Weak islands, individuals, and scope *

Anna Szabolcsi
University of California at Los Angeles

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 other, wh-phrases to extract. E.g.,

(1) Which man$_i$ are you wondering [whether to invite -$_i$]?
(2) * How$_i$ are you wondering [whether to behave -$_i$]?

Two proposals will be considered. The first was developed in Szabolcsi and Zwarts (1990, 1991). It took the syntactic theory of Relativized Minimality as a point of departure, and reinterpreted its generalizations as follows:

(3) Wh-phrases that are sensitive to weak islands are the ones that range over partially ordered domains, rather than discrete individuals.
(4) Weak islands are environments in which the interveners between the wh-phrase and its trace cannot be composed into an upward monotone function.
(5) The link between (3) and (4) is that only upward monotone functions preserve partial ordering.

This proposal, in particular (4) and (5), has certain descriptive and conceptual shortcomings, however. These point to a revision in terms of scope. The characterization of island-sensitive phrases in (3) is retained, but its significance is explicated in terms of the impoverished structure semi-lattices have.

(6) The weak island effect comes about when the wh-phrase should take wide scope over some operator but it is unable to.
(7) For a wh-phrase to take wide scope over another scope bearing element SBE means that operations associated with SBE need to be performed in the domain of wh in order to compute an answer. When a wh-phrase ranges over discrete individuals, these can be collected into unordered sets. All Boolean operations can be performed on sets. When a wh-phrase does not range over discrete

* This research is part of a larger project with Frans Zwarts. I wish to thank him, D. Ben-Shalom, I. Heim, J. Higginbotham, M. Krifka, A. Ojeda, B. Partee, D. Pesetsky, B. Schein, H. de Swart, Z. Szabó, and members of the UCLA Linguistics community for recent discussions. All errors are my own.

individuals, only a smaller set of operations (possibly none) are available in its domain, hence answers cannot be computed in the general case.

(8) Harmless interveners are harmless only in that they can give rise to at least one reading that presents no scopal conflict of the above sort: they can 'get out of the way'. But even they create an island effect if the sensitive wh-phrase has to expressly take scope over them.

The paper is organized as follows. Section **1** reviews the core weak islands data, and outlines the accounts in Rizzi (1990) and Cinque (1991) on the one hand and in Szabolcsi and Zwarts (1990, 1991) on the other. Section **2** explains the notion of individuals on which both the latter proposal and its subsequent revision are based. Section **3** summarizes the monotonicity account and points out its problematic aspects. 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. **5.1** is concerned with claim (7). **5.2** 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. **5.3** establishes a connection between event structure and whether the predicate denotes an ordered or an unordered set.

1.1 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 (1991) 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:

(9) a. Which man$_i$ are you wondering [whether to invite -$_i$]?
b.* How$_i$ are you wondering [whether to behave -$_i$]?

(10) a. Welke man$_i$ vraag jij je af [of je -$_i$ uit moet nodigen]?
'Which man are you wondering whether you should invite?'
b.* Hoe$_i$ vraag jij je af [of je -$_i$ moet gedragen]?
'How are you wondering 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):

(11) ?/* Which man$_i$ are you wondering [who saw -$_i$]?
(12) * Welke man$_i$ vraag jij je af [wie -$_i$ gezien heeft]?
'Which man are you wondering who saw?'
(13) Melyik embert$_i$ találgattad, [hogy ki látta -$_i$]
which man-acc were-you-guessing that who saw

Although the variation in (11) through (13) is not well-understood, I will follow standard practice both in assuming that the strong islandhood of certain wh-complements is syntactic in nature and in restricting my attention to examples that qualify as weak islands in the given dialect or language.

Drawing from work by Obenauer and Ross, Rizzi (1990) and Cinque (1991) observe that the same kind of selectivity is exhibited by many further environments: the presence of *beaucoup*, negation or negative quantifiers, *only*phrases, adversative and factive predicates, and extraposition all create weak islands:

(16) a. Quel livre as-tu beaucoup consulté -?
what book have-you a lot consulted
b.* Combien as-tu beaucoup consulté - de livres?
how-many have-you a lot consulted of books
(17) a. Which man didn't you / did no one think that I invited -?
b.* How didn't you / did no one think that I behaved -?
(18) a. Which man did only John think that I invited -?
b.* How did only John think that I behaved -?
(19) a. Which man did you deny / regret that I invited -?
b.* How did you deny / regret that I behaved -?
(20) 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:

(21) How did everyone / two men think that I behaved -?

They propose the following uniform explanation for the contrasts in (9) and in (16) through (21):

(22) 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.
(23) 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.
(24) REFERENTIAL wh-phrases are those that both bear a thematic role like Agent, Patient, etc. and are Discourse-linked; NON-REFEREN-

TIAL 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 (23i). 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. A-bar specifiers are phrases in a (non-adjoined) position where they do not receive thematic role and/or case. They contrast with complements, A-specifiers, heads, and adjoined phrases. 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 (23ii).

As regards referentiality, Rizzi draws the crucial line between those phrases that refer to participants of 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 Discourse-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:

(25) 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 -?

1.2 Reinterpreting Relativized Minimality in semantic terms

Szabolcsi and Zwarts (1990, 1991) -- henceforth Sz&Z -- accept the above empirical generalizations and propose to reinterpret them in semantic terms. The motivation is three-fold. First of all, the data have a tempting semantic flavor. Second, certain problems with Rizzi's and Cinque's specific claims can be overcome if the generalizations are stated in semantic terms. Third, there is a theoretical challenge. The Relativized Minimality account makes essential reference to traces. Does the weak island phenomenon indeed make traces indispensable, or can it be accounted for in a way that is neutral between theories with and without traces?

