

A MINIMALIST THEORY OF WEAK ISLANDS

M. RITA MANZINI

*Dipartimento di Linguistica
Università di Firenze
Firenze, Italy*

Chomsky (1995) takes Relativized Minimality effects, in the sense of Rizzi (1990), as the core locality property in syntax, and correspondingly builds a Minimal Link Condition (MLC) intended to capture them into the definition of the operation Move. Thus the contrast between wh-extraction classes that are and are not sensitive to weak islands cannot be explained at the interface, making most current analyses unusable, to the extent that they are based on interface notions such as referentiality, specificity, and D-linking. In this chapter, I argue for a model where it is possible to account for the two classes of wh-phrases, those that are and are not sensitive to weak islands, in purely syntactic terms. The generalization that emerges is that sensitivity to wh-islands correlates with lack of pied-piping, or, equivalently, with reconstruction. A similar generalization is arrived at by Erteschik-Shir (this Volume); no recourse to syntactic extensions of Focus is however invoked here.

In section 1 I illustrate the data that fall within the empirical domain of weak islands; in section 2 I provide a sketch of the minimalist framework of Chomsky (1995) and of the role of the MLC within it. In section 3, I tackle the first major empirical problem for the MLC, namely its extension from wh-islands to weak islands in general. I argue that the latter requires a reformulation of the MLC, which I shall provide. Finally, in section 4 I return to the problem of the two extraction classes and provide a solution which adheres to the guidelines already sketched.

1. WEAK ISLANDS: SYNTAX VERSUS MEANING

The prototypical example of weak islands are *wh*-islands, exemplified for the case of *wh*-movement in (1); data of this type are known since at least Chomsky (1964):

- (1) *How do you wonder [what to repair t_{what} t_{how}]

In (1) *how* cannot be extracted out of the interrogative sentence introduced by *what*, or in general a *wh*-phrase cannot be extracted from a sentence introduced by another *wh*-phrase. This descriptive constraint has been refined by further work in at least three different directions. First, not all *wh*-extractions are sensitive to *wh*-islands, or at least not equally sensitive. Thus (1) is certainly worse than (2), no matter what the absolute grammaticality judgment associated with (2) is:

- (2) What do you wonder [how to repair t_{what} t_{how}]

The contrast between (1) and (2) immediately suggests, as theorized by Huang (1982) and Chomsky (1986a), that only the extraction of *wh*-adjuncts is sensitive to *wh*-islands, as opposed to the extraction of *wh*-arguments, which is not. More careful inspection of the data, however, reveals that some adjuncts are more easily extractable than others, and similar contrasts hold within the class of arguments. Compare first (3) and (4):

- (3) *Why do you know [who they fired t_{who} t_{why}]

- (4) For which of these reasons do you know [who they fired t_{who}
 $t_{\text{for which of these reasons}}$]

Extraction of *why* in (3) is ungrammatical, exactly as extraction of *how* in (1). However, as first noted by Kroch (1989) and Szabolcsi and Zwart (1990), if a context is provided which specifies the existence of a finite set of possible reasons, as in (4), extraction of the causal adjunct improves markedly. A difference between the *in situ* behavior of two *wh*-classes defined in similar terms has been observed by Pesetsky (1987), who refers to them as the D(iscourse)-linked and non-D-linked class respectively. As before, the absolute grammaticality judgments are taken to be essentially irrelevant, and the relative ones paramount.

