What's in an island?

HANS VAN DE KOOT & ERIC MATHIEU

Abstract

The primary aim of this paper is to argue that a unified approach to strong and weak islands is misguided: strong and weak islands are different phenomena, to be accounted for at different levels of representation. Our central claim is that a theory of locality that represents these islands in one level cannot capture the pattern of island sensitivity exhibited by the various grammatical dependencies. The relevant empirical difficulties are easily overcome in an approach that treats strong islands as syntactic and weak islands as scope-related. We then attempt to answer the question whether scope is represented at LF or later. In order to maintain the view that scope is an LF phenomenon, one has to demonstrate that strong islands are not due to properties of LF but have their origin in the order of syntactic operations in a derivational syntax. We show that such derivational theories of strong island phenomena have major shortcomings. This strongly suggests that scope is not an LF phenomenon.

1 Introduction

As is well-known, certain syntactic relations are sensitive to strong but not to weak islands. Thus, while WH movement of an argument out of an adjunct or subject is impossible (1a,b), moving such a constituent across a weak island inducer, such as negation, is generally acceptable (1c).

(1) a. *Which book did you suggest a movie to John [after reading *t*]
b. *Which book is [reading *t* to children] fun
c. Which book didn't you think that John had read *t*

There are also relations that exhibit weak but lack strong island effects. The dependency between a negative polarity item and its licenser is a case in point. The contrast between (2a) and (2b) shows that *anything* requires an operator of a certain type. In (2c) the intervention of *fewer than three beggars* gives rise to a weak island violation (Honcoop 1998), while the relation between *no one* and *any intermission* freely crosses an adjunct boundary in (2d). The ungrammaticality of (2e) confirms that we are dealing with an NPI and not with free choice *any*.

(2) a. *Someone gave Mary anything.
b. No one gave Mary anything.

c. *No one gave fewer than three beggars anything.
d. No one gave Mary a drink during any intermission.
e. *Someone gave Mary a drink during any intermission.

A third class of relations, exemplified by the dependency between a WH adjunct and its trace, is sensitive to both strong and weak islands:

(3) a. How did John say [that the mechanic had repaired his car *t*]
b. *How didn't John say [that the mechanic had repaired his car *t*]
c. *How did John thank the mechanic [after he had repaired his car *t*]

The final possibility, a relation that is sensitive to neither type of locality condition, is also attested. For instance, the dependency between an anaphor and its binder may cross an adjunct boundary (4a), a weak island inducer like negation (4b), or indeed both at the same time (4c).

(4) a. The professors fell asleep [during each other's syntax lecture]
b. The professors didn't like each other's syntax lecture
c. The professors didn't fall asleep [during each other's syntax lecture]

The situation is summarized in (5).

(5)

	STRONG ISLAND SENSITIVE	WEAK ISLAND SENSITIVE
WH ARGUMENT	+	–
WH ADJUNCT	+	+
NPI	–	+
ANAPHORS	–	–

There is ongoing debate about the nature of strong and weak islands. Some researchers aim for a theory of locality that is entirely syntactic (Rizzi 1990, Manzini 1992, Starke 2001, and others), while another school of thought considers weak islands a semantic phenomenon, belonging entirely to the theory of scope (De Swart 1992, Szabolsci & Zwart 1993, Honcoop 1998, and others). Within that group, at least some researchers would be inclined to treat all island phenomena as essentially semantic (Szabolsci & Den Dikken 2002). However, no matter one's preferences in this regard, every theory of locality faces the challenge of explaining the island sensitivity of the four types of dependencies in (5). To the best of our knowledge, no unified theory of locality can claim to be able to do so.

Indeed, the primary aim of this paper is to argue that a unified approach to strong and weak islands is misguided: strong and weak islands are different phenomena, to be accounted for at different levels of representation. Our central claim is that a

theory of locality that represents these islands in one level cannot capture the pattern of island sensitivity in (5). The relevant empirical difficulties are easily overcome in an approach that treats strong islands as syntactic and weak islands as scope-related (see Lechner 1998 for similar conclusions based on different arguments).

The paper is organized as follows. Section 2 presents a minimalist encoding of grammatical dependencies that lives up to the stringent requirements on phrase structure imposed by Inclusiveness (Neeleman and Van de Koot 2002). Section 3 shows that this proposal makes available a characterization of strong islands that manages to unify (i) strong island phenomena, (ii) the existence of structures containing parasitic gaps and secondary predication, and (iii) the anti-c-command requirement on such structures. The theory achieves this by relating all these phenomena to the effects of a single LF constraint.

Section 4 reviews evidence showing that weak islands arise when certain scopal elements intervene between a bare operator and its restriction (Beck 1995, Honcoop 1998). Crucially, such configurations can be base generated but may also result from scope reconstruction. A scope-based theory of weak islands allows a unified account of a variety of so-called intervention effects, including those found in partial WH movement constructions (see Mathieu 2002 for extensive discussion and further references).

Section 5 takes up the question whether strong and weak islands can be represented in a single level of representation (namely LF). Such a theory requires this single level to represent syntactic dependencies and the output of scope reconstruction simultaneously (as in Hornstein 1995, Fox 1999, and many others). We argue that such a unified approach to strong and weak island phenomena is empirically inadequate.

This leaves us with two possibilities. It could be that strong islands are an LF phenomenon but that scope reconstruction takes place at a later level. Alternatively, scope *is* represented at LF but strong island phenomena are not. On this latter view, these islands are not due to properties of LF but have their origin in the order of syntactic operations in a derivational syntax (Uriagereka 1999, Stepanov 2001, Johnson 2002).

Section 6 compares such derivational approaches to strong islands with the theory of strong islands of section 3 and shows that the latter is superior in a number of respects. This leads us to conclude in section 7 that strong islands are an LF phenomenon and that scope (and hence scope reconstruction) is represented post-LF.

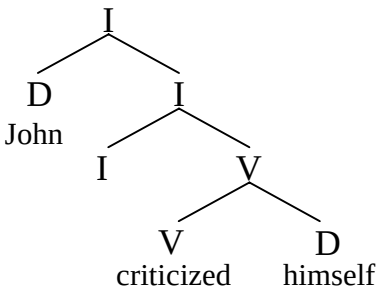
2 Inclusiveness and syntactic dependencies
2.1 The syntactic encoding of dependencies

The basic assumption underlying the theory of bare phrase structure (Chomsky 1995) states that properties of syntactic structure must derive from information stored in the lexicon:

(6) *Inclusiveness*: The syntactic properties of a nonterminal node are fully recoverable from its daughters; the syntactic properties of a terminal node are fully recoverable from the lexicon.

Inclusiveness has profound implications for the theory of grammatical dependencies. In particular, the standard analysis of such dependencies is incompatible with it. Consider the relation between the reflexive *himself* and its antecedent *John* in (8).

(7)



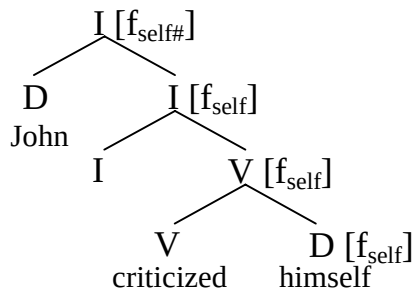
In this binding relation, the antecedent *John* caters to a property of the anaphoric element. As a result, the latter acquires new properties that 'neutralize' its dependent status (so that it cannot enter into other binding relations). However, these new properties cannot be recovered from the internal structure of the anaphoric DP. Indeed, they are not detectable at all until we consider the structure dominated by the root node, because this is the smallest structure that contains both the anaphor and its antecedent. It follows that any subtree that contains the anaphor but not its antecedent is noninclusive in the sense of (6).

Neeleman and Van de Koot (2002) propose an encoding of grammatical dependencies that is compatible with Inclusiveness. Consider again the relation between a reflexive and its binder. The relevant dependency is obligatorily established whenever a reflexive is present in the structure. It is therefore natural to see it as resulting from a selectional requirement associated with this element. Henceforth, we refer to such a selectional property as a (syntactic) function. Although the functions that mediate syntactic dependencies (binding, θ-marking,

movement, etc.) generally receive some interpretation at the LF interface, they are syntactic objects that undergo syntactic operations.[1]

Neeleman and van de Koot show that Inclusiveness dictates that the satisfaction of selectional requirements is mediated by two primitive operations: upward copying and downward function application. On this view, a binding relation is associated with a syntactic structure as in (8).

(8)



Here the reflexive introduces the selectional function f_{self} as a lexical property. The function is repeatedly copied upward until it directly dominates its argument and is satisfied (informally indicated by '#'). The resulting structure satisfies Inclusiveness: the content of each nonterminal node is recoverable from its daughters, while the properties of each terminal node are recoverable from the lexicon. Crucially, Inclusiveness also has consequences for function application: the satisfied status of a function must be recoverable from a daughter of the node that hosts the function. It follows that the satisfaction of selectional requirements must be downward.

Although function application is downward, copying can only be upward. This is so because downward copying gives rise to a node whose properties can only be recovered from its mother, in violation of Inclusiveness.

Copying transfers information from one node to another. It may therefore apply recursively, with the result that the upward trajectory of a function is in principle unbounded. By contrast, function application is restricted to immediate domination. Hence, this theory of dependencies explains the generalization that the antecedent in a syntactic dependency invariably c-commands the dependent element.

---

[1] C-selection appears to be mediated by a function that has no interpretation at LF. See Neeleman, Van de Koot and Doetjes (to appear) for an argument that c-selection cannot be reduced to semantic selection.

2.2 The syntactic encoding of order

The theory just outlined treats nodes as collections of features and functions. Neeleman and Van de Koot make the following assumption concerning the distinctness of features and functions in a node:

(9) *Distinctness*: The syntax interprets attributes of a node that cannot be distinguished as one and the same.