The main claims in Sz&Z are as follows. The distinction between good ex-

tractors and bad extractors can be characterized in denotational terms. The former range over INDIVIDUATED domains, the latter over domains whose elements exhibit a PARTIAL ORDERING (inclusion relations). Discourse context plays a role only in that it may individuate things that are normally partially ordered; this is an important but ancillary role. I will call good extractors 'individuals', regardless whether their individuation is inherent or contextual. The characterization of weak islands can be given in terms of the MONOTONICITY properties of the items intervening between the extractor and its trace. Downward monotone and non-monotone interveners block the extraction of non-individuals; upward monotone ones are harmless. The connection lies in the fact that only the upward monotone environments PRESERVE partial ordering. Since individuals are not partially ordered, they are not interested in whether order is preserved: they must be insensitive to weak islands. Non-individuals are partially 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: [1]

(26)

How much milk	did(*n't) you drink
$S/(S/NP)_{MON\uparrow}$	----------compose
	$(S/NP)_{MON\uparrow}$
----------apply	
S	

Assume that *how much milk* is marked to apply to an expression of category S/NP only if it denotes an upward monotone function. This assumption is methodologically analogous to (in fact, is inspired by) Zwarts' (1986) claim that negative polarity items must be arguments of downward monotone functions. Categorial grammar assembles form and meaning simultaneously. Since monotonicity properties are inherited under composition, *did you drink* will be upward monotone, whereas *didn't you drink* will inherit downward monotonicity from *n't*.

In the following sections I discuss the empirical motivation for these claims in some detail, and then go on to point out some problematic aspects.[2]

---

[1] If the extracted phrase is an adjunct, a functor looking for it is created by lifting the category to be modified by it.

[2] The combinatory grammars in Steedman (1987) and Szabolcsi (1992) 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,

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 extractors and, in particular, what role contextualization plays. The arguments to be put forth in this section are consonant with S7&Z but are significantly more elaborate.

I argue that the crucial distinction is between wh-phrases that range over individuals and those that do not. I 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 overlaps and complements we expressly choose to ignore. INDIVIDUALS can naturally be collected into UNORDERED SETS. ('Unordered' sets of course contrast with sets whose members exhibit inclusion relations and not with ordered tuples.) NON-INDIVIDUALS are then characterized by the fact that they exhibit a PARTIAL ORDERING, and this ordering is in fact taken seriously.

The present notion of individuals is as in Szabolcsi (1983), a discussion of the focusing of common nouns in Hungarian in Montague Grammar. The construction in (27) represents a split indefinite, rather than a partitive (see van Riemsdijk (1987) for German). The sentences are equally good with or without a numeral; in the latter case they mean 'at least one of'.

(27) a. **Mit** látott Mari (hármat)?
what+acc saw Mary three+acc
'What did Mary see (three of)?'

b. **Biciklit** látott Mari (hármat).
bike+acc saw Mary three+acc
'It was bikes and nothing else that Mary saw (three of)'

Common nouns denote properties with natural overlaps and complements. If these all are honestly taken into account, (27b) cannot be interpreted as a sensible ex-

---

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.

haustive answer.[3] Instead, properties need to be individuated: a set of relevant properties must be distinguished and Boolean operations disallowed. (In real life, the felicity of (27) does not require an explicit list of relevant properties; the existence of some criterion of relevance is sufficient.) Notice that once we make this move, it does not matter whether the co-relevant properties actually overlap: as we cannot intersect them, we cannot 'see' their overlaps.

This procedure might be subsumed under contextualization: *mi* 'what' now ranges over members of some salient set. What I wish to stress here is that *mi* 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 ordering that is otherwise vital to them. The same semantic change takes place when we have strong D-linking in the sense that the salient set is given as a checklist. Thus both pragmatics and semantics are involved in individuation. My explanation of the weak island phenomenon will rest on the semantic aspect.

Let us first see how the core examples can be described in terms of the individual versus non-individual (ordered) distinction. Wh-phrases like *which person* can easily be taken to range over individuals (plural *which persons*, too, as long as the predicate is distributive; certain non-distributive cases will be taken up later). *Who* and, especially, *what* can range over not only individuals but also properties; the latter are ordered, see above. *How many N*-phrases have an individual interpretation but, like *how many pounds* and *how much attention*, also an amount interpretation. Amounts can only be made sense of in terms of an ordering. The individual/amount ambiguity of numeral phrases is highlighted by the presence or absence of copula agreement in Italian clefts (I owe the observation to Filippo Beghelli). The agreeing version (a) is insensitive to weak islands, the non-agreeing version (b) is sensitive:

(28) a. Sono cinque donne che non ho invitato.
'There are **five** women who I didn't invite'

b.* È cinque donne che non ho invitato.
'The number 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:

(29) a. Combien as-tu (*beaucoup) consulté de livres?
b. Combien de livres as-tu (beaucoup) consulté?
c. Combien de cercles as-tu (*beaucoup) dessiné?
'How many circles did you draw a lot? [* unless types]'

[3] Rooth (1984) recognizes the same problem in connection with verbal focus.