As for arguments, Cinque (1990) was the first to note the contrast between (5) and (6) in Italian:

- (5) *Che cosa non sai* [*a chi regalare* $t_{\text{che cosa}}$ $t_{\text{a chi}}$]
 What (thing) don't you know to whom to give

- (6) **Che non sai* [*a chi regalare* t_{che} $t_{\text{a chi}}$]
 What don't you know to whom to give

(5) and (6) have essentially the same meaning, but the *wh*-phrase is represented by the bare *wh*-element in (6), while it includes a non *wh*-component in (5), namely *cosa*. In the same vein, Manzini (1990, 1992) notices the contrast between (7) and (8), again in Italian:

- (7) *Quanti posti non sai [a chi abbiano già assegnato t_{quanti posti} t_{a chi}]*
How many jobs don't you know to whom they already gave out
- (8) **Quanti non sai [a chi ne abbiano già assegnati t_{quanti} t_{a chi}]*
How many don't you know to whom they already gave out of them

In the absence of islands, it is possible both to extract the bare *wh*-element *quanti*, stranding a partitive, and to pied-pipe non-*wh*-material along with it. However while a *quanti*-NP can be extracted across a *wh*-island, as in (7), extraction of the bare *wh*-element is blocked, as in (8). Interestingly, it appears that similar data can to some extent be reproduced in English (Peter Culicover and Louise McNally, personal communication); thus extraction of the bare *wh*-element across an embedded declarative is clearly better than extraction across an embedded question, as in (8'):

- (8') a. ?How many did they give of the jobs that have been funded to union members
b. *How many must they decide to whom to give of the jobs that have been funded

Provisionally, then, we can conclude that the *wh*-island constraint must be refined by an appropriate characterization of the *wh*-extraction classes that are and are not sensitive to it. Going by the data in (3)–(8) the basic contrast appears to involve bare *wh*-elements versus *wh*-phrases including non-*wh*-material. As anticipated, however, there are two other respects in which the *wh*-island constraint, even so modified, proves inadequate. First, and most important, Rizzi (1990) shows that it is not only an intervening *wh*-phrase that blocks the extraction of *wh*-phrases of the relevant class. In particular, all and only the *wh*-phrases subject to *wh*-islands are sensitive to what Ross (1984) calls inner islands. In practice, they cannot cross negation, as illustrated in the first instance by a contrast such as (9)–(10):

- (9) **Why don't you believe [they fired him t_{why}]*
- (10) *Who don't you believe [they fired t_{who}]*

In the intended reading of (9), *why* is extracted from the embedded sentence across the negation *not*. This yields ungrammaticality, in contrast with (10), where *who* is extracted across *not*. On the basis of this contrast, one generally expects inner islands to discriminate between bare and nonbare *wh*-phrases. Indeed (11)

appears to be better than (9); similarly there appears to be a contrast between Italian (12) and (13), parallel to the contrast between (7) and (8):

- (11) For which of these reasons don't you believe [they fired him
 $t_{\text{for which of these reasons}}$]
- (12) *Quanti lavori non credi [che abbiano ancora assegnato $t_{\text{quanti lavori}}$]*
 How many jobs don't you believe that they have yet assigned
- (13) **Quanti non credi [che ne abbiano ancora assegnati t_{quanti}]*
 How many don't you believe that they have yet assigned of them

In brief, the same class of wh-extractions which are blocked by an intervening wh-phrase, are also blocked by an intervening negation. In fact, there is a whole set of configurations which turn out to block wh-extractions of the relevant class. Thus extraction from factive sentences displays the relevant contrast between bare and nonbare wh-phrases as in (14)–(15) or Italian (16)–(17):

- (14) *Why do you regret [that they fired him t_{why}]
- (15) For which of these reasons do you regret [that they fired him
 $t_{\text{for which of these reasons}}$]
- (16) *Quanti posti ti dispiace [che abbiano già assegnato $t_{\text{quanti posti}}$]*
 How many jobs do you regret that they have already given out
- (17) **Quanti ti dispiace [che ne abbiano già assegnati t_{quanti}]*
 How many do you regret that they have already given out of them

The term of weak islands was used by Cinque (1990) precisely to refer to the whole set of environments that block the extraction of only a subset of wh-elements. In this sense, weak islands contrast with strong islands which are observed by all wh-movement. Of course, having noted that wh-islands, inner islands, and factives are weak islands, the question arises whether there is a unitary characterization of such contexts. Rizzi's (1990) initial answer to this problem has proven very influential; according to it, a weak island is created in essence by an A'-specifier. This is immediately true of wh-islands created by wh-phrases in the [Spec, CP] position. A prima facie problem is posed, however, by the impossibility to extract wh-phrases of the relevant class from embedded interrogatives introduced by an *if* complementizer rather than by a wh-phrase. This is illustrated by the contrasts of the familiar type in (18)–(19) and (20)–(21):

