toutes	vues		
all(fem)(DO)	seen(fem)		

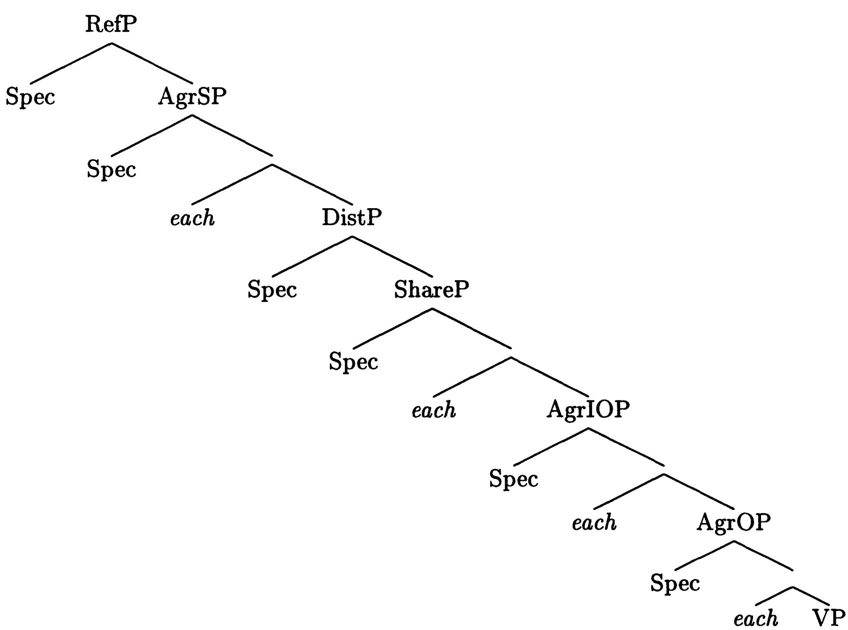
b. *

Ils	les	ont	toutes	tous	vus
they(masc)	them	have	all(fem)	all(masc)	seen(masc)

The data match the findings of syntactic literature on Germanic, where there is evidence that the Case/Agreement position of indirect objects is higher than the Case/Agreement position of the direct object.[13]

The following tree representation summarizes the possible positions of silent *each*:

(48) **Distribution of silent** ***each***



The distribution of the PD pattern is therefore predicted to be as follows.

(49) **Distribution of PD**

a. PD involves the use of a (silent) distributive adverbial *each,* which is distinct from the distributive operator in Dist^0.

[13] Similar assumptions are also made by Pica and Snyder (1995).

b. This adverbial is only available in AgrXP (X=S, O, IO) and in ShareP.

c. A GQP can antecede the adverbial if it has a trace in the Spec position of one of these projections (the GQP itself may have moved on to [Spec, RefP]).

d. The QP that serves as distributee in the PD relation must be within the c-command domain of silent *each* at LF.

To see how this proposal accounts for the PD pattern, recall that GQPs have distinct scope positions in LF depending on their interpretation (i.e., the logico-semantic features associated wih them). The following table summarizes the scope positions available to GQPs:

(50) **Scope positions for GQPs**

	POSITION	LOGICO-SEMANTIC FEATURE	INTERPRETATION
a.	Spec of RefP	[+subject of predication]	presuppositional
b.	Spec of ShareP	[+group referent]	presuppositional
c.	Spec of AgrXP	none	counting

GQPs that have a presuppositional interpretation are at least as high as Spec of ShareP at LF: this is because they contribute a group referent to the interpretation of the sentence. If, additionally, they are the logical subject of predication, they move higher, to [Spec, RefP]. Partitive GQPs (*two of the(se) students, ...*) are only interpreted as presuppositional. Bare GQPs (like *two students, ...*) can be interpreted either as presuppositional or as counting. (This ambiguity holds mostly for VP-internal GQPs; as subjects, GQPs typically resist counting interpretations).

It follows that presuppositional GQPs can always be PD distributors unless the distributee QP is a GQP in subject position. As an illustration, consider the following derivations. When distributor and distributee are VP-internal arguments, a presuppositional GQP can always function as distributor:

(51) John showed two books to five of the students
OK/? 'five of the students are such that to each of them, John showed a (possibly different) group of two books'
$[_{\text{AgrSP}}$ John $[_{\text{ShareP}}$ five of the students$_3$ $[_{\text{AgrIOP}}$ t_3 $[$*each* $[_{\text{AgrOP}}$ two books$_2$ $[_{\text{VP}}$ showed t_2 to t_3 $]]]]]]$

(52) John showed five of the books to two students
OK/? 'five of the books are such that for each of them, John showed it to a (possibly different) group of two students'
$[_{\text{AgrSP}}$ John $[_{\text{ShareP}}$ five of the books$_2$ $[$*each* $[_{\text{AgrIOP}}$ two students$_3$ $[_{\text{AgrOP}}$ t_2 $[_{\text{VP}}$ showed t_2 to t_3 $]]]]]]$

If the distributor GQP is a subject, we predict it will support PD over any VP-internal argument, even a presuppositional one; a presuppositional GQP cannot, however, support PD when the distributee is a subject GQP:

(53) Two students read three of the(se) books

a. OK 'two students each read a (possibly different) group of three books'

a′. [$_{\text{RefP}}$ two students$_1$ [$_{\text{AgrSP}}$ t_1 [*each* [$_{\text{ShareP}}$ two of the(se) books$_2$ [$_{\text{VP}}$ read t_2]]]]]

b. * 'three of the(se) books were each read by a (possibly different) group of two students'

b′. [$_{\text{RefP}}$ three of the(se) books$_2$ [$_{\text{AgrSP}}$ two students$_1$ [AgrOP t_2 [each [$_{\text{VP}}$ read t_2]]]]]

When GQPs receive a counting interpretation, which is the most natural reading in examples like the following, which feature GQPs as VP-internal arguments, the positional hierarchy S > IO > DO determines the availability of PD, as illustrated in the following derivations. (The examples are repeated from Section 3.2).

(54) a. I showed two papers to three students
[ok/? PD three students > two papers]
[$_{\text{AgrSP}}$ I [$_{\text{AgrIOP}}$ three students$_3$ [each [$_{\text{AgrOP}}$ two papers$_2$ [$_{\text{VP}}$ showed t_2 to t_3]]]]]

b. I showed five books to a student
[?(?) PD five books > a student]
[$_{\text{AgrSP}}$ I [$_{\text{AgrIOP}}$ a student$_3$ [$_{\text{AgrOP}}$ five books$_2$ [each [VP showed t_2 to t_3]]]]]

The pattern can only be reversed if we force a presuppositional interpretation on one of the VP-internal arguments, in which case the prediction is that the readings will be as in (51)–(52) above.[14]

[14] Under this approach, DQPs, even when not strongly distributive, still must support some kind of distributivity (namely, PD). In fact, it seems that collective readings with DQPs are always excluded, even under negation:

(i) ?? John didn't compare every man

It is an interesting question, then, what prevents DQPs from having collective readings. One line of explanation that could be suggested is that collective readings are incompatible with the agreement features of DQPs, which are singular; i.e. that collective interpretation requires plural agreement with plural QPs. A second line would be to say that collective readings crucially rely on the fact that the QP introduces a group in the form of a plural individual. DQPs introduce a set referent, hence not the right type of object. On the other hand, the fact

5 FAILURES OF SD IN THE SCOPE OF NEGATION

The PD pattern discussed in the preceding section will be of central importance to our analysis of PL: I will argue that interactions between WhQPs and QPs built with *every* in matrix questions essentially follow this pattern. *Each,* on the other hand, will be shown to follow the SD pattern.

Before we can claim that this is the relevant generalization, however, we must explain how it is possible for a quantifier that builds DQPs, and that has been shown to follow the SD pattern in declarative clauses, to switch to the pattern of GQPs in matrix questions.

To this end, I will consider some independent evidence which shows that QPs built with *every* do not always support SD, unlike those that are built with *each,* which do. The relevant data comes from interactions between DQPs and negation or NQPs. This topic is the subject of another paper in this collection, Beghelli and Stowell (1996), to which the reader is referred. Here I will only review the data which is directly relevant.

It is well-known (cf. Aoun and Li 1993, Hornstein 1995) that the scope of *every* is blocked by c-commanding negation; the same effect is produced by a c-commanding NQP. Consider:

(55) a. John didn't read every book
  * 'for every book x, John didn't read x (=no book was read by John)'
 b. No student(s) read every book
  * 'for every book x, there are no students who read x (=no book was read by any students)'

Aoun and Li derive this effect from the Minimal Binding Requirement: the NQP/negation are closer potential binders for the variable left by QR of *every N*. For Hornstein, this fact follows from a Relativized Minimality violation.

These explanations meet with some empirical problems, however. First, it is not clear why other types of QPs are immune from the effect: as is well known, GQPs can freely scope out of these negative environments.

(56) Every student didn't read one book
 OK 'for every student x, there is one (possibly different) book y that x did not read'

Second, the blocking effect observed in (55) is lifted with *each* in certain configurations. Examples like the following allow (albeit marginally) a reading

that CQPs also participate in collective readings is an entirely different matter. Szabolcsi (1996b) argues that they themselves do not denote collectives, they only count the atoms of the collective denoted by the predicate.

where *each submission* takes scope over the subject *one reviewer* by crossing over negation.

(57) One reviewer didn't support each submission
?/OK 'for each submission x, there is one (possibly different) reviewer who didn't support x'

5.1 The treatment of negation

In the face of these difficulties, we should look for an alternative account. Let's begin with why *every* cannot scope over negation in examples like (55).

The premise of the account presented in Beghelli and Stowell (1996) is that the negative operator in Neg^0, which is activated by either clausal negation or an NQP, is not to be seen as a propositional operator (as in classical logic), but rather as a (negative) quantifier over events ('for no events ... '). A negative statement like *John didn't come* accordingly receives a logical translation like 'there is no event of coming of which John was the agent' (cf. Krifka 1989, Schein 1993).

Syntactically, we can execute this by assuming that in negative clauses a negative quantifier over events (*NO:e*) is inserted in [Spec, NegP], from where it binds the position of the event argument.

(58) a. John came
[$_{\text{AgrSP}}$ John [$_{\text{ShareP}}$ $\exists_k$ [$_{\text{VP*}}$ *event*$_k$ [$_{\text{VP}}$ read$_k$]]]]
'John is such that there is an event of coming, of which he is the agent'

b. John didn't come
[$_{\text{AgrSP}}$ John [$_{\text{NegP}}$ NO$_k$ [$_{\text{VP*}}$ *event*$_k$ [$_{\text{VP}}$ read$_k$]]]]
'John is such that there are no events of coming of which he is the agent'

The effect observed in (55) follows directly from this premise and from our treatment of SD. Recall that DistP and ShareP are higher than NegP in the hierarchy of functional projections of the clause, and that SD is only satisfied when both distributor and distributee are in their respective positions, Spec of DistP and Spec of ShareP. In (55a), *every book* cannot move to [Spec, DistP] because there is no distributee to occur in [Spec, ShareP]. Given that there is no overt GQP in this sentence, only the event argument can fill the role of distributee. But clausal negation binds the event argument, which is accordingly unavailable to raise to Spec of ShareP (a move that would place it outside the scope of its binder, the operator in Neg^0).

The treatment of (55b) is analogous, asuming that NQPs move to (or through) [Spec, NegP] to check their negative features (cf. Moritz and Valois 1994, Haegeman 1994), and thus "activate" the operator in Neg^0.

The binding conflict in (55a) can be represented as in the diagram below (irrelevant details aside):

(55a′) * [$_{\text{AgrSP}}$ John [$_{\text{DistP}}$ every book$_{\text{i}}$ [$_{\text{ShareP}}$ [$_{\text{NegP}}$ NO$_{\text{k}}$ [$_{\text{VP*}}$ *event*$_{\text{k}}$ [$_{\text{VP}}$ read t_i]]]]]]

\# 'for every book x, there is no event e such that John read x (at e)'

5.2 Differences between *every* and *each*

This analysis derives the blocking effect observed in (55), plus other facts about the interaction of DQPs and negation, as discussed in Beghelli and Stowell (1996). Moreover, by pinning the source of the effect on the specific requirements of SD, it can discriminate between *every N* and GQPs. However, the account provided so far is incomplete: it remains to be explained why the examples in (55) are grammatical.

If DQPs are driven to move to Spec,DistP to check their morphological and semantic features, we should expect these examples to be deviant, given that *every book,* by staying within the scope of negation, cannot have moved in LF to the higher Spec of DistP.

Interestingly, these expectations turn out to be realized with *each.* First, QPs built with each do not occur in contexts like (55); unlike (59a) (=55a), (59b) is deviant:

(59) a. John didn't read every book

b. ?? John didn't read each book

Second, as (57) shows, *each* **can** scope out of c-commanding negation, provided there is an overt distributee to fill the [Spec, ShareP] position. The relevant example is repeated below:

(60) One reviewer didn't support each submission

?/OK 'for each submission x, there is one (possibly different) reviewer who didn't support x'

The factual conclusions we can draw from the data are as follows: DQPs built with *each* behave just like our theory predicts so far. They must move to [Spec, DistP] and satisfy SD. DQPs built with *every,* on the other hand, can be licensed outside DistP in negative sentences.

To account for the latter's behavior, we should modify the featural specification of *every.* Let's assume that both *every* and *each* are endowed with common morphological features, such as [+singular agreement], which allow access to [Spec, DistP].[15] In addition, let's assume that *each* is endowed with

[15] It seems plausible to assume that only QPs that have singular agreement can access [Spec, DistP], since the operator in Dist^0 is defined to apply only to individuals.

a [+distributive] feature, which **requires** movement to [Spec, DistP]. *Every* is underspecified for this feature. It therefore may, but does not have to, move to [Spec, DistP].

The next step is to clarify the relation between *every* and negation. Recall that, as in Szabolcsi (1996b), DQPs introduce discourse referents, in the form of a set variable. (GQPs also introduce discourse referents, but in the form of individual (group) variables.) Variables of course need to be bound. In non-negative, declarative sentences, the existential operator in Ref^0 binds the set variable. In this configuration, we observe that DQPs built with *every* invariably support SD. We must therefore conclude that when its set variable is bound by an existential operator, *every* is driven to move to [Spec, DistP].

The fact that *every* does not support SD when in the scope of negation suggests the following hypothesis. Assume that the negative operator can (unselectively) bind the set variable. This ensures that *every* will be licensed, while forcing it to remain in the scope of NegP (i.e., to remain in its Case/Agreement position), and thus not move on to DistP.

The LF representations below, with the indicated reading, will be assigned on the basis of the theory outlined above ('i' is the index that identifies the set variable introduced by *every book*):

(61) John read every book
$[_{\mathrm{RefP}}\ \exists_i\ [_{\mathrm{AgrSP}}$ John $[_{\mathrm{DistP}}$ every $\mathrm{book}_{2/i}$ $[_{\mathrm{ShareP}}\ \exists_k\ [_{\mathrm{VP*}}\ \mathit{event}_k$
$[_{\mathrm{VP}}$ read t_2]]]]]]

(55a′) John didn't read every book
$[_{\mathrm{AgrSP}}$ John $[_{\mathrm{NegP}}\ \mathrm{NO}_{k/i}\ [_{\mathrm{AgrOP}}$ every $\mathrm{book}_{2/i}$ $[_{\mathrm{VP*}}\ \mathit{event}_k$
$[_{\mathrm{VP}}$ read t_2]]]]]

5.3 PD behavior of *every*

Having outlined an account of the LF licensing of *every* in the scope of negation, we need to consider how the facts tie in with our account of distributivity. Consideration of the distributive properties of *every* under negation has already provided some independent evidence that our approach might be on the right track.

Under the theory of scope we have adopted, the quantificational properties of QPs are related to their entering into particular configurations in LF. For example, DQPs become strong distributors by moving to [Spec, DistP]. Therefore, when for some reason a DQP is licensed outside of this position, we should expect its distributive properties to be affected.

This prediction is borne out. It can be shown that when under the scope of the negative operator, and thus in [Spec, AgrXP] (instead of [Spec, DistP]), QPs built with *every* behave like pseudo-distributors. Take the examples in (62):

(62) a. John didn't show every book to some student/one of the students
b. John didn't show some book/one of the books to every student

Consider in particular the scopal relation between *every N* and *some N/one of the N.* Under normal intonation, we easily interpret (62a) as talking about a singular student, but cannot naturally assign (62a) a distributive reading paraphrasable as 'it is not the case that for every book x, John showed x to a (possibly different) student.' This is surprising: *every N,* as a strong distributor, can distribute over a clausemate indefinite (even a presuppositional one) from any position, including the object position (recall for example the data presented in (27)).

With (62b), too, the most natural interpretation seems to be that we are talking about a single book, though a distributive construal over the indefinite becomes available when we substitute *a book* for *some book/one of the books,* as in (62c):

(62) c. John didn't show a book to every student
OK/? 'it is not the case that for every student x, there a possibly different book that John showed x'

Substituting *a student* for *some student* in (62a) seems to have less of an effect on the availability of a distributive reading.

In sum, as regards (62a, b), we find no distributivity when the direct object is the distributor and the indirect object is the distributee (cf. 62a); or when the distributor is the indirect object, but the distributee is presuppositional (cf. 62b). We find that distributivity is supported only when the distributor is the indirect object, and the distributee is a non-specific direct object (cf. 62c).

As we saw in Sections 3.2 and 4.2.2, this is exactly what we get in the PD pattern (when the distributor is a bare GQP). Our conclusion with respect (62a, b) is thus the following: *every N* in the scope of clausal negation becomes a pseudo-distributor (i.e., shifts to the PD pattern).

A final observation supports this claim: if we substitute *a different book/student* for *some book/student* in (62), the sentences appear degraded (under the distributive reading of *a different*). These sentences are only acceptable on the relevant reading if interpreted as denials of their affirmative counterparts, given in (62c–d). The latter of course are fine, since there is no negation.

(63) a. John didn't show every book to a different student [?? unless denial]
b. John didn't show a different book to every student [?? unless denial]

c. John showed every book to a different student

d. John showed a different book to every student

6 THE DISTRIBUTION OF PL IN MATRIX QUESTIONS WITH DQPS

We are now in a position to tackle our first task: deriving the distribution of PL with *every.* This will be done in Sections 6.2 and 6.3. Then, in Section 6.4, I will consider the distribution of PL with *each.* But first, I will briefly indicate the semantic commitments of the analysis.

6.1 A brief excursus in the semantics of PL

Before we enter into the details of the derivation of PL, it is necessary to give some indications as to the semantic treatment that the LF account that I present in this section is supposed to interface to. I will not present a discussion of the semantics of questions, as this would go beyond the scope of the present paper. All the more so, since the semantic proposal that I rely on is presented in Szabolcsi (1996a), to which the reader is referred. Some familiarity with her paper is recommended for an understanding of the present one.

As pointed out by Chierchia (1993), there are two basic approaches to the semantics of PL. One consists in devising a technique for quantifying NPs into questions. Both Higginbotham (1991) and Groenendijk and Stokhof (1984) are representative of this tradition. The other, which Chierchia adopts (and considerably develops) is the functional approach: PL is viewed as a subcase of individual readings, where the individual is actually a function, spelled out "extensionally" as a set of pairs.

Szabolcsi's proposal, which I follow, does not assume that PL is a subcase of the functional reading. One feature of her analysis is that it distinguishes matrix questions and complements of *wonder*-type verbs from complements of *find out*-type verbs, primarily on the basis of what quantifiers support PL in each context.

In matrix questions, Szabolcsi adopts a version of Groenendijk and Stokhof's interpretation schema, which employs a device known as Domain Restriction. Crucially, however, Szabolcsi's schema does not incorporate lifting. Whereas for Groenendijk and Stokhof PL readings denote generalized quantifiers over individual questions (which amounts to "lifting" the interpretation of questions), as represented in a schematic format in (64c), Szabolcsi assigns the interpretation in (64b) to a PL question such as (64a). (The format given in (64b) and (c) abstracts away from the choice of a particular semantics for individual questions.)

(64) a. Who did every dog bite?

b. which $x \in A$, which y [x bit y]
where A is the unique set determined by the quantifier

c. $\lambda P \exists X$ (X a set determined by the quantifier & P(which$x \in X$ bit whom))

The difference between (64b) and (64c) is thus just one of lifting. The motivation for eliminating lifting is empirical. Lifting is necessitated by the desire to account for apparent PL readings in questions with indefinite (numeral) QPs, as in the example below:

(65) Who did two dogs bite?
? Fido bit John and Spot bit Mary

These readings are known as "choice" readings, since the answerer must choose a group referent (in the above example, a particular pair of dogs) for the QP prior to answering the question. Szabolcsi argues that such readings, to the extent they are available at all, are semantically distinct from PL readings, and should accordingly not fall under the same semantic interpretation schema. Since, as noted in Section 1, neither CQPs nor NQPs support PL in matrix questions, PL is under this view a distinctive property of DQPs. With DQPs, which are universal terms, there is no choice of discourse referent, as the set introduced by these QPs is unique. Having excluded GQPs, NQPs, and CQPs from participating in PL in matrix interrogatives, the motivation for lifting disappears, and the simpler format in (64b) can be used instead of (64c).

Aside from the issue of lifting, the use of Domain Restriction (in either (64b) or (64c)) deserves attention, since it will play a significant role in the LF analysis to be proposed later in this section. The intuition behind Groenendijk and Stokhof's Domain Restriction is that in PL questions the quantifier does not operate in its "usual" way, i.e., in the way it operates in a declarative sentence. It does not quantify over the whole question—this would be semantically problematic, owing to a difference in semantic type. Rather, the QP simply "lends" its witness set (which for *every dog* is the set containing all dogs and nothing else) to the interpretation of the question. The domain of the *wh*-question operator is thus "augmented" and spans over pairs. This mechanism is reminiscent of the syntactic operation of "absorption," where two QPs with overlapping scope are merged into a complex or "binary" quantifier.

With respect to PL readings arising in *wh*-complements of verbs like *find out, know, ...* , Szabolcsi proposes an interpretation where the QP operates in its usual "quantificational" way, i.e., a semantic schema that incorporates lifting and does not invoke Domain Restriction:

(66) a. John found out who every dog bit

b. $\lambda P[\text{every-dog}'(\lambda x[P(\text{which } y\ [x \text{ bit } y])])]$
where $P =$ the property that p has iff John found out p

I will come back to the semantics of complement clause PL in Sections 7 and 8.

6.2 The PL pattern with *every* as a subcase of the PD pattern

I return now to our central concerns, the derivation of PL at the syntactic level of LF. The first claim that I present is that the distribution of PL readings in matrix questions with *every*-QPs (henceforth, the "PL pattern with *every*"), is a subcase of the PD pattern. In other words, I claim that (i) PL is a distributive dependency; and (ii) the distribution of PL in matrix questions with *every N* follows the same pattern supported by GQPs in declarative contexts.