Dobrovie-Sorin (1991) provides interesting evidence from Romanian that a *how many*-phrase has three distinct interpretations: amount, non-D-linked individual, and D-linked individual. D-linked animate direct objects in Romanian are clitic-doubled (c). The following contrasts show that 'how many women' on the amount interpretation cannot extract from a factive island (a), but on the individual interpretation it can extract even if it is not D-linked, i.e., not clitic doubled (b):[4]

(30) Cîte femei regreţi că ai iubit?

a.* 'For what number, you regret having loved that number of women?' (answer: Three.)

b. 'How many women are there such that you regret having loved them?' (answer: There are three such women.)

c. Pe cîte femei regreţi că le ai iubit?
'How many [=which] of the women do you regret having loved?' (answer: Three, namely, A, B, and C.)

Why requires a propositional answer, and propositions are ordered by entailment, a special case of inclusion. It seems 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. This intuition can be corroborated by using *only* as a test. *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 the two:

(31) a. John visited London only in 1953 and 1961.
'exclusively; German *nur*'

b. John's son was born only in 1990.
'merely, recently; German *erst*'

(32) a. Er zijn *alleen* drie stoelen in de kamer.
'There are only three chairs (and nothing else) in the room'

b. Er zijn *slechts* drie stoelen in de kamer.
'There are only three chairs (and no more) in the room'

Frans Zwarts observes that Dutch *alleen* means 'exclusively' and *slechts* 'merely'. It can thus serve to diagnose adverbs:

(33) a.* Hij loste het probleem om 2:00 alleen elegant op.
'He solved the problem at 2:00 only [=exclusively] elegantly'

---

[4] Dobrovie-Sorin (1991) makes the crucial distinction in terms of restricted versus non-restricted quantification.

b.? Hij loste het probleem om 2:00 slechts elegant op.
'He solved the problem at 2:00 only [=merely] elegantly'
c. Hij loste het probleem om 2:00 slechts met tegenzin op.
'He solved the problem at 2:00 only [=merely] reluctantly'
d. Zijn hele leven, loste hij problemen alleen/*slechts elegant op.
'In all his life, he solved problems only [=exclusively] elegantly'

(33a) 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'. (33b) with *slechts elegant* is somewhat strange, since elegance is towards the high end of the scale; (33c) with *slechts met tegenzin* is fine, since reluctance is towards the low end. (33d) 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 extracting abilities:

(34) a.* In what way didn't you solve the problem at 2:00?
b. In what way did you never solve problems?

(34a) 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 each in the answer.

The next question to ask is whether there are other 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 that their counterparts, e.g.,

(35) a. Who are you wondering whether to invite?
b.?? Who the hell are you wondering whether to invite?

I wish to argue that D-linkability is not a minimal difference between *wh-the-hell* expressions and their plain counterparts. Consider the following pair:

(36) 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 (36b). Even so, the contexts in which the two questions are usable are not the same. The existential presupposition wh-questions carry does not prevent (36a) from being an open question,

readily answerable by *Nobody*. (36b) 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 I 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 the institution of 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. (35b). This provides an explanation of why such questions are notoriously bad. On the other hand, in some situations we do have such evidence. E.g., seeing someone madly searching through the dictionary, we may ask (37) or, one thief, seeing another trying to smuggle an item back to a house just robbed, may ask (38):

(37) What the hell do you still not know how to spell?
(38) What the hell are you upset that you took?

Tetsuya Sano (p.c.) informs me that these intuitions are parallelled by the interpretation and behavior of *ittai*-phrases in Japanese. I interpret these data as indicating that D-linking is not the critical factor in the behavior of *wh-the-hell* expressions; they are bad extractors 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.[5]

3.1 Weak islands and monotonicity

Let us now turn to weak islands themselves. Sz&Z observe that the contexts Rizzi

[5] Two remarks on the data. (i) There are significant cross-linguistic differences in the behavior of wh-phrases. E.g., Dutch *welk* and Hungarian *melyik* are much less D-linked than *which* (but extract just as well). Or, Hungarian *mikor*, in contrast to *when*, ranges over individuals well and is a good extractor. (ii) In the Eighties *why* was the paradigmatic island-sensitive example, but it seems quite atypical: it is captured by any 'interesting' thing in sentence. The same holds for German *warum*, even though it can stay in situ (T. Kiss 1991, H. van Riemsdijk p.c.).

and Cinque characterize as weak islands share some simple monotonicity properties: they are all either DOWNWARD MONOTONE or NON-MONOTONE.

(39) a. A function f is upward monotone if for every A, B in its domain, if $A \subseteq B$, then $f(A) \subseteq f(B)$.
b. A function f is downward monotone if for every A, B in its domain, if $A \subseteq B$, then $f(A) \supseteq f(B)$.
c. A function f is non-monotone if neither (a) nor (b).

Let us briefly review how the material in **1.1** fits these notions. *Not*, *no one* and *deny* are clearly downward monotone; 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 non-monotone. Since some of these items are focus-sensitive, I try to keep the focus structure of the examples constant:[6]

(40) a. [I know the answer to the question] who/whether he exercises ←/→
b. [I know the answer to the question] who/whether he does pushups
(41) a. John regrets that I exercise ←/→
b. John regrets that I do pushups
(42) a. Only John exercises ←/→
b. Only John does pushups
(43) a. John exercises a lot ←/→
b. John does pushups 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.,

---

[6] Some comments on (41) and (43). (41) is clearly invalid in the b→a direction. The a→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 monotone if taken extensionally. See Ladusaw (1980) on both points. In (43), the non-monotone analysis of *beaucoup, a lot*, etc. is inspired by Westerståhl (1985), who proposes four interpretations for *many*, two of which are non-monotone due to context-dependence. Suppose John does nothing but pushups for exercise. What he does may count as a lot of pushups but not as a lot of exercise, if the norms associated with the two are different. De Swart (1991) points out that on this view *seldom* would be non-monotone, 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 non-monotone.

which do not create weak islands, are UPWARD MONOTONE. (It is difficult to find a good sample of extraposition islands that are not also factive islands; no proposal is made for them in Sz&Z.)