The syntax distinguishes properties trivially if they have inherently different properties. However, some additional mechanism is required if a node is to contain the same function more than once, as in the case of a verb containing more than one thematic function. To this end, functions in a node may be linked to a so-called ordering tier, thereby establishing a partial order in a node. As with any other syntactic property, the distribution of linking in a structure is constrained by Inclusiveness. This has the important consequence that linking in a nonterminal node must be inherited from a daughter, while linking in a terminal node must be recoverable from the lexicon. Finally, the ordering tier is assumed to be an intrinsic property of a head. This is formalized by assuming that the ordering tier contains a head's categorial feature. Since linking is only preserved if the tier is copied to the dominating node, it follows that linking is only preserved if projection takes place.

2.3 Licensing functions

The functions introduced by dependent elements come in two types: some have a licensing capacity, others do not. An argument, for example, can only be combined with a predicate if it satisfies a θ-function (f_θ). In this sense, then, a predicate licenses the presence of another phrase in the tree structure. Similarly, landing sites of movement must be dominated by the function mediating movement (f_{move}). Since a structure with movement contains one more constituent than the same structure without movement, the trace must have a licensing capability with respect to its antecedent.

The effects of licensing are shown in (10) and (11). In (10b), either *John* or *Mary* fails to satisfy a θ-function (since *sleep* introduces only one), while in (11b) *what* fails to satisfy f_{move} (since this sentence does not contain an element that introduces this function).

(10) a. John sleeps.
b. *John sleeps Mary.

(11) a. What did John buy t_{DP}
b. *What did John buy the book.

By contrast, the antecedent of a binding relation is not structurally licensed by the function f_{self}. Hence, the grammaticality of (12a) is not affected when *himself* is replaced by the non-dependent element *Mary* in (12b).

(12) a. John criticized himself.
b. John criticized Mary.

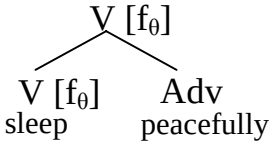
There is something rather mystifying, however, about the claim that a constituent that needs no special licensing when it occurs on its own should need licensing when present in a bigger structure. Neeleman and Van de Koot explore an alternative that takes a licensing function to sanction the structure that accommodates the constituent, rather than the constituent itself. Assume that adjunction creates a multi-segmented category, while the sister of an argument and its mother belong to different categories (Chomsky 1986). A licensing function can then be thought of as licensing the creation of a category:

(13) *Categorial licensing*: A category (as opposed to a segment) is licensed by the application of a licensing function.

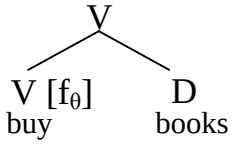
(14) *Category*: Let α be an immediate projection of β. Let n_α be the number of unsatisfied licensing functions in α copied from its daughters. Let n_β be the number of unsatisfied licensing functions in β. α and β are segments of the same category iff n_α equals n_β.

According to (14), two nodes in a projection line form segments of category if the arity information in the higher segment is identical to that copied from the lower segment. The tree in (15a) illustrates this. By contrast, merging an argument with a predicate creates a new category, as in (15b). As required by (13), the creation of a category in (15b) is sanctioned by the application of a licensing function.

(15) a.



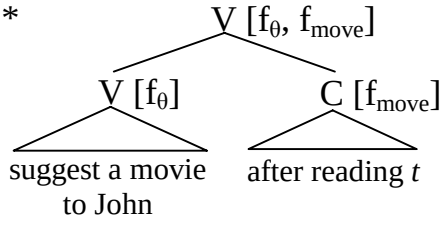
b.



3 Characterizing strong islands
3.1 Categorial licensing and the adjunct island condition

At first sight, the theory of categorial licensing expressed by (13) and (14) seems entirely circular: the creation of a category requires the satisfaction of a licensing function, while the satisfaction of a licensing function gives rise to the creation of a category. But (14) leaves open the possibility that the creation of a category is the result of an increase in the number of licensing functions. Consider the structure associated with WH movement out of an adjunct:

(16) *



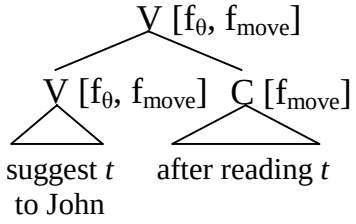
This is the partial representation associated with the ungrammatical example in (1a), repeated here as (17).

(17) *Which book did you suggest a movie to John [after reading *t*]

In (16), the higher V-node is a category. Hence, Categorial Licensing requires that a licensing function be satisfied in this node. Since this has not happened, the structure is ruled out.

If the function f_{move} copied from the adjunct is copied into a verbal projection line that already contains such a function, as in (18), the two functions 'collapse' into one on the receiving node, because – by Distinctness – they cannot be distinguished in the absence of ordering. It is impossible to introduce 'ad hoc' order on the higher verbal node in (18), as that would lead to a violation of Inclusiveness: the order is not inherited from either daughter. In the resulting structure, the two verbal nodes constitute segments of a single category. Since no new category has been created, Categorial Licensing does not apply and the structure is well-formed.

(18)



(18) is a partial structure for the sentence in (19), involving a parasitic gap.

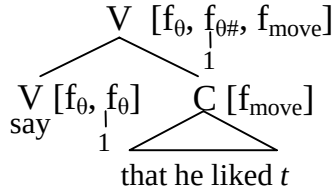
(19) Which book did you suggest to John after reading?

We see, then, that the same principle that blocks movement relations across an adjunct boundary permits such movement if it is parasitic on an identical relation outside the adjunct.

Although Categorial Licensing prevents a Move function from being copied out of an adjunct, as in (16), the copying of such a function out of a complement is allowed. Consider the partial structure in (21) for the example in (20), which involves extraction from a complement CP.

(20) Which book did you say that he liked *t*

(21)



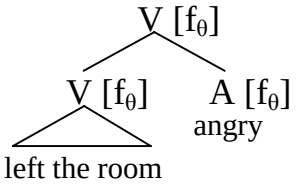
The higher V-node in (21) is a category and, as required by categorial Licensing, the creation of this category is accompanied by the satisfaction of a licensing function.

Since θ-functions are also licensing functions, one would expect to be able to construct examples parallel to those above, but involving predication. This is indeed the case. Consider the well-known contrast in (22), demonstrating that secondary predication is parasitic on primary predication in much the same way that a parasitic gap is parasitic on a primary gap.

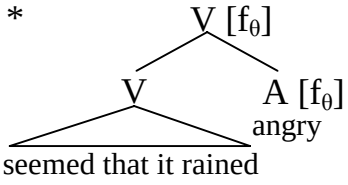
(22) a. John left the room angry.
b. *John [seemed [that it rained] angry]

Their associated (partial) structures are given in (23a) and (23b), respectively.

(23) a.



b. *



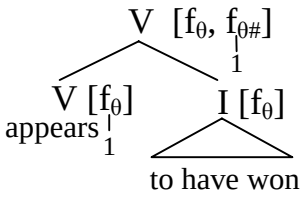
If the θ-function introduced by the secondary predicate can collapse with an identical function in the verbal projection, no category is created and the result is grammatical. However, if the function introduced by the depictive is copied to the

verbal projection and thereby gives rise to the creation of a category, as in (23b), Categorial Licensing is triggered and the structure is ruled out.

In the same way that Categorial Licensing does not prevent the copying of a Move function from a complement, it also does not block the copying of a θ-function in this environment. Consider (24), a case of NP-raising, and its partial structure in (25).

(24) John appears to have won.

(25)



The higher V-node in (25) is a category and, as required by Categorial Licensing, the creation of this category is accompanied by the satisfaction of a licensing function.[2]

The explanation offered for the ungrammaticality of (17) and (22b) receives support from the fact that it also captures the anti-c-command requirement on parasitic gaps and secondary predicates. The mere presence in a structure of a primary dependency is never enough to license the secondary, parasitic, dependency: the function introduced by the primary dependent must originate low enough in the structure to enable the function originating in the parasitic dependent to collapse with it. This implies that the primary dependent may not c-command the parasitic dependent. The anti-c-command requirement for parasitic gaps and secondary predication is illustrated by the data in (26).

In (26a), the primary gap c-commands the parasitic gap and therefore, when the adjunct is merged with β, the function introduced by the parasitic gap fails to be identified with the function originating in the trace. This is caused by the fact that the primary gap is not contained in β. Hence, β does not contain the function f_{move} that is copied into α from the adjunct. This gives rise to a violation of Categorial Licensing. The resulting structure is identical to (16) in the relevant respects.

Similarly, in (26b), attachment of the secondary predicate *dronken* 'drunk' leads to a violation of Categorial Licensing, because β is thematically complete and does not carry the θ-function that is copied into α from the secondary predicate. Hence, the resulting structure is identical to (23b) in the relevant respects.

---

[2] Of course, copying of a θ-function from a complement clause should be ruled out in the presence of an external θ-function in the verb. Neeleman and Van de Koot (2002) attribute the ungrammaticality of such structures to their version of the θ-criterion (Exclusivity).

(26) a. *A man who [*t* [$_{\alpha}$ [$_{\beta}$ looks old] [whenever I meet *t*]]]
b. *Jan$_{i}$ arriveerde nadat [$_{\alpha}$ dronken$_{i}$ [$_{\beta}$ hij gefeest had]]
John arrived after drunk he partied had

Categorial Licensing has several further effects, but it would take us too far afield to discuss these here (see Neeleman and Van de Koot 2002 for further discussion).