Let's proceed to show that the distribution of PL with *every*-QPs conforms to the pattern outlined in (29). Both the type of WhQP and its syntactic position with respect to *every* are relevant to the availability of PL. Like with GQPs, there are two types of WhQPs, which I consider in turn: (i) BARE WhQPs (*who, what, ...*); (ii) D-LINKED OR PARTITIVE WhQPs (*which (of the(se) men)*).

I assume that WhQPs can undergo reconstruction in LF, thus lowering their scope from [Spec, CP]. Let's then view the pattern of PL as one case of the distributivity pattern.

Bare WhQPs interacting with *every N* show a pattern that parallels that of PD when both distributor and distributee are bare GQPs: we find that PL is essentially available when the distributor (=*every N*) is higher than the distributee (=the reconstructed WhQP) according to the LF Case Hierarchy: subject > indirect object > direct object. So, subject *every N* is best in supporting PL; as indirect object, *every N* is still somewhat able to yield PL readings with object Wh; the converse configuration, *every N* object and *wh* indirect object, is less likely as PL; finally, no PL is available when *wh* is the subject.

Note that these judgements, which come from my own field work, are slightly different from those reported by Chierchia. He notes that when *every* and *wh* are both VP internal, PL is sometimes possible either way, though he predicts that as direct object, *every* should be generally better in supporting PL than as indirect object. (This matches the distribution of WCO). I find that for most of the speakers I interviewed, PL is more easily available when *every* is an indirect, rather than a direct, object. The judgements can be summarized in the following pattern:

(67) a. What did every student write about? [ok PL]
b. What did you show to every man? [? PL]
c. To whom did you show every picture? [?? PL]
d. Who wrote about every book? [* PL]
e. Who showed this book to every man? [* PL]

A somewhat different pattern is found with QPs built with *which* (see also Szabolcsi 1994). With D-linked WhQPs, PL readings are only available when *every* is in subject position:

(68) a. Which man did every dog bite? [(?)PL]
b. Which dog bit every man? [*PL]
c. Which picture did you show to every man? [*PL]
d. To which man did you show every picture? [*PL]

Partitive WhQPs (e.g. *which of the(se) men*) behave likewise. If they support PL, it is only when *every N* is the subject.

The pattern in (68) has special theoretical significance. First, it is not predicted under May's or Aoun and Li's proposals, which do not distinguish between types of WhQPs. These accounts say that subject *every N* should behave, with respect to a WhQP in any other position, just like direct object *every N* behaves with respect to indirect object WhQP. The data in (67)–(68) show that this prediction is incorrect, both with bare and D-linked *wh*.

The pattern in (68) cannot be captured under Chierchia's WCO account either. Chierchia's prediction is that PL should be more easily available when both *every N* and *wh* are VP-internal arguments, given that WCO effects are generally very weak (if present at all) in such configurations. With bare *wh*, judgements are perhaps simply not so clear as to falsify or confirm this prediction. But with D-linked and partitive *wh* we have cases where the speakers' intuitions point in the opposite direction from WCO. These data therefore seem problematic for the WCO story. On the other hand, they provide significant confirmation for the suggestion that the distribution of PL in matrix questions follows the PD pattern.

Thus, the data reviewed in (67)–(68) support the hypothesis that the PL pattern with *every* is a subcase of the PD pattern as outlined in Section 3.

6.3 Deriving the distribution of PL with *every*

Given the similarity between the PL pattern with *every* and the PD pattern, I turn to providing an explanation for this convergence. The core questions, in this respect, are: (i) why does *every N* revert to PD, and (ii) how does this fit in with the behavior of *every* in the scope of negation, which we have observed in the previous section.

To answer these questions, I submit that it is the interrogative operator that triggers the PD behavior of *every N* in matrix *wh*-questions. I propose that the interrogative operator (like negation) can license *every N* by binding its set variable. Since the question operator in C^0 is a closer binder than the existential operator in Ref^0 (CP is below RefP), when present, the question operator will take precedence in binding the set variable introduced by *every.* This amounts to extending the treatment offered in the preceding section for *every* in the scope of negation to the case where *every* is in the scope of a question operator.

Next, I assume that this pre-empts movement of *every N* to [Spec, DistP]. I assume that *every* is only strongly distributive when its set variable is bound by an existential operator. Under the analysis that has been developed thus far (cf. especially Section 5), *every* is not lexically endowed with the property of being a strongly distributive quantifier. Rather, it is semantically underspecified. Its logico-semantic (quantificational) properties derive from the LF configuration in which it occurs. *Every* is interpreted as a strongly distributive quantifier when its set variable is bound by an existential operator. In interrogative (or negative) contexts, it simply lends its domain of quantification (i.e. its restrictor set) to the question operator (or to the negation). This represents a direct LF implementation of Groenendijk and Stokhof's Domain Restriction, reviewed in Section 6.1.

The proposal amounts to saying, effectively, that in interrogative/negative contexts, *every N* "merges" or "absorbs" with the interrogative/negative quantifier. (Unlike Chierchia, I do not implement syntactic absorption; rather, I invoke binding. But the two notions serve similar functions). Neither the negative nor interrogative operator is associated with strong distributivity. *Every* accordingly relinquishes its strong distributor behavior, functioning more like an interrogative/negative QP.

On the basis of this account, given that *every N* does not move to [Spec, DistP] when its set variable is bound by an interrogative or negative operator, it will not support SD. Instead, it will revert to the weaker type of distributivity supported by the other QP-types, namely PD.

Let's now consider more closely the structural conditions under which the set variable introduced by QPs built with *every* can be bound by the question operator. This will allow us to refine the analysis given above.

6.3.1 Conditions on the binding relation between every *and the question operator*

Three observations, which I consider in turn in this subsection, will be helpful in characterizing the conditions on the emergence of PL readings with *every*-QPs, and will thus fine-tune the analysis. The first is that simply being in the scope of an interrogative operator is not sufficient to license PD behavior on the part of *every N.* To begin with, in yes/no questions, *every N* clearly shows the SD pattern. Consider:

(69) a. Did every student read a different book?
 b. Did you show every book to a (different) student?
 c. Did a different student write about every book?

(69a) can be quite naturally interpreted with *every* taking *a different book* as distributee. In (69b) there is distribution between *every book,* in direct object position, and *a different student,* the indirect object. (69c) offers an example of inverse distributive scope where an object is the distributor and the subject is the distributee. These are all distinctive properties of SD (recall that they do not obtain with PD). In yes/no questions, therefore, *every* behaves as in non-negative declarative contexts.

The second observation is that in matrix *wh*-questions *every N* is not always prevented from behaving as a strong distributor. Rather, what we observe is that in matrix *wh*-questions, *every N* either supports SD or PL, but not both. Consider the examples in (70).

(70) a. Who showed every book to a different student? [* PL, ok SD]
 b. Who showed a different book to every student? [* PL, ok SD]

Since in (70a, b) the WhQP is a subject, these questions do not have a PL construal (cf. the generalization on the distribution of PL given above). However, on the individual answer construal, *every* behaves as a strong distributor: we can construe both (70a, b) in such a way that we have distribution of *every book/student* over *(a different) student/book.* This is something not predicted by our analysis. These examples prompt a more careful formulation of the licensing conditions on PL laid out in the previous section: they show that the mere presence of a [+Wh] Question operator in C^0 is not sufficient to bind the set variable of a DQP; being the closer binder is not enough.

A further observation can however be drawn from these examples, our third. What characterizes these examples is the relative scope of WhQP and DQP. Insofar as the WhQP scopes higher than the (Case) position of *every,* the latter can show its "default" SD behavior.

These three observations force us to characterize more precisely the conditions under which *every N* can be bound by the question operator. I propose the following analysis of these conditions.

In interpreting *wh*-questions with quantifiers, we can in principle construe the WhQP as either (i) wide scope with respect to the QP, or (ii) within the scope of the QP. In the latter case, I assume that the WhQP is reconstructed into its Case position, if it is a bare WhQP, or into [Spec, ShareP] if it is partitive or D-linked (it is plausible to assume that partitive WhQPs introduce a group referent). More exactly, I assume that whereas the WhQP can be reconstructed, the interrogative operator remains in CP.[16]

Reconstruction of the *wh*-phrase is thus the first condition on the appearance of PL in matrix questions with *every N*. This accounts for our third observation: there is no PL if the WhQP is assigned scope above the LF position of the DQP.

It follows that matrix questions with DQPs can always support an individual answer interpretation: if the *wh*-phrase does not reconstruct, it will be interpreted as taking scope in [Spec, CP], and thus as taking wide scope over the QP. A PL interpretation is only possible when, by reconstructing the WhQP into its Case position, the *wh* finds itself in the scope of *every*.

The two possible scope relations between *wh* and *every* that were considered in (i)–(ii) above can be represented as in (71). PL is not possible in (71i) but it is okay in (71ii). Only the scope positions of *every* and *wh* are indicated in the LF diagrams, and reconstructed elements are in curly brackets (a convention that I follow in the rest of the paper):

(71) i. $[_{\text{CP}}\ [\text{Q-Op}_\text{i} + \text{WhQP}_\text{i}\ ]\ [\ \ldots\ [_{\text{AgrXP}}\ [\textit{every N}]_\text{k}\ \ldots\ ]\ \ldots\ ]\ ]$

ii. $[_{\text{CP}}\ [\text{Q-Op}_\text{i} + t_i\ ][\ \ldots\ [_{\text{AgrXP}}\ [\textit{every N}]_{\text{k/i}}\ \ldots\ ]\ \ldots\ [_{\text{AgrYP}}\ \{\text{WhQP}\}_\text{i}\ \ldots\ ]\ \ldots\ ]]$

Given that the first condition on the emergence of PL requires the WhQP to take narrow scope, the second condition has to do with the relative positions of the reconstructed *wh*-phrase and *every N*. Since we have seen that the PL pattern appears to be a subcase of the PD pattern, we should expect, for PL to be supported, that the relative position of DQP and WhQP conform to the configurations that license PD.

Collecting all of these conditions together, the derivation of PL proceeds as follows. The WhQP is reconstructed either into [Spec, ShareP] or into its Case position, depending on its interpretation. The relevant type of LF configuration is given in (71ii). The reconstructed *wh* is bound by the interrogative (Q)

[16]The hypothesis that I am making here is simply that WhQPs, whether in *wh*-movement languages or *in-situ* languages, are a type of QP which can be bound by the Question Operator. When an interrogative WhQP is reconstructed at LF, it is like a *wh*-in-situ in multiple interrogation constructions.

operator (as shown by their sharing the 'i' index). Since *every N* falls within the binding path of the same Q operator, it gets unselectively bound along with the WhQP. In the configuration that supports only the individual reading, (71i), the WhQP remains in [Spec, CP], and *every N* is outside the binding path of the question operator; thus *every N* is not bound by it.

Thus, it is not enough that the question operator in C^0 be the closest binder for *every N*; since a [+Wh] question operator (plausibly) binds a reconstructed WhQP (which has, prior to reconstruction, entered into Spec-Head agreement with the operator), the set variable introduced by *every N* can only be unselectively bound when the DQP intervenes between the question operator and the reconstructed WhQP.

Binding by the interrogative operator enforces the PD pattern on *every N*—like we saw that binding by negation did. PL will be available iff the position of *every N* (the distributor) and the position of the reconstructed WhQP (the distributee) fall within a configuration that supports PD, as characterized in Section 3.

6.3.2 Yes/no questions

I return now to accounting for the first observation in this section, i.e. to why yes/no questions cannot bind the variable of DQPs. The reason for this is that the yes/no Q operator is not a variable binding operator, or at least not a variable binding operator of the type that could bind the set variable of a DQP.

First, the yes/no Q operator does not bind the event argument: in a question like *Did John graduate?,* the speaker is not asking "what is the event in which John graduated." If the yes/no operator should bind anything, this would likely be a sentential operator or adverb. The only way to turn a yes/no question into a *wh*-question is to paraphrase the question *Did John graduate?* as *How true is it that John graduated?.* This question can receive an individual answer like True or False.

Now, pursuing the hypothesis that the yes/no operator does bind some suitable adverbial (i.e. assuming this is possible), the plausible conclusion is that the set variable of a DQP could not be unselectively bound along with the adverbial because they wouldn't, semantically, be variables of the same kind.

Some supporting evidence for this conclusion comes from looking at multiple interrogation. If looking at multiple interrogation provides us with a comparable situation (semantically, at least, if not structurally, as actually argued by Hornstein 1995 and Williams 1994), we see that the semantic variable introduced by adverbials like *how* and *why* cannot be questioned along with individual-denoting WhQPs in a multiple interrogation. As is well-known, adverbial *wh*-elements like *how* and *why* cannot remain in-situ:

(72) a. * Who left why?

b. * Who left how?

6.3.3 *Summary of the analysis*

At this point, let's tie in the analysis of *every* in *wh*-questions with the analysis given in the previous section of *every* under negation.

Having formulated the conditions under which the set variable introduced by *every*-QPs is bound by the question Operator, we can see that these conditions, summarized in the configuration in (71ii) above, are not different from the conditions under which *every N* can be bound by negation. Both cases belong under the same generalization. In either case, (i) the binding operator must be the closest (potential) binder; in addition, (ii) binding of the set variable of the DQP is done unselectively, by the DQP intervening in LF in the binding path connecting the binder to the variable that it normally binds. In the case of the question operator, the variable is located with the reconstructed WhQP; in the case of negation, the variable bound by negation is the event argument. In both cases, binding by an operator other than the existential inhibits movement of *every* to [Spec, DistP], and thus the emergence of the PD pattern.

Thus, the cases where *every* is in the scope of c-commanding negation are indeed an instance of the same configuration as (71ii). We have the following parallel configurations, as given by the diagrams below (73b = 71ii):

(73) a. What did every student read? [under PL reading]

b. $[_{\text{CP}}$ $[\text{Q-Op}_{\text{i}} + t_i]$ $[\ldots$ $[_{\text{AgrXP}}$ $[\textit{every } N]_{\text{k/i}} \ldots]$ $\ldots$ $[_{\text{AgrYP}}$ $\{\text{WhQP}\}_{\text{i}}$ $\ldots]$ $\ldots]]$

(74) a. John didn't read every book

b. $[\ldots$ $[_{\text{NegP}}$ NO_{i} $[\ldots$ $[_{\text{AgrXP}}$ $[\textit{every } N]_{\text{k/i}} \ldots$ $[_{\text{VP*}}$ $\text{event}_{\text{i}} \ldots$ $[_{\text{VP}}$ V $\ldots]]]]]]$

To summarize our proposal so far, PL will only appear when (i) the WhQP reconstructs in the c-command domain of *every N,* and (ii) the relative positions of *every N* and of the reconstructed *wh* support PD. I give below the LF representations for the pattern of bare *wh* presented in (67) above:

(75) a. What did every student write about? [ok PL]
$[_{\text{CP}}$ $\text{Q-Op}_{\text{i}} + t_i$ $[_{\text{AgrSP}}$ every $\text{student}_{\text{k/i}}$ $[\textit{each}$ $[_{\text{AgrOP}}$ $\{\text{what}\}_{\text{i}}$ $[_{\text{VP}}$ t_k wrote about $t_i]]]]]$

b. What did you show to every man? [? PL]
$[_{\text{CP}}$ $\text{Q-Op}_{\text{i}} + t_i$ $[_{\text{AgrSP}}$ you_{j} $[_{\text{AgrIOP}}$ every $\text{man}_{\text{k/i}}$ $[\textit{each}$ $[_{\text{AgrOP}}$ $\{\text{what}\}_{\text{i}}$ $[_{\text{VP}}$ t_j showed t_i to $t_k]]]]]]$

c. To whom did you show every picture? [?? PL]
[$_{\text{CP}}$ Q-Op$_i$ + t_i [$_{\text{AgrSP}}$ you$_j$ [$_{\text{AgrIOP}}$ {who}$_i$ [$_{\text{AgrOP}}$ every picture$_k$ [$_{\text{VP}}$ t_j showed t_k to t_i]]]]]

d. Who wrote about every book? [* PL]
[$_{\text{CP}}$ Q-Op$_i$ + t_i [$_{\text{AgrSP}}$ {who}$_i$ [$_{\text{AgrOP}}$ every book$_k$ [$_{\text{VP}}$ t_i wrote about t_k]]]]

e. Who showed this book to every man? [* PL]
[$_{\text{CP}}$ Q-Op$_i$ + t_i [$_{\text{AgrSP}}$ {who}$_i$ [$_{\text{AgrIOP}}$ every man$_k$ [AgrOP this book$_j$ [$_{\text{VP}}$ t_i showed t_j to t_k]]]]]

Note that the schema in (71ii) generalizes directly to the cases where *wh*, thematically originating in an embedded clause, has moved to the matrix CP:

(76) a. What do you think that every student read? [ok PL]

b. [$_{\text{CP}}$ [Q-Op$_i$ + t_i] do [you think [$_{\text{CP}}$ that [$_{\text{AgrSP}}$ every student$_{k/i}$ [$_{\text{AgrOP}}$ {what}$_i$] [$_{\text{VP}}$ t_k read t_i]]]]

Unlike May (1985), who raises the DQP to matrix IP, we account for these complex PL questions by reconstructing the WhQP into the scope of *every N* in the embedded sentence.

The licensing conditions for DQPs built with *every* are summarized below:

(77) **Licensing of *every N***

a. QPs built with *every* are endowed with the feature [+singular agr.], being underspecified for [distributive]. This featural specification may, but doesn't have to be discharged in [Spec, DistP].

b. *Every*-QPs must move to [Spec, DistP] only when the closer binder of their set variable is an existential operator in [Spec, RefP]. When in [Spec, DistP] in LF, *every*-QPs activate the distributive operator in Dist^0, and support SD.

c. When the closer binder is a negative or question operator, QPs headed by *every* do not move to [Spec, DistP].

Before concluding this section, I should point out that our account is considerably different from May's and Aoun and Li's, but bears some resemblance to Chierchia's, as mentioned above. The mechanisms whereby PL is generated are different, WCO vs. PD; this, as was noted, makes different predictions on the distribution of PL. But there is a partial convergence in the assumption that PL is the result of a special operation, absorption for Chierchia, binding of the discourse referent here, whereby the *wh* "merges" with the QP. The following section will show that, beyond this convergence, there remain substantial empirical differences between the two accounts.

6.4 PL in matrix questions with *each*

The syntactic distribution of PL readings in matrix questions with *each*-QPs provides one of the main empirical arguments for the analysis presented in this section. The difference between the PL pattern with *every* and with *each* represents a distinctive feature of the present approach, since it cannot be (easily) incorporated in either May's/Aoun and Li's or in Chierchia's/Hornstein's treatments of PL.

For most speakers, *each* appears to support PL not only in the configurations where PL is possible with *every*-QP, but also in configurations where PL with *every* is excluded (78a, b) or difficult (78c, d). In (78a), *each book* is a direct object and the WhQP is a subject, a configuration where *every book* does not support PL (78b). In (78c) *each book* is also a direct object, whereas the WhQP is a prepositional indirect object: in this configuration, it was found that *every*-QPs would support PL only with difficulty.[17]

(78) a. Who read each book? [ok PL]
b. Who read every book? [* PL]
c. Who did you assign each book to? [ok PL]
d. Who did you assign every book to? [?* PL]

The data presented above show, in essence, that there are no syntactic asymmetries in the distribution of PL with *each*. This observation goes back at least to Williams (1986, 1988). (Note also that Karttunen and Peters 1980 use *each* throughout their paper without remarking on any asymmetry of the type shown by *every*.)

None of the other approaches to PL reviewed here seem to be able to account for these data. May (1985) and Aoun and Li (1993) treat all quantifiers

[17]The same data are available in Italian, which displays a distinction which seems to match that between *each* and *every*: unlike French and Spanish, which have only one active DQP (cf. French *chaque,* Spanish *cada*), Italian has a distinction between *ogni* 'every' and *ciascun(o)* 'each.' In Italian, PL with *ciascuno* appears to be available in the same configurations where *each* licenses PL. PL with *ogni* on the other hand, has a more restricted distribution, comparable to that found with *every*-QPs:

(i) Chi ha letto ciascun libro? [ok PL]
who has read each book?

(ii) Chi ha letto ogni libro? [* PL]
who has read every book?

(iii) A chi hai mostrato ciascun libro? [ok PL]
to who have-you shown each book

(iv) A chi hai mostrato ogni libro? [?* PL]
to who have-you shown every book

alike, hence it is unlikely that their account could be made to distinguish two otherwise so similar quantifiers like *every* and *each.* Neither can Chierchia's (1993) account for the different behavior of *every* and *each,* since WCO effects are just as strong with *each* as with *every*:

(79) a. ?? His_i mother accompanied every child_i

b. ?? His_i mother accompanied each child_i

The behavior of *each* in PL questions can however be accounted for by our proposal. Recall the licensing conditions on *each*-QPs:

(80) **Licensing of** *each N*

a. QPs built with *each* are endowed with the features [+singular, +distributive]. Since these features can only be discharged in [Spec, DistP], *each*-QPs must move there in LF.

b. Therefore, *each*-QPs always support SD.

c. Semantically, the set variable introduced by QPs headed by *each* must be bound by an existential or Question Operator.

Given that *each*-QPs are endowed with a [+distributive] feature, they have to check it off in [Spec, DistP]. This requirement prevents *each*-QPs from being bound by negation, since DistP is higher than NegP, the site of the negation operator. In fact, as observed in Section 5.2, there is a basic incompatibility between *each* and negation. Sentences where negation c-commands *each* are somewhat deviant, or at least show very restricted distribution and/or require special intonation (cf. 60).

Getting bound by the question operator does not, however, prevent *each*-QPs, in principle, from moving to [Spec, DistP], since the site of the operator (C^0) is higher than DistP. Given that *each*-QP **must** check its features in [Spec, DistP], it moves there even when its set variable is bound by the question operator.

The difference between *each* and *every* is thus that the latter supports SD only when its set variable is bound by the existential operator. DQPs built with *every* thus show a sort of "polarity" behavior, whereby their quantificational properties are affected by the type of their binder.

In detail, the derivation of PL readings with *each*-QPs proceeds as follows. The QP headed by *each* is driven to move to [Spec, DistP]; the WhQP is reconstructed as usual (in [Spec, ShareP] or [Spec, AgrXP]). The set variable of the *each*-QP is bound by the question operator in C^0. The respective thematic or Case positions of *each N* and *wh* are irrelevant to the availability of PL with *each,* just as the relative Case positions of the distributor and the distributee were irrelevant in the SD pattern. The *each*-QP is driven to move to [Spec, DistP] for distributivity checking from any argument position. The distribution of PL with *each* as in (78)–(79) is therefore derived.