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: I am 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 trace back the cross-linguistically uniform effect to the uniform semantics of negation.

The most important question is why downward monotone and non-monotone contexts constitute weak islands. The definitions in (39) make it clear that upward monotonicity means simply that the function PRESERVES PARTIAL ORDERING; downward monotone functions reverse it and non-monotone ones obliterate it. Now recall that in the previous section I argued 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 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.

In sum, there seems to be a very natural connection between the semantic properties of islands and island-sensitive phrases.

3.2 Problems

The problems with the above proposal come in two varieties: descriptive and conceptual.

(44) There are downward monotone and non-monotone interveners that for many speakers do not create weak islands.

(45) There are upward monotone interveners which do create weak islands.

(46) Two downward monotone items in the path do not (regularly) cancel out.

(47) The explanation of why downward monotone and non-monotone paths are islands is not as strong as it should be.

Let us consider these in turn.

First, Sz&Z predict that all non-upward monotone interveners are equally bad. But many speakers report a contrast between (48a) and (48b,c):

(48)	a.*	How did few people think that you behaved?	MON↓
	b.	How did exactly five people think that you behaved?	−MON
	c.	How did at most five people think that you behaved?	MON↓

Second, Sz&Z predict that all upward monotone interveners are harmless.[7] De Swart (1991) 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 monotone ones. She also reanalyzes *beaucoup, veel* 'a lot' as upward monotone; this may be a matter of debate, cf. note 6, but 'twice' alone is sufficient to establish her case:

(49)	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

Third, the most natural implementation of Sz&Z's proposal, as was mentioned in **1.2**, 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 monotone interveners are grammatical, since the composition of two downward monotone functions is upward monotone. Now, there is at least one case when this is borne out:

(50)	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-extrac-

---

[7] Szabolcsi and Zwarts (1991) has a chapter on 'gradience', but its data are not built into the theory. I will return to this below.

tion, too, but it does not prove significant under closer scrutiny:

(51) 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, Sz&Z point out that the link between the partially ordered nature of sensitive extractors and the non-upward monotone 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. (52), 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. (53) is also assigned a single structure, but *no one* occupies an A-bar specifier position in it, whence it must block *how*-extraction.

(52) How did everyone behave?
(53) * How did no one behave?

Sz&Z 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 monotone paths do not preserve partial order -- but the assumption that upward monotone interveners QUA INTERVENERS are harmless became part and parcel of the theory.

Results by É.Kiss' (1991) and de Swart (1991) indicate that this assumption is wrong. In addition to pointing out the island creating effect of iterative adverbs, cf. (49), de Swart also notes that sentences like (54) are potentially ambiguous, and they are ungrammatical on the narrow scope universal reading. And, as we shall see, (52) is also ungrammatical on the same reading.

(54) 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'

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' (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 Sz&Z account was based.[8]

4 Weak islands and scope

In what follows I will assume that weak islands are a scope phenomenon. That is, I adopt the following informal version of É.Kiss' (1991) and de Swart's (1991) proposals as a point of departure:

(55) The weak island effect comes about when the wh-phrase should take wide scope over some operator but it is unable to.

(56) Harmless interveners are only harmless in that they can give rise to at least one reading 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.[9] Developing a formal semantic explanation, at least two questions need to be asked:

(57) Why are non-individual wh-phrases restricted in their scope-taking abilities?

(58) What interveners are able to 'get out of the way', and how?

---

[8] Whether Relativized Minimality can be restated to cope with these data is left as an open question. The restatement would involve a modified concept of LF and/or a modified definition of relevant interveners.

[9] (i) 'Specificity Filter: If Op_i is an operator which has scope over Op_j and binds a variable in the scope of Op_j, then Op_i must be specific' [in the sense of Enç (1991)]. (Kiss 1991)

(ii) 'A quantifier Q_1 can only separate a quantifier Q_2 from its restrictive clause if Q_1 has wide scope over Q_2.' (De Swart 1991)

In the following sections I will focus on (57). An answer to (58) is to be developed in Szabolcsi and Zwarts (1992), within the context of how scope behavior is determined by the meanings of the specific quantifiers. Before turning to (57), however, I provide a brief overview of some results in the literature that pertain to (58), and indicate their relation to the monotonicity hypotheses in Sz&Z.

An intervener[10] is harmless iff (i) it is scopeless, or (ii) it can take wide scope over the wh-phrase (family of questions reading), or (iii) it can participate in a scope independent reading with the given wh-phrase (branching, cumulative, etc. readings). The reason why (Relativized Minimality and) Sz&Z's proposal could be descriptively almost correct is that typically, though not all and only, upward monotone 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 are SCOPELESS. As regards (ii), both Groenendijk and Stokhof (1984) and Higginbotham (1991) claim that all quantifiers that are not downward monotone can give rise to a FAMILY OF QUESTIONS reading. These quantifiers have non-empty minimal elements (=sets of individuals); 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 (see also Chierchia (1992)). Indefinites have 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. G&S call this a choice reading.