In summary, Categorial Licensing provides a unified account of (i) strong island phenomena, (ii) the existence of structures containing parasitic gaps and secondary predicates, and (iii) the anti-c-command requirement on such structures.

3.2 Categorial licensing and the subject island condition

Within the framework we have adopted, there is a natural way in which the account of adjunct islands can be extended to subjects. In this section we briefly present a proposal to that effect.

If, as is generally assumed, subject DPs are always θ-marked, then one would expect them to behave like object DPs in the relevant respect. It is true that subjects are sometimes transparent for extraction. Consider (27a), where the embedded subject is exceptionally case-marked by the matrix verb and extraction is possible. Such examples become completely ungrammatical if the embedded clause is tensed, as in (27b). (See Chomsky 1986, pp. 22-23, for related discussion.)

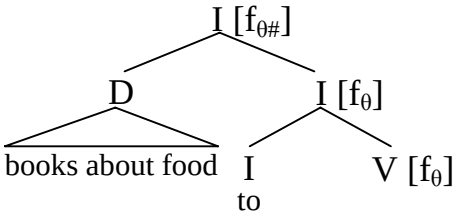
(27) a. ?Which topics do you expect [$_{IP}$ [$_{DP}$ books about *t*] to sell well]
b. *Which topics do you expect that [$_{IP}$ [$_{DP}$ books about *t*] will sell well]

We can account for this contrast if the embedded subject satisfies a θ-function in (27a) but not in (27b). This would turn the subject DP in (27b) into an adjunct and hence Categorial Licensing would prevent the copying of any licensing function from this DP (unless function identification can take place). (See Ackema 2002 for related ideas.)

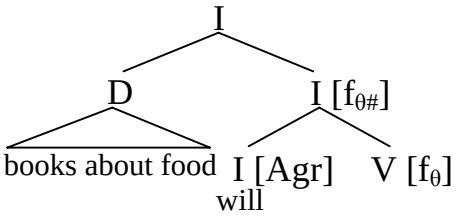
Suppose we assume that the verb's external θ-function is satisfied by agreement – if present – and by a DP otherwise. Then we could tie the distinction between the examples above to the observation that English tensed clauses exhibit agreement (so that the specifier of IP is an adjunct) but infinitival clauses do not (so that the same specifier is an argument). Thus, the IPs in (28a) and (28b) have the partial structures in (29a) and (29b).

(28) a. I expect [$_{IP}$ [$_{DP}$ books about food] to sell well]
b. I expect that [$_{IP}$ [$_{DP}$ books about food] will sell well]

(29) a.



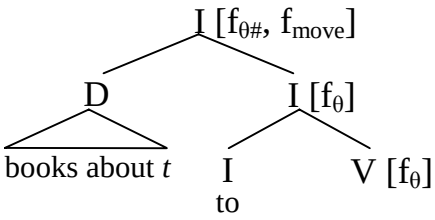
b.



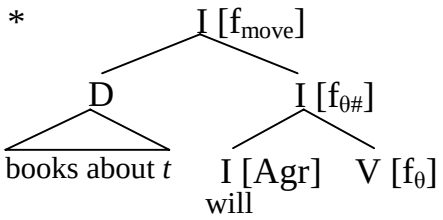
In (29a), the I-node fails to provide an argument for the θ-function that directly dominates it and this function is therefore copied to the higher I-node, where it is satisfied by the DP. This turns the higher I-node into a category. In (29b), by contrast, the head I contains agreement and therefore satisfies the external θ-function. The intermediate I-node and I's maximal projection contain the same number of unsatisfied licensing functions and therefore form segments of a single category.

The structure in (29a) is identical in all relevant respects to that in (15b), which depicts a θ-marked complement, whereas the structure in (29b) mirrors that of adjunction (compare (15a)). Hence, Categorial Licensing allows extraction from the argument DP in (29a) but not from the adjunct DP in (29b). The corresponding structures appear below:

(30) a.



b. *



Because the maximal projection of I in (30a) is a category, this node may receive the function f_{move} without violating Categorial Licensing. The subject DP is therefore transparent for extraction. By contrast, in (30b), copying of f_{move} to the maximal projection of I creates a category. Categorial Licensing requires that this be accompanied by the satisfaction of a licensing function, but this does not happen. Hence, the DP in this structure is an island for extraction.

One question raised by this proposal is why English tensed clauses should have a DP subject at all. After all, if the agreement features in I satisfy the external θ-function, then the presence of the DP is not required by θ-theory. Following a suggestion by Ackema (2002), we may assume that in English the DP is a true specifier and that it must be present to support the weak agreement specification of I. In languages with rich agreement, such as Greek and Italian, the DP is optional and, when present, does not fulfil the role of a specifier but behaves like a Clitic Left Dislocated DP. Indeed, Greek and Italian subject DPs differ from their English

counterparts in having A'-properties (see Speas 1995, Alexiadou and Anagnostopoulou 1998 and Ackema 2002 for related ideas).

3.3 Strong islands are transparent for non-licensing functions

Because the notion of category is defined in terms of licensing functions, the copying of a non-licensing function never gives rise to the creation of a new category. The implication is that such copying never violates Categorial Licensing. This accounts for the grammaticality of (31a) and (31b).

(31) a. The professors fell asleep [$_{PP}$ during each other's syntax lecture]
b. No one gave Mary a drink [$_{PP}$ during any intermission]

Recall that the function mediating reflexive binding is not licensing (see the discussion surrounding (12)). It is therefore predicted that strong islands should be transparent for binding. This is correct, as demonstrated by (31a), where binding penetrates an adjunct PP. The function introduced by an NPI is non-licensing as well: a sentence can contain an NPI-licensing operator even when no NPI is present. It is therefore to be expected that negation can license an NPI across an adjunct boundary. This prediction is borne out by the grammaticality of (31b). The ungrammaticality of (32a,b) confirms that the PPs in question are indeed adjuncts.

(32) a. *What did the professors fall asleep [$_{PP}$ during *t*]
b. *What did no one give Mary a drink [$_{PP}$ during *t*]

We see then that adjuncts are opaque for dependencies involving licensing functions but transparent for those mediated by non-licensing functions. This captures the pattern of strong island sensitivity in table (5).

The unified characterization of subjects and adjuncts presented here has the advantage that it correctly predicts that subjects, too, are islands for relations mediated by licensing functions only. Consider the examples in (33) involving an NPI and a reciprocal, respectively. In each case, the antecedent for the dependent element is external to the subject. In (33a), the NPI *any child* is licensed by negation (or, alternatively, by the question operator; see Laka 1990 and Giannakidou 1997[0] for discussion of similar data), while in (33b) the reciprocal *each other* is bound by the matrix subject *the men*.

(33) a. Hasn't the mother of any child come yet?
b. The men thought that each other's pictures were on sale.

The grammaticality of these examples is in stark contrast with the severely degraded status of extraction from a subject, illustrated earlier in (27b).

Perhaps somewhat surprisingly, subjects and adjuncts are also transparent for partial WH questions. In such questions the WH scope marker may be outside a strong island, while the contentive WH phrase is contained in it. (34a) is a Hungarian partial question in which the scope marker in the root clause is related to a contentive WH phrase across an adjunct boundary (from Horvath 1997). (34b) is a German partial question in which a scope marker in the root clause is related to a contentive WH phrase in the subject (from Haider 1993).[3]

(34) a. **Miért** voltál szomorú mert **hogy** viszonyultak hozzád *t*
why were-2SG sad because how related-3PL to-you
Literally: 'Why were you sad because how they had related to you?'

b. **Was** sagt sie [**welches Klavier** *t* zu spielen] ihm
what says she which piano to play him
Freude bereiten wuerdes
pleasure give would
Literally: 'What would playing which piano please him?'

While binding, NPI licensing and partial WH questions may hold across a strong island boundary, there is every reason to believe that these relations are established in the syntax. This has long been an uncontroversial assumption for binding (Chomsky 1981a,b and much subsequent work) and, although to a lesser extent, for the licensing of NPIs (Progovaç 1994). While the idea that the relation between scope marker and contentive in partial WH movement is syntactic is less widely accepted, it can easily be shown that partial WH movement patterns with binding and NPI licensing in many respects.

For instance, binding and NPI licensing are obligatory: if the dependent is present, then the antecedent must be present as well. This is shown for binding in (33a) and for NPI licensing in (33b).

(35) a. *John liked herself.
b. *John bought any books.

[3] For reasons that are poorly understood, German disallows partial questions of the type in (34a) (across an adjunct boundary), whereas Hungarian partial questions show no island effects with either subjects or adjuncts (see Horvath 1997). We will assume that the grammaticality of (34b) indicates the absence of strong island effects in German partial questions and that the ungrammaticality of such questions across an adjunct boundary in that language is due to a language-particular factor (for instance, the unavailability of a landing site for WH movement just to the right of a complementizer).

The relation between the contentive WH phrase and the scope marker in partial WH movement is obligatory as well:

(36) a. *Szomorú voltál mert **hogy** viszonyultak hozzád *t*
sad were-2SG because how related-3PL to-you
b. *Sie sagt [**welches Klavier** *t* zu spielen] ihm Freude bereiten
she says which piano to play him pleasure give
wuerdes
would

The antecedent of an anaphor (37a) and the truth-conditional operator licensing an NPI (37b) must c-command the element that depends on it.[4]

(37) a. *John's mother loves himself
b. *The professor [who didn't like students] interviewed any applicants

The following examples establish that the c-command requirement also applies in partial WH movement. In (38b), a partial question based on (38a), the scope marker *was* does not c-command the contentive WH phrase *wen*. The ungrammaticality that results from this is comparable to that found with sideward full WH movement, as in (38c).