- (18) *Why do you wonder [if they fired him t_{why}]
- (19) For which of these reasons do you wonder [if they fired him
 $t_{\text{for which of these reasons}}$]

- (20) *Quanti posti ti chiedi* [se abbiano già assegnato $t_{\text{quanti posti}}$]
How many jobs do you wonder if they have already given out
- (21) **Quanti ti chiedi* [se ne abbiano già assegnati t_{quanti}]
How many do you wonder if they have already given out of them

Rizzi (1990) argues that the parallelism between canonical wh-islands and interrogative islands of the type in (18)–(21) can be maintained by postulating an empty counterpart to the wh-phrase in the [Spec, CP] position of the latter. Much the same holds in the case of inner islands, where negation is overtly realized by a head, possibly in English and certainly in Italian. Crucially, the Neg head in Italian can co-occur with negative phrases such as *nessuno* ‘nobody’ without giving rise to a double negation reading. Haegeman and Zanuttini (1991) take this to indicate that the Neg head and the Neg phrase are in an agreement configuration (i.e., a Spec-head one), in abstract syntax. If so, we have some measure of independent justification for the postulation of an empty Neg phrase in [Spec, NegP] in (12)–(13). Finally, Rizzi (1990) does not explicitly discuss factive islands, but an empty operator in the A'-Specifier of a factive C is postulated by Melvold (1991).

To summarize, the generalization corresponding to the data reviewed so far could be expressed by a Weak Island Constraint, to the effect that A'-configurations create islands for a subset of wh-extractions, as already described. There is, however, one final way in which even this fairly detailed generalization must be further refined. So far, only cases of wh-extractions have been considered. As it turns out, however, other syntactic dependencies can be argued to be equally sensitive to weak islands. Thus Manzini (1993a) argues that in Italian no weak island can intervene on the path between the Neg head *non* and an associated negative phrase. The argument can be roughly schematized as follows. First, morphologically negative phrases such as *nessuno* can be licensed not only by negation but also by an interrogative complementizer, as seen in the simple cases in (22)–(23):

- (22) *Non vedo nessuno.*
I don't see anybody.
- (23) *Vedi nessuno?*
Do you see anybody.

In this respect of course *nessuno* behaves more like a polarity item than like a negative quantifier. Next, consider the two sentences in (24)–(25):

- (24) *Non so che veda nessuno.*
I do not know that s/he see anybody.
- (25) *Non so se veda nessuno.*
I do not know whether s/he see anybody.

In (24), it is possible for *non* to license *nessuno* across sentence boundaries, where the complementizer is declarative *che* and the verb is inflected in the subjunctive. In (25) on the other hand, the only possible reading for *nessuno* is associated with the interrogative C, *se*, and therefore roughly existential. The negative reading associated with *non* is completely impossible. If we construe the licensing of *nessuno* by *non* as subject to weak islands, then this result follows, in that Q acts as a block for the dependency (Neg, . . . , *nessuno*). The correctness of this approach is verified by the existence of a similar contrast between (26) and (27):

(26) *Sai che veda nessuno?*
Do you know that s/he see anybody.

(27) *Sai che non veda nessuno?*
Do you know that s/he do not see anybody.