(59) Who did every man see?
Man$_1$ saw Mary, man$_2$ saw Susan, etc.

(60) Who did two men see?
For instance, John saw Susan, and Bill saw Jill.

It is remarkable that according to G&S, both *exactly five men* and *at most five men*, which were found problematic in (48), give rise to the choice reading (the latter does because it is supposed to allow for an upward monotone group reading). Downward monotone 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

[10] The notion 'intervener' needs to be made more precise. Cases of the type **Who didn't destroy this city?* will show 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.

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, e.g., in Hungarian. But even in this preliminary form they provide a partial explanation of why downward monotone 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 (van der Does 1992, Verkuyl 1992).

(61) 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'

(62) 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, 1991) conducted an empirical study of what noun phrases participate in branching readings in sentences like (61). She identifies a subset of noun phrases denoting upward monotone 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. (No similar study of the restrictions on cumulative readings has been carried out yet.) Questions that may be analyzed as exhibiting these readings are as follows:

(63) How many circles did everyone draw? (branching)
'Everyone drew the same number of circles -- how many was it'

(64) How many circles did these two people draw? (cumulative)
'altogether how many circles'

In a chapter on GRADIENCE, Szabolcsi and Zwarts (1991) observed that downward monotone interveners create the most robust weak islands, while Liu's G-specific noun phrases are the most innocent even among the upward monotone 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. (Below I return to the question why independent readings are good.)

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 I take it that the global plausibility of the scope account is established.

5 Individuality and scope

This section addresses the question why certain wh-phrases cannot take wide scope and are thus sensitive to weak islands. I will adhere to the claim, advanced in Sz&Z and elaborated in **2**, that the crucial property these wh-phrases have is that they do not range over individuals, i.e., over members of unordered sets. The essence of the claim to be put forth is as follows:

(65) For a wh-phrase to take wide scope over another scope bearing element SBE means that operations associated with SBE need to be performed in the domain of wh in order to compute an answer. When a wh-phrase ranges over discrete individuals, these can be collected into unordered sets. All Boolean operations can be performed on sets. When a wh-phrase does not range over discrete individuals, only a smaller set of operations (possibly none) are available in its domain, hence answers cannot be computed in the general case.

Discussion will proceed in three steps. Section **5.1** justifies claim (65). **5.2** provides new empirical support for the claim that it is precisely this notion of individuality that plays a role here. 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. **5.3** argues that whether a domain consists of sums or unordered sets depends on whether the predicate is iterative and summative in the pertinent respect.

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:

(66) Who did Fido see?
(67) Who didn't Fido see?
(68) Who did every dog see?
(69) Who did at least two dogs see?

I assume that the interpretation of questions, whatever it should precisely be, includes that an exhaustive list is determined by the answer. I will be concerned with how such a list can be computed or verified.

The steps to be described will of course be unsurprising. For (66) we form the set of people that Fido saw, and list its members. For (67), we form the complement of this set. For (68), we take the sets of individuals that each dog saw, intersect them, and list the members of the intersection. If (69) 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. (The algorithm is of no interest here.) These contrast with the pair-list reading of (68), for instance, where no Boolean operation needs to be performed.

The moral is that for a wh-phrase to take WIDE SCOPE OVER some scope-bearing element SBE means that the computation/verification of the answer involves specific OPERATIONS associated with SBE.

Notice three respects in which this definition is rather general. First, it does not require for the narrow scope SBE to become referentially dependent on the wide scope taker, hence SBEs like negation are covered. Second, it does not require for the wide scope taker to 'bind a variable' in the scope of the narrow scope SBE; it is structurally neutral. Third, the definition would easily extend from wh-phrases to arbitrary quantifiers.

In this paper I restrict my attention to simple cases as above, where the operations associated with SBE are just the BOOLEAN operations. I will not try to explicate what operations are involved when a wh-phrase takes scope over a factive verb, for instance.

A simple consequence of the above is that a particular wh-phrase is able to take scope over some SBE iff the requisite operations are available in the domain the wh-phrase ranges over. In (67)-(69) this was no problem. *Person that Fido saw* denotes a set of individuals; an UNORDERED SET. Unions, intersections, and complements are defined for this domain. But are these operations available in the domains of all wh-phrases?

In section 2 I argued that island-sensitive wh-phrases do not range over individuals but, rather, elements of PARTIALLY ORDERED domains. To be more specific, their domain will be said to have the structure of a join semilattice. The following are standard definitions:

(70) a. A BOOLEAN ALGEBRA is a partially ordered set closed under unions, intersections, and complements.

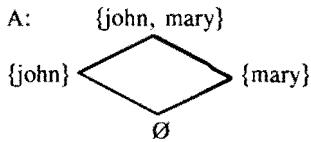
b. A LATTICE is a partially ordered set closed under unions and intersections.

c. A JOIN SEMILATTICE is a partially ordered set closed under unions.

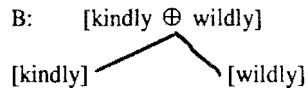
Join semilattices have been proposed for various domains, e.g. mass terms and plurals. Manner expressions resemble mass terms in most respects. For instance, if John behaved kindly but wildly, then the components of his behavior do not form a set {kindly, wildly} but, rather, a single object that is made up of kindness and wildness. The conjunction can be written as *kindly* ⊕ *wildly*, where ⊕ is interpreted as the sum -- union -- operation. Note that it must be union, not intersection. A kind but wild behavior is not one that has just the features common to both kindness and wildness; it is what you get when you put the two together. It is another matter whether the people who behaved kindly but wildly are in the intersection of those who behaved kindly and those who behaved wildly. (The answer depends on whether we regard *behave* as a distributive predicate; if we do not, only a weaker relation will hold.)