(38) a. [Die Entsliessung [um diesen Film zu sehen]] nötigte den Hans um
the decision for this film to see forced the Hans for
sie mit zu fragen
her with to ask
'The decision to see this movie forced Hans to ask her along.'
b. *[Die Entsliessung [was diesen Film zu sehen]] nötigte den Hans
the decision what this film to see forced the Hans
wen *t* mit zu fragen
who with to ask
'Who did the decision to see this movie force Hans to ask along.'
c. *[Deciding [who to see that new movie next]] makes t_{WH} very happy

[4] Apparent violations of c-command with NPI licensing occur in examples like (i). These fall into line with the general picture outlined in the text if the negative feature of *no professor* percolates to the DP *no professor's mother*, so that c-command in fact holds.

(i) No professor's mother likes any students.

Of course, this raises the question why the negative feature of *didn't* in example (37b) cannot percolate high enough to license the NPI in that example. We will not attempt to answer this question here.

Finally, all three grammatical dependencies under consideration exhibit a locality effect that is typical of relations mediated by functions. This effect is due to the fact that Economy considerations cause a function to be satisfied at the earliest opportunity. For example, in (39) the function f_{self} introduced by *himself* finds a suitable antecedent in the embedded clause and therefore *himself* cannot be related to *John* across *Mary*.

(39) *John said that Mary likes himself

For the same reasons, full WH movement cannot cross a WH complementizer, such as *if* or *whether*. The movement function introduced by the trace in (40) is satisfied at the earliest opportunity, namely by the WH complementizer, and the WH phrase *what* therefore fails to be licensed.

(40) ??What does John wonder whether Mary likes *t*

The degraded status of (40) does not appear to be a weak island effect. WH movement may cross negation and other scopal elements without giving rise to an island violation, at least under a wide scope interpretation of the WH phrase (see section 4.2 for further discussion). But an example like (41) remains degraded on the reading in which *what* has scope over *everybody*, indicating that some other factor is at play.

(41) ??What does everybody wonder whether Mary likes *t*

The ungrammaticality of these examples is readily accounted for in terms of the Economy constraint sketched above: the function introduced by the trace is satisfied when it reaches the WH complementizer and therefore is unable to license the presence of *what*.

As one would expect, the same effect can be observed in German:

(42) ??Welches Buch fragte Karl ob Petra gesagt hätte dass Marie kaufen wird
which book asked Karl if Petra said had that Mary buy will

Crucially, an example with partial WH movement, where the WH complementizer intervenes between the contentive WH phrase and the WH scope marker is equally degraded:

(43) ??Was fragte Karl ob Petra gesagt hätte welches Buch Marie kaufen wird
what asked Karl if Petra said had which book Mary buy will

A similar locality effect involving NPI licensing is provided by examples like (44), where the relation between an NPI and its licenser is interrupted by another truth-conditional operator, namely *and*.

(44) *No professor likes (any) books and any records.

This example can be accounted for on the assumption that the truth-conditional operator *and* prevents the functions introduced by the NPIs from reaching the negative subject. In other words, *and* has the syntactic properties that satisfy this type of function, but since it lacks negation, semantic licensing of the NPIs fails.[5]

Let us take stock. The preceding discussion has established the following. Partial WH movement exhibits the diagnostic properties of syntactic dependencies that are also found with anaphoric binding and the licensing of NPIs. Just like these dependencies, partial WH movement does not exhibit strong island effects. We may therefore conclude that the contentive WH phrase is related to its scope marker by a function and furthermore that this function is non-licensing.

Thus, in an example like (34b), the function f_{move} introduced at the root of the WH chain travels up the tree until it directly dominates the contentive WH phrase (see the tree diagram in (45)). This phrase lacks a suitable WH operator but may still satisfy f_{move}, provided it introduces a function that will find the missing operator. Let us call this function f_{bare_op}. In other words, the intermediate WH phrase may satisfy f_{move} because it 'promises' to find the bare operator that it fails to supply itself. The non-licensing function f_{bare_op} that expresses this promise can be copied from the embedded subject without this resulting in a violation of Categorial Licensing (compare the structure in (30b)).[6]

---

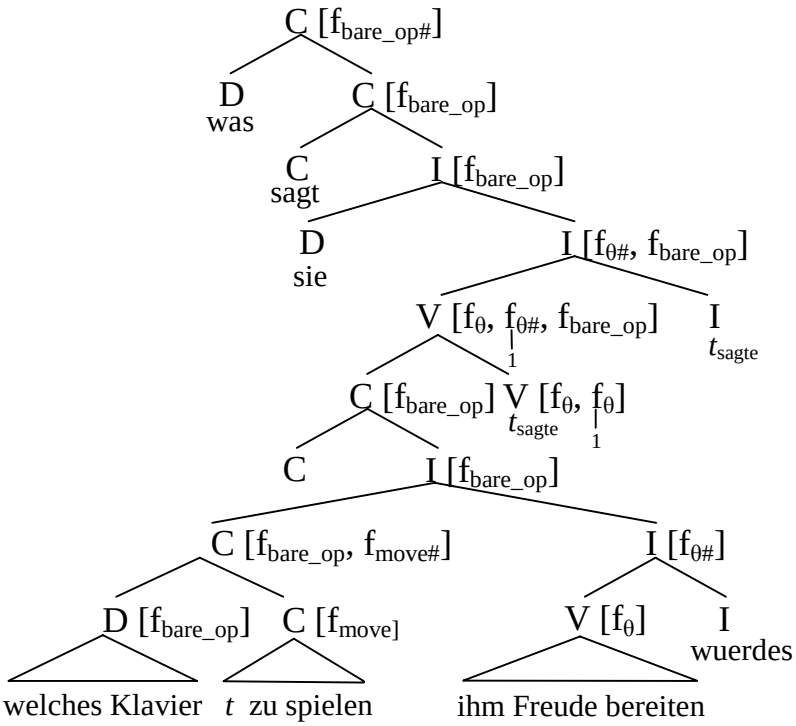
[5] On this view, the grammaticality of (i) is due to the fact that *or* can mean *nor*, as shown by (ii):

(i) No professor likes (any) books or any records
(ii) I like neither books (n)or records.

See Hoeksema 2000 and Chierchia *to appear* for further discussion of these phenomena.

[6] It is irrelevant for the present discussion whether the subject clause has undergone vacuous movement to the specifier of the complement CP.

(45)



3.4 Summary

We have shown that the strong island sensitivity of the dependencies in table (5) is accounted for if we assume that the creation of new categories requires the satisfaction of a licensing function. This condition, Categorial Licensing, also explains the existence of parasitic gap phenomena and secondary predication, as well as the anti-c-command requirement that they must obey.

The contrast between licensing and non-licensing functions seems to us to be clearly syntactic: a licensing function allows the creation of a bit of syntactic structure that a non-licensing does not allow. Therefore, to the extent that the proposed theory of strong islands is empirically supported, it also supports a syntactic characterization of strong islands.

4 Characterizing weak islands

4.1 The scope of indefinites

While there is strong evidence for a purely syntactic theory of strong islands, we argue in this section that an account of weak islands must make reference to the theory of scope. Although this may not be a particularly controversial position to adopt, it will be useful to review some of the key evidence for it.

Let us begin by considering (46a). This sentence is ambiguous: the indefinite may be interpreted as taking scope over the universal or as dependent on it. This

ambiguity is traditionally expressed in terms of QR, in which case the LFs corresponding to the wide and narrow scope of the existential may be represented as in (46b') and (46c'), respectively.

(46) a. Every student read a book
b. $\exists f$ [CH($f_{\text{books} \rightarrow \text{1\_book}}$) & $\forall x$ [student(x) → x read f(books)]] $\exists > \forall$
b'. $\exists y$ [book(y) & $\forall x$ [student(x) → x read y]
c. $\exists f$ [SK($f_{\text{students} \rightarrow \text{books}}$) & $\forall x$ [student(x) → x read f(x)]] $\forall > \exists$
c' $\forall x$ [student(x) → $\exists y$ [book(y) & x read y]]

An alternative view of the relevant ambiguity is that the indefinite may be interpreted as introducing either a choice function, as in (46b), or a skolem function, as in (46c) (on choice functions, see Reinhart 1997, 1998; Winter 1997, Kratzer 1998 and Matthewson 1999; on Skolem functions, see von Heusinger 2000 and Winter 2002). The choice function selects a single book from the set of books, so that (46b) corresponds to a wide scope reading: the choice of book does not depend on the choice of student. By contrast, the skolem function is a function from students to books, so that (46c) expresses a narrow scope interpretation for the indefinite, according to which the choice of book depends on the choice of student. We assume, for now without motivation, that choice functions and skolem functions are very different from the functions mediating dependencies discussed in section 2. While the latter are syntactic objects, we take the former to be semantic in nature. In section 5 we explore the consequences of using syntactic functions to encode scope.

Although the representation in (46b) is derived without movement of the indefinite, it is interpreted as if the restrictive part of the indefinite was adjacent to the existential quantifier that binds the choice function. It is this property of choice functions that underlies the ability of an indefinite to take arbitrarily wide scope. This behaviour of indefinites is a problem for the QR analysis, because it cannot explain the differential scope properties of universals and existentials. For example, a universal contained in a tensed clause cannot take scope outside that clause, but an indefinite in the same position can:

(47) a. Someone said that every student had left. $\exists > \forall$; $*\forall > \exists$
b. Every student said that someone had left. $\exists > \forall$; $\forall > \exists$

This contrast between universals and indefinites can even be observed when these expressions are contained in strong islands:

(48) a. Some student left after meeting every professor. $\exists > \forall$; $*\forall > \exists$
b. Every student left after meeting some professor. $\exists > \forall$; $\forall > \exists$

As has been pointed out by Reinhart (1997, 1998), it would be a mistake to devise a theory of scope that would allow the indefinite in (48b) to take apparent wide scope over the universal while leaving it in the adjunct (for instance, by interpreting it as specific and hence scopeless). Such a theory faces the problem that in some cases truth conditions require an existentia to be interpreted outside the strong island in which it is contained. For instance, in (49), *which philosopher* cannot remain in the scope of the conditional. If it did, this question would admit as an adequate answer that Mary will be unhappy if we invite Donald Duck (Donald Duck is not a philosopher).