The derivation of (78a) below summarizes the proposal for PL with *each*:

(78a′) Who read each book?
$[_{\text{CP}}\ \text{Q}_{\text{i}} + t_i\ [_{\text{AgrSP}}\ t_i\ [_{\text{DistP}}\ \text{each book}_{\text{k/i}}\ [_{\text{ShareP}}\ \exists_{\text{j}}\ [_{\text{AgrOP}}\ t_k\ [_{\text{VP*}}\ \textit{event}_{\text{j}}\ [_{\text{VP}}\ \{\text{who}\}_{\text{i}}\ \text{read}\ t_k\ ]]]]]]]$

7 PL READINGS WITH *EVERY*- AND *EACH*-QPS IN *WH*-COMPLEMENTS

I come now to considering another asymmetry in the distribution of PL, namely the contrast between the distribution of PL in matrix and embedded questions with DQPs. The main focus will be on so-called EXTENSIONAL *wh*-complements (e.g. those selected by ***know, find out,*** etc.). Other types of *wh*-clauses will be considered only in passing.

The premise of the account is that what distinguishes matrix from (extensional) *wh*-clauses is that in the latter, WhQPs don't have interrogative force. Following Groenendijk and Stokhof (1984), Szabolcsi (1996a) and others, I assume that the *wh*-operator in C^0 is distinct from the question operator found with matrix *wh*-questions. Given that the distribution of PL in matrix questions hinges, on the analysis presented here, on the question operator acting as binder for the set variable of DQPs, we would expect its absence in *wh*-complements to have effects on the distribution of PL. This expectation is borne out, as will be seen below.

7.1 Interactions between DQPs and non-interrogative WhQPs

As pointed out by Szabolcsi (1994), in embedded *wh*-clauses the asymmetries in the availability of PL that are found in matrix questions with *every* seem to disappear. The annotations in brackets under each example highlight the differential availability of PL in the two configurations.

(81) a. By tomorrow, we'll find out which book every student read [ok PL, also ok in matrix case]

b. By tomorrow, the committee will decide which applicant will fill every position [ok PL, * in matrix case]

c. By tomorrow, we'll find out which topic the teacher has assigned to every student [ok PL, ? in matrix case]

d. The teacher has already decided to which student he'll assign every book on the list [ok PL, * in matrix case]

Speakers are able to construe *every student* in (81a) as distributing over *which book,* arriving at construals like 'we'll find out, for every student x, which book x read.' Such interpretations are on a par with the distributed (matrix) question readings that I labeled as PL. I refer to them as PL-like interpretations (in the glosses however, I simply indicate 'PL', as in (81) above).

That PL-like construals are possible in the configuration (81a) where *every N* is in subject position is unsurprising, given that *every N* as subject generally supports PL in matrix questions. What is surprising is that speakers don't have difficulty in obtaining PL-like construals uniformly throughout the paradigm in (81), contrary to the matrix case.

Let's start by considering the semantics of PL-like readings with embedded interrogatives. Extensional *wh*-clauses are complements of verbs like *know, find out,* etc., which denote relations between an individual and the **answer** to a question (cf. Groenendijk and Stokhof 1984). These are distinct from intensional *wh*-complements of verbs like *wonder,* which denote relations between an individual and a **question**.

Under Szabolcsi's (1996a) analysis, which I adopt here, PL is not a unitary phenomenon. Szabolcsi argues, partly on the basis of the different distribution of PL-like readings observed above in (81), that these readings with extensional *wh*-complement clauses are to be treated differently than matrix PL. She proposes that they are obtained, semantically, by direct quantification into lifted questions (which is what she argues against in the matrix case). A "lifted question" is a set of properties such that the set of true answers to the question has those properties.

By ordinary, or direct, quantification Szabolcsi means that the DQP behaves in its own regular way (as opposed to providing a domain restriction, as in matrix questions). In ordinary declarative contexts, DQPs take a property of individuals (=a set of individuals) as their nuclear scope: e.g. the property "sleep" in *Every student sleeps.* In an extensional *wh*-complement, a DQP which supports a PL-like reading takes a property of answers as its scope. So in *John figured out what every student needed,* the scope of *every student* is the property P such that a student x has it iff John figured out what x needed. Quantification works identically to the declarative case.

There are both conceptual and empirical reasons for this approach. As to the former, extensional *wh*-complements are not direct questions; the matrix clause provides the extra material that we need to build the property to serve as nuclear scope for the quantifier. In the case of matrix interrogatives, there is no such material, unless we adopt the hypothesis of a "silent performative" verb.

The empirical reasons are given (i) by the different distribution of PL-like reading in embedded *wh*-complements with DQPs; and more importantly (ii) by the fact, to be considered in Section 8, that non-DQPs (such as CQPs) also support PL-like readings in embedded interrogatives. These QPs, however, do not support PL in matrix questions, even when they occur as subjects. Since QP-types do not essentially differ in their ability to support PL-like readings in extensional *wh*-complements, it must be concluded that this is an instance of their basic behavior, i.e., plain quantification.

Thus there are grounds to assume that PL in the matrix and in (extensional) *wh*-complements are distinct semantic phenomena. Szabolcsi's analysis of the semantics of PL-like readings in *wh*-complement clauses matches, essentially, the conclusions that can be reached when the syntactic (=LF) analysis developed in the previous sections is applied to the problem at hand.

Recall that we have assumed that there is a fundamental difference between matrix questions and (embedded) extensional *wh*-complements. The relevant syntactic difference is that only the CP of matrix interrogatives contains a question operator. This has been argued in Groenendijk and Stokhof (1984) and Munsat (1986) for interrogatives in general and in Szabolcsi (1996b) as regards PL in particular. The CP of extensional *wh*-complements does not contain a question operator. The CP of other types of embedded *wh*-clauses similarly lacks a question operator: relative clauses, comparative clauses, etc. I assume that, in the relevant respects, the C^0 operator which occurs with extensional *wh*-complements and relative/comparative clauses is akin to an existential operator.

The analysis of matrix PL (Section 5) relied on the assumption that *every N* gets bound by the question operator, which in turn pre-empts its moving to [Spec, DistP], enforcing a shift from the SD pattern of distributivity to the PD pattern. Since the examples in (81) do not involve a question operator, we expect *every N* to be bound by an existential operator; consequently, not to shift to the PD pattern. That is, we expect *every N* in extensional complements to behave like *each N* does in matrix questions (though for different reasons).

This expectation is borne out since PL-like readings in extensional *wh*-complements with *every*-QPs do not show the asymmetries found in matrix questions with *every*-QPs. In extensional complements, *every N* is thus able to satisfy SD by moving to [Spec, DistP], and distribute over the (reconstructed) WhQP. This accounts for the lack of asymmetries found with (81) above.

PL-like readings with extensional *wh*-complements can thus be derived as follows. Since it is bound by an existential operator, *every N* moves to [Spec, DistP]. The WhQP reconstructs in its scope (in [Spec, ShareP] or in its Case position). Therefore the SD pattern becomes instantiated. I present below the derivation of examples like (81b). Similar derivations obtain for the other examples in the paradigm in (81):[18]

(82) John found out who will present every paper
$[_{\text{AgrSP}}$ John $[_{\text{VP}}$ found out $[_{\text{RefP}}\ \exists_j\ [_{\text{CP}}\ t_i\ [_{\text{AgrSP}}\ t_i$ [DistP every book$_{k/j}$ $[_{\text{ShareP}}$ {who}$_i$ $[_{\text{AgrOP}}\ t_k$ [VP t_i present t_k]]]]]]]]

Consider, now, (intensional) *wh*-complements of *wonder,* which selects only for a [+Wh] complement. Szabolcsi demonstrates that the semantics of PL readings in this type of construction is parallel to the case of matrix questions. Interestingly, we find that the PL pattern of matrix questions with *every*-QPs re-emerges, too. This is shown in (82). (82b, c) can only be construed with the WhQP taking wide scope over *every*; but (82a, d) appear to allow for the reverse scoping:

(83) a. I wonder what book every student will write about [ok PL]
b. I wonder who read every book [* PL]
c. I wonder to whom you will assign every problem [?? PL]
d. I wonder who John introduced to every girl [? PL]

There are some other respects in which complements of *wonder* behave, syntactically, like matrix clauses (cf. Szabolcsi 1996a): one is that they can license negative polarity items, in the same rather restricted way in which *wh*-questions do (cf. Munsat 1986). Extensional *wh*-complements do not.

(84) a. When will I ever get a raise?
b. I wonder when I will ever get a raise
c. * I found out when I will ever get a raise

(85) a. Why did anybody bother to come?
b. I wonder why anybody bothered to come
c. * I know/found out why anybody bothered to come

[18]It is a separate question why *wh*-clauses introduced by *whether* do not show PL-like readings, not even in the argument configurations that normally support PL (cf. Karttunen and Peters 1980, Szabolcsi 1996a, and references therein):

(i) John found out whether everyone left
* 'John found out about everyone whether (s)he left'

Moltmann and Szabolcsi (1994) suggest that this is a subcase of a general prohibition against quantification into clauses lacking a variable-binding operator. *Whether*-complements are, semantically, the embedded version of yes/no questions. Referring back to our discussion of *every* in yes/no questions, our conclusion was that the yes/no question operator was not a variable-binding Operator, or at least that it couldn't bind the kind of variable introduced by DQPs. This explanation can be extended to *whether*-complements.

For these reasons, I assume that complements of *wonder* contain a question operator in C^0. This accounts for the distribution of PL readings in (82), which proceeds as outlined in Section 6 for matrix questions with DQPs.

8 PAIR-LIST READINGS WITH GQPS AND CQPS

In the discussion of PL so far, I have concentrated on DQPs. I now consider how the analysis can be extended to account for the availability of PL-like readings (and the lack thereof) with other QP types. This discussion provides another illustration of the diverse behavior of QP-types.

I begin by discussing CQPs and GQPs in matrix questions (Section 8.1, 8.2), then move on to embedded interrogatives (Section 8.3).

8.1 CQPs in matrix questions

It is well known that in matrix questions, CQPs do not support PL readings. Neither of (86a, b) can be answered as in (86c).

(86) a. What did more than two students write about? [* PL]

b. What did few students write about? [* PL]

c. * John wrote about *War and Peace,* Susan about *Buddenbrooks,* Bill about *Death in Venice,* ...

These facts follow directly from the account presented here. Since CQPs are interpreted as generalized quantifiers, they do not introduce a variable, and thus cannot be bound by the question operator in C^0.

As argued by Groenendijk and Stokhof (1984), Chierchia (1993), and Szabolcsi (1994) direct quantification into matrix questions is not an available option, for semantic reasons. Thus, some alternative mechanism must become available for a given QP type to support PL in matrix questions. This mechanism is taken here to be binding, by the question operator, of the variable introduced by the QP; for Chierchia, this mechanism consisted of absorption. It follows that a CQP is not expected to support PL even if the *wh*-element should reconstruct in its scope, as CQPs do not introduce a discourse referent.

Chierchia must stipulate that CQPs do not undergo absorption. Our account offers a more principled account of why it should be so, based on the theory of QP types of Section 4.

8.2 GQPs in matrix questions

The approach developed so far excludes the possibility that GQPs behave exactly like DQPs with respect to PL in matrix questions. This is because only DQPs introduce a set variable that can be bound by operators other than the existential; as claimed above, the group variable contributed by GQPs can only be bound by an existential operator. Thus the claim is that if PL-like readings are available with GQPs at all, these should be derived via a distinct mechanism.

Let's consider definite GQPs first. There is general consensus in the literature that "true" PL readings are not available in matrix questions with definite GQPs. There are a number of properties that distinguish the "apparently PL-like" answers supported by definites from PL answers with DQPs. First, as observed by Pritchett (1990), there are no asymmetries (of the type shown by *every*-QPs) in the availability of such answers. Pritchett's examples are as follows:

(87) a. What did the boys rent last night?
b. Who rented these movies last night?
c. OK/? John rented *Casablanca,* Moe rented *Red,* ...

Pritchett points out that both (87a, b) may be answered as in (87c).

The answers supported by definite GQPs further differ from those available with DQPs as follows: (i) answers to questions with definite GQPs allow for vagueness as to the actual pairings, and do not require exhaustive listing; (ii) definites do not seem to support list answers with singular WhQPs like *which/what book* (bare WhQPs, like *who/what* may be interpreted as plural).

Krifka (1991) and Srivastav (1992) argue that the PL-like answers in (87) are to be viewed as resulting from a reading of the question which is not PL, but rather, as arising from a type of "dependent plural" reading. PL-like answers with definite GQPs should be regarded as "cooperative" attempts to spell-out a dependency between two plural terms. This observation is found already in Szabolcsi (1983). An adequate answer to a question like (87a) would be "The boys (John, Moe, ...) rented *Casablanca, Red,* ... (between them)." But the answerer can choose to be more cooperative, i.e. informative, and spell-out some of the pairs ⟨boy, movie rented⟩.

Crucially, as remarked by Krifka and Srivastav, answers to questions as in (87) tolerate vagueness as to the actual pairings and do not require exhaustiveness. These observations argue that these answers do not spell-out a distributive dependency; but rather, something similar to a "cumulative" dependency (cf. Scha 1981).

It is unclear at present if cumulative dependencies correspond to a particular LF configuration; it seems that they do not require the QPs to assume a particular scopal order. They appear to fall outside our concerns here, therefore.

Let's turn now to indefinite GQPs. It has been claimed (cf. Groenendijk and Stokhof 1984, Higginbotham 1991) that indefinite GQPs support a type of list reading which is referred to as a "choice" reading, and that this is to be assimilated to PL-readings with DQPs. Consider:

(88) a. What (book) did two students write about?

b. ? (Well, for example) John wrote about *War and Peace* and Mary wrote about *Buddenbrooks.*

The question in (88a) can, marginally, be answered as in (88b). But unlike a question containing a DQP, like *What (book) did every/each student write about?,* (88a) does not have a unique and complete answer. To answer (88a), it has been claimed that one has to choose one group of two students (e.g. John and Mary) and answer about these. Answering a choice question amounts thus to answering a disjunction of questions by choosing one.

Szabolcsi argues that apparent list readings supported by indefinite GQPs are a separate phenomenon from PL readings. She offers both a speculative and an empirical argument. The reader is referred to Szabolcsi (1996a) for details. Her conclusion is that there is no actual choice, in the logical sense of answering a disjunction of questions by answering one of the disjuncts; rather, the answer is of one the "mention-some" type discussed by Groenendijk and Stokhof (1984).

On the basis of Szabolcsi's work, I conclude that neither CQPs nor GQPs support actual PL readings in matrix questions. This is consistent with our LF account.

8.3 CQPs and GQPs in embedded questions

I consider now PL-like readings with CQPs and GQPs in non-matrix questions; the focus, as in Section 7, will be on extensional *wh*-complements. The basic data comes again from Moltmann and Szabolcsi (1994).

Let's begin with CQPs. These authors note that although (subject) CQPs support PL-like answers in extensional complements (89), the syntactic distribution of such answers is narrower than with DQPs in the same syntactic environment, as shown by the contrast in (90):

(89) John found out which book/what more than five boys needed.

OK 'John found out about more than five boys which book/what each needed'

(90) a. I know which boy John introduced every girl to.

OK 'I know about every girl which boy John introduced her to'

b. I know which boy/who John introduced more than five girls to.

* 'I know for more than five girls which boy/who John introduced her to'

These data do not present a problem for our analysis. We may assume that in (89) *more than five boys* in AgrSP scopes over the *wh*-phrase reconstructed into its Case position, AgrOP, or in [Spec, ShareP], in accordance with our general assumptions about PL-like readings in extensional *wh*-complements (cf. Section 7).

In Section 4, I proposed that CQPs take scope *in situ,* and hence that distributive readings with CQPs are possible whenever the LF (Case) position of the CQP is higher than that of the indefinite which is the distributee. PL in (90b) is out, then, because the Case position where the WhQP reconstructs, i.e. [Spec, AgrIOP], is higher in the hierarchy than the LF scope position of the CQP, i.e. [Spec, AgrOP].

Similar observations apparently hold of GQPs in extensional *wh*-complements. The pattern of PL-like readings with GQPs in these contexts appears to match the PL pattern with *every N* in matrix clauses, both with bare and D-linked *wh*-phrases. Consider the following data:

(91) a. I found out what three students need [ok PL]
b. I found out who needs three books [* PL]
c. I found out to who(m) John showed five books [?* PL]
d. I found out who John introduced to five girls [?(?) PL]

(92) a. I found out which book three students need [?/ok PL]
b. I found out which student needs three books [* PL]
c. I found out to which student John showed five books [* PL]
d. I found out which boy John introduced to five girls [* PL]

The same account can be extended to these cases: when the WhQP is reconstructed to its Case position (if it is a bare WhQP, or to Spec of ShareP for a D-linked WhQP), PL-like readings will be supported (according to the PD pattern) insofar as the Case position of the GQP is higher than that of *wh.*

9 CONCLUSION

In this paper, I have focused on what I have taken to be a basic inadequacy of current LF theories of scope in natural language: the tenet that "all quantifiers are created equal." This empirical inadequacy is already present in standard accounts of QP/QP interactions (cf. Section 3). It is equally manifest in the account that these theories offer of Wh/QP interactions.

I have shown that on the basis of an alternative LF theory of scope (as proposed in Beghelli and Stowell 1994, 1996, and Beghelli 1995) which is in part based on Szabolcsi (1996b), we can account for a number of asymmetries in the distribution of PL. These asymmetries relate to contrasts between matrix and embedded questions, between DQPs and other types of QPs, and even between two apparently very similar quantifiers such as *every* and *each.* None of the current syntactic approaches to PL can straightforwadly account for this diversity.

The proposal presented in this paper relies also on a novel approach to the notion of distributivity (cf. Beghelli 1995), which has been shown to be a more complex phenomenon than standardly thought. In particular, this paper has tried to show that distributivity has a specific syntactic encoding in LF.

REFERENCES

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics* 2(2):83–135.

Aoun, Joseph, and Yen-hui Audrey Li. 1993. *Syntax of Scope.* Cambridge: MIT Press.

Barss, Andrew, and Howard Lasnik. 1986. A note on anaphora and double-objects. *Linguistic Inquiry* 17(2):347–354.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In Amy J. Schafer (Ed.), *Proceedings of the North East Linguistic Society 23, Volume One*, 65–80, Amherst. GLSA Publications, University of Massachussetts.

Beghelli, Filippo. 1995. *The Phrase Structure of Quantifier Scope.* PhD thesis, UCLA.

Beghelli, Filippo, and Tim Stowell. 1994. The direction of quantifier movement. *GLOW Newsletter* 32:56–57.

Beghelli, Filippo, and Tim Stowell. 1996. Distributivity and negation. In this volume.

Chierchia, Gennaro. 1992. Functional WH and weak crossover. In Dawn Bates (Ed.), *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 75–90, Stanford. CSLI.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1(2):181–234.

Choe, Jae-Woong. 1987. *Anti-Quantifiers and a Theory of Distributivity.* PhD thesis, University of Massachussetts, Amherst.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale and Samuel J. Keyser (Eds.), *The View from Building 20*, 1–52. Cambridge: MIT Press.

Cinque, Guglielmo. 1994a. Lectures presented at the Girona International Summer School in Linguistics.

Cinque, Guglielmo. 1994b. On the relative position of some lower adverbs in French and Italian. Unpublished manuscript.

Diesing, Molly. 1992. *Indefinites.* Cambridge: MIT Press.

Farkas, Donka. 1996. Evaluation indices and scope. In this volume.

Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers.* PhD thesis, University of Amsterdam.

Haegeman, Liliane. 1994. *The Syntax of Negation.* Cambridge University Press.

Higginbotham, James. 1991. Interrogatives I. In Lisa Cheng and Hamida Demirdash (Eds.), *More Papers on WH-Movement (MIT WPL 15)*, 47–76, Cambridge. MIT.

Higginbotham, James, and Robert May. 1981. Questions, quantifiers and crossing. *Linguistic Review* 1(1):41–80.

Hornstein, Norbert. 1995. *Logical Form: From GB to Minimalism.* Oxford: Blackwell.

Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar.* PhD thesis, MIT.

Kamp, Hans, and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Karttunen, Lauri, and Stanley Peters. 1980. Interrogative quantifiers. In Christian Rohrer (Ed.), *Time, Tense and Quantifiers: Proceedings of the Stuttgart Conference on the Logic of Tense and Quantification*, 181–206. Tübingen: Niemeyer.

É. Kiss, Katalin. 1987. *Configurationality in Hungarian.* Dordrecht: Kluwer.

Krifka, Manfred. 1989. Nominal reference, temporal constitution, and quantification in event semantics. In Renate Bartsch et al. (Eds.), *Semantics and Contextual Expression*, 75–115. Dordrecht: Foris.

Krifka, Manfred. 1991. Definite Noun-Phrases are not quantifiers. *Linguistic Inquiry* 23(2):156–163.

Larson, Richard. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21(4):589–632.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Rainer Bäuerle et al. (Eds.), *Meaning, Use, and Interpretation of Language*, 302–323. Berlin: Mouton de Gruyter.

May, Robert. 1985. *Logical Form: Its Structure and Derivation.* Cambridge: MIT Press.

Milsark, Gary. 1974. *Existential Sentences in English.* PhD thesis, MIT.

Moltmann, Friederike, and Anna Szabolcsi. 1994. Scope interactions with pair-list quantifiers. In Mercé González (Ed.), *NELS 24: Proceedings of the North East Linguistic Society, Volume Two*, 381–395. GLSA Publications, University of Massachussetts, Amherst.

Moritz, Luc, and Daniel Valois. 1994. Pied-piping and Specifier-Head agreement. *Linguistic Inquiry* 25(4):667–707.

Munsat, Stanley. 1986. Wh-complementizers. *Linguistics and Philosophy* 9(2):191–217.

Pesetsky, David. 1982. *Paths and Categories.* PhD thesis, MIT.

Pica, Pierre, and William Snyder. 1995. Weak crossover, scope and agreement in a minimalist framework. In Raul Aranovich et al. (Eds.), *The Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, 334–349, Stanford. CSLI.

Pritchett, Bradley L. 1990. A note on scope interaction with definite plural NPs. *Linguistic Inquiry* 21(4):646–654.

van der Putten, F. 1989. *Bepaaldheidm, Onbepaaldheid, Genericiteit en Specifiteit.* PhD thesis, University of Leiden.

Roberts, Craige. 1987. Distributivity. In Jeroen Groenendijk et al. (Eds.), *Proceedings of the 6th Amsterdam Colloquium*, 291–310. University of Amsterdam.

Ruys, Eddy Godehard. 1993. *The Scope of Indefinites.* PhD thesis, Research Institute for Language and Speech, Utrecht.

Safir, Ken, and Tim Stowell. 1989. Binomial *each.* In *NELS 18: Proceedings of the North East Linguistic Society.* GLSA Publications, University of Massachussetts, Amherst.

Scha, Remko. 1981. Distributive, collective, and cumulative quantification. In Jeroen Groenendijk et al. (Eds.), *Formal Methods in the Study of Language*, 483–512. Amsterdam: Mathematisch Centrum.

Schein, Barry. 1993. *Plurals and Events.* Cambridge: MIT Press.

Srivastav, Veneeta. 1992. Two types of universal terms in questions. In *NELS 22: Proceedings of the North East Linguistic Society*, 443–457. GLSA Publications, University of Massachussetts, Amherst.