The following diagrams illustrate the structures of a Boolean algebra and a join semilattice:

(71) *that who Fido saw* ∈ A



the way how John behaved ∈ B



The semilattice B is not closed under complements and intersections. Therefore, the computation/verification of an answer must not depend on these operations: an expression that ranges over the elements of a join semilattice cannot take scope over negation or a universal quantifier (or over any more complex operator whose definition inescapably incorporates at least one of these).

The union operation that is available in the join semilattice does allow for some computations, though. For instance, whether entity A is part of entity B, or whether two entities are identical, can be checked using only unions:

(72) a. If $A \subseteq B$, then $A \cup B = B$
b. If $A \subseteq B$ and $B \subseteq A$, then $A = B$

Thus, for instance, the answer to a question like *Did John pay at least three cents?* can be computed by checking whether the amount John paid is bigger than, or equal to, three cents.

(73) has no intersective reading: if Mary behaved clumsily but nicely, Peter loudly but nicely, and John quietly and nicely, (73) cannot be asking for 'Nicely' as an answer. It does have the wide scope universal (family of questions) reading. But, as É.Kiss (1991) observed, it has an additional good interpretation:

(73) How did everyone behave?
'What was the uniform behavior exhibited by everyone?'

This presupposes, rather than asserts, that everyone behaved in the same way. This is an extremely strong presupposition, whence (73) cannot be used as an innocent question. But once this presupposition is granted, we do not need to take intersections. Since everyone behaved in the same way, it is sufficient to inspect an arbitrary person's behavior and report that. This reading can also be regarded as a special instance of branching quantification, with a slight extension of the notion. The two fully connected entities are the set of persons and the unique (full) behavior that characterizes every person. More important to us now is the fact that this reading does not require the performance of intersections.

At this point the question arises why we 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 (67) through (69) 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 scope-bearing 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, context forced us to individuate loudness and twenty pounds by requiring, for instance, that the presence of these particular properties be checked in every dog.) 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.

I am convinced that 'look-up' plays an important role in a pragmatic/procedural model. But it does not eliminate the need for computation, and hence it does not eliminate the vulnerability of wh-phrases that denote in an impoverished domain.

It may be important to point out a difference between the roles this proposal and Sz&Z 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,

even an abstract kind, then Sz&Z 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 answer computation.

5.2 Unique arguments and weak islands

Although the analysis sketched in **5.1** is logically plausible, one may still be uncertain whether the lack of Boolean structure alone is sufficient to lead to island-sensitivity. In this section I will discuss a set of extractors which, as far as I can see, share nothing else but this aspect 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, destroy this building,* and *win the Rimet Cup in 1978* are 'one time only' predicates: it is not possible to get the same letter (token), or to destroy the same building (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 I 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 (74) examples the distributive answer *John did and Bill did* is as acceptable as *John and Bill did*. In 'one time only' (75), the former is unacceptable: John and Bill must form a collective recipient . Similarly, in (74) the short (exhaustive) answer *Bill* can be modified by *only*. In (75) it cannot or, more precisely, if *only* is acceptable, it must mean 'alone' and not 'exclusively'. The effect disappears in (76).

(74) a. Who showed this letter 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.

(75) a. Who got this letter from Mary?
John and Bill did / * John did and Bill did.
Bill did / (*) Only Bill did.

b. Who destroyed this building?

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.

(76) 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.,

(77) From whom did you get this letter?
From Mary / (*) Only from Mary.

(78) When did you get this letter?
Yesterday / Only yesterday [=not earlier].

This phenomenon, together with its consequences for scope, was observed in Szabolcsi (1986: 334-7). In what follows I will somewhat enlarge the set of data and spell out the explanation in terms of the present proposal.

(79) and (80) 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.) (81) and (82) 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.

(79) a. Who didn't show this letter to Mary?
To whom didn't you show this letter?

b.* Who didn't get this letter from Mary?
From whom didn't you get this letter?

(80) a. Who showed every letter to Mary?
To whom did you show every letter?

b.* Who got every letter from Mary?
From whom did you get every letter?
[unless pair-list or same person for every letter]

(81) a. To whom do you regret having shown this letter?

b.* From whom do you regret having gotten this letter?

(82) a. To whom do you wonder whether I showed this letter?

b.* From whom do you wonder whether I got this letter?

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 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 from Mary* denotes a sum of whose parts the predicate does not hold independently:

(83) ιx[get the 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 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 to perform certain operations in the computation/verification of the answer, rather than to the wide scope taker's binding a variable within the syntactic scope of the other operator.

5.3 Event structure and set formation

In this section I 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 answer computation will also be made more precise.

The basic idea derives from Carlson's (1984: 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 John and Bill as its unique Agent, and in *Bob washed the car*, the car is the Theme, and its parts are not.

I dub events characterized by thematic uniqueness MINIMAL EVENTS ($e_{m/i}$):

(84) 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. I 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? I 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:[11]

(85) 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*. I 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:

(86) 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 I adhere): the same city cannot be destroyed more than once.