(49) Who will be unhappy if we invite which philosopher.

It is precisely to avoid this problem that Reinhart introduces existential closure over choice functions, because that analysis achieves the wide scope of existentials without QR.

4.2 The nested scope constraint

With this background in place, let us turn to (50), in which a negative operator intervenes between the universal and the indefinite. As before, the wide scope reading of the indefinite is encoded using a choice function (cf. (50b)). The narrow scope reading, which requires the more complex skolem function, turns out to be unavailable in this case.[7] We return shortly to the question of what blocks this interpretation.

(50) a. Every student didn't read a book
b. $\exists f\,[\mathrm{CH}(f_{\text{books} \rightarrow \text{1\_book}})\ \&\ \forall x\,[\text{student}(x) \rightarrow \neg\,[x \text{ read } f(\text{books})]]]$ $\exists > \forall$
c. $^{*}\exists f\,[\mathrm{SK}(f_{\text{students} \rightarrow \text{books}})\ \&\ \forall x\,[\text{student}(x) \rightarrow \neg\,[x \text{ read } f(x)]]]$ $\forall > \exists$

The following two example sets demonstrate that the same intervention effect occurs when the indefinite is an interrogative expression. Thus, (51a) is ambiguous

[7] The intervention effect observed here is often referred to as a 'weak' island. Arguably, this is because it is lessened by the effects of D-linking. For instance, if we set up a context in which it had been agreed that student A would read book A, student B would read book B and student C would read book C, and it subsequently materializes that none of the students have read any books, then we could perhaps answer the question *Which book hasn't every student read?* with *Student A hasn't read book A, student B hasn't read book B and student C hasn't read book C.* Throughout this paper we abstract away from these effects and all the judgments are for the non-D-linked interpretations of the sentences in question.

in precisely the same way as (46a), while (52a) lacks the wide scope reading for the universal that (50a) also lacks.

(51) a. Which book did every student read *t*
b. WHf [CH($f_{\text{books} \rightarrow \text{1\_book}}$) & $\forall x$ [student(x) → x read f(books)]] $\exists > \forall$
c. WHf [SK($f_{\text{students} \rightarrow \text{books}}$) & $\forall x$ [student(x) → x read f(x)]] $\forall > \exists$

(52) a. Which book didn't every student read *t*
b. WHf [CH($f_{\text{books} \rightarrow \text{1\_book}}$) & $\forall x$ [student(x) → ¬ [x read f(books)]]] $\exists > \forall$
c. *WHf [SK($f_{\text{students} \rightarrow \text{books}}$) & $\forall x$ [student(x) → ¬ [x read f(x)]]] $\forall > \exists$

In (52c) negation could be taken to interrupt the relation between WHf and the skolem function variable it binds or the relation between the universal operator and the individual variable it binds. However, the contrast between (53) and (54) establishes that it is the binding of the skolem function variable that is adversely affected, because here negation does not intervene between the universal and the variable it binds.

(53) a. Which book did you say that every student read *t*
b. WHf [CH($f_{\text{books} \rightarrow \text{1\_book}}$) & you say that $\forall x$ [student(x) → x read f(books)]] $\exists > \forall$
c. WHf [SK($f_{\text{students} \rightarrow \text{books}}$) & you say that $\forall x$ [student(x) → x read f(x)]] $\forall > \exists$

(54) a. Which book didn't you say every student read *t*
b. WHf [CH($f_{\text{books} \rightarrow \text{1\_book}}$) & ¬ [you say that $\forall x$ [student(x) → x read f(books)]]] $\exists > \forall$
c. *WHf [SK($f_{\text{students} \rightarrow \text{books}}$) & ¬ [you say that $\forall x$ [student(x) → x read f(x)]]] $\forall > \exists$

As discussed earlier, a skolem function and a choice function are not of the same type. A choice function is function from a set to a member of that set, whereas a skolem function relates members of one set to members of another (the skolem function in the examples above relates members of the set of students to members of the set of books). It achieves this by introducing a variable that can be bound by a c-commanding scopal element (the universal quantifier in the examples considered so far). Apparently, only the latter type of functional variable is sensitive to intervention by certain types of scopal elements, such as negation. After all, we find a systematic contrast between the grammaticality of examples in which negation intervenes between an operator and a choice function variable – (50b), (52b) and (54b) – and the ungrammaticality of examples in which negation

intervenes between a skolem function variable and the operator that binds it – (50c), (52c) and (54c).

It seems reasonable to attribute this contrast to the effect that each type of function has on the interpretation of the restrictive part of the indefinite. As discussed earlier, an indefinite containing a choice function is interpreted as if it has undergone QR. Thus, its restrictive material is interpreted outside the scope of any scopal element that is c-commanded by the operator that binds the choice function variable. Very much the opposite is true of the restrictive part of an indefinite that introduces a skolem function variable. The effect of introducing such a variable is precisely to allow the restrictive part of an indefinite to depend on an operator in whose scope it occurs. Therefore, this restrictive part is necessarily interpreted in its base position.

Crucially, the ability of the restrictive part to be interpreted in its base position appears to depend on other scopal elements. In particular, the restrictive part may only enter into a relation with a c-commanding operator across another scopal element if it is able to enter into a skolem dependency with that element. Williams (1994) dubs this the Nested Scope Constraint:

(55) *Nested Scope Constraint (NSC)*
In $[XP_i [\ldots YP_k [_{\text{scope-k}} \ldots t_i [_{\text{scope-i}} \ldots]]]]$ t_i depends on YP_k.

On this view, the intervention effects observed above are due to the inability of the restrictive part of the indefinite to depend on negation.

That this characterization of choice functions and skolem functions is on the right track is corroborated by the behaviour of interrogative adjuncts. In (56a), the WH adjunct why has moved across a universal quantifier. The representation in (56b), in which the trace is interpreted as introducing a choice function, is incompatible with the scope requirements of the restrictive part of the WH expression: this component is a modifier that must take VP-scope. In other words, the trace of *why* cannot introduce a choice function, because the effect of this function is to assign its restriction wide scope. By contrast, the trace of the WH adjunct can introduce a skolem function variable, as such a function causes the modifier part of *why* to be interpreted in the root of the WH chain where it takes scope over the embedded VP. Hence, (56a) allows the pair-list reading expressed by (56c).

(56) a. Why did you say that every student $[_{VP} [_{VP}$ read Barriers] $t]$

b. *WHf [CH($f_{\text{reasons} \rightarrow \text{1\_reason}}$) & you say that $\forall x$ [student(x) → x read Barriers for f(reasons)]] $\exists > \forall$

c. WHf [SK($f_{\text{students} \rightarrow \text{reasons}}$) & you say that $\forall x$ [student(x) → x read Barriers for f(x)]] $\forall > \exists$

We now make the following prediction: given that a skolem function variable obeys the NSC, the addition to (56a) of an intervening negative operator should result in complete ungrammaticality. This prediction is borne out:

(57) a. *Why didn't you say that every student [$_{VP}$ [$_{VP}$ read Barriers] *t*]
b. *WH*f* [CH($f_{\text{books} \rightarrow \text{1\_book}}$) & ¬ [you say that ∀*x* [student(*x*) → *x* read Barriers for *f*(reasons)]]] ∃ > ∀
c. *WH*f* [SK($f_{\text{students} \rightarrow \text{books}}$) & ¬ [you say that ∀*x* [student(*x*) → *x* read Barriers for *f*(*x*)]]] ∀ > ∃

As before, (57b) is incompatible with the scope requirement of the VP-modifier part of *why*, while in (57c) a bare operator is related to the restrictive part of an indefinite across negation.

We could summarize our observations as follows. Whenever a bare operator is related to the restrictive part of an indefinite in such a way that that restrictive part is interpreted in situ, the relevant relation exhibits sensitivity to the intervention of other scopal elements, as expressed by the NSC. Because the restrictive part of a WH adjunct is obligatorily interpreted in the foot of the chain, adjunct movement always displays weak island sensitivity. WH arguments may escape the effect of the NSC, but only if their trace is interpreted as a choice function. As a result, movement of a WH argument across negation only allows a wide scope reading of the WH phrase. This outcome implies that we must refine the table in (5) to reflect the weak island sensitivity of WH arguments under their narrow scope reading:

(58)

	STRONG ISLAND SENSITIVE	WEAK ISLAND SENSITIVE
WH ARGUMENT (WIDE SCOPE)	+	–
WH ARGUMENT (NARROW SCOPE)	+	+
WH ADJUNCT	+	+
NPI	–	+
ANAPHORS	–	–

The NSC may also account for the intervention effects found with NPI licensing and for the absence of such effects with anaphors. An NPI introduces a variable that must be bound by negation or some other (bare) truth conditional operator. If the relevant variable cannot depend on *fewer than three beggars*, then the sharply degraded status of (2c), repeated here as (59) can be understood (see also Krifka 1995 and Chiercia *to appear*).

(59) *No one gave fewer than three beggars anything.

By contrast, anaphors are not licensed by operators and therefore their insensitivity to the NSC is expected.

Another instance of a relation between a bare operator and an indefinite expression that is sensitive to other scopal elements is presented by partial WH movement, as exemplified by the following German example.