In (26) it is possible for the N-phrase to be licensed with a roughly existential reading by the abstract interrogative C of the matrix sentence. The same reading, however, is blocked in (27), where the N-phrase can only be associated with the negation, yielding the equivalent of a negative quantifier. The latter result of course follows if we take all licensing of N-words in Italian to be sensitive to weak islands. If so, the impossible reading of (27) represents just an instance of an inner island, i.e., the blocking of a dependency (Q, . . . , *nessuno*) by Neg. Similarly, Aoun and Li (1993) show that the effects just reviewed for Italian hold in Chinese, where the same lexical item *in situ* can be interpreted as a wh-phrase, as a negative phrase, or as a polarity existential, depending on its licenser. In short, the data in (22)–(27) indicate that N-words, like wh-phrases, are sensitive to weak islands. Unlike wh-phrases, however, all N-words display such sensitivity. The question then is how a generalization across the two relevant classes can be obtained.

The presentation of the basic empirical observations pertaining to weak islands and the theoretical questions they raise can be considered concluded at this point. Before providing my own solutions to these questions, I shall briefly go through the various theoretical systematizations that have been proposed for them. Descriptively, what I have called the Weak Island Constraint holds of the dependencies entered by an appropriately defined subset of wh-phrases as well as by what have been provisionally dubbed N-phrases and perhaps more. The constraint itself prevents such dependencies from crossing configurations which include questions, negations, factives, and perhaps more. Correspondingly, there are two major desiderata for any theory of this complex of phenomena, namely (i) a correct characterization of the elements/processes that are subject to weak islands, including the split between the two wh-classes, and (ii) a correct characterization of the island environments themselves.

One major reason for interest in this theoretical problem is that on the basis of

the descriptive generalizations presented so far, both semantic/pragmatic or syntactic accounts can in principle be entertained. Semantic/pragmatic accounts began to emerge in particular once it became clear that the adjunct versus argument asymmetry, still advocated by Chomsky (1986a), was not empirically adequate to account for the behavior of *wh*-extractions with respect to weak islands. Thus Cinque (1990), who wants otherwise to maintain a syntactic construal of weak islands, in terms of Rizzi's (1990) Relativized Minimality, resorts to syntactic encoding of the relevant semantic–pragmatic distinction between the two *wh*-extraction classes, by means of the feature $\pm$ referential. The same treatment is essentially adopted by Rizzi (1990), who treats $\pm$ referential as a feature of theta-roles. But the encoding of semantic–pragmatic properties by nonindependently motivated syntactic features is not the most desirable of outcomes. This appears to provide motivation for a fully semantic theory such as Szabolcsi and Zwart's (1990), where not only the two (*wh*-)extraction classes, but also the island contexts are characterized in purely semantic terms, namely in terms of individual variables and monotonic decreasing quantifiers, respectively. A similar line appears to be taken by Kiss (1993), who not only characterizes the two extraction classes in terms of specificity, but also takes the weak island constraint to reduce to a requirement that a nonspecific *wh*-phrase must be in the scope of a specific one and not vice versa.

By contrast, the most consistent exponent of a purely syntactic theory of weak islands has been Chomsky (1964, 1973, 1986a, 1995), whose basic intuition appears to be that there is no problem in associating *wh*-island violations with a perfectly well-formed meaning: the violation is purely formal and computational. Indeed, at a suitably abstract level, some version or other of the weak island constraint falls together with completely unrelated descriptive constraints from the point of view of meaning. Thus weak islands fall under some version of the A-Over-A (family of) constraint(s) for Chomsky (1964) and under Superiority for Chomsky (1973). Chomsky (1986a) attempts an at least partial reduction of weak islands to strong islands, as well as to locality constraints on A-movement and head-movement, in terms of barriers; he assumes the distinction between the two *wh*-extraction classes to be simply that of argument versus adjunct, or to be more precise, of theta-governed position (object) versus non-theta-governed position (subject and adjunct). Most recently, within the minimalist framework, Chomsky (1995) takes weak islands to fall together with locality constraints on A- and head movement under the MLC, while leaving the problem of the two (*wh*-)extraction classes completely open.

The fundamental aim of this chapter is to show that a purely syntactic approach to weak islands is correct. I shall do so by taking Chomsky's (1995) theory of weak islands as my basis and illustrating its empirical problems. I shall then proceed to show that the theory can be modified, within the conceptual bounds of minimalism, to accommodate the problems.