NON-ITERABILITY means that the predicate describes a biunique relation between slobjects and minimal events. I encode this by writing the event parameter as a FUNCTION of that other parameter with respect to which the event is not iterative. (The agent may be so written, too, but it does not seem necessary.)

(87) $\text{destroy(Rome)(Bob)}(f_e(\text{Rome}))$

Prior to proceeding to the description of events involving manners and amounts, let us see how the above assumptions are utilized in set formation. I will use 'set' to mean unordered set, unless otherwise specified.

I stipulate that SET FORMATION takes place iff the predicate is both summative and iterable. On the basis of (84) we can form the standard denotation of the predicate *visit Rome*, the set of those who visit Rome, as follows:

(88) $\lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \wedge x \in \bigcup_{i \in I}\{x\colon [\text{visit(Rome)}(x)(e_{m/i})]\}]]$

---

[11] This definition as well as (86) and (99) are borrowed from Krifka (1990).

The first step is to form the set of those who visited Rome at some particular minimal event. This will be a singleton. Then, for all minimal events $i \in I$, we union the singletons of those whose visited Rome at $e_{m/i}$, and thus form the set of those who ever visited Rome (within the event range defined by I). At the same time, the minimal events are lumped together into one big event e (here union amounts to sum formation).

The empirical claim that is being made here is that non-iterable and/or non-summative 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 I will NOT say that. The denotation of a non-iterable predicate remains a slobject. The intuition behind this is that a predicate denotes a set iff it can in principle hold of more than one thing independently. Empirical support for this intuition comes from the data reviewed in **5.2**, i.e., the fact that the questioning of a unique parameter is sensitive to weak islands.

The computation of an answer to *Who visited Rome?* now involves (88), but that of an answer to *Who destroyed Rome?* can involve only (89):

(89) $\quad \iota x[\text{destroy(Rome)}(x)(f_c(\text{Rome}))] = ?$

As regards *Who didn't visit Rome?*, *Who visited every city?*, and *Who visited a(ny) city?*, the reasoning in **5.1** can be reproduced as follows. If we have sets, as in (88), we can form their complements, or we can intersect and union them with others. The outputs also feed the Boolean operations.

(90) $\quad U - \lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \ \wedge \ x \in \bigcup_{i \in I}\{x\colon [\text{visit(Rome)}(x)(e_{m/i})]\}]] = ?$

(91) $\quad \bigcap_{n \in N}(\lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \ \wedge \ x \in \bigcup_{i \in I}\{x\colon [\text{visit(city}_n)(x)(e_{m/i})]\}]]) = ?$

(92) $\quad \bigcup_{n \in N}(\lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \ \wedge \ x \in \bigcup_{i \in I}\{x\colon [\text{visit(city}_n)(x)(e_{m/i})]\}]]) = ?$

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. Similarly, *Who destroyed every city?* cannot have a reading parallel to (91). 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 person destroyed every city, cf. (73). This latter will be expressed roughly as follows:

(93) $\quad \iota x \forall z[\text{destroy(city } z)(x)(f_c(z))] = ?$

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:

(94) $\quad \bigcap_{i \in I}\{x\colon [\text{destroy(city}_i)(x)(f_c(\text{city}_i))]\} = ?$

However, this interpretation asserts, rather than presupposes, that the same person destroyed every city, which seems counterintuitive. Furthermore, it would predict that a complement can be formed in the next step: *Who didn't destroy every city?* This is wrong, so (93) is 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. Due to the first fact we cannot use (92). 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:

(95) $\lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \ \wedge \ x \in \bigcup_{i \in I}\{x\colon [\text{destroy}(\text{city}_i)(x)(e_{m/i})]\}]] = ?$

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:

(96) $\text{behave}([\text{kindly} \oplus \text{stupidly}])([\text{John}])([e_{m/i}])$

Second, while both the *behaved kindly and/but stupidly* and the *John behaved* relations are iterable, summativity fails (we never get cumulative readings):

(97) John behaved kindly at event e and John behaved stupidly at event e´
/→ John behaved kindly and stupidly at e∪e´

(98) John behaved kindly at event e and Bob behaved stupidly at event e
/→ 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 (93).

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´.

(99) 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, I will only examine the *two cities* type. Following Krifka (1990), I 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:

(100) John visited six cities at e and John visited five cities at e´
→ John visited eleven cities at e∪e´

(101) John visited six cities at e and Bob visited five cities at e
→ 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:

(102) $\mu(\lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \wedge x \in \bigcup_{i \in I}\{c: [visit(c)(John)(e_{m/i})]\}]]) = ?$

Similarly, a good reading can be computed for *How many cities didn't you visit?*, etc. by measuring the complement of the set of cities visited:[12]

(103) $\mu(U - \lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \wedge x \in \bigcup_{i \in I}\{c: [visit(c)(John)(e_{m/i})]\}]]) = ?$

For the cumulative reading of *John and Bob visited eleven cities*, the two sets of cities are unioned before measuring (I do not provide a general algorithm here):

(104) $\mu((\lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \wedge x \in \bigcup_{i \in I}\{c: [visit(c)(John)(e_{m/i})]\}]]) \bigcup$
$(\lambda x[\exists e \exists I[e = \bigcup_{i \in I} e_{m/i} \wedge x \in \bigcup_{i \in I}\{c: [visit(c)(Bob)(e_{m/i})]\}]])) = ?$

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, and the other is that our sets were subjected to an operation whose value is itself not a set.