Szabolcsi, Anna. 1983. Focussing properties, or the trap of first order. *Theoretical Linguistics* 10(2/3):125–145.

Szabolcsi, Anna. 1994. Quantifiers in pair-list readings and the non-uniformity of quantification. In Paul Dekker and Martin Stokhof (Eds.), *Proceedings of the Ninth Amsterdam Colloquium*, 645–664. ILLC/Department of Philosophy, University of Amsterdam.

Szabolcsi, Anna. 1996a. Quantifiers in pair-list readings. In this volume.

Szabolcsi, Anna. 1996b. Strategies for scope taking. In this volume.

Verkuyl, Henk. 1988. Aspectual asymmetry and quantification. In Veronika Ehrich and Heinz Vater (Eds.), *Temporalsemantik: Beiträge zur Linguistik der Zeitreferenz*, 220–259. Tübingen: Niemeyer.

Williams, Edwin. 1986. A reassignment of the functions of LF. *Linguistic Inquiry* 17(2):265–299.

Williams, Edwin. 1988. Is LF distinct from S-structure? Reply to May. *Linguistic Inquiry* 19(1):135–146.

Williams, Edwin. 1994. *Thematic Structure in Syntax.* Cambridge: MIT Press.

Zanuttini, Rafaella. 1991. *Syntactic Properties of Sentential Negation: A Comparative Study of Romance Languages.* PhD thesis, University of Pennsylvania.

11

QUESTIONS AND GENERALIZED QUANTIFIERS*

Javier Gutiérrez Rexach

Department of Linguistics
UCLA

1 INTRODUCTION

In this paper I present the basic elements of an analysis of matrix interrogatives within the theory of Generalized Quantifiers (Barwise and Cooper 1981, Keenan and Stavi 1986, Keenan and Westerståhl 1994). Current syntactic theories defend a uniform treatment of declarative quantifiers such as *every*, *some* and *most*, and interrogative quantifiers such as *which* or *what*. For instance, theories of quantification within the Government and Binding framework and later developments assume that movement of declarative and interrogative quantifiers is subject to the same constraints on well-formedness (subjacency, government, both types of quantifiers bind variables, etc.) at the relevant level of representation. Thus, at least from a syntactic point of view, it seems desirable to give a uniform account of the different types of quantification. From a semantic point of view, we also have strong arguments in favor of the proposed connection. It can be easily noticed that interrogative quantifiers have the property of Conservativity, a property satisfied by all declarative determiners (Keenan and Stavi 1986). This semantic property explains why sentence (1a) is equivalent to (1b), and sentence (1c) is equivalent to (1d).

(1) a. Some students are vegetarians.
 b. Some students are both students and vegetarians.
 c. Which students are vegetarians?
 d. Which students are both students and vegetarians?

*I wish to thank Ed Keenan for hours of discussions and extensive comments on several versions of this paper. I would also like to thank Anna Szabolcsi, Irene Heim and the audiences at the Fourth CSLI Workshop on Logic, Language and Computation and the Tenth Amsterdam Colloquium, especially Johan van Benthem, William Ladusaw and Barbara Partee for their comments.

A. Szabolcsi (ed.), Ways of Scope Taking, 409–452.

A seminal paper in the analysis of interrogative quantifiers from the perspective of Generalized Quantifiers (GQ) theory is Higginbotham and May (1981), and here I will develop some of their insights. A topic that has recently received broad attention is the problem of quantification into questions (Groenendijk and Stokhof 1984, Chierchia 1993, Szabolcsi 1994). Nevertheless, no general attempt has been made to explore whether the tools and assumptions behind GQ theory can be extended to the domain of interrogatives. Perhaps this is because some aspects of the semantics of questions seem inherently intensional, something which is orthogonal to the standard GQ approach to the semantics of natural language determiners. In other words, questions are a type of semantic object whose nature or complexity goes beyond the tools and strategies used in the GQ research program. However, on the approach to the semantics of interrogatives adopted here, the former claim loses part of its force, since we will be focusing on the properties of questions that pertain to interrogative quantifiers and their interactions with other quantifiers. The approach has two distinctive features: it is extensional and it combines insights of the categorial and the propositional approaches to the semantics of interrogatives.

This paper is organized as follows: in Section 2 the basic notions of GQ theory are introduced. Sections 3 to 7 explore the semantics of argument and modifier questions in more detail. In Section 8 the problem of how to extend the logical characterization of declarative determiners to the interrogative domain is discussed and monotonicity and entailment patterns are analyzed in detail. Section 9 deals with multiple questions and, finally, Section 10 presents some results on the interactions between declarative and interrogative quantifiers.

2 GENERALIZED QUANTIFIERS AND DETERMINERS

The theory of Generalized Quantifiers was born in the early eighties as a development of Montague's treatment of the semantics of noun phrases (see Westerståhl 1989 for a survey of the historical origins of the theory). The research in the first half of the decade focused on the linguistic and logical properties of declarative determiners in the set $[\mathcal{P}(E) \to [\mathcal{P}(E) \to 2]]$, i.e., functions from sets of individuals to functions from sets of individuals to truth values.[1] Later developments brought into the picture dynamic properties of determiners and other properties specific to polyadic determiners and plurals (see van Benthem and ter Meulen 1994) for a complete survey of current trends). In a global perspective, a declarative determiner or a generalized quantifier in

[1] It is common to refer to such determiners as type $\langle 1, 1\rangle$ determiners, following Lindström's notation. In general we write $[A \to B]$ for the set of functions from A to B.

$[\mathcal{P}(E) \to [\mathcal{P}(E) \to 2]]$ is a functional D that maps each universe E to a local quantifier $D_E \in [\mathcal{P}(E) \to [\mathcal{P}(E) \to 2]]$.[2]

DEFINITION 1 (PROPERTIES OF DECLARATIVE DETERMINERS' DENOTATIONS) *For an arbitrary universe E, let* $D \in [\mathcal{P}(E) \to [\mathcal{P}(E) \to 2]]$. *Then, writing* D *for* D_E *when no confusion results:*

a. D *is* conservative *iff* $\forall A, B, B' \subseteq E$, *if* $A \cap B = A \cap B'$ *then* $D(A)(B) = D(A)(B')$.

b. D *satisfies* extension *iff* $\forall A, B \subseteq E \subseteq E', D_E(A)(B) = D_{E'}(A)(B)$.

c. D *is* permutation invariant *iff for every permutation of E,* $\pi(E), \forall A, B \subseteq E, D(\pi(A))(\pi(B)) = D(A)(B)$.

d. D *is* intersective *iff* $\forall A, B, A', B' \subseteq E$, *if* $A \cap B = A' \cap B'$ *then* $D(A)(B) = D(A')(B')$.

e. D *is* co-intersective *iff* $\forall A, B, A', B' \subseteq E$, *if* $A - B = A' - B'$ *then* $D(A)(B) = D(A')(B')$.

f. D *is* cardinal *iff* $\forall A, B, A', B' \subseteq E$, *if* $|A \cap B| = |A' \cap B'|$ *then* $D(A)(B) = D(A')(B')$.

g. D *is* co-cardinal *iff* $\forall A, B, A', B'$, *if* $|A - B| = |A' - B'|$ *then* $D(A)(B) = D(A')(B')$.

h. D *is* proportional *iff* $\forall A, B \subseteq E, |A \cap B|/|A| > m/n$ *(or* $|A \cap B|/|A| \geq m/n$*) iff* $D(A)(B)$.

The reader is referred to Keenan and Westerståhl (1994) for a more detailed discussion of the constraints above and many others. In GQ theory, the denotations of natural language declarative determiners are given as follows:

DEFINITION 2 (DECLARATIVE DETERMINERS IN $[\mathcal{P}(E) \to [\mathcal{P}(E) \to 2]]$)

a. $[\![\textit{all}]\!](A)(B) = [\![\textit{every}]\!](A)(B) = 1$ *iff* $A \subseteq B$.

b. $[\![\textit{some}]\!](A)(B) = 1$ *iff* $A \cap B \neq \emptyset$.

c. $[\![\textit{no}]\!](A)(B) = 1$ *iff* $A \cap B = \emptyset$.

d. $[\![\textit{three}]\!](A)(B) = 1$ *iff* $|A \cap B| = 3$.

[2] When GQs of this type are isomorphism invariant, a global perspective is required. For the sake of simplicity, we are going to restrict ourselves here to a local perspective ignoring the parameter E.

e. $[\![\textit{all but three}]\!](A)(B) = 1$ *iff* $|A - B| = 3$.

f. $[\![\textit{most}]\!](A)(B) = 1$ *iff* $|A \cap B| > |A - B|$.
etc.

The determiners *some* and *no* are intersective, *all* is co-intersective. The determiner *three*, besides being intersective, is cardinal, and *all but three* is co-intersective and co-cardinal. Finally, *most* is a proportional determiner. It is a well known fact that in some cases the interpretation of quantifiers and especially certain determiners is strongly context dependent. Westerståhl (1985) introduces the notion of a *context set* to formalize the idea within the GQ framework. Since context restriction turns out to be essential for the interpretation of some interrogative quantifiers, we are going to study its characterizing properties in more detail.

DEFINITION 3 (CONTEXT RESTRICTION) *If D is a determiner and $A, B, X \subseteq E$, then D_E^X is the restriction of D to X defined as follows: $D_E^X(A)(B) = D_E(X \cap A)(B)$.*

Van Benthem (1986) characterizes a determiner as *logical* if it satisfies the properties of conservativity, extension and permutation invariance. Restricted determiners are in general not logical, in this sense, since they satisfy conservativity and extension but they are not in general permutation invariant.

FACT 4 *Context restriction preserves conservativity and extension*
Proof: For conservativity: let D_E be conservative. We show that for all $X \subseteq E$, D_E^X is conservative. $D_E^X(A)(B) = 1$ iff $D_E(X \cap A)(B) = 1$ iff $D_E(X \cap A)(X \cap A \cap B) = 1$ (by conservativity), iff $D_E(X \cap A)(A \cap B) = 1$ (again by conservativity), iff $D_E(A)(A \cap B) = 1$ (def. restriction). For X a set, we say that D^X satisfies extension iff for all $X, A, B \subseteq E \subseteq E'$, $D_E^X(A)(B) = D_{E'}^X(A)(B)$. Let D_E satisfy extension. We have to show that D_E^X satisfies extension. $D_E^X(A)(B) = D_E(X \cap A)(B)$ (def. restriction) $= D_{E'}(X \cap A)(B)$ (D_E satisfies extension) $= D_{E'}^X(A)(B)$ (def. restriction). □

FACT 5 *For $|E| > 1$, context restriction does not in general preserve permutation invariance (PI)*
Proof: For example, let $E = \{a,b\}$. Then D_E given by $D_E(A)(B) = 1$ iff $A = \emptyset$ is PI (and conservative). But D_E^X is not PI when $X = \{a\}$. Let $\pi(a) = b$, $\pi(b) = a$. Then, $D^{\{a\}}(\{a\})(\emptyset) = D(\{a\})(\emptyset) = 0$. But $D^{\{a\}}(\pi(\{a\})(\pi(\emptyset)) = D^{\{a\}}(\{b\})(\emptyset) = D(\{a\} \cap \{b\})(\emptyset) = D(\emptyset)(\emptyset) = 1$. □

We can formulate a weaker version of permutation invariance which is preserved under context restriction (see also Westerståhl 1985):[3]

DEFINITION 6 (LOCALITY) *Given $E, X \subseteq E$, D a determiner function, we say that D is PI at X iff $\forall \pi \in PERM(E)$ (i.e. for every permutation π of E), if $\pi(X) = X$ then $D^X(A)(B) = D^X(\pi(A))(\pi(B))$.*

Westerståhl claims that only the weaker property of Locality is preserved under context restriction. This is not a very surprising property, if we realize that the choice of a context set is determined directly or indirectly by previous context sets, the universe of discourse, the situation in which the sentence is uttered, etc. Restricted determiners can be considered logical constants in a local sense.

3 QUESTIONS AS FUNCTIONS

In the extensional theory of questions that we present here, questions will be defined as functions from sets of objects to truth values. Matrix interrogative sentences denote such functions. Specifically, a question of type $\langle\langle\alpha, t\rangle, t\rangle$ is a function (of a certain sort) from sets of objects in type α to truth values. This corresponds to the intuition that a speaker, in asking a question of type $\langle\langle\alpha, t\rangle, t\rangle$, is asking the hearer to identify a unique object of type $\langle\alpha, t\rangle$, i.e. a unique set of objects of type α.

DEFINITION 7 *A* question *of type $\langle\langle\alpha, t\rangle, t\rangle$ is a function $f \in [\mathcal{P}(\alpha) \rightarrow 2]$ such that $\exists! x \in \mathcal{P}(\alpha)$ such that $f(x) = 1$. We will call such an x the* answer A_f *of f. We write $[_q\mathcal{P}(\alpha) \rightarrow 2]$ for the set of questions of type $\langle\langle\alpha, t\rangle, t\rangle$.*

From the above definition, it follows that for an arbitrary question f its answer exists and is unique (no question has more than one answer). The intuition here is that the unique x a question f is true of is the complete true answer to f. Therefore, we follow Groenendijk and Stokhof (1984) in assuming that an essential ingredient of the definition of a question is that it is strongly exhaustive . Consider the question in (2a):

(2) a. Who came to the party?

b. John, Mary and Bill.

[3]The central idea behind the notion of locality is that D^X is PI if we restrict ourselves to permutations which fix X, since $D^X_E = D^X_X$ (Proof: $D^X_E(A)(B) = D_E(X \cap A)(B)$ (by def. restriction) $= D_E(X \cap A)(X \cap A \cap B)$ (by conservativity) $= D_X(X \cap A)(X \cap A \cap B)$(by extension) $= D_X(X \cap A)(B) = D^X_X(A)(B)$. $\square$

c. John and Mary.

At this point and for the sake of simplicity, we think of (2b) as denoting a three element set, and (2c) a two element set. In a state of affairs in which John, Mary and Bill are exactly the individuals who came to the party, the constituent response in (2b) denotes the answer set of the question that (2a) denotes. So the question denoted by the interrogative sentence (2a) maps {JOHN,MARY,BILL} to True, and the rest of the objects in $\mathcal{P}(E)$ to False. A *partial answer*, as the one denoted by the expression in (2c), is a subset of the answer set of the question. There are also other responses to (2a) which provide some pragmatic or semantic information about the answer set but do not constitute proper or even partial answers. Here are some:

(3) a. I have no idea.

b. You already know it.

c. Wow, what a question that is!

Although partial, non-canonical, and uninformative answers are possible answers to a question, they should not be considered as equal in status to complete true answers. The intuition that my approach builds on is that only complete true answers are the objects that fulfill the information gap represented by the question. On many concrete occasions of everyday life, we are forced to give partial or uninformative answers to questions, either because we do not have enough information or we want to hide something, etc., but in doing so we actually are not logically answering the question.[4]

An additional motivation for treating questions as functions from sets of objects (for instance, individuals) to truth values rather than as merely sets of individuals is that otherwise we are conflating the denotations of questions and relative clauses or free relatives (Cooper 1983, Jacobson 1995). In the NP *the students who came to the party* we may treat the relative clause *who came to the party* as denoting a function mapping the set of students to the set of

[4]The notion of a partial answer can be captured with a slight modification of the definition, as suggested above. A partial answer is a non-empty proper subset of the answer set. We could easily provide a three-valued semantics for questions if our intent were to consider also partial answers as basic. The definition of a question would be then as follows:

(i) $$f(X) = \begin{cases} 1 & \text{if } X = A_f \\ \% & \text{if } X \subset A_f \ \& \ X \neq \emptyset \\ 0 & \text{otherwise} \end{cases}$$

The requirement that a partial answer is non-empty is important because, otherwise, *nobody* would be as satisfactory as a partial answer to (2a) as (2c). On the other hand, with the 'proper subset' requirement we also predict that there are no partial answer sets of an empty answer set.

students who came to the party. But the unit constituted by a question and its answer is of a propositional nature, it should extensionally denote a truth value (or proposition) rather than a set. Interestingly, the fact that we are one type higher does not mean that we have an increase in expressive power, as the following observation shows:

FACT 8 $|[_q\mathcal{P}(\alpha) \to 2]| = |\mathcal{P}(\alpha)|$

An interrogative sentence may denote one question in a possible world (situation), and a different question in another possible world. For example, the denotation of the expression in (2c) is one element of the answer space of the question which can constitute its answer set in a different world. Let I be an index set. An *indexed question* f is a function from indices to questions: $f \in [I \to [_q\mathcal{P}(\alpha) \to 2]]$. There are other ways to let possible worlds enter the picture, such as those proposed by Karttunen (1977) and Groenendijk and Stokhof (1984). In this paper we will restrict ourselves to non-indexed questions. Taking English as our object language we are going to consider two basic types of questions: argument questions and modifier questions.

4 ARGUMENT QUESTIONS

DEFINITION 9 *An element of* $[_q\mathcal{P}(E) \to 2]$ *is a* (unary) argument question.

In general, argument interrogative quantifiers are functions from n-ary relations to questions. Unary argument interrogative quantifiers are functions from sets to unary argument questions. Unary argument interrogative determiners are functions from sets to unary argument interrogative quantifiers.

DEFINITION 10 (ARGUMENT INTERROGATIVE GQS)
$[\mathcal{P}(E) \to [_q\mathcal{P}(E) \to 2]]$ *is the set of* (unary) argument interrogative quantifiers. $[\mathcal{P}(E) \to [\mathcal{P}(E) \to [_q\mathcal{P}(E) \to 2]]]$ *is the set of* (unary) argument interrogative determiners.

For example, consider the following interrogative sentences:

(4) a. Who is smoking?

b. What student is smoking?

The *wh*-word *who* denotes an argument interrogative quantifier, as illustrated in (5a, b). *What* denotes an argument interrogative determiner (5c, d).

(5) a. $[\![\text{Who}]\!] \in [\mathcal{P}(E) \to [_q\mathcal{P}(E) \to 2]]$

b. $[\![\text{Who}]\!]([\![\lambda x.\text{Smoke}(x)]\!]) \in [_q\mathcal{P}(E) \to 2]$

c. $[\![\text{What}]\!] \in [\mathcal{P}(E) \to [\mathcal{P}(E) \to [_q\mathcal{P}(E) \to 2]]$

d. $[\![\text{What}]\!]([\![\lambda x.\text{Student}(x)]\!])([\![\lambda x.\text{Smoke}(x)]\!]) \in [_q\mathcal{P}(E) \to 2]$

Using uppercase letters to represent denotations in a fixed universe E we can define the following English argument interrogative quantifiers:

DEFINITION 11 (ENGLISH INTERROGATIVE QUANTIFIERS) *For all* $Z, Y, X \subseteq E$:
WHO(Y)(X) $= 1$ *iff PERSON* $\cap Y = X$
WHAT(Y)(X) $= 1$ *iff* $E \cap Y = X$
WHICH_n$^Z(Y)(X) = 1$ *iff* $Z \cap Y = X$ & $|X| = n$
WHICH_ONES$^Z(Y)(X) = 1$ *iff* $Z \cap Y = X$ & $|X| \geq 2$

Applying the above definition, we see that sentence (6a) denotes a function that maps the set PERSON $\cap$ IN_THE_CORRIDOR to 1 and any other set of individuals to 0. The calculation of the truth conditions of the interrogative sentence/response pair in (6) is as in (7):

(6) a. Who is in the corridor?
b. Fred and Bill

(7) WHO($\{x|x \in$ IN_THE_CORRIDOR$\}$)($\{[\![\text{Fred}]\!], [\![\text{Bill}]\!]\}$) $= 1$ iff
PERSON $\cap$ ($\{x|x \in$ IN_THE_CORRIDOR$\}$) $= \{[\![\text{Fred}]\!], [\![\text{Bill}]\!]\}$

The functions WHICH_*n* and WHICH_ONES are inherently restricted to context sets. Therefore, they cannot be uttered in "out of the blue situations". Consider the interrogative sentence (8a) and the answer (8b) in a context where we are talking about female students in our deparment. The relevant context set is $Z = \{x|x \in$ FEMALE_STUDENT$\}$ and the interpretation as in (9).

(8) a. Which three like tacos?
b. Jill, Jodie and Jenny

(9) WHICH_THREE^Z($\{x|x \in$ LIKE_TACOS$\}$)($\{[\![\text{Jill}]\!], [\![\text{Jodie}]\!], [\![\text{Jenny}]\!]\}$) $=$
1 iff
$\{x|x \in$ FEMALE_STUDENT$\} \cap \{x|x \in$ LIKE_TACOS$\}$
$= \{[\![\text{Jill}]\!], [\![\text{Jodie}]\!], [\![\text{Jenny}]\!]\}$

The quantifier WHAT normally takes the complement set of the set of persons in the model ($E \upharpoonright -$PERSON) as its restriction. Nevertheless, the range of WHAT seems to be wider. For instance, the following interrogative sentences can be answered with event denoting expressions:

(10) a. What is he doing?
b. He is reading a book.
c. What do you learn at school?
d. I learn to play the piano.

Another interesting difference between WHAT and WHO is that WHAT admits more easily a "de dicto" or "kind" reading of the answer (Cooper 1983, Heim 1987). Consider the following dialogues:

(11) a. What do you want?
b. A unicorn.

(12) a. Who do you need?
b. A secretary.

The "de dicto" / "kind" reading of (11b) is available or even preferred, whereas Heim (1987) observes that *a secretary* has to be specific or "de re" as an answer to (12a). However, Cooper warns against the conclusion that the availability of a "de dicto" reading depends on the nature of the *wh*-phrase in addition to there being an intensional verb. In (13b), in contrast to the previous example, the "de dicto" reading can be obtained.

(13) a. Who do you need to see?
b. A psychiatrist.

In order to account for event and kind answers to WHAT questions, one can extend the domain of quantification of WHAT to include not only first-order individuals but also kinds and events, as done in current theories of plurals. I leave the issue open for further research. We move now to the semantics of interrogative determiners, defined as follows:

DEFINITION 12 (ENGLISH INTERROGATIVE DETERMINERS)
For all $Z, Y, X, W \subseteq E, x, y \in E, m \in \mathcal{N}$:
$WHAT_{sg}(Z)(Y)(X) = 1$ *iff* $Z \cap Y = X \;\&\; |X| = 1$
$WHAT_{pl}(Z)(Y)(X) = 1$ *iff* $Z \cap Y = X \;\&\; |X| \geq 2$
$WHICH^{W}_{sg}(Z)(Y)(X) = 1$ *iff* $(W \cap Z) \cap Y = X \;\&\; |X| = 1$
$WHICH^{W}_{pl}(Z)(Y)(X) = 1$ *iff* $(W \cap Z) \cap Y = X \;\&\; |X| \geq 2$
$WHICH\_n^{W}(Z)(Y)(X) = 1$ *iff* $(W \cap Z) \cap Y = X \;\&\; |X| = n$
$HOW\_MANY(Z)(Y)(\{m\}) = 1$ *iff* $|Z \cap Y| = m$
$WHOSE(Z)(Y)(X) = 1$ *iff* $Z \cap Y = X \;\&\; \forall x \in X\, \exists y[Poss(y, x)]$[5]

[5] Here we understand Poss as a possession relation between individuals, the possessor and the possessed thing.