(60) **Was** glauben alle Kinder **mit wem** Maria gesprochen hat
what believe all children with whom Maria spoken has
'With whom do all the children believe that Mary has spoken?'

The relation between the WH word that marks the scope of the question and the WH expression in the lower clause exhibits intervention effects (Beck 1995):

(61) ***Was** glauben alle Kinder nicht **mit wem** Maria gesprochen hat
what believe all children not with whom Maria spoken has
'With whom don't all the children believe that Mary has spoken?'

Similar data can be found in other languages with partial WH movement, such as Hungarian:[8]

(62) ***Mit** nem gondolsz hogy **kivel** beszélt
what-ACC not think-2SG-INDEF.DO that who-with spoke-3SG
Mari?
Mari-NOM
'With whom don't you think that Mary has spoken?'

It is widely assumed that the contentive part of the higher WH element in such sentences remains uninterpreted (see, for instance, Brody 1997, Horvath 1997). Instead, the WH phrase in the intermediate specifier position spells out the restrictive part of the WH expression. Therefore, at LF, we have a representation in which a bare operator is separated from its semantic restriction, just as in the cases considered earlier.

4.3 Summary

In this section we have seen that weak island effects occur when an indefinite that is interpreted in situ is bound by a bare operator across another scopal element on

[8] Horvath (1997) claims that Hungarian partial WH questions show weak island effect only with non-factive matrix verbs. It seems reasonable to assume that with factive verbs the effects of D-linking obviate any intervention effects. See note 7 for related discussion.

which it cannot depend, as expressed by the NSC. The NSC captures the weak island sensitivity of NPIs and WH adjunct, and correctly predicts that WH arguments exhibit weak island effects whenever their restriction must be interpreted in the chain root.

5 The copy theory, reconstruction and strong islands

5.1 Introduction

The condition responsible for weak island effects (the Nested Scope Constraint) seems to have nothing in common with Categorial Licensing, the constraint that captures strong islands. While this leaves open the possibility that both of these constraints apply at LF, we argue in this section that Categorial Licensing and the Nested Scope Constraint cannot hold at the same level of representation. The conclusion we arrive is at variance with much minimalist work, because it is incompatible with the view that scope is represented at LF (see especially Fox 1999).

5.2 Strong islands and reconstruction

As was demonstrated earlier, a moved WH phrase may require reconstruction for the purposes of scope. For instance, (51a), repeated here as (63a), allows the narrow scope reading (63c) for the restrictive part of *which book*.

(63) a. Which book did every student read *t*
b. WHf [CH($f_{\text{books} \rightarrow \text{1\_book}}$) & $\forall x$ [student(x) → x read f(books)]] $\exists > \forall$
c. WHf [SK($f_{\text{students} \rightarrow \text{books}}$) & $\forall x$ [student(x) → x read f(x)]] $\forall > \exists$

In the copy theory of movement (Chomsky 1993, Hornstein 1995) this scope ambiguity is represented in terms of whether the restrictive part of the indefinite is deleted in the higher or the lower copy:

(64) a. Which book did every student read *t*
b. Which book did every student read ~~which book~~ $\exists > \forall$
c. Which ~~book~~ did every student read ~~which~~ book $\forall > \exists$

Thus, these LF representations attempt to express simultaneously the representation associated with movement and its interpretation in terms of scope. In terms of the theory of grammatical dependencies developed in section 2, this amounts to an attempt to integrate the representation of scope into the representation of syntactic dependencies. Our aim in this section is to show that such an integrated theory

cannot adequately capture strong island phenomena. In order to develop this argument, we modify the theory of grammatical dependencies in section 2 so as to make it capable of capturing the effects of reconstruction.

Recall that we accounted for the ability of partial WH movement to penetrate strong islands by assuming that a dependent element that only looks for a bare operator introduces the non-licensing function f_{bare_op}. This suggests a variant of the theory of dependencies that uses the opposition between the licensing function f_{move} and the non-licensing function f_{bare_op} to encode where the restrictive part of a moved phrase is to be interpreted. On this view, the function relating the lower copy to the higher copy is mediated by the licensing function f_{move} in (64b) and by the non-licensing function f_{bare_op} in (64c), as indicated in (65a) and (65b), respectively. After all, the relation in (65b) is identical in the relevant respects to that between the scope marker and the contentive WH phrase in a WH partial question. Therefore, if the restrictive part of the indefinite in (64c) is truly interpreted in the foot of the chain, the lower copy should introduce a function that is looking for a bare operator only.

(65) a. f_{move}

Which book did every student read ~~which book~~

b. f_{bare_op}

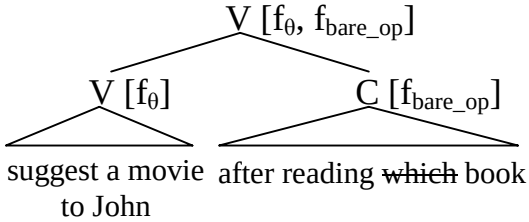
Which ~~book~~ did every student read ~~which~~ book

Now consider once more a standard adjunct island violation, as in (1a), but interpreted with scope reconstruction of book, as in (66).

(66) *Which ~~book~~ did you suggest a movie to John [after reading ~~which~~ book]

Since the restriction is interpreted in the foot of the chain, the lower copy introduces a function looking for bare operator only (f_{bare_op}), as shown in the partial structure below:

(67)

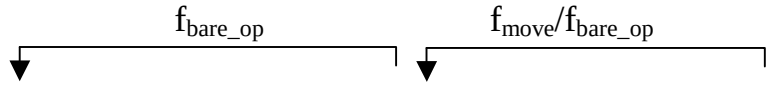


Because f_{bare_op}, the function copied from the adjunct, is non-licensing, no violation of Categorial Licensing is incurred and the structure is predicted to be grammatical on the scope-reconstructed reading, contrary to fact. In other words, this variant of the theory of grammatical dependencies predicts, incorrectly, that a strong island violation can be avoided through scope reconstruction.

The key shortcoming of the approach just sketched is that it cannot distinguish full movement with scope reconstruction from partial movement. Recall that Hungarian allows partial WH movement from an adjunct:

(68) **Miért** voltál szomorú mert **hogy** viszonyultak hozzád?
why *were-2SG sad* *because* *how* *related-3PL* *to-you*
Lit. 'Why were you sad because how they had related to you?'

According to the variant of the theory of grammatical dependencies we are considering here, the functions involved in deriving this example are as follows:

(69) 

Just as in (67), since the function that is copied out of the adjunct is non-licensing, no violation of Categorial Licensing occurs.

The problem is straightforward: in (69) the restriction of the contentive WH phrase is never moved out of the adjunct. Thus the positioning of the restrictive material of the WH phrase as determined by the overt syntax happens to coincide with that implied by the representation of scope. In (66), by contrast, the WH phrase has undergone movement across the adjunct boundary. Hence in this case the positioning of the restrictive material of the WH phrase as determined by the overt syntactic derivation does not coincide with that implied by the representation of scope. In short, this translation of the copy theory cannot simultaneously represent that restrictive material has moved out of the strong island and is interpreted inside it. This creates a problem for the determination of CED violations, for which only the position of restrictive material as determined by the overt syntax is relevant.

If we combine the copy theory of movement, or rather its translation to the theory of relations mediated by functions, with a QR analysis of the scope properties of indefinite expressions, then we get the mirror image of the problem just discussed. Consider again (48b), repeated here as (70).

(70) Every student left after meeting some professor. $\exists > \forall$; $\forall > \exists$

The indefinite expression *some professor* can take matrix scope. If this scope is achieved through QR, a copy of the indefinite is created in the matrix clause. Since the restriction in this higher copy is interpreted, the restriction of the copy in the adjunct is not. If this must be expressed through the type of function involved, then the lower copy of the indefinite introduces a function looking for a restriction. Since, by hypothesis, this is a *licensing* function and since this function is copied out of the adjunct, a violation of Categorial Licensing should result. Hence, the sentence should be ungrammatical on the intended reading, contrary to fact.

The case at hand is another one in which the positioning of the restrictive material as determined by the syntactic derivation does not coincide with that implied by the representation of scope. And again we must draw the conclusion that, as far as CED phenomena are concerned, what matters is only the position of restrictive material as determined by the overt syntax.

There would seem to be a potential way around the dilemma presented by the unacceptable example (66) and its well-formed representation (67). If scope reconstruction, although generally available, could be prevented from reconstructing material into a strong island, then the problematic representation in (67) could never be derived. But the following example shows that this route cannot be taken:

(71) a. How many gadgets will each genius invent *t* [without PRO ever being likely to patent *t*]

b. WHf [SK($f_{\text{geniuses} \rightarrow \text{number of gadgets}}$) & $\forall x$ [genius(x) → x will invent $f(x)$ & ¬[x is likely to patent $f(x)$]]

(71a) admits (indeed requires) the reading in (71b) where *many gadgets* is in the scope of *likely* (which is itself contained in an adjunct). An interpretation of (71a) without scope reconstruction presupposes that there exist as yet uninvented gadgets about which we can entertain a likelihood of them being patented. But such a presupposition gives rise to a virtual contradiction. As pointed out by Fox (1999), "[W]e think about the objects of invention as being created at the time of invention, and we therefore can't talk about these objects at earlier moments – hence the oddness of *#John will invent this story*, *#Which of these stories is John likely to invent?*". By contrast, we *can* ask, for each genius, what is the number such that he is unlikely to patent that number of as yet uninvented gadgets. (See Longobardi 1987, Heycock 1995 and Cresti 1995 for additional discussion of *how many* questions).

Even if a solution of the type suggested above could be made to work, it would not extend to the mirror image problem illustrated by (70), since a stipulation to the effect that restrictive material contained in an adjunct cannot be interpreted outside that adjunct is untenable on empirical grounds (see also section 4.1).