---

[12] 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. I suggest that to capture this, and the behavior of pure amount readings in general, we measure non-iterable events directly. The elaboration of this suggestion goes beyond the scope of this paper, however.

References

Carlson, Gregory. 1984. Thematic roles and their role in semantic interpretation. Linguistics 22. 259-279.

Chierchia, Gennaro. 1992. Functional wh and weak crossover. Proceedings of the Tenth WCCFL. ed. by Dawn Bates, 75-91. Stanford: CSLI.

Cinque, Guglielmo. 1991. Types of A'-dependencies. Cambridge: MIT Press.

Comorovski, Ileana. 1989. Discourse and the syntax of multiple constituent questions. Cornell University. Doctoral dissertation.

Dobrovie-Sorin, Carmen. 1991. What does QR raise? Paris: CNRS. Manuscript.

van der Does, Jaap. 1992. Applied quantifier logics. Amsterdam University. Doctoral dissertation.

Groenendijk, Jeroen and Martin Stokhof. 1984. Studies on the semantics of questions and the pragmatics of answers. Amsterdam University. Doctoral dissertation.

Harada, Yasunari and Naohiko Noguchi. 1992. On the semantics and pragmatics of *dake* (and *only*). In this volume.

Hepple, Mark. 1990. The grammar and processing of order and dependency: a categorial approach. Edinburgh University. Doctoral dissertation.

Higginbotham, James. 1991. Interrogatives I. More papers on wh-movement (MIT WPL 15). ed. by Lisa Cheng and Hamida Demirdash. 47-76. Cambridge: MIT.

É.Kiss, Katalin. 1991. Wh-movement and specificity. To appear in Natural Language and Linguistic Theory.

Kiss, Tibor, 1991. Some properties of negation in German. Stuttgart: IBM. Manuscript.

Krifka, Manfred. 1990. Four thousand ships passed through the lock. Linguistics and Philosophy 13. 487-520.

Kroch, Anthony. 1989. Amount quantification, referentiality and long wh-movement. University of Pennsylvania. Manuscript.

Ladusaw, William, 1980. Affective OR, factive verbs, and negative polarity items. Proceedings of the Sixteenth CLS. 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. UCLA. Doctoral dissertation.

Liu, Feng-hsi, 1991. Branching quantification and scope independence. Generalized quantifier theory and applications. ed. by Jaap van der Does and Jan van Eijck. 297-331. Dutch Network for Language, Logic and Information.

May, Robert. 1985. Logical form: its structure and derivation. Cambridge: MIT Press.

Melvold, Janis. 1991. Factivity and definiteness. More papers on wh-movement (MIT WPL 15). ed. by Lisa Cheng and Hamida Demirdash. 97-117. Cam-

bridge: MIT.
Pesetsky, David. 1987. Wh-in-situ: movement and unselective binding. The representation of (in)definiteness. ed. by Eric Reuland and Alice ter Meulen. 98-130. Cambridge: MIT Press.
Progovac, Ljiljana. 1988. A binding approach to polarity sensitivity. University of Southern California. Doctoral dissertation.
van Riemsdijk, Henk. 1987. Movement and regeneration. Tilburg University. Manuscript.
Rizzi, Luigi. 1990. Relativized minimality. Cambridge: MIT Press.
Rizzi, Luigi. 1992. Argument/adjunct (a)symmetries. To appear in the Proceedings of NELS 21.
Rooth, Mats. 1984. How to get even with domain selection. Proceedings of NELS 14. ed. by Charles Jones and Peter Sells. 337-402. Amherst.
Steedman, Mark. 1987. Combinatory grammar and parasitic gaps. Natural Language and Linguistic Theory 5. 403-440.
de Swart, Henriette. 1991. Islands, monotonicity and scope. Groningen: RUG. Manuscript.
Szabolcsi, Anna. 1983. Focussing properties, or the trap of first order. Theoretical Linguistics 10. 125-145.
Szabolcsi, Anna. 1986. From the definiteness effect to lexical integrity. Topic, focus and configurationality. ed. by Werner Abraham and Sjaak de Meij. 321-348. Amsterdam: John Benjamins.
Szabolcsi, Anna. 1992. Combinatory grammar and projection from the lexicon. Lexical matters. ed. by Ivan Sag and Anna Szabolcsi. 241-268. Stanford: CSLI.
Szabolcsi, Anna and Frans Zwarts. 1990. Semantic properties of composed functions and the distribution of wh-phrases. Proceedings of the Seventh Amsterdam Colloquium. ed. by Martin Stokhof and Leen Torenvliet. 529-555. Amsterdam: ITLI.
Szabolcsi, Anna and Frans Zwarts. 1991. Unbounded dependencies and algebraic semantics. 3rd European Summer School in Language, Logic and Information. Saarbrücken. Lecture notes.
Szabolcsi, Anna and Frans Zwarts. 1992. Weak islands. UCLA and Groningen University. Manuscript.
Verkuyl, Henk. 1992. Some issues in the analysis of multiple quantification with plural Nps. Utrecht University. Manuscript.
Westerståhl, Dag. 1985. Logical constants in quantifier languages. Linguistics and Philosophy 8. 387-413.
Zimmermann, Ede. 1991. Scopeless quantifiers and operators. Generalized quantifier theory and applications. ed. by Jaap van der Does and Jan van Eijck. 207-223. Dutch Network for Language, Logic and Information.
Zwarts, Frans. 1986. Categoriale grammatika en algebraïsche semantiek. Groningen: RUG. Doctoral dissertation.