Which determiner expressions denote context dependent functions. The interrogative sentence (14) is anomalous in an "out of the blue" situation but becomes felicitous if a context set is provided by the previous discourse, as in (15). The difference between the argument interrogative determiners WHICH_{sg} and WHICH_{pl} lies in the additional condition imposed on the cardinality of their answer sets. Here we treat the difference as the semantic correlate of grammatical number in parallel to the contrast between singular and plural declarative determiners (THE_{sg} vs. THE_{pl}). Informally, the question denoted by (14) either poses a query about the set of students in the model or about a subset of those students that the speaker has in mind.

(14) Which students came to the party?

The latter reading is sometimes called a "partitive" reading. The descriptive intuition behind the term relies on the equivalence between the interpretation of *which students* and the interpretation of *which of the students*, as noted by Heim (1987). Let us consider first the interpretation of (14):

(15) Let $W = \{x | x \in \mathrm{LINGUIST}\}$
$\mathrm{WHICH}^{W}_{pl}(\mathrm{STUDENT})(\mathrm{COME})(X) = 1$ iff
$\{x|x \in \mathrm{LINGUIST}\} \cap \{x|x \in \mathrm{STUDENT}\} \cap \{x|x \in \mathrm{COME}\} = X$ &
$|X| \geq 2$

As observed in Barwise and Cooper (1981) and Keenan and Stavi (1986), declarative partitive determiners obey the restriction that only definite plural determiners can follow the preposition *of*. The same constraint surfaces in the interrogative domain.

(16) a. * Which of some/most/many/every/three students came to the party?
b. Which of the (ten)/John's (ten)/these ten students came to the party?

This fact suggests an analysis of *which of the* as a complex determiner, along the lines proposed by Keenan and Stavi for the declarative counterpart. One can check immediately that for $\alpha \in \{sg, pl\}$, $\mathrm{WHICH\_OF\_THE}^{W}_{\alpha}(Z)(Y)$, as defined below, and $\mathrm{WHICH}^{W'}_{\alpha}(Z)(Y)$ are the same question function when the contextual restrictions of the determiners are equal (W = W').

(17) a. $\mathrm{WHICH\_OF\_THE}^{W}_{sg}(Z)(Y)(X) = 1$ iff $(W \cap Z) \cap Y = X$ & $|X| = 1$
b. $\mathrm{WHICH\_OF\_THE}^{W}_{pl}(Z)(Y)(X) = 1$ iff $(W \cap Z) \cap Y = X$ & $|X| \geq 2$

The sentence in (14) and its counterpart with *which of the students* are not equivalent only if the contexts sets are different or if *which students* is interpreted as context neutral and *which of the students* as context dependent. In general English speakers tend to interpret *which Z* as context dependent, contrasting it to the context neutral *what Z*. There are other languages where this contrast has also a morphological reflection. For instance, in Spanish there are two different interrogative determiners: *qué*, which is equivalent to *what*, and *cuál*, which is equivalent to the context dependent determiner *which*. As predicted, *qué* cannot be used when a context set has been introduced (18a), and it cannot act as a partitive determiner (18b):

(18) a. Hay manzanas en la caja. ¿Cuál/*Qué quieres?
there-are apples in the box Which/What you-want

'There are apples in the box. Which one do you want?'

b. ¿Cuál/*Qué de los estudiantes?
which/what of the students

5 PLURAL QUESTIONS

We say that an interrogative sentence denotes a plural question iff it maps a collection of sets of individuals to 1 and any other collection to 0.

DEFINITION 13 $[_q\mathcal{P}(\mathcal{P}(E)) \rightarrow 2]$ *is the set of* (unary) plural argument questions

Consider the following sentences:

(19) What students gathered in the plaza?

(20) Who carried the piano upstairs?

As a plural question, sentence (19) is asking for the groups of students that gathered in the plaza. Similarly, (20) can be interpreted as a question about the group(s) of persons that collectively carried the piano upstairs. If we analyze plural determiners as functions from sets of individuals (properties) to functions from collections of sets of individuals (plural properties) to truth values, we can extend the same treatment to the analysis of the plural readings associated with interrogative quantifiers. This line of analysis has been proposed for declarative determiners, among others, by van Benthem (1991) and van der Does (1992). Its extension to the interrogative domain seems to be straightforward. In other words, plural interrogative quantifiers do not seem to exclude any of the readings associated with declarative ones (distributive, collective or neutral, see van der Does 1992). Let us consider first the collective lifts of interrogative quantifiers and determiners. For any quantifier Q or determiner D we write $\mathbf{C}(Q)$,$\mathbf{C}(D)$ for the collective lift of the interrogative quantifier and determiner respectively. Here are two examples:

DEFINITION 14 (COLLECTIVE LIFTS) *Let* $Z, W \subseteq E, Y, X \subseteq \mathcal{P}(E)$. *Then,*

(i) $\mathbf{C}(\mathit{WHO}) \in [\mathcal{P}(\mathcal{P}(E)) \to [_q\mathcal{P}(\mathcal{P}(E)) \to 2]]$
$\mathbf{C}(\mathit{WHO})(Y)(X) = 1$ *iff* $X = \{W | W \subseteq \mathit{PERSON}\ \&\ W \in Y\}$

(ii) $\mathbf{C}(\mathit{WHAT}_{pl}) \in [\mathcal{P}(E) \to [\mathcal{P}(\mathcal{P}(E)) \to [_q\mathcal{P}(\mathcal{P}(E)) \to 2]]]$
$\mathbf{C}(\mathit{WHAT}_{pl})(Z)(Y)(X) = 1$ *iff* $X = \{W | W \subseteq Z\ \&\ W \in Y\}$

The collective lift of WHO, $\mathbf{C}$(WHO), is a function from collections of sets of individuals to plural questions. The collective lift of WHAT$_{pl}$, $\mathbf{C}$(WHAT$_{pl}$), is a function that maps a property Z to a collective interrogative quantifier $\mathbf{C}$(WHAT$_{pl}$)(Z). The truth conditions of (19) and (20) are as follows:

(21) a. $\mathbf{C}(\mathrm{WHAT}_{pl})(\mathrm{STUDENT})(\mathrm{GATHER})(X) = 1$ iff
$X = \{W | W \subseteq \mathrm{STUDENT}\ \&\ W \in \mathrm{GATHER}\}$

b. $\mathbf{C}(\mathrm{WHO})(\mathrm{CARRY\_THE\_PIANO})(X) = 1$ iff
$X = \{W | W \subseteq \mathrm{PERSON}\ \&\ W \in \mathrm{CARRY\_THE\_PIANO}\}$

The answer set of the plural question denoted by (19) is the collection of sets of students in the extension of the plural property GATHER. Therefore, in a situation where John got together with Bill and Sam, and Susan got together with Pam and Joe, the collection {{JOHN, BILL, SAM},{SUSAN, PAM, JOE}} would be the answer set of (19). Consider now (22):

(22) What students ate pizza?

The distributive interpretation of sentence (22) is a plural question true of the collection of singletons of students who ate pizza. Again, for Q an interrogative quantifier and D an interrogative determiner function, we write $\mathbf{D}$(Q) and $\mathbf{D}$(D) for the distributive lifts of Q and D respectively.

DEFINITION 15 (DISTRIBUTIVE LIFTS) *Let* $Z, W \subseteq E,\ Y, X \subseteq \mathcal{P}(E)$, *and* $AT(Z) = \{W | W \subseteq Z\ \&\ |W| = 1\}$. *Then,*

(i) $\mathbf{D}(\mathit{WHO}) \in [\mathcal{P}(\mathcal{P}(E)) \to [_q\mathcal{P}(\mathcal{P}(E)) \to 2]]$
$\mathbf{D}(\mathit{WHO})(Y)(X) = 1$ *iff* $X = \mathit{AT}(\mathit{PERSON}) \cap Y$

(ii) $\mathbf{D}(\mathit{WHAT}_{pl}) \in [\mathcal{P}(E) \to [\mathcal{P}(\mathcal{P}(E)) \to [_q\mathcal{P}(\mathcal{P}(E)) \to 2]]]$
$\mathbf{D}(\mathit{WHAT}_{pl})(W)(Y)(X) = 1$ *iff* $X = \mathit{AT}(Z) \cap Y$

The intended interpretation of (22) is:

(23) $\mathbf{D}(\mathrm{WHAT}_{pl})(\mathrm{STUDENTS})(\mathrm{EAT\_PIZZA})(X) = 1$ iff
$X = \mathrm{AT}(\mathrm{STUDENTS}) \cap \mathrm{EAT\_PIZZA}$

6 ANSWERS AND LINGUISTIC RESPONSES

6.1 Questions in different types and exhaustivity

In principle, there are no logical restrictions as to what types can be suitable answer spaces of a question. One can construct languages in which there are questions over all the denotable types of the language. This is clearly not the case in natural languages, where only a subset of the denotable types are suitable answer sets. Furthermore, the problem arises as to whether pursuing an "orthodox" categorial approach is the best strategy, namely whether different linguistic responses should give rise to questions in different types despite the uniformity in the *wh*-word used. Let us pursue this line of reasoning to see its disadvantages. A sensible, but stronger than needed, working hypothesis would be that only the lexical categories of a natural language are questionable, whereas functional categories are not. It seems that in general, there are no questions over the denotations of prepositions, modals or verbal auxiliaries. Nevertheless, the generalization is not completely accurate. Consider the category of determiners. This category of expressions is regarded as a functional category by current syntactic theories. But one could argue that *how many*-questions can be conceived as questioning determiners, since numerals are canonical responses to them and numerals denote determiners. A similar argument can be constructed for the case of *whose* questions. In other languages, the denotations of other determiner expressions are suitable answers. Therefore we should propose at least the following question types:

DEFINITION 16 (DETERMINER AND GENERALIZED QUANTIFIER QUESTIONS)

$[_q[\mathcal{P}(E) \rightarrow [\mathcal{P}(E) \rightarrow 2]] \rightarrow 2]$ *is the set of determiner questions.*
$[_q[\mathcal{P}(E) \rightarrow 2] \rightarrow 2]$ *is the set of generalized quantifier questions.*

Nevertheless, this would not be the end of the story, since there also seem to be prosodic restrictions constraining the set of questionable types. Only strings which are "prosodic phrases" (Zec and Inkelas 1990) are good constituent responses and, therefore, their denotation types are answer sets. Prepositions, auxiliaries and clitics, for example, cannot constitute prosodic phrases and thus their denotation types are *not* answer sets. Conversely, one can observe that in some types of echo questions or correction statements even bound morphemes can become prosodic phrases. As a consequence, the following generalization seems to emerge:

GENERALIZATION 17 (PROSODIC EFFABILITY) *A type $[_q\alpha \rightarrow 2]$ is an interrogative type in a language $\mathcal{L}$ if for an arbitrary question $f \in [_q\alpha \rightarrow 2]$, the answer A_f can be expressed as a prosodic phrase in $\mathcal{L}$.*

We have reached a point in which it seems evident that we are letting prosodic and syntactic factors determine the semantic type of a question. It is also clear that some cross-linguistic variation is to be expected. Nevertheless, our claim is that the orthodox categorial approach is missing something important, since we have the intuition that no matter whether the answer is an expression whose denotation is a determiner, a generalized quantifier or a proposition, the question belongs to a unique semantic type, or what we have called an argument question. In addition, it can be shown that there is only one case where determiner questions or generalized quantifier questions will be well defined as questions. In other words, there will be a unique determiner or generalized quantifier mapped to true, satisfying the exhaustivity requirement incorporated in the definition of a question. Consider the following question/answer pairs:

(24) a. How many apples are in the bag?

b. Six /??At least six /*Most.

The example in (24) illustrates the fact that *how many*-questions can only be answered with cardinal determiners. Moreover, only cardinal determiners of the form EXACTLY_n constitute genuine complete true answers. To see this point, consider *at least six* as the answer of (24a). As discussed at the beginning of the paper answers of this sort do not resolve the question properly since they are compatible with there being exactly six apples in the bag or two thousand. In this respect, they are partial answers and do not resolve the question completely.[6] Now we define HOW_MANY$(Z)(Y)$ as a determiner question:

[6] A more complex case are questions like the following: *How many men do you need to carry the piano?* Here, a response such as *at least three* seems to be more felicitous than as an answer to (24a). A possible explanation of the felicity of monotone cardinal determiners as answers to questions of this sort is that the alleged anomaly is caused by the presence of a modality operator: in all possible worlds accessible from the speaker's world only three men or more would be able to lift the piano. Similar questions are:

(i) a. How many people fit in the room?

b. No more than twelve.

(ii) a. How many days are necessary to build this barn?

b. At least two.

Answers (ib) and (iib) provide evaluations for potential alternative situations, and in fact this is what a speaker is looking for when asking (ia) and (iia). Therefore, sentences of this sort cannot be taken as an argument against strong exhaustivity. Since there is always a modal predicate or operator in this type of sentence, the non-exhaustivity effect can be attributed to this fact. Actually, in all potential alternative situations the question would be exhaustive (the answer set would be unique).

DEFINITION 18 *For all determiners* $D \in \{EXACTLY\_n : n \in \mathcal{N}\}$, *all* $Z, Y \subseteq E$:
$HOW\_MANY(Z)(Y) \in [_q\mathcal{P}(E) \to [\mathcal{P}(E) \to 2]]$ *and*
$HOW\_MANY(Z)(Y)(D) = 1$ *iff* $D(Z)(Y) = 1$

In general, questions of this class are not limited to answers in the form of determiner expressions. They can also be answered with noun phrases, like *six apples.* The function HOW_MANY as defined above does not have generalized quantifers in its domain. Therefore, we have to extend it to a function HOW_MANY* whose domain includes also generalized quantifiers of the form EXACTLY_$n(Z)$. In this case the strong exhaustivity condition is also satisfied by HOW_MANY*.

DEFINITION 19 *Let* $GQ^{EX} = \{EXACTLY\_n(Z) | Z \subseteq E\}$ *Then,*
$Dom(HOW\_MANY^*(Z)(Y)) = Dom(HOW\_MANY(Z)(Y)) \cup GQ^{EX}$,
$HOW\_MANY^*(Z)(Y)(D) = HOW\_MANY(Z)(Y)(D), \forall D \in$
$Dom(HOW\_MANY(Z)(Y)) HOW\_MANY^*(Z)(Y)(Q) = 1$ *iff*
$\exists n[Q = EXACTLYn(Z)\ \ \&\ \ HOWMANY(Z)(Y)(EXACTLYn) = 1]$

Consider now the case of *whose*-questions, where also determiner and quantifier expressions are good constituent responses (25). The determiner question function WHOSE and its extension WHOSE* would be defined in a similar fashion:

(25) a. Whose cats are on the mat?
b. John's / His / John's cats / *Every cat.

DEFINITION 20 *For all* $D \in POSS$[7] , *all* $Z, Y \subseteq E$,
$WHOSE(Z)(Y) \in [_q\mathcal{P}(E) \to [\mathcal{P}(E) \to 2]]$ *and*
$WHOSE(Z)(Y)(D) = 1$ *iff* $D(Z)(Y) = 1$

DEFINITION 21 *Let* $GQ^{POSS} = \{D(Z) | Z \subseteq E\ \ \&\ \ D \in POSS\}$. *Then,*
$Dom(WHOSE^*(Z)(Y)) = Dom(WHOSE(Z)(Y)) \cup GQ^{POSS}$
$WHOSE^*(Z)(Y)(D) = WHOSE(Z)(Y)(D),\ \forall D \in Dom(WHOSE(Z)(Y))$
$WHOSE^*(Z)(Y)(Q) = 1$ *iff* $\exists D[Q = D(Z)\ \ \&\ \ WHOSE(Z)(Y)(D) = 1]$

In the case of WHOSE and WHOSE* questions strong exhaustivity is not satisfied. Consider a situation in which John and Bill possess the same things ($\{y | Poss(\text{JOHN}, y)\} = \{y | Poss(\text{BILL}, y)\}$). Then, the determiner functions JOHN'S and BILL'S are true answers to (25a) and the strong exhaustivity condition is violated. Therefore, WHOSE(A)(B) is not well defined as a determiner question and neither is WHOSE* as a generalized quantifier question.

[7]Let POSS = $\{x's | x \in E\}$, where x's (A)(B) = 1 iff $\forall y \in A[Poss(x, y)]\ \ \&\ \ \text{THE}(A)(B)$ (see Keenan and Stavi 1986).

6.2 Question resolution

We can also consider the relation between a question and its linguistic answer as indirect. It is mediated by a resolution relation. Question resolution is a natural mechanism since from the denotation of noun phrases we can recover sets (a noun phrase denotes a set of sets). Therefore, one can recover answer sets in $\mathcal{P}(E)$ from answers in $[\mathcal{P}(E) \rightarrow 2]$, and also from answers in $[\mathcal{P}(E) \rightarrow [\mathcal{P}(E) \rightarrow 2]]$. The three function types correspond to a single class of expressions: argument interrogatives, i.e., those that question one argument of the relation.[8] Given a question f, we need to recover the answer set A_f from an expression ϕ, the linguistic answer, whose type does not match the type of the domain of f. All that answers do is to resolve the question by providing its answer set. A generalized quantifier $D(Z)$—the denotation of an NP constituent response—resolves a question f iff one its elements which is a subset of the restrictor set (a witness in Barwise and Cooper's terminology) is the answer set of the question.

DEFINITION 22 *Let* $D(Z) \in [\mathcal{P}(E) \rightarrow 2]$. *Then, for all* $f \in [_q\mathcal{P}(E) \rightarrow 2]$, $\text{Resolve}(D(Z), f)$ *iff* $A_f \subseteq Z \ \ \& \ \ D(Z)(A_f) = 1$

As an illustration of the process involved, consider the question-answer pair in (26):

(26) a. What did you put on the table?
b. Three forks

Applying the above definition, the generalized quantifier denoted by *three forks* resolves the question if and only if one of its elements is the answer set of the question:

(27)
$$\begin{aligned}
&\text{WHAT}([\![\lambda x.\text{You put } x \text{ on the table}]\!])(X) = 1 \text{ iff}\\
&E \cap [\![\lambda x.\text{You put } x \text{ on the table}]\!] = X\\
&\text{Resolve}([\![\text{three forks}]\!], \text{WHAT}([\![\lambda x.\text{You put } x \text{ on the table}]\!])) \text{ iff}\\
&A_{\text{WHAT}([\![\lambda x.\text{You put } x \text{ on the table}]\!])} \subseteq \text{FORK} \ \&\\
&\text{THREE (FORK)}(A_{\text{WHAT}([\![\lambda x.\text{You put } x \text{ on the table}]\!])}) = 1 \text{ iff}\\
&[\![\lambda x.\text{You put } x \text{ on the table}]\!] \subseteq \text{FORK} \ \&\\
&\text{THREE(FORK)}([\![\lambda x.\text{You put } x \text{ on the table}]\!]) = 1 \text{ iff}\\
&[\![\lambda x.\text{You put } x \text{ on the table}]\!] \subseteq \text{FORK} \ \&\\
&|\text{FORK} \cap ([\![\lambda x.\text{You put } x \text{ on the table}]\!])| = 3
\end{aligned}$$

[8]The need for several types in relation to a single expression suggests a treatment in terms of type-shifting operations (Partee 1987). The question types in the above definitions are related as follows:

$[_q[\mathcal{P}(E) \rightarrow 2] \rightarrow 2]$ $\qquad$ $[_q[[\mathcal{P}(E) \rightarrow 2] \rightarrow 2] \rightarrow 2]$

$[_q\mathcal{P}(E) \rightarrow 2]$

Notice that more than one generalized quantifier can resolve the same question, as long as the answer set is an element of the resolving quantifiers and a subset of their respective restrictors. The definition of question resolution is also related to exhaustivity.

FACT 23 (EXHAUSTIVITY AND RESOLUTION)
If $D(Z)$ resolves f with X, then $\neg\exists Y$ such that $X \subset Y$ and $D(Z)$ resolves f with Y

Question resolution by determiners seems to pose a problem, since they are not sets of sets. Therefore, one cannot recover a set from their denotation and check whether this set is the answer set of the question. We claim that this is precisely the reason why determiner responses are so scarce. In English, only *how many-* and *whose-*interrogatives admit them clearly. In Spanish and other Romance languages, there is a wider variety of determiners that can occur as constituent responses:

(28) a. ¿Quiénes vinieron a la fiesta?
'Who came to the party?'

b. Algunos/ Muchos/ todos...
some-pl. many-pl. all
'Some people/ many people/ Everybody/...'

Only context-dependent determiners occur as constituent responses to argument interrogatives. These determiners are relativized to context sets and behave like generalized quantifiers in disguise. A type lowering operation of pronominalization(Pron) provides the restrictor of the generalized quantifier: for D a determiner, A a context set, $Pron(D) = D^A(A) = D(A)$. A question f is resolved by a pronominalized determiner *Pron(D)* iff the answer set of f, A_f, is a subset of the context set A and an element of $Pron(D)$. As in the case of standard GQs, we will say then that the pronominalized determiner resolves the question f.

DEFINITION 24 *Let $Pron(D) = D(A)$, for A a context set, and $f \in [_q\mathcal{P}(E) \rightarrow 2]$. Then,*
$Resolve(Pron(D), f)$ iff $A_f \subseteq A$ & $Pron(D)(A_f) = 1$

7 MODIFIER QUESTIONS

The standard analysis of modifiers, for instance in Keenan and Faltz (1985), is to treat them as denoting functions in $[\mathcal{P}(E^n) \rightarrow \mathcal{P}(E^n)]$, for $n \geq 0$. Nevertheless, modifer interrogative quantifier expressions like *where, when*, etc. behave more like true quantifiers. They quantify over different domains (times, places, manners) and are treated as variable binding operators in grammatical theories that posit logical form representations. Therefore, it seems that what is needed is to conceive of modifiers not as maps from n-ary relations to n-ary relations but as arguments of the relation that we can question or quantify over (McConnell-Ginet 1982). Our representation language needs to be extended to a many sorted language with models $\mathcal{M} = \langle\langle E, \langle\mathcal{D}_j\rangle_{j\in S}\rangle, \mathcal{I}\rangle$, where S is an index set of sorts and for each $j \in S$,$\mathcal{D}_j$ is a non-empty set. For instance, $\mathcal{D}_l$ = PLACE is the set of locations in the model) , $\mathcal{D}_t$ = TIME is the set of times in the model, etc. It seems reasonable to assume that these domains have a rich underlying structure (see Szabolcsi and Zwarts 1993 for manners and times, Nam 1995 for locatives). Therefore, we have to shift the type of i-ary relations to $E^i \times \prod_j \mathcal{D}_j$. Writing s for $|S|$ here and later, relation-denoting expressions now denote sets of i-tuples of individuals and s-tuples of modifiers. By adopting this view we make the so-called adjuncts or modifiers into arguments. Therefore, adding modifier variables or constants to a relation increases its arity as follows:

Definition 25 (Argument extension of a relation by modifiers) *For all* $R \subseteq E^i$, $M \subseteq \prod_j \mathcal{D}_j$, *M is the* argument extension *of a relation R to a i+s-ary relation* $R' \subseteq E^i \times \prod_j \mathcal{D}_j$ *iff*
$R' = \{\langle\alpha_1, ..., \alpha_i, \mu_1, ..., \mu_s\rangle | \langle\alpha_1, ..., \alpha_i\rangle \in R \ \& \ \langle\mu_1, ..., \mu_s\rangle \in M\}$

We are now in a position to introduce the notion of an argument extended question, and the definitions of argument extended interrogative quantifiers and determiners in the new type. For brevity we will also call this type of quantifier and determiner modifier interrogative quantifiers and modifier interrogative determiners respectively.