Since we have established that scope reconstruction can penetrate strong islands, the conclusion arrived at earlier seems unavoidable: strong and weak islands cannot both be represented at LF.

5.3 Strong islands first – reconstruction later

The data discussed in the previous section fall into place if the constraint responsible for strong island effects applies to a representation that encodes grammatical dependencies but not scope reconstruction. So let us then revert to our original theory of grammatical dependencies, in which functions contained in LF-representation do not encode the effects of reconstruction; we assume that scope is regulated at some later level. How does such a theory account for the data in the previous section?

Consider first cases of scope reconstruction that do not involve strong islands. The LF representation of such examples is schematized in (72a). This LF is mapped to a representation of scope. At that level the trace may be interpreted as introducing a skolem function. This allows the restrictive material to be interpreted in the foot of the chain, as in (72b).

(72) a. ***LF representation:***

f_{move}

WH-restriction . . . *t*]

b. ***Scope representation (trace interpreted as a skolem function):***
WH-~~restriction~~ . . . ~~WH~~-restriction]

This type of derivation is blocked if the trace is contained in a strong island. The problem is not that scope reconstruction is impossible: if the post-LF scope rules interpret the trace as a skolem function, then the restrictive material will be interpreted in the foot of the chain (see (73b)). But irrespective of what happens at the level that deals with scope, such examples are associated with an LF representation as in (73a), which violates Categorial Licensing.

(73) a. ****LF representation:***

f_{move}

WH-restriction . . . [$_{\text{strong island}}$ . . . *t*]

b. ***Scope representation (trace interpreted as a skolem function):***
WH-~~restriction~~ . . . [$_{\text{strong island}}$ ~~WH~~-restriction]

We also considered cases in which the scope rules interpret restrictive material higher than would be warranted on the basis of the structure that surfaces. In

particular, non-WH indefinites can take very wide scope, although they do not have to. We repeat example (48b), which demonstrates this, as (74a). Since no movement takes place, the LF representation satisfies Categorial Licensing. At the level of scope, the indefinite may be assigned wide or narrow scope depending on whether it is interpreted as introducing a choice function or a skolem function:

(74) a. Every student left after meeting some professor.
b. ***LF representation:***
Every student left [$_{\text{strong island}}$ after meeting some professor]
c. ***Scope representation 1 (trace interpreted as a choice function):***
$\exists f$ [CH($f_{\text{professors} \rightarrow \text{1\_professor}}$) & $\forall x$ [student(x) → x left after x meeting f(professors)]]
c'. ***Scope representation 2 (trace interpreted as a skolem function):***
$\exists f$ [SK($f_{\text{students} \rightarrow \text{professors}}$) & $\forall x$ [student(x) → x left after x meeting f(x)]]

Let us finally turn to partial WH movement across a strong island boundary. In this case a violation of Categorial Licensing is avoided because the contentive WH phrase introduces the non-licensing function $f_{\text{bare\_op}}$ (see (75a)). The scope rules interpret the trace in the adjunct as a skolem function, yielding the scope reading indicated in (75b).

(75) a. ***LF representation:***
$f_{\text{bare\_op}}$ f_{move}
WH . . . [$_{\text{strong island}}$ WH-restriction . . . *t*]
b. ***Scope representation (trace interpreted as a skolem function):***
WH . . . [$_{\text{strong island}}$ ~~WH-restriction~~ . . . ~~WH~~-restriction]

As was demonstrated by example (61) in section 4.2, the relation between the scope marker and the contentive WH in partial WH questions shows intervention effects. This indicates that the restrictive part of the contentive WH can be interpreted in the specifier of the strong island but not higher. Hence, the LF-representation in (75a) cannot be mapped to the scope representation in (75b).

(76) a. ***LF representation:***
$f_{\text{bare\_op}}$ f_{move}
WH . . . [$_{\text{strong island}}$ WH-restriction . . . *t*]
b. ****Scope representation (trace interpreted as a choice function):***
WH-restriction . . . [$_{\text{strong island}}$ ~~WH-restriction~~ . . . ~~WH-restriction~~]

Why should this be so? The situation is reminiscent of what we found earlier with VP-adjuncts (see the discussion surrounding (56)): while an element in an A-position can be assigned arbitrarily wide scope, the scope of an element in an adjoined position is fixed. If the same is true of an element in an A'-position, then the structure in (75b) cannot be generated. Instead, interpreting the trace in (75a) as a choice function yields the uninterpretable representation in (77), where the same restriction must be interpreted in two different positions.

(77) ****Scope representation (trace interpreted as a choice function):***
WH-restriction . . . [$_{\text{strong island}}$ ~~WH~~-restriction . . . ~~WH-restriction~~]

6 Against a derivational theory of strong islands

We have argued that strong islands and scope reconstruction cannot be represented at the same level. In theory, this leaves us with two possibilities. It could be that strong islands are an LF phenomenon but that scope is not represented there (it is regulated at a later level). Alternatively, scope *is* represented at LF but strong island phenomena are not. On this view, these islands are not due to properties of LF but have their origin in the order of syntactic operations in a derivational syntax. We conclude the paper by showing that the theory of strong islands of section 3 is superior in a number of respects to such a derivational theory of strong islands.

Several proposals have recently been put forward that seek to provide a derivational account of strong islands (Uriagereka 1999, Stepanov 2001 and Johnson 2002). Uriagereka suggests that when two phrases are in a relation of sisterhood, the first one is spelled out as a head, making it entirely opaque to the derivation.[9] Stepanov proposes that adjuncts are adjoined to their hosts after the derivation is otherwise complete. This makes it impossible for movement operations to target any constituent contained in the adjunct.[10] According to Johnson, merger is asymmetric: it selects a host from the numeration and then proceeds by selecting further items from the numeration for merger to this host. Sometimes the derivation can only proceed if an intermediate structure formed by

[9] As Johnson (2002) points out, this proposal faces serious empirical difficulties because it would seem to make a VP with a right-adjoined adjunct an island for extraction. If the relevant order of VP and adjunct is derived by leftward movement of the VP instead, then again we would expect that VP to be an island in virtue of being in a derived position.

[10] There are a number of problems with this idea, but the most obvious one in the present context is that it fails to account for the (obligatory) VP-scope of the restriction part of *why* in examples like (56a), repeated here as (i).

(i) Why did you say that every student [VP [VP read Barriers] t]

merger is put back in the numeration, so that a new host can be selected. It is the phrases that undergo this process of 'renumeration' that are strong islands. The islandhood of renumerated phrases is due to the fact that upon re-entering the numeration they undergo spell-out. This in turn forces all of the terms within renumerated phrases to have their linear position fixed.

The proposals just discussed differ in a number of respects from the theory of strong islands presented in section 3.

First, the theories put forward in Uriagereka 1999 and Stepanov 2001 are unable to explain why movement and predication should be sensitive to strong islands (see examples (17) and (22b), repeated here as (78)), while other grammatical dependencies, such as binding and the licensing of NPIs, are not (see examples (31a,b), repeated here as (79)).

(78) a. *Which book did you suggest a movie to John [after reading *t*]
b. *John seemed [that it rained] angry.

(79) a. The professors fell asleep [$_{PP}$ during each other's syntax lecture]
b. No one gave Mary a drink [$_{PP}$ during any intermission]

This is an important shortcoming, because all these dependencies share a cluster of properties (Koster 1987), which strongly suggests that they form a natural class. As mentioned earlier, Neeleman and Van de Koot (2002) argue that this cluster has its basis in the workings of syntax: it results from restrictions imposed by Inclusiveness. Johnson's (2002) proposal would seem to fare better than Uriagereka's and Stepanov's in this respect, because it could in principle distinguish movement and predication from other dependencies by treating the latter as involving movement as well (for instance, by assuming that predication always involves movement to a case-checking position).

Second, the constraints derived by derivational theories of strong islands are too strong: it is not the case that subjects and adjuncts are always islands for movement and predication. In sentences with parasitic gaps a relation resembling movement is established between a null element inside a strong island and a WH phrase outside it. Similarly, while adjuncts are islands for predication (see (78b)), they apparently cease to be so if the predicative relation is parasitic on primary predication as in (22a), repeated here as (80).

(80) John left the room angry.

Derivational theories seem ill-equipped to explain these contrasts. The only option they have is to deny that the relation between the dependent element in the island

and its antecedent outside it is formed in the syntax. Instead, this relation must be established after the derivation is complete.

Thus, parasitic gaps could be accommodated in terms of null operator movement internally to the adjunct followed by the application of chain composition in the mapping from LF to syntax-external systems (Chomsky 1986). (Although it remains unclear – in a minimalist theory – what would trigger null operator movement in the absence of a +WH C).

Similarly, the analysis of secondary predication could rely on the presence of a PRO subject in the secondary predicate, followed by the application of an interpretive rule that relates PRO to its controller; there seems to be no other way to establish a relation between the primary and the secondary predicate.

But both the in the case of parasitic gaps and in the case of secondary predication this move comes at a cost: the relevant interpretive rules (chain composition and control) exhibit all the hallmarks of grammatical dependencies: the relation is obligatory, it requires c-command (between the two operators in the case of parasitic gaps; between the overt subject and PRO in the case of secondary predication), the antecedent is unique (but the dependent need not be) and the relation must be local.