Definition 26 (Unary modifier questions)
$\bigcup_{j\in S}[_q\mathcal{P}(\mathcal{D}_j) \to 2]$ *is the set of* (unary) modifier questions.

Definition 27 (Modifier interrogative quantifiers) *Let* $\mu_l \in$ *PLACE*, $\mu_t \in$ *TIME*, $\mu_m \in$ *MANNER*, $\mu_c \in$ *CAUSE*, $\mu_r \in$ *REASON*. *Then, for all* $n \geq 1$, *all* $R' \subseteq E^i \times \prod_j \mathcal{D}_j$, *and all* $X \subseteq \bigcup_{j\in S} \mathcal{D}_j$:
$WHERE(R')(\alpha_1, ..., \alpha_i)(X) = 1 \text{ iff } \{\mu_l | \langle\alpha_1, ..., \alpha_i, ..., \mu_l, ..., \mu_s\rangle \in R'\} = X$
$WHEN(R')(\alpha_1, ..., \alpha_i)(X) = 1 \text{ iff } \{\mu_t | \langle\alpha_1, ..., \alpha_i, ..., \mu_t, ..., \mu_s\rangle \in R'\} = X$
$HOW(R')(\alpha_1, ..., \alpha_i)(X) = 1 \text{ iff } \{\mu_m | \langle\alpha_1, ..., \alpha_i, ..., \mu_m, ..., \mu_s\rangle \in R'\} = X$
$WHY_c(R')(\alpha_1, ..., \alpha_i)(X) = 1 \text{ iff } \{\mu_c | \langle\alpha_1, ..., \alpha_i, ..., \mu_c, ..., \mu_s\rangle \in R'\} = X$
$WHY_r(R')(\alpha_1, ..., \alpha_i)(X) = 1 \text{ iff } \{\mu_r | \langle\alpha_1, ..., \alpha_i, ..., \mu_r, ..., \mu_s\rangle \in R'\} = X$

For $j \in S$, and $X, Y \subseteq \mathcal{D}_j$, let $X \cap_j Y$ be the meet (glb) of X and Y in the lattice with domain $\mathcal{P}(\mathcal{D}_j)$.

DEFINITION 28 (MODIFIER INTERROGATIVE DETERMINERS) *For all* $Z \subseteq E$, *all* $R' \subseteq E^i \times \prod_j \mathcal{D}_j$, *and all* $X \subseteq \bigcup_{j \in S} \mathcal{D}_j$:
$IN\_WHICH(Z)(R')(\alpha_1, ..., \alpha_i)(X) = 1$ *iff*
$\rho(Z) \cap_l \{\mu_l | \langle \alpha_1, ..., \alpha_i, ..., \mu_l, ..., \mu_s \rangle \in R'\} = X$[9]
$FOR\_WHAT(REASON)(R')(\alpha_1, ..., \alpha_i)(X) = 1$ *iff*
$REASON \cap_r \{\mu_r | \langle \alpha_1, ..., \alpha_i, ..., \mu_r, ..., \mu_s \rangle \in R'\} = X$
$AT\_WHAT(TIME)(R')(\alpha_1, ..., \alpha_i)(X) = 1$ *iff*
$TIME \cap_t \{\mu_t | \langle \alpha_1, ..., \alpha_i, ..., \mu_t, ..., \mu_s \rangle \in R'\} = X$
$IN\_WHICH(MANNER)(R')(\alpha_1, ..., \alpha_i)(X) = 1$ *iff*
$MANNER \cap_m \{\mu_m | \langle \alpha_1, ..., \alpha_i, ..., \mu_m, ..., \mu_s \rangle \in R'\} = X$

The truth conditions of the interrogative sentence (29a), in which *when* denotes in the type of modifier (argument extended) interrogative quantifiers, are in (29b):

(29) a. When did John arrive?
b. WHEN (ARRIVE')(JOHN)(X) = 1 iff
$\{\mu_k | \langle \text{JOHN}, \mu_1, ..., \mu_k, ..., \mu_s \rangle \in \text{ARRIVE}'\} = X$

A qualification on exhaustivity is needed at this point. The fact that not only *at 10 am* but also *on Monday* or *last week* are also proper responses if John arrived at 10am on Monday last week does not count as evidence against exhaustivity, since all these responses are more or less specific descriptions of the moment of time in which John's arrival took place. Otherwise, they do not resolve the question. Not considered in the previous definition are degree questions, a special class of modifier questions. Assuming that degree predicates like *tall* denote functions from individuals to degrees, we say that a property G is *gradable* iff $G \subseteq E \times \mathcal{D}_d$, where $\mathcal{D}_d$ is a domain of degrees δ.

DEFINITION 29 (DEGREE QUESTIONS) *For all* $G \subseteq E \times \mathcal{D}_d$, *all* $R' \subseteq E^i \times \prod_j \mathcal{D}_j$, *all* $\delta \in \mathcal{D}_d$ *and all* $X \subseteq \bigcup_{j \in S} \mathcal{D}_j$:
$HOW(G)(R')(\alpha_1, ..., \alpha_i)(X) = 1$ *iff*
$\{\delta | \exists x. \langle x, \delta \rangle \in G\} \cap_d \{\delta | \langle \alpha_1, ..., \alpha_i, ..., \delta, ..., \mu_s \rangle \in R'\} = X$

Sentence (30a) is a degree question. Its answer set is the set of degrees in the meet(glb) of $\{\delta | \exists x[\langle x, \delta \rangle \in \text{FAST}]\}$ and $\{\delta | \langle \text{JOHN}, ..., \delta, ..., \mu_s \rangle \in \text{RUN}'\}$.

(30) a. How fast did John run?

[9] I take ρ to be an operator mapping sets of individuals to the region of space (location) occupied by those individuals. For example, ρ(BOX) is the region of space occupied by the denotation of *box* in the model. Note that, for all Z, $\rho(Z) \subseteq$ PLACE. See Nam (1995) for further details.

b. $\text{HOW}(\text{FAST})(\text{RUN}')(\text{JOHN})(X) = 1$ iff
$\{\delta | \exists x.\langle x, \delta\rangle \in \text{FAST}\} \cap_d \{\delta | \langle \text{JOHN}, ..., \delta, ..., \mu_s\rangle \in \text{RUN}'\} = X$

The approach to questions that I am developing gives an adequate semantics to interrogative sentences in which preposition stranding has taken place. The descriptive generalization for the contrast in (31) is that in sentence (31a) the preposition stays in its place and the *wh*-word "moves" to sentence initial position, leaving the preposition stranded. In sentence (31b) the whole PP has moved to the initial position.

(31) a. Which box did you put my shoes in?
b. In which box did you put my shoes?

Question (31a) is an argument question, whereas question (31b) is a modifier question. As a matter of fact, one can observe that their linguistic answers are different. (32a) is an adequate answer to (31a) and (32b) is an adequate answer to (31b), but they cannot be interchanged.

(32) a. This box.
b. In this box.

The truth conditions of the interrogative sentences in (31) are as follows:

(33) a. $\text{WHICH}([\![\lambda x.\text{Box}(x)]\!])([\![\lambda\, x.\text{You put my shoes in}(x)]\!])(X) = 1$ iff
$\text{BOX} \cap \{x | \text{You put my shoes in } (x)\} = X$
b. $\text{IN\_WHICH}([\![\lambda x.\text{Box}(x)]\!])([\![\lambda \mu_l.\text{You put my shoes } (\mu_l)]\!])(X) = 1$ iff
$\rho([\![\lambda x.\text{Box}(x)]\!]) \cap_l \{\mu_l | \langle [\![\text{you}]\!], [\![\text{my shoes}]\!], ..., \mu_l, ..., \mu_s\rangle \in \text{PUT}'\} = X$

The equivalence of the stranded preposition and non-stranded preposition interpretations is not immediate. It comes from the general equivalence between *which*-questions and *in which*-questions when the ρ operator (Nam 1995) is applied to the answer set of the first one: $\rho(A_{\text{WHICH}(Z)(Y)}) = A_{\text{IN\_WHICH}(Z)(Y)}$. In other words, answers (32a) and (32b) are spatially equivalent.

8 PROPERTIES OF INTERROGATIVE DETERMINERS AND QUANTIFIERS

8.1 Conservativity and intersectivity

Some characterizing properties of declarative quantifiers seem to hold of interrogative quantifiers. Here we will restrict ourselves to argument interrogative determiners and quantifiers, but most of the claims hold also for modifier interrogative determiners. Keenan and Westerståhl (1994) observe that interrogative quantifiers satisfy conservativity and extension, and give the following examples to illustrate this fact:

(34) a. Which roses are red? = Which roses are roses and are red?

b. Whose cat can swim? = Whose cat is a cat that can swim?

The claim holds not just for WHICH and WHOSE but for any argument interrogative determiner.

DEFINITION 30 (GENERALIZED CONSERVATIVITY) *Let E be a universe and X any set. Then,* $D \in [\mathcal{P}(E) \rightarrow [\mathcal{P}(E) \rightarrow X]$ *is* conservative *iff for all* $A, B, B' \subseteq E$, *if* $A \cap B = A \cap B'$ *then* $D(A)(B) = D(A)(B')$. *Equivalently:* $D(A)(B) = D(A)(A \cap B)$

Conservativity of declarative and interrogative determiners follows from the above definition, since in the case of declarative determiners X is the set of truth values and in the case of argument interrogative determiners X is the set of argument questions. It also follows that if an interrogative determiner D satisfies conservativity, then $D(A)(B)$ and $D(A)(A \cap B)$ are the same question function. Applying the definition to (34a) we see that WHICH(ROSE)(RED) = WHICH(ROSE)(ROSE ∩ RED).

FACT 31 *All argument interrogative determiners are conservative*

Argument interrogative determiners all satisfy the property of extension.

DEFINITION 32 (GENERALIZED EXTENSION) *For all* $D \in [\mathcal{P}(E) \rightarrow [\mathcal{P}(E) \rightarrow X]]$, *D satisfies* extension *iff for all* $A, B \subseteq E \subseteq E', D_E(A)(B) = D_{E'}(A)(B)$

With respect to the property of permutation invariance (PI), it is interesting to note that WHO respects a local notion of it (Westerståhl 1985), namely when we fix the set PERSON in all permutations. Context restricted determiners are also invariant under permutations that satisfy Locality, defined as follows:

DEFINITION 33 (LOCALITY) *Let* $C \subseteq E$, *PERM(E) be the set of permutations of E and* $D \in [\mathcal{P}(E) \rightarrow [\mathcal{P}(E) \rightarrow X]]$. *We say that D is* PI at C *iff* $\forall \pi \in$ *PERM(E), if* $\pi(C) = C$ *then* $D^C(A)(B) = D^C(\pi(A))(\pi(B))$.

The determiner WHOSE satisfies a more specific condition that van Benthem (1986) calls "quality". It requires that all permutations π preserve the possession relation induced by *Poss* (a possession relation).

8.2 Generalized existential interrogative functions and context neutrality

Argument interrogative determiners satisfy a stronger invariance condition than conservativity. They are all intersective or generalized existential (Keenan 1987, Keenan 1993).

DEFINITION 34 (GENERALIZED INTERSECTIVITY) *Let X be any set. Then, $D \in [\mathcal{P}(E) \to [\mathcal{P}(E) \to X]]$ is* intersective *iff for all $A, B, A', B' \subseteq E$, if $A \cap B = A' \cap B'$ then $D(A)(B) = D(A')(B')$. Equivalently: $D(A)(B) = D(A \cap B)(A \cap B)$*

If a determiner is intersective, then the denotation of D(A)(B) depends only on the intersection of the arguments. In the interrogative domain, we have seen that to determine the answer set of an argument question we only have to know the intersection of A and B. This set is precisely the answer of the question. Keenan (1987) and Lappin (1988) show that a conservative binary determiner is intersective iff it is existential iff it is symmetric. The generalized definition of the two latter notions for $\langle 1, 1\rangle$ determiners is as follows (their application to interrogative determiners follows again as a special case):

DEFINITION 35 *(i) For all $D \in [\mathcal{P}(E) \to [\mathcal{P}(E) \to X]]$, D is* generalized existential *iff for all $A, B \subseteq E, D(A)(B) = D(A \cap B)(E)$.*
(ii) For all $D \in [\mathcal{P}(E) \to [\mathcal{P}(E) \to X]]$, D is symmetric *iff for all $A, B \subseteq E$, $D(A)(B) = D(B)(A)$.*

The fact that argument interrogative determiners satisfy intersectivity is equivalent to being existential. The question (35b) denoted by the interrogative sentence (35a) is equivalent to the one denoted by (35c).[10]

(35) a. How many students are vegetarians?

b. HOW MANY (STUDENT)(VEGETARIAN)

[10]There are some exceptions, though. Heim (1987) notices the marginal status of the following existential constructions with *which* and *which of the* determiners.

(i) a. ??Which one of the three men was there in the room?

b. ??Which actors were there in the room?

Notice that the functions WHICH_{pl} and WHICH_ONE_OF_THE_TWO are inherently context restricted. Therefore, they are related to a set of entities already present in the discourse and are not compatible with presentational/existential predicates that require "discourse novel" answer sets. Furthermore, the determiner function WHICH_*n*_OF_THE_*m* defined as $\text{WHICH\_}n\text{\_OF\_THE\_}m(A)(B)(X) = 1$ iff $|A| = m \;\;\&\;\; X = A \cap B \;\;\&\;\; |X| = n$ is not intersective.

c. How many students who are vegetarians are there/exist?

d. HOW MANY(STUDENT $\cap$ VEGETARIAN)(E)

Since argument interrogative determiners and declarative determiners like SOME are generalized existential functions in their respective domain, it is not surprising that in many languages, like Chinese, Greek, Latin, Romance, etc., the declarative quantifier is derived from the interrogative one by attaching a morpheme to it. In other languages the same word is used for some declarative and argument interrogative determiners. In some Australian languages like Diyari, Maruthunira and Panyjima, no affix is attached at all and disambiguation is reached by positional differences. Here are some examples from Panyjima (Cheng 1991):

(36) a. ngatha ngnanhalu nhantha-nnguli-nha
1sg.nom. something-instr. bit-pass-pst.

'I was bitten by something'

b. nganha ma-rna nyinta ngunhalku
what caus.-pst. 2sg-nom. that-acc.

'What have you done to him'

The equivalence between intersectivity and symmetry apparently poses some problems. Consider the interrogative sentence in (37a):

(37) a. How many vegetarians are students?

b. HOW MANY (VEGETARIAN)(STUDENT)

Since intersective determiners are symmetric, the questions in (35b) and (37b) should be equivalent. In fact, they are. The answer set of (37b) is the intersection set of the denotations of *student* and *vegetarian* which is also the answer set of (35b). However, the intuition remains that the two questions are "about" different things. Imagine a situation in which a school's cook wants to know the number of students who are vegetarians. Sentence (35a) would be felicitous in that situation whereas (37a) would not be so, despite the fact that their respective answer sets are the same. Higginbotham (1993) relates the contrast to the property of domain restriction. Since domain restriction is formally defined as conservativity + extension (Keenan and Westerståhl 1994) and symmetry is a property satisfied only by a subset of the determiners that satisfy domain restriction, namely those which are intersective, it seems more reasonable to relate the "aboutness" problem with this latter property of interrogative determiners. The problem also surfaces with declarative determiners. In the school's cook scenario described above, only sentence (38a) would be felicitous.

(38) a. Some students are vegetarians.

b. Some vegetarians are students.

From an information-based perspective one can easily conclude that the difference between the two sentences is that their respective topics or themes are different. In other words, they are not context neutral. We can generalize this new property as follows:

DEFINITION 36 (CONTEXT NEUTRALITY) *For all context sets $C \subseteq E$, all determiners D, D is* C-neutral *iff for all $A, B, C \subseteq E$, $D^C(A)(B) = D(A)(B)$.*

When a determiner is context neutral in a given context C its arguments can be inverted preserving truth values. This property only makes a difference in the case of intersective determiners. Co-intersective determiners (Keenan 1993) are not symmetric nor are non-intersective determiners in general. Therefore, the sentences in (39) are only equivalent if STUDENT = VEGETARIAN.

(39) a. Every student is a vegetarian.

b. Every vegetarian is a student.

Since all argument interrogative determiners are intersective, context neutrality becomes a relevant issue. The property captures the idea that when a symmetric determiner is relativized to a non-empty context set then its arguments cannot be flipped in general, i.e., the determiner is not context neutral.

8.3 Monotonicity, entailment and negative polarity items (NPIs) licensing

Argument interrogative determiners can be characterized as continuous in their monotonicity behaviour. Continuous functions are meets of increasing and decreasing functions.

DEFINITION 37 *An argument interrogative quantifier $Q \in [\mathcal{P}(E) \rightarrow [_q\mathcal{P}(E) \rightarrow 2]]$ is continuous iff for all $A, B, C \subseteq E$, if $A \subseteq B \subseteq C$ and $Q(A) = Q(C) = f$ then $Q(B) = f$.*

FACT 38 (ARGUMENT INTERROGATIVE QUANTIFIERS ARE CONTINUOUS)
Proof: Let $D(A)$ be an argument interrogative quantifier. Assume for arbitrary sets W,Y,Z that $W \subseteq Y \subseteq Z$ and $D(A)(W) = D(A)(Z) = f$. Therefore, there is an X such that $D(A)(W)(X) = D(A)(Z)(X) = 1$. Then, $A \cap W = A \cap Z = X$ and, since $W \subseteq Y \subseteq Z$, $A \cap Y = X$ so $D(A)(Y)(X) = 1$ □

There is a grammatical fact associated with monotonicity which seems not to follow from this characterization. It is a fairly common observation that negative polarity items are licensed in interrogative sentences.

(40) a. Which student has *ever* been to Moscow?
b. Do you have *any* money?
c. Who has ever *lifted a finger* to help us?

The piece of data above does not follow from the properties of questions studied so far, since as observed in pioneering work by Fauconnier (1975) and Ladusaw (1979) negative polarity items are licensed in decreasing environments. Nevertheless, although interrogative determiners are continuous, questions can be considered as downward entailing with respect to their answer sets. Consider the following questions:

(41) a. Which guests smoked?
b. Which guests smoked cigars?
c. In which state do you have relatives?
d. In which state of the West Coast do you have relatives?
e. How many cars are parked in the garage?
f. How many red cars are parked in the garage?

There is a natural information-based relation between (41a) and (41b) above. Namely, a true complete answer to (41b) is a partial (and possibly complete) answer to (41a). Informally, (41b) asks for more specific information than (41a). In other words, if A_f is the answer set of (41a), then a subset of A_f is the answer set of (41b). The same applies to (41c) with respect to (41d) and to (41e) with respect to (41f). Let us call this relation between questions *subsumption*:

DEFINITION 39 *Question f* subsumes *question g ($f \leq g$) iff $A_g \subseteq A_f$.*

Clearly, the subsumption relation is a partial order (reflexive, antisymmetric and transitive). Then, if we allow not only the monotonicity behaviour of the quantifier but also the subsumption relations between questions to enter the picture, interrogative determiners will exhibit the entailment pattern of declarative NO. As noted above, if question f subsumes question g, then a complete true answer to g is a partial or complete true answer to f but not necessarily vice versa.[11]

[11] The subsumption relation presented here is apparently different from the relation of entailment between questions in Groenendijk and Stokhof (1989). For them the entailment relation holds between propositions and here subsumption holds between questions (it is the subset relation between answer sets). Notice, however, that if question f subsumes question g, then question f entails question g in Groenendijk and Stokhof's (1989) sense, so the notion of subsumption could also be captured in their terms. Notice also that the notion of subsumption is identical to Higginbotham's (1993) notion of downward entailment for interrogatives.

DEFINITION 40 *(i) An interrogative quantifier Q is* decreasing *iff* $\forall A, B \subseteq E$ *if* $A \subseteq B$ *then* $Q(B) \leq Q(A)$
(ii) An interrogative determiner D is decreasing *iff* $\forall A, B, C \subseteq E$ *if* $A \subseteq B$ *then* $D(B)(C) \leq D(A)(C)$

FACT 41 (ARGUMENT INTERROGATIVE QUANTIFIERS Q ARE DECREASING)
Proof: Let $A, B, C \subseteq E$, $A \subseteq B, Q = D(C)$ *and* $D =$ *WHICH, WHAT, etc. We have to show that for arbitrary X,Y, if* $Q(B)(X) = Q(A)(Y) = 1$, *then* $Y \subseteq X$. *Assume* $Q(B)(X) = Q(A)(Y) = 1$. *Since* $A \subseteq B$, $Y = C \cap A \subseteq C \cap B = X$.
□

FACT 42 (ARGUMENT INTERROGATIVE DETERMINERS D ARE DECREASING)
Proof: Let $A, B, C \subseteq E$ *and* $A \subseteq B$. *We have to show that* $D(B)(C) \leq D(A)(C)$. *Let X, Y be such that* $D(B)(C)(X) = 1$ *and* $D(A)(C)(Y) = 1$. *Then,* $Y = A \cap C \subseteq B \cap C = X$. □

The notion of subsumption given above predicts entailments between questions arising from their monotonicity pattern as the ones illustrated in (41). A complete (partial) answer to question (41b) will be a partial (complete) answer to (41a) since the answer set of (41b) is a subset of the answer set of (41a). Fact 42 also predicts that negative polarity items can occur in the first argument of interrogative determiners.

(42) Which students that have *ever* been to Moscow want to go back there?

The presence of NPIs in interrogative environments triggers a peculiar phenomenon observed, among others, by Linebarger (1991). In all the examples above involving NPIs the interpretetation of the questions as rhetorical is either available or strongly preferred.[12] A rhetorical question is not a "well-behaved" question. The speaker already knows the answer and he asks it for rhetorical purposes (irony). For instance, in question (40c) the speaker knows already that the answer set of the question is empty but he asks the question to highlight precisely this fact: that the set of persons who have done something to save us is empty. A sentence like (43) uttered as a rhetorical question has an empty answer set. Assume that the speaker knows that nobody came: $\text{PERSON} \cap \text{COME} = \emptyset$. Then, he would ask this question for rhetorical reasons.

(43) Who came?

[12]There are cross-linguistic differences with respect to this fact. In English, when NPIs occur in yes/no questions, the rhetorical reading is available but not preferred, in contrast to constituent questions where it is preferred. In Chinese, the question is totally ambiguous and the situation disambiguates the rhetorical/non rhetorical interpretation (Zhang 1991). In Spanish, Catalan or Hindi, the presence of NPIs makes the rhetorical reading strongly preferred.

Sentence (44) presents the opposite case. Assume that the speaker knows that everybody went to the party: PERSON $\subseteq$ COME, i.e. PERSON $\cap$ COME = PERSON. Therefore, for rhetorical reasons and with the characteristic intonation associated to rhetorical interrogatives he would ask:

(44) Who didn't come?

The answer set of (44) is PERSON $\cap$ $\neg$COME = $\emptyset$, since everybody went to the party. In sum, for a speaker to be able to ask a rhetorical question he has to be able to go over the whole entailment set of a question, pick its smallest element and ask a question about it. The presence of the NPI signals precisely this calculation (Fauconnier 1975). Nevertheless, we are not claiming that rhetorical interpretations arise only when there are NPIs in the sentence. As observed in the literature, practically any question can be interpreted as rhetorical, depending on the circumstances and the speaker's intentions. What needs to be stressed is the close relationship between answer set entailment and the calculation of rhetorical questions. If question f is rhetorical (Rhet(f)), then $A_f = \emptyset$.

DEFINITION 43 (SUBSUMPTION SET OF A QUESTION) $SU_f = \{g | f \leq g\}$

FACT 44 *If Rhet(f) then* $SU_f = \{f\}$.

Further evidence for the semantic treatment of rhetorical questions comes from the behaviour of *why* and *how*-questions. The occurrence of NPIs in these sentences does not trigger rhetorical readings (Lawler 1971).

(45) a. Why did you tell *anybody* about us?

b. How did *anybody* buy that house?

Sentence (45a) presupposes that the addressee told somebody about them and (45b) presupposes that somebody bought the house. Neither of the questions has empty answer sets. Szabolcsi and Zwarts (1993) claim that manners and reasons are structured as join semilattices with no least element. They are closed under joins but not under complements. Being semilattices without a bottom element, they cannot constitute proper denotations of rhetorical questions (there is no empty set of manners or reasons). Therefore, the explanation of why there are no proper rhetorical *why* and *how* questions is mainly semantic. Since *why* and *how*-questions cannot have empty answer sets, they do not meet the essential denotational requirement to be a rhetorical question. See Gutiérrez Rexach (1996) for a cross-linguistic study of the licensing of NPIs in questions and the semantics of rhetorical readings.

9 MULTIPLE QUESTIONS

Interrogative quantifiers do not only occur in sentence initial position. In English, the so-called echo questions are characterized by the occurrence of the interrogative quantifier in its canonical or non-displaced position, as in (46).

(46) John saw what?

Other constructions in which interrogative quantifiers are forced to stay in their canonical position are multiple questions. The sentences in (47) illustrate the fact that only one quantifier can be "moved", whereas the others remain in-situ.

(47) a. Which boy kissed which girl?

b. Who saw what?

In other languages, such as Chinese, Korean, Hindi, Japanese, etc., interrogative quantifiers remain in their canonical position, as the Chinese examples in (48) illustrate. By contrast, in Polish, Hungarian or Bulgarian all interrogative quantifiers are preposed. The examples in (49) are from Bulgarian (Rudin 1988).[13]

```
(48) a. John mai le   shenme?
        John buy Asp. what
        'What did John buy?'

     b. Shei kanjian le   shenme?
        who  see     Asp. what
        'Who saw what?'

     c. Nage      nanhai qin  le   nage      nhai?
        which-one boy    kiss Asp. which-one girl?
        'Which boy kissed which girl?'

(49) a. Koj kakvo na kogo  e   datt?
        who what  to whom  has given
        'Who gave what to whom?'
```

Syntactic theories within the GB framework have proposed that the diversity that we have hinted at has to be accounted for positing a syntactic level of

[13]There are cross-linguistic restrictions on the syntactic order of the quantifiers in the prefix. It is also possible that in a language some quantifiers have to be preposed and others do not have to be. See Rudin (1988) and É. Kiss (1992).

representation Logical Form (LF) in which all languages of the world are identical. The representation at this level is reached by a series of overt or covert operations that front the quantifiers forming a prefix. With respect to multiple questions, there is the specific claim that *wh*-quantifiers absorb, forming a polyadic operator that binds n variables (Higginbotham and May 1981, May 1989). The LFs of the sentences in (47) are given in (50).

(50) a. Which_{xy} [Boy(x) & Girl(y) & Kiss(x,y)]

b. $(\text{Who,What})_{xy}$ [See(x,y)]

What the LFs in (50) show is that the interrogative quantifiers absorb yielding an n-ary quantifier, a pair quantifier in the above examples (May 1989). Keenan (1992) also conjectures that multiple questions are unreducible. Two points seem to be of interest from the semantic point of view: (i) Is it impossible to give a semantics of interrogative quantifiers that interprets them "in situ", in other words, that is faithful to the surface syntax of the language?, and (ii) Does the sequence of n interrogative quantifiers in a prefix always form a resumptive or unreducible polyadic quantifier, as claimed by Higginbotham, May, Keenan with respect to which quantifiers, and assumed by many others for multiple questions in general? Contra what is commonly sustained in the literature, the answer to the two questions above is negative. As it will be shown, it is straightforward to give an "in situ" semantics for interrogative quantifiers, something that has also been argued recently from a syntactically minimalist point of view (Reinhart 1995). The second negative answer is more interesting. It can be shown that a sequence of n interrogative quantifiers is reducible to iterations of the quantifiers forming the sequences. As a consequence, the rule of quantifier absorption is vacuous for the majority of multiple questions—for the cases in which the absorbed prefix is "separable" in May's (1989) terms—and thus not obligatory in this particular domain. Consider the following two sentences:

(51) a. What did John buy?

b. John bought what?

The semantic difference between (51a) and (51b) is that when a speaker utters (51b), he is asking about a set of objects that has already been introduced in the previous discourse. Therefore, *what* in (51b) denotes WHAT^X, the interrogative quantifier WHAT relativized to the context set X. In English, the interpretation of *what* is disambiguated by prosodic and syntactic means. In Chinese, the disambiguation is not syntactic: sentence (48a) can be interpreted as an echo or a non-echo question.

In previous sections we have only analyzed sentences with one interrogative quantifier. In order to give a proper semantics of multiple questions—sentences where more than one interrogative quantifier interact—we have to define the nominative, accusative and dative extensions of an interrogative quantifier. This will allow us to give a surface compositional semantics of English interrogative VPs like *bought what* or *bought for whom.*

DEFINITION 45 (EXTENSIONS OF INTERROGATIVE QUANTIFIERS)
Let $R \subseteq E^n, \alpha_1, ..., \alpha_n \in E, X \subseteq E, Q \in [\mathcal{P}(E) \to [_q\mathcal{P}(E) \to 2]]$. *Then,*
(i) A nominative interrogative quantifier *(or the nominative extension of Q) is a function* $Q_1 \in [\mathcal{P}(E^n) \to [E^{n-1} \to [_q\mathcal{P}(E) \to 2]]]$ *defined as follows:*
$Q_1(R)(\alpha_2, ..., \alpha_n)(X) = 1$ *iff* $Q(\{\alpha_1 | \langle\alpha_1, ..., \alpha_n\rangle \in R\})(X) = 1$
(ii) An accusative interrogative quantifier *(the accusative extension of Q) is a function* $Q_2 \in [\mathcal{P}(E^n) \to [E^{n-1} \to [_q\mathcal{P}(E) \to 2]]]$ *defined as follows:*
$Q_2(R)(\alpha_1, \alpha_3, ..., \alpha_n)(X) = 1$ *iff* $Q(\{\alpha_2 | \langle\alpha_1, ..., \alpha_n\rangle \in R\})(X) = 1$
(iii) A dative interrogative quantifier *(the dative extension of Q) is a function* $Q_3 \in [\mathcal{P}(E^n) \to [E^{n-1} \to [_q\mathcal{P}(E) \to 2]]]$ *defined as follows:*
$Q_3(R)(\alpha_1, \alpha_2, \alpha_4, ..., \alpha_n)(X) = 1$ *iff* $Q(\{\alpha_3 | \langle\alpha_1, ..., \alpha_n\rangle \in R\})(X) = 1$

The interpretation of the VP *buy what* is the following:

(52) $\mathrm{WHAT}_2(\mathrm{BUY})(\alpha)(X) = 1$ iff $\mathrm{WHAT}(\{\beta | \langle\alpha, \beta\rangle \in \mathrm{BUY}\})(X) = 1$ iff $X = E \cap \{\beta | \langle\alpha, \beta\rangle \in \mathrm{BUY}\}$

The Chinese version of the VP is interpreted as above. In English, it would receive an "echo" interpretation, arising from the fact that *what* denotes WHAT^X when it does not occur in its canonical fronted position. There is a significant difference between declarative and interrogative argument quantifiers. Declarative Qs behave as "arity reducers" (van Benthem 1986, Keenan and Westerståhl 1994).. They take an n-ary relation as input and return an n-1-ary relation. Interrogative quantifiers are not arity reducers. They take an n-ary relation and return another n-ary relation. The output n-ary relation is not the same as the input one. The argument that has been "queried" is turned into an answer set. Consider the following sentence:

(53) Which men love which women?

In its most natural reading, question (53) asks for the sets of pairs in the love relation whose first coordinate is a member of the set of men and its second coordinate is a member of the set of women.[14]

(54) $(\mathrm{WHICH}_{pl}\mathrm{MAN}, \mathrm{WHICH}_{pl}\mathrm{WOMAN})(\mathrm{LOVE})(S) = 1$ iff $S = R \cap \mathrm{MAN} \times \mathrm{WOMAN}$

[14]In what follows we will be ignoring the context set parameter.

Sentence (53) denotes a binary question, i.e., a function mapping a binary relation S to true iff $S = R \cap \text{MAN} \times \text{WOMAN}$. Generalizing to the n-ary case, we define first the notion of a n-ary argument question and afterwards the polyadic WHICH_{pl} interrogative quantifier induced by a sequence of n WHICH_{pl} quantifiers:

DEFINITION 46 *For* $n \geq 1, [_q\mathcal{P}(E^n) \to 2]$ *is the set of* n-ary argument questions.

DEFINITION 47 (POLYADIC RESUMPTION OF WHICH_{pl} QUANTIFIERS)
Let $R, S \subseteq E^n$, $Z_1, \dots, Z_n \subseteq E$. *Then,* $\text{Res}(WHICH_{pl(1)}Z_1, \dots, WHICH_{pl(n)}Z_n) \in [\mathcal{P}(E^n) \to [_q\mathcal{P}(E^n) \to 2]]$ *is defined as follows:*
$\text{Res}(WHICH_{pl(1)}Z_1, \dots, WHICH_{pl(n)}Z_n)(R)(S) = 1$ *iff* $S = R \cap Z_1 \times \cdots \times Z_n$

All interrogative quantifiers can participate in multiple questions. We treat first the resumptions of arbitrary argument interrogative quantifiers except WHICH_{sg}.

DEFINITION 48 (RESUMPTION OF ARGUMENT INTERROGATIVE QUANTIFIERS)
Let $R, S \subseteq E^n$, $Z_1, ..., Z_n \subseteq E$, *for all* Q_i,

$$\text{let } Z_i = \begin{cases} Z & \text{if } Q_i = WHAT(Z_i) \text{ or } Q_1 = WHOSE(Z_i) \\ PERSON & \text{if } Q_i = WHO \\ E & \text{if } Q_i = WHAT \end{cases}$$

$\text{Res}(Q_1, ..., Q_n)(R)(S) = 1$ *iff* $S = R \cap Z_1 \times ... \times Z_n$

Other examples of multiple questions in which different argument interrogative quantifiers interact are given below:

(55) a. Who bought which tables for whom?
b. Who saw what?
c. Which students ate what?

The reducibility result that we prove states that the answer set that we get by an application of the prefix in definitions 47 and 48 and the one that we get by successively applying the n interrogative quantifiers in the prefix are equivalent in a sense we make precise below. Therefore, what we really prove is equivalence of answer sets or, in less formal terms, the questions that we get with an absorbed interrogative quantifier and the iterated application of the members of the quantificational prefix are querying "about" the same object.

FACT 49 *The polyadic* $\text{Res}(Q_1, ..., Q_n)$ *is reducible to iterations*

Instead of a full proof we give here a worked out example. Consider the following sentence:

(56) Who gave what to whom?

Applying definition 45(iii), we see that $\mathrm{WHO}_3(\mathrm{GIVE})$ is a function in $[E \to [E \to [_q\mathcal{P}(E) \to 2]]]$. Then, $\mathrm{WHO}_3(\mathrm{GIVE})(\alpha,\beta)(X) = 1$ iff $X = \mathrm{PERSON} \cap \{\gamma | \langle\alpha,\beta,\gamma\rangle \in \mathrm{GIVE}\}$. In the second step of the calculation, the accusative extension of WHAT, WHAT_2 applies to $\mathrm{WHO}_3(\mathrm{GIVE})$, and we get the function $\mathrm{WHAT}_2(\mathrm{WHO}_3(\mathrm{GIVE})) \in [E \to [_q\mathcal{P}(E^2) \to 2]]$ defined as $\mathrm{WHAT}_2(\mathrm{WHO}_3(\mathrm{GIVE}))(\alpha)(S) = 1$ iff $S = \{\langle\beta,\gamma\rangle | \beta \in E \ \& \ \gamma \in \mathrm{WHO}_3(\mathrm{GIVE})(\alpha,\beta)\}$ (writing $\gamma \in \mathrm{WHO}_3(\mathrm{GIVE})(\alpha,\beta)$ for $\gamma \in A_{\mathrm{WHO}_3(\mathrm{GIVE})(\alpha,\beta)}$). Finally, the nominative extension of WHO, WHO_1, applies to $\mathrm{WHAT}_2(\mathrm{WHO}_3(\mathrm{GIVE}))$ yielding the function $\mathrm{WHO}_1(\mathrm{WHAT}_2(\mathrm{WHO}_3(\mathrm{GIVE}))) \in [_q\mathcal{P}(E^3) \to 2]$ such that $\mathrm{WHO}_1(\mathrm{WHAT}_2(\mathrm{WHO}_3(\mathrm{GIVE})))(S) = 1$ iff $S = \{\langle\alpha,\beta,\gamma\rangle | \alpha \in \mathrm{PERSON} \ \& \ \langle\beta,\gamma\rangle \in \mathrm{WHAT}_2(\mathrm{WHO}_3(\mathrm{GIVE}))(\alpha)\}$. In general, $Q_1(...(Q_n(R))) \in [_q\mathcal{P}(E^n) \to 2]$. Let $Q_i = D(Z_i)$, for some $Z_i \subseteq E$. Then, $Q_1(...(Q_n(R)))(S) = 1$ iff $S = \{\langle\alpha_1,...,\alpha_n\rangle | \alpha_1 \in Z_1 \ \& \ \langle\alpha_2,...,\alpha_n\rangle \in Q_2(...(Q_n(R)))(\alpha_1)\}$.

Resumptions of WHICH_{sg} determiners deserve a more detailed analysis. Higginbotham and May (1981) claim that multiple WHICH_{sg} questions have two interpretations: a singular interpretation and a bijective interpretation. Consider the following sentence:

(57) In *Gone with the wind*, which character admires which character?

The singular interpretation of the sentence is the one rendered by the answer in (58a). The answer in (58b) corresponds to the bijective reading.

(58) a. Ashley Wilkes admired Rhett Butler

b. Ashley Wilkes admired Rhett Butler and Melanie Wilkes admired Scarlett O'Hara.

Under the singular interpretation of the question, the answer set is a singleton, i.e., it consists of a unique pair. The bijective interpretation asks for a bijection between the restriction sets of the interrogative quantifiers.[15] The following definition characterizes Higginbotham and May's intuition:

DEFINITION 50
SINGULAR READING OF $(\mathrm{WHICH}_{sg},...,\mathrm{WHICH}_{sg})$ QUANTIFIERS

[15] According to Higginbotham and May the bijective reading is only available when "the domain of quantification in the subject NP is disjoint from that of the object"(p. 46). Nevertheless, the claim does not seem to be completely correct for all English dialects (Ed Keenan, p.c.). We assume, then, that the two readings are generally available, though accomodating Higginbotham and May's disjointness condition is straightforward.

$\mathrm{Sg}(WHICH_{sg(1)}(Z_1),..., WHICH_{sg(n)}(Z_n))(R)(S) = 1$ *iff*
$S = R \cap Z_1 \times, ..., \times Z_n$ & $|S| = 1$
(The disjointness condition Higginbotham and May 1981): $\bigcap_i Z_i = \emptyset$)

The singular reading is the one that we get by iterated application of $\mathrm{WHICH}_{sg}(Z)$ quantifiers. Therefore, it is reducible and does not lie "beyond the Frege boundary". We are able to get the interpretation above by applying a $\mathrm{WHICH}_{sg(i)}(Z_i)$ quantifier to the function $\mathrm{WHICH}_{sg(i+1)}(Z_{i+1})(\ldots$ $(\mathrm{WHICH}_{sg(n)}(Z_n)(R))$. It can be conceived of as the polyadic resumption of $\mathrm{WHICH}_{sg}(Z)$ quantifiers, so definition 50 corresponds to $Res(\mathrm{WHICH}_{sg(1)}(Z_1),$ $\ldots, \mathrm{WHICH}_{sg(n)}(Z_n))$. The fact that this resumption is also reducible to iterations confirms the claim made in van Benthem (1989) concerning the general reducibility of resumptive polyadic quantifiers.

FACT 51 *The singular reading is derived from iterations of* $WHICH_{sg}(Z)$ *quantifiers.*
Proof: Let $R \subseteq E^n, Z_1, ..., Z_n \subseteq E$. *Then,* $WHICH_{sg(n)}(Z_n)(R)(\alpha_1, ..., \alpha_{n-1})(X)$ $= 1$ *iff* $X = Z_n \cap \{\alpha_n | \langle \alpha_1, ..., \alpha_n\rangle \in R\}$ & $|X| = 1$. *For all i* $(1 \leq i \leq n-1)$,
$WHICH_{sg(i)}(Z_i)(...(WHICH_{sg(n)}(Z_n)(R)))(\alpha_1, ..., \alpha_{i-1})(S) = 1$ *iff*
$S = \{\langle \alpha_i, ..., \alpha_n\rangle | \alpha_i \in Z_i$ & $\langle \alpha_{i+1}, ..., \alpha_n\rangle \in$
$WHICH_{sg(i+1)}(Z_{i+1})(...(WHICH_{sg(n)}(Z_n)(R)))(\alpha_1,$
$..., \alpha_i)\}$ & $|S| = 1$.□

In the bijective reading there is an apparent loss of the uniqueness condition imposed by WHICH_{sg}, due to the fact that the polyadic is not reducible to iterations of $\mathrm{WHICH}_{sg}(Z)$ quantifiers.

DEFINITION 52 *Let* $R, S \subseteq E^n, Z_1, ..., Z_n \subseteq E$. *Then,*
$\mathrm{Bij}(WHICH_1(Z_1), ..., WHICH_n(Z_n)) \in [\mathcal{P}(E^n) \to [_q\mathcal{P}(E^n) \to 2]]$ *and*
$\mathrm{Bij}(WHICH_1(Z_1), ..., WHICH_n(Z_n))(R)(S) = 1$ *iff* $S = R \cap Z_1 \times ... \times Z_n$ &
$\forall \alpha_i \exists! \langle \alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_n\rangle \in Z^1 \times ... \times Z^{i-1} \times Z^{i+1} \times ... \times Z^n$
such that $\langle \alpha_1, ..., \alpha_n\rangle \in R$

10 INTERACTIONS OF DECLARATIVE AND INTERROGATIVE QUANTIFIERS

In the previous section we analyzed how interrogative quantifiers are combined (multiple questions). In this section we treat the combinations of interrogative and declarative quantifiers. Different interactions give rise to different readings of the interrogative sentence. The first reading to be considered is the *individual* (Groenendijk and Stokhof 1984) or "single constituent" reading (Chierchia

1993). For example, the individual reading of sentence (59a) is reflected in the response (59b). In terms of the behaviour of the quantifiers, this reading corresponds to the iteration of the declarative and the interrogative quantifier, as (60) shows.

(59) a. What did every boy read?

b. *Tom Sawyer* and *The Jungle Book*

(60) $\mathrm{WHAT}_2(\mathrm{EVERYBOY}_1(\mathrm{READ}))(X) = 1$ iff
$\mathrm{WHAT}(\{\beta|\mathrm{BOY} \subseteq \{\alpha|\langle\alpha,\beta\rangle \in \mathrm{READ}\}\})(X) = 1$ iff
$X = \{\beta|\mathrm{BOY} \subseteq \{\alpha|\langle\alpha,\beta\rangle \in \mathrm{READ}\}\}$

Sentence (59a) denotes a unary question in its individual reading, as does (61a). In (59a) the accusative extension of the interrogative quantifier combines with the nominative extension of the declarative quantifier. In (61a) the nominative extension of the interrogative quantifier combines with the accusative extension of the declarative quantifier.

(61) a. Which students read more than three books?

b. John and Sam

(62) $\mathrm{WHICH\,STUDENTS}_1(\mathrm{MORE\_THAN\_THREE\_BOOKS}_2(\mathrm{READ}))(X) = 1$ iff
$\mathrm{WHICH\,STUDENTS}_1(\{\alpha|\,|\mathrm{BOOK}\cap\{\beta|\langle\alpha,\beta\rangle \in \mathrm{READ}\}| > 3\,\})(X) = 1$ iff
$X = \mathrm{STUDENT} \cap \{\alpha|\,|\mathrm{BOOK} \cap \{\beta|\langle\alpha,\beta\rangle \in \mathrm{READ}\}| > 3\}\ \ \&\ \ |X| \geq 2$

A second type of reading is called by Groenendijk and Stokhof a *pair-list* reading, since the answer specifies a set of pairs. The response in (63c) would be a pair-list answer for sentences (63a) and (63b).

(63) a. Which book did each boy read?

b. Which book did these boys (each) read?

c. Bill read *Tom Sawyer* and Joe read *The Jungle Book*

The problem with pair-list readings is that they are not freely available with all declarative quantifiers. Groenendijk and Stokhof (1984), Chierchia (1993) and Szabolcsi (1994) have observed that only quantifiers that denote principal filters give rise to pair-list readings,[16] as the examples in (64) illustrate:

(64) a. Which book did every student read? (pair-list o.k.)