Finally, it is remarkable that both parasitic gap structures and those involving secondary predication exhibit an anti-c-command requirement. This was illustrated in (26), repeated here as (81).[11]

(81) a. *A man who [*t* [$_{\alpha}$ [$_{\beta}$ looks old] [whenever I meet *t*]]]

b. *Jan$_i$ arriveerde nadat [$_{\alpha}$ dronken$_i$ [$_{\beta}$ hij gefeest had]]
John arrived after drunk he partied had

Approaches to these parasitic structures that rely on interpretive rules, such as those just discussed, could invoke principle C to account for the ungrammaticality of (81a) (Chomsky 1982, 1986), but this mode of explanation does not carry over to (81b). It seems unavoidable, in other words, that the anti-c-command requirement for secondary predication must be stipulated as a component of the relevant interpretive rule. This in turn implies that a unified account of strong islands and the anti-c-command requirement on parasitic gap structures and secondary predication remains out of reach.

We conclude that the theory of grammatical dependencies developed in section 2 allows a unified account of (i) strong island effects, (ii) the existence of parasitic gaps and secondary predicates and (iii) the fact that an anti-c-command condition

[11] Example (81b) is marginally acceptable if *dronken* is interpreted with focus, in which case it has focus scrambled from a lower position. We disregard this reading here.

holds of these structures. By contrast, a derivational theory of strong islands cannot relate these phenomena at all.

7 Concluding remarks

We began this paper by presenting what we take to be the best available theories of strong and weak islands. Strong islands were argued to be syntactic in nature and fall out from an LF licensing constraint on the creation of categories. Weak islands, by contrast, lend themselves most naturally to an account in terms of the theory of scope.

The rest of the paper has been devoted to making the case that this hybrid locality theory is superior to a unified account of strong and weak islands that represents these islands in one level. Such a unified theory would integrate the representation of scope (on which weak islands are computed) into the representation of syntactic dependencies (on which strong islands are computed). We argued that this integrationist approach is faced with serious empirical shortcomings, because it predicts that strong island violations can be avoided through scope reconstruction. From this, we concluded that strong and weak islands cannot both be represented at LF.

We then attempted to answer the question whether scope is represented at LF or later. In order to maintain the view that scope is an LF phenomenon, one has to demonstrate that strong islands are not due to properties of LF but have their origin in the order of syntactic operations in a derivational syntax. However, derivational theories of strong island phenomena were shown to have three major shortcomings.

First, they seem unable to distinguish the island behaviour of grammatical dependencies mediated by licensing functions from those mediated by non-licensing functions (with the possible exception of Johnson 2002). This presents serious conceptual and empirical problems. Grammatical dependencies share a cluster of properties and should therefore be treated as a natural class. At the same time dependencies do not behave uniformly with respect to strong islands. It is not clear how a derivational account of strong islands can come to grips with this dilemma.

Second, derivational theories of strong islands have difficulties explaining the properties of parasitic gaps and secondary predicates. Both phenomena involve a relation across a strong island boundary. Since a derivational approach does not allow this relation to be established in the syntax, it must rely on post-derivational interpretive rules. But since the relevant relations have the defining characteristics of syntactic dependencies, this misses an important generalization.

Third, a derivational approach to strong islands cannot tie together strong island phenomena, the existence of parasitic structures (parasitic gaps and secondary

predicates) and the fact that parasitic structures exhibit the anti-c-command requirement.

Taken together, these conclusions strongly suggest that the theory of strong islands based on the LF principle of Categorial Licensing is preferable to its derivational alternatives. If so, scope cannot be represented at LF and must instead be regulated at a later level of representation.

References

Ackema, Peter (2002). A morphological approach to the absence of expletive PRO. In *UCLWPL* 14, Ad Neeleman and Reiko Vermeulen (eds.).

Alexiadou, Artemis and Elena Anagnostopoulou (1998). Parametrising Agr: word-order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16, 451-539.

Beck, Sigrid (1995). Negative islands and reconstruction. In *On Extraction and Extraposition in German*, Uli Lutz and Jürgen Pafel (eds.), 121-143, Amsterdam: Benjamins.

Brody, Michael (1997). Mirror theory. Ms. University College London.

Chierchia, Gennaro (to appear). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In *Structures and Beyond*, Adriana Belletti (ed.). Oxford: Oxford University Press.

Chomsky, Noam (1981a). On binding. *Linguistic Inquiry* 11, 1-46.

Chomsky, Noam (1981b). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam (1986). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Chomsky, Noam. (1986). *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. (1993). A minimalist program for linguistic theory. In *The view from Building 20: Essays on Linguistics in Honor of Sylvain Bromberge*r, Kenneth Hale and Samuel Jay Keyser (eds.), 1–52. Cambridge, MA: MIT Press.

Cresti, Diana (1995). Extraction and reconstruction. *Natural Language Semantics* 3, 79-122.

Fox, Danny (1999). Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30.2, 157–196

Giannakidou, Anastasia (1997). The landscape of polarity items. PhD dissertation. University of Groningen.

Haider, Hubert (1993). *Deutsche Syntax, generativ - Vorstudien zur Theorie einer projektiven Grammatik*. Tübingen: Ginter Narr Verlag.

Caroline Heycock (1995). Asymmetries in reconstruction. *Linguistic Inquiry* 24, 547-570.

Hoeksema, Jack (2000). Negative polarity items: Triggering, scope, and c-command. In *Negation and Polarity: Syntactic and Semantic Perspectives*, Laurence R. Horn and Kato Yasuhiko (eds.). Oxford: Oxford University Press.

Honcoop, Martin (1998). *Dynamic Excursions on Weak Islands*. Holland Institute of Generative Linguistics.

Hornstein, Norbert (1995). *Logical Form: From GB to Minimalis*m. Oxford: Blackwell.

Horvath, Julia (1997). The status of 'WH expletives' and the partial WH movement construction of Hungarian. *Natural Language and Linguistic Theory* 15, 509-572.
Johnson, Kyle (2002). Towards an etiology of adjunct islands. Ms. University of Massachusetts. Amherst.
Koster, Jan (1987). *Domains and dynasties*. Dordrecht: Foris.
Kratzer, Angelika (1998). Scope or pseudoscope? Are there widescope indefinites? In *Events in Grammar*, Susan Rothstein (ed.), 163-196, Dordrecht: Kluwer.
Krifka, Manfred (1995). The semantics and pragmatics of polarity items. *Linguistic Analysis* 25, 209--257.
Laka, Itziar. (1990). Negation in syntax: On the nature of functional categories and projections. PhD dissertation. MIT.
Lechner, Winfried (1998). Two kinds of reconstruction. *Studia Linguistica* 52, 276-310.
Longobardi, Giuseppe (1987). Extraction from NP and the proper notion of Head Government. In *The Syntax of Noun Phrase*s, Alessandra Giorgi and Giuseppe Longobardi (eds.), 57–112. Cambridge: Cambridge University Press.
Manzini, Rita Maria (1992). *Locality*. Cambridge: MIT Press.
Matthewson, Lisa (1999). On the interpretation of wide scope indefinites. *Natural Language Semantics* 7, 79-134.
Mathieu, Eric (2002). *The Syntax of Non-Canonical Quantification: A Comparative Study*. PhD dissertation. University College London.
Neeleman, Ad and Hans van de Koot (2002). The configurational matrix. *Linguistic Inquiry* 33.4, 529-574.
Neeleman, Ad, Hans van de Koot and Jenny Doetjes (to appear). Degree expressions. *The Linguistic Review*.
Progovaç, Ljiljana (1994). Negative and positive polarity. A binding approach. Cambridge: Cambridge University Press.
Reinhart, Tanya (1997). Quantifier scope: How labor is divided between QR and choice-functions. *Linguistics and Philosophy* 20.4, 335-397.
Reinhart, Tanya (1998). WH-in situ in the framework of the minimalist program. *Natural Language Semantics* 6, 29-56.
Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge: MIT Press.
Ross, John R. (1984). Inner islands. In *Proceedings of the Tenth Annual Meeting, BLS*, Berkeley Linguistic Society, Berkeley, California.
Speas, Margaret (1995). Economy, agreement and the representation of null arguments. Ms. University of Massachusetts, Amherst.
Starke, Michal (2001). Move dissolves into merge: A theory of locality. PhD dissertation. Université de Genève.
[http://theoling.auf.net/papers/starke_michal]
Stepanov, Arthur (2001). Late Adjunction and Minimalist Phrase Structure. *Syntax* 4.2, 94-125.
Swart, Henriëtte de (1992). Intervention effects, monotonicity and scope. In *Proceedings of the Second Conference on Semantics and Linguistic Theory* (OSU WPL40), Chris Barker and David Dowty (eds.), 387-406. The Ohio State University. Columbus.
Szabolsci, Anna en Marcel den Dikken (2002). Islands. *The second Glot International State-of-the-Article book*, Rint Sybesma and Lisa Cheng (eds.). Berlin: Mouton.
Szabolsci, Anna & Frans Zwart (1993). Weak Islands and an Algebraic Semantics for Scope Taking. *Natural Language Semantics* 1, 235-284.
Uriagereka, Juan (1999). Multiple Spell-Out. In *Working Minimalism*, Samuel Epstein and Norbert Hornstein (eds.), 251-282. Cambridge, MA: MIT Press.

Von Heusinger, Klaus (2000). The reference of indefinites. In *Reference and Anaphoric Relations*, Klaus von Heusinger and Urs Egli (eds.), 247-265. Dordrecht: Kluwer.

Williams, Edwin (1994). *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.

Winter, Yoad (1997). Choice-functions and the scopal semantics of indefinites. *Linguistics and Philosophy* 20, 399-471.

Winter, Yoad (2002). Functional quantification. In *Proceedings of the Workshop "Choice Functions and Natural Languages Semantics" Arbeitspapier 110 FB Sprachwissenschaft,* Klaus von Heusinger, Ruth Kempson and Wilfried Meyer-Viol (eds.), 225-258. Universität Konstanz.