[16] Szabolcsi claims that this is only the case in matrix interrogatives and in complements of verbs of the *wonder* type.

b. Which books did the students read? (pair-list o.k.)

c. Which books did two students read? (pair-list ok only if TWO STUDENTS is a principal filter)

A generalized quantifier Q is a principal filter iff it has a generator set GSET(Q) defined as follows:

DEFINITION 53 *X is a* generator set *for Q ($GSET(Q) = X$) iff* $Q(X) = 1$ & $\forall Y[Q(Y) = 1 \text{ iff } X \subseteq Y]$

Declarative quantifiers such as MORE_THAN_THREE BOYS or FEW BOYS are not principal filters and, as expected, sentences (65a) and (65b) lack pair-list readings.

(65) a. Which book did more than three boys read?

b. Which book did few boys read?

c. *Bill read *Tom Sawyer* and Joe read *The Jungle Book*

An interrogative sentence with one interrogative and one declarative quantifier denotes a binary question in its pair-list reading. It maps a unique set of pairs to true, the set specified by responses such as (63c). When the interrogative sentence has n declarative quantifiers and m interrogative quantifiers, the "$n+m-tuple$"-list reading is a $n+m$-ary question. For example, the triple-list readings of sentences (66a) and (66b) are ternary questions.

(66) a. Which book did each boy put on each desk?

b. Which book did each boy put on which desk?

Combinations of a declarative quantifier which is a principal filter and a modifier interrogative quantifier can be also conceived as binary questions. Therefore these combinations have pair-list readings (67b).

(67) a. When did each of your relatives arrive?

b. Uncle John arrived on Monday, Grandma arrived on Tuesday and my parents on Christmas eve.

Here we will treat only pair-list readings (binary questions) arising from the combination of an argument interrogative quantifier, a declarative quantifier and the denotation of a transitive verb. There are two cases to consider: when the nominative extension of the interrogative quantifier combines with the accusative extension of the declarative quantifier and the opposite case.

DEFINITION 54 (PAIR-LIST LIFT)

(i) *Let* $Q_1 \in [\mathcal{P}(E^2) \rightarrow [E \rightarrow [_q\mathcal{P}(E) \rightarrow 2]]]$, $Q_2 \in [\mathcal{P}(E) \rightarrow 2]$, *and* $R, S \subseteq E^2$. *Then,*
Pair-list$(Q_1, Q_2) \in [\mathcal{P}(E^2) \rightarrow [_q\mathcal{P}(E^2) \rightarrow 2]]$ *and* Pair-list$(Q_1, Q_2)(R)(S)$ $= 1$ *iff*
$S = \{\langle\alpha, \beta\rangle | \beta \in GSET(Q_2)\ \&\ \alpha \in Q_1(R)(\beta)\}$
(where $\alpha \in Q_1(R)(\beta)$ *abbreviates* $\alpha \in X$ *such that* $Q_1(R)(\beta)(X) = 1$*)*

(ii) *Let* $Q_1 \in [\mathcal{P}(E) \rightarrow 2]$, $Q_2 \in [\mathcal{P}(E^2) \rightarrow [E \rightarrow [_q\mathcal{P}(E) \rightarrow 2]]]$, *and* $R, S \subseteq E^2$. *Then,*
Pair-list$(Q_2, Q_1) \in [\mathcal{P}(E^2) \rightarrow [_q\mathcal{P}(E^2) \rightarrow 2]]$ *and* Pair-list$(Q_2, Q_1)(R)(S)$ $= 1$
iff $S = \{\langle\alpha, \beta\rangle | \alpha \in GSET(Q_1)\ \&\ \beta \in Q_2(R)(\alpha)\}$

Not all polyadic pair-list lifts with English universal quantifiers seem to be denotable. Karttunen and Peters (1980) found an asymmetry in the interpretation of the following questions:

(68) a. Which customer is each clerk now serving?
b. Which clerk is now serving each customer?

Given a situation like the one depicted in (69), and the answer in (70), they observe that such an answer constitutes a correct answer only to question (68a), but not to (68b).

(69)

		A	E	I
	customers	B	F	J
		C	G	K
		D <clerk 1	H <clerk 2	L <clerk 3
			check-out counters	

(70) Clerk 1 is serving customer D, clerk 2 is serving customer H, and clerk 3 is serving customer L.

Their explanation of the contrast is that "in this particular situation, [(70)] implicates something false, namely, that all customers are being served by some clerk, when in fact only three of them are" (p. 198). May (1985) attributes the contrast to a difference in scope, similar to the one arising in (71).

(71) a. What did everyone buy for Max?
b. Who bought everything for Max?

The question in (71a) can be answered with an individual answer (72a) or a pair-list answer (72b), whereas a pair-list answer seems not to be adequate for (71b). Thus, only (73) is an appropriate answer to (71b).

(72) a. Everyone bought Max a Bosendorfer piano.

b. Mary bought Max a tie, Sally a sweater, and Harry a piano.

(73) Oscar bought everything for Max.

According to May, the subject-object asymmetry in the interpretation of the interrogative sentences in (71) shows that the quantifier in the subject position is able to form a Sigma-sequence with the wh-operator and, by the Scope principle, can participate in any possible scope interaction with it. The fact that we do not get a pair-list answer when the QP is in the object position follows from the assumption that operators that do not belong to the same Sigma-sequence cannot participate in scope interactions. The declarative quantifier phrase *everything* in (71b) has to adjoin to the VP to avoid a violation of the Path Containment Condition. Therefore, it cannot form a Sigma-sequence with the wh-operator and is forced to get narrow scope with respect to it. May's explanation and many subsequent theories posit syntactic constraints on the semantic availability of the pair-list readings (see other chapters in this volume on related issues). What is of interest here is that ALL does not participate in pair-list polyadics at all, only the nominative extension of EVERY seems to participate in them, and EACH shows a much wider flexibility. In the end it is a question of economy: why have polyadic pair-list lifts denotable with EACH, EVERY and ALL when the function denoted would be the same? There is a clear specialization of morphological resources taking place: ALL is used for individual readings and EACH is used for pair-list constructions. Therefore, only distributive-key universal determiners (Gil 1995) participate in this construction.[17] In the case of of definites, pair-list polyadics are not denotable in English with WHICH_{sg} interrogative determiners, as noted by Srivastav (1992) and shown by the following examples:

(74) a. Which book did the students read?

b. * John read *El Quijote* and Bill read *Magic Mountain*

(75) Which student admires those two professors? (*pair-list)

The situation is similar in Romance languages (Gutiérrez Rexach 1995), but the presence of a floated distributive quantifier makes accessible the pair-list reading, as illustrated in the following contrast:

[17] The position of EVERY is intermediate, at least for some dialects. Crosslinguistically, the most common scenario is that one universal determiner participates in individual readings (iterations) and a different one, the distributive-key universal determiner, in pair-list readings (polyadic).

(76) a. ¿Cuál libro leyeron los estudiantes? (*Pair-list)
Which book read the students

'Which book did the students read?'

b. ¿Cuál libro leyeron los estudiantes cada uno? (Pair-list ok)
Which book read the students each one

'Which book did each one of the students read?'

From the analysis of the empirical data it follows that the necessity of a distributive interpretation of the declarative quantifier plays an important role in the availability of pair-list questions. This suggests treating them as plural n-ary questions, where the polyadic is formed by a distributive lift of the declarative and the interrogative quantifier.

DEFINITION 55 (PLURAL N-ARY QUESTIONS)
$[_q\mathcal{P}(\mathcal{P}(E)^n) \to 2]$ *is the set of* plural n-ary questions.

DEFINITION 56 (PAIR-LIST LIFT, REVISED) *Let* $i \neq j \in \{1,2\}$, $Q_i \in [\mathcal{P}(E^2) \to [E \to [_q\mathcal{P}(E) \to 2]]], Q_j \in [_q\mathcal{P}(E) \to 2]$, *and* $R \subseteq E^2$, $S \subseteq \mathcal{P}(E)^2$ *(i.e.* $S \subseteq \mathcal{P}(E) \times \mathcal{P}(E)$*). Then,* $\text{Pair-list}(Q_i, Q_j) \in [\mathcal{P}(E^2) \to [_q\mathcal{P}(\mathcal{P}(E)^2) \to 2]]$ *and*

(i) if $i = 2, j=1$ *then* $\text{Pair-list}(Q_i, Q_j)(R)(S) = 1$ *iff*
$S = \{\langle A, B\rangle | A \in AT(GSET(Q_j))\ \ \&\ \ B = \{\beta | \beta \in Q_i(R)(\alpha)$, *for (the unique)* $\alpha \in A\}\}$

(ii) if $i = 1, j = 2$ *then* $\text{Pair-list}(Q_i, Q_j)(R)(S) = 1$ *iff*
$S = \{\langle A, B\rangle | B \in AT(GSET(Q_j))\ \ \&\ \ A = \{\alpha | \alpha \in Q_i(R)(\beta)$, *for (the unique)* $\beta \in B\}\}$

In some languages like Hungarian or Turkish, the pair-list reading is not expressible with a combination of a declarative and an interrogative quantifier. Only a combination of interrogative quantifiers can express it, as observed by É. Kiss (1992) with respect to Hungarian. The existence of the two alternatives is due to the fact that, for $i, j \in \{1,2\}, Q_j$ a declarative quantifier and R a binary relation, when the answer set of the polyadic *Pair-list*$(\text{WHICH}_{sg(i)}(Z_i), Q_j)(R)$, in the lower type characterized in definition 54, is a bijective function (a set of pairs $\langle\alpha, \beta\rangle$ where $\alpha \in \text{GSET}(Q_j)$, $\beta \in \text{WHICH}_{sg(i)}(Z_i)(R)(\alpha)$ and such that it defines a function of that sort), then *Pair-list*$(\text{WHICH}_{sg(i)}(Z_i), Q_j)(R) =$ *Bij*$(\text{WHICH}_{sg(i)}(Z_i), \text{WHICH}_j(Z_j))$ (R). In this case, (76a) and (76b) denote the same question: they have the same answer set (76c).

(77) a. Which boy likes which girl?

b. Which girl does each boy like?

c. John likes Mary, Bill likes Pam, Joe likes Sue.

Functional readings (Groenendijk and Stokhof 1984, Engdahl 1985) of questions are characteristic of combinations of interrogative quantifiers and decreasing declarative quantifiers. In these cases the pair-list response is anomalous.

(78) a. Which book did no student like to read?

b. His last week's assignment.

(79) a. Who do few Italian married men like? (Chierchia 1993)

b. Their mother in law.

Functional readings are intensional renderings of pair-list answers (Groenendijk and Stokhof 1984, Chierchia 1993). In clarification contexts, though, the pair-list answer can be explicitly obtained. Consider (79) as an answer to (77a):

(80) Their last week's assignment. More explicitly, Bill did not like to read *El Quijote* and Joe did not like to read *Magic Mountain.*

How is this "extensionalization" possible? Since NO BOY (B) = EVERY BOY($\neg B$) and EVERY BOY($\neg B$) has a generator, namely BOY, then we may define the corresponding pair-list lift as follows:

(81) $\textit{Pair-list}(\mathrm{NOBOY}_1, \mathrm{WHICHBOOK}_2)(R)(S) = 1$ iff
$S = \{\langle\alpha,\beta\rangle | \alpha \in \mathrm{GSET}(\mathrm{EVERYBOY}\neg_1)\ \ \&$
$\beta \in \mathrm{WHICHBOOK}_2(\neg\mathrm{LIKETOREAD})(\alpha)\}$

Interrogatives like (81a) have an additional reading which has been called "cumulative" (Srivastav 1992). A standard cumulative answer for (81a) would be one like (81b), whereas the pair-list answer would be (81c).

(82) a. Which books did the boys read?

b. They read *Tom Sawyer* and *The Jungle Book.*

c. Bill read *Tom Sawyer* and Joe read *The Jungle Book.*

Srivastav (1992) and Szabolcsi (1994) claim that the pair-list reading can be considered as the cooperative spell-out of the cumulative reading when the latter is the preferred interpretation. I perceive more of a semantic difference than of a Gricean phenomenon. Cumulative questions are plural n-ary questions, functions that in a situation map a relation between sets to true. Notice also that there does not seem to be a branching lift in the interrogative domain.

DEFINITION 57 ("CUMULATIVE" LIFT) *Let* $i \neq j \in \{1,2\}$, $Q_i \in [\mathcal{P}(E^2) \to [E \to [_q\mathcal{P}(E) \to 2]]]$, $Q_j \in GQ^{DEF}$ *(the set of definite generalized quantifiers), and* $R \subseteq E^2, S \subseteq \mathcal{P}(E)^2$ *(i.e.* $S \subseteq \mathcal{P}(E) \times \mathcal{P}(E)$*). Then,* $\mathrm{Cum}(Q_i, Q_j) \in [\mathcal{P}(E^2) \to [_q\mathcal{P}(\mathcal{P}(E)^2) \to 2]]$ *and*

(i) *if* $i = 2, j=1$ *then* $\mathrm{Cum}(Q_i, Q_j)(R)(S) = 1$ *iff*
$S = \{\langle A, B\rangle | A = GSET(Q_j) \ \& \ B = \{\beta | \exists \alpha \in A[\beta \in Q_i(R)(\alpha)]\}\}$

(ii) *if* $i = 1, j = 2$ *then* $\mathrm{Cum}(Q_i, Q_j)(R)(S) = 1$ *iff*
$S = \{\langle A, B\rangle | B = GSET(Q_j) \ \& \ A = \{\alpha | \exists \beta \in B[\alpha \in Q_i(R)(\beta)]\}\}$

Another type of question considered by Groenendijk and Stokhof (1984) are choice questions. The name comes from the fact that sentences like (82a) can be paraphrased as "for two boys of your choice, which book did each read".

(83) a. Which books did two boys read?

b. Steve read *A Tale of Two Cities* and Mark read *The Never Ending Story*

There is common agreement that this reading is quite marginal in normal discourse. They are only natural in contests or quizzes and are also called "quiz" questions. The reason for its marginality might be that in the definition of the corresponding polyadic, the relevant domain set cannot be the generator of the declarative quantifier (it does not denote a principal filter) but one of its elements.

DEFINITION 58 ("CHOICE" LIFT)
Let $i \neq j \in \{1,2\}$, $Z_j, W \subseteq E$, $Q_i \in [\mathcal{P}(E) \to [_q\mathcal{P}(E) \to 2]]$, $D_j \in [\mathcal{P}(E^2) \to [E \to [\mathcal{P}(E) \to 2]]]$, $Q_j = D_j(Z_j)$ *and* $R, S \subseteq E^2$*. Then,* $\mathrm{Choice}(Q_i, Q_j) \in [\mathcal{P}(E^2) \to [_q\mathcal{P}(E^2) \to 2]]$ *and*

(i) *if* $i = 2, j = 1$ *then* $\mathrm{Choice}(Q_i, Q_j)(R)(S) = 1$ *iff* $\exists W$ $[Q_j(W) = 1 \ \& \ W \subseteq Z_j \ \& \ S = \{\langle \alpha, \beta\rangle | \alpha \in W \ \& \ \beta \in Q_i(R)(\alpha)\}]$

(ii) *if* $i = 1, j = 2$ *then,* $\mathrm{Choice}(Q_i, Q_j)(R)(S) = 1$ *iff* $\exists W$ $[Q_j(W) = 1 \ \& \ W \subseteq Z_j \ \& \ S = \{\langle \alpha, \beta\rangle | \beta \in W \ \& \ \alpha \in Q_i(R)(\beta)\}]$

The explanation of why choice questions are only possible with non-filter denoting declarative quantifiers is the result of the fact that if $|\{W | Q_j(W) = 1 \ \& \ W \subseteq Z_j\}| = 1$ then $Choice(Q_i, Q_j) = Pair\text{-}list(Q_i, Q_j)$.

REFERENCES

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2):159–219.

van Benthem, Johan. 1986. *Essays on Logical Semantics.* Dordrecht: Reidel.

van Benthem, Johan. 1989. Polyadic quantifiers. *Linguistics and Philosophy* 12(4):437–465.

van Benthem, Johan. 1991. *Language in Action.* Amsterdam: North Holland Press.

van Benthem, Johan, and Alice ter Meulen (Eds.). 1994. *Handbook of Logic and Linguistics.* Amsterdam: Elsevier. Forthcoming.

Cheng, Lisa. 1991. *On the Typology of Wh-Questions.* PhD thesis, MIT.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1(2):181–234.

Cooper, Robin. 1983. *Quantification and Syntactic Theory.* Dordrecht: Reidel.

van der Does, Jaap. 1992. *Applied Quantifier Logics.* PhD thesis, University of Amsterdam.

Engdahl, Elisabet. 1985. *Constituent Questions.* Dordrecht: Reidel.

Fauconnier, Gilles. 1975. Polarity and the scale principle. In Robin Grossman et al. (Eds.), *Papers from the Eleventh Regional Meeting, Chicago Linguistics Society*, 188–199.

Gil, David. 1995. Universal quantifiers and distributivity. In Emmon Bach et al. (Eds.), *Quantification in Natural Languages, Volume 1*, 321–362. Dordrecht: Kluwer.

Groenendijk, Jeroen, and Martin Stokhof. 1984. *Studies on the Semantics of Questions and the Pragmatics of Answers.* PhD thesis, University of Amsterdam.

Groenendijk, Jeroen, and Martin Stokhof. 1989. Type-shifting rules and the semantics of interrogatives. In Gennaro Chierchia et al. (Eds.), *Properties, Types and Meaning, Volume 2: Semantic Issues*, 21–68. Dordrecht: Kluwer.

Gutiérrez Rexach, Javier. 1995. The scope of universal quantifiers in Spanish interrogatives. In Karen Zagona (Ed.), *Grammatical Theory and Romance Languages*, 87–98, Amsterdam. John Benjamins.

Gutiérrez Rexach, Javier. 1996. Negative polarity licensing and the rhetorical interpretation of questions. In Vida Samiian (Ed.), *Proceedings of the 25th Western States Conference on Linguistics.* California State University at Fresno.

Heim, Irene. 1987. Where does the definiteness restriction apply? Evidence from the definiteness of variables. In Eric Reuland and Alice ter Meulen (Eds.), *The Representation of (In)definiteness*, 21–42.

Higginbotham, James. 1993. Interrogatives. In Ken Hale and Samuel J. Keyser (Eds.), *The View from Building 20*, 195–228. Cambridge: MIT Press.

Higginbotham, James, and Robert May. 1981. Questions, quantifiers and crossing. *Linguistic Review* 1(1):41–80.

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In Emmon Bach et al. (Eds.), *Quantification in Natural Languages*, 451–486. Dordrecht: Kluwer.

Karttunen, Lauri. 1977. The syntax and semantics of questions. *Linguistics and Philosophy* 1(1):3–44.

Karttunen, Lauri, and Stanley Peters. 1980. Interrogative quantifiers. In Christian Rohrer (Ed.), *Time, Tense and Quantifiers: Proceedings of the Stuttgart Conference on the Logic of Tense and Quantification*, 181–206. Tübingen: Niemeyer.

Keenan, Edward. 1987. A semantic definition of indefinite NP. In Eric Reuland and Alice ter Meulen (Eds.), *The Representation of (In)definiteness*, 286–317. Cambridge: MIT Press.

Keenan, Edward. 1992. Beyond the Frege boundary. *Linguistics and Philosophy* 15(2):199–221.

Keenan, Edward. 1993. Natural language, sortal reducibility and generalized quantifiers. *Journal of Symbolic Logic* 58(1):314–325.

Keenan, Edward, and Leonard Faltz. 1985. *Boolean Semantics for Natural Language.* Dordrecht: Reidel.

Keenan, Edward, and Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9(3):253–326.

Keenan, Edward, and Dag Westerståhl. 1994. Generalized quantifiers in linguistics and logic. In Johan van Benthem and Alice ter Meulen (Eds.), *Handbook of Logic and Linguistics.* Amsterdam: Elsevier. Forthcoming.

É. Kiss, Katalin. 1992. Wh-movement and specificity. *Natural Language and Linguistic Theory* 11(1):85–120.

Ladusaw, William. 1979. *Polarity Sensitivity as Inherent Scope Relations.* PhD thesis, University of Texas at Austin.

Lappin, Shalom. 1988. The semantics of 'many' as a weak determiner. *Linguistics* 26(6):977–998.

Lawler, John. 1971. Any questions? In *Papers from the Seventh Regional Meeting, Chicago Linguistics Society*, 163–173. University of Chicago.

Linebarger, Marcia. 1991. Negative polarity as linguistic evidence. In Lise Dobrin et al. (Eds.), *Papers from the 27th Regional Meeting of the Chicago Linguistics Society. Part Two: The Parasession on Negation*, 165–188.

May, Robert. 1985. *Logical Form: Its Structure and Derivation.* Cambridge: MIT Press.

May, Robert. 1989. Interpreting Logical Form. *Linguistics and Philosophy* 12(4):387–437.

McConnell-Ginet, Sally. 1982. Adverbs and logical form: A linguistically realistic theory. *Language* 58(1):144–184.

Nam, Seungho. 1995. *The Semantics of Locative Phrases.* PhD thesis, UCLA.

Partee, Barbara. 1987. Noun phrase interpretation and type-shifting principles. In Jeroen Groenendijk et al. (Eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, 115–143. Dordrecht: Foris.

Reinhart, Tanya. 1995. Interface strategies. Unpublished manuscript, Utrecht University.

Rudin, Catherine. 1988. On multiple questions and multiple WH fronting. *Natural Language and Linguistic Theory* 6(4):445–501.

Srivastav, Veneeta. 1992. Two types of universal terms in questions. In *NELS 22: Proceedings of the North East Linguistic Society*, 443–457. GLSA Publications, University of Massachussetts, Amherst.

Szabolcsi, Anna. 1994. Quantifiers in pair-list readings and the non-uniformity of quantification. In Paul Dekker and Martin Stokhof (Eds.), *Proceedings of the Ninth Amsterdam Colloquium*, 645–664. ILLC/Department of Philosophy, University of Amsterdam.

Szabolcsi, Anna, and Frans Zwarts. 1993. Weak islands and an algebraic semantics of scope taking. *Natural Language Semantics* 1(3):235–284. Reprinted in this volume.

Westerståhl, Dag. 1985. Determiners and context sets. In Johan van Benthem and Alice ter Meulen (Eds.), *Generalized Quantifiers in Natural Language*, 45–71. Dordrecht: Foris.

Westerståhl, Dag. 1989. Quantifiers in formal and natural languages. In Dov Gabbay and Franz Guenthner (Eds.), *Handbook of Philosophical Logic, Volume IV: Topics in the Philosophy of Language*, 1–131. Dordrecht: Reidel.

Zec, Draga, and Sharon Inkelas. 1990. Prosodically constrained syntax. In Sharon Inkelas and Draga Zec (Eds.), *The Phonology-Syntax Connection*, 365–378. Chicago: University of Chicago Press.

Zhang, Jiannan. 1991. Questions and negation in Chinese. In Lise Dobrin et al. (Eds.), *Papers from the 27th Regional Meeting of the Chicago Linguistics Society. Part Two: The Parasession on Negation*, 345–358.

AUTHOR INDEX

SUBJECT INDEX