2. THE MINIMALIST BACKGROUND

In Chomsky's (1995) Minimalist Program each linguistic expression is characterized by two representations, a P(honological) F(orm) representation, interfacing with the articulatory-perceptual system(s) and a L(ogical) F(orm) representation, interfacing with the conceptual-intentional system(s). This represents a clear departure from previous models, even within the Principles and Parameters framework, which are characterized by at least two additional syntactic representations, namely Deep-Structure (DS) and Surface-Structure (SS). Everything else being equal, the minimalist model is obviously preferable on conceptual grounds, since in it linguistic representations are seen to simply match the interfaces between language and other cognitive systems. A central property of the minimalist model is what Chomsky (1995) terms *inclusiveness*. In other words, not only inputs but also outputs of linguistic computations are entirely expressed in terms of properties of lexical items. This implies the impossibility of adding information in the course of the derivation and/or in the output representation in the form of such standard devices of previous models as bar levels, indices, and the like.

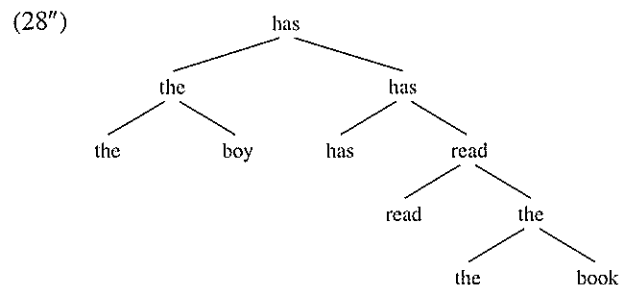
The basis of each linguistic computation is represented by a set of lexical items. One of the operations of grammar, *Select*, applies to this set, selecting one of its members and introducing it into the set of syntactic objects of which a derivation consists at each of its stages. New syntactic objects are formed by the operation *Merge*, which takes a pair of given syntactic objects (S_i , S_j) and creates a single syntactic object out of them, namely a labeled set ($S(S_i, S_j)$), where the label S is either S_i or S_j . A simple example of a computation involving only *Merge* is shown in (28)–(28'). (28) provides the numeration set for the computation; note that separate instances of the same lexical item are listed separately. (28') represents each individual step in the computation. In particular, the labeled sets in (28'a) and (28'd) are formed by the merger of two lexical items selected from (28), (28'b) and (28'c) by the merger of an item selected from (28) with a previously formed labeled set. Finally (28'e) is created by the merger of two labeled sets, which for readability's sake I have presented on separate lines:

(28) (the, boy, has, read, the, book)

- (28') a. (the, (the, book))
 b. (read (read, (the (the, book))))
 c. (has (has, (read (read, (the (the, book)))))
 d. (the (the, boy))
 e. (has
 ((the (the, boy)),
 (has (has, (read (read, (the (the, book)))))

A fairly obvious correspondence can be defined between the complex object in (28'e), resulting from the computation and a familiar tree structure of the type in

(28''). In essence, lexical items correspond to terminal nodes in the tree, while labels correspond to nonterminals:



The other fundamental rule of the computational system is *Move*. Suppose we have a complex object *S* containing the objects *A* and *K*. *Move* takes *A*, copies it and merges it with *K*. A chain is the ordered pair consisting of *A* in the position resulting from this merger, and of its copy, technically known as its trace t_A , in the original position. This is also easily illustrated. Consider the numeration set in (29) and the computation in (29') that takes it as its basis. The steps in (29'a-c) are entirely comparable to those in (28'a-c). In (29'd), however, the entire object in (29'c) is merged with a copy of (29'a), which it contains. In other words, movement applies, yielding two copies of the same object (29'a) in its output. One of the two copies, in fact the trace, is then subjected to the operation *Delete*, which simply marks it as not interpreted at PF; this is informally indicated in (29'd) by overstrike:

(29) (the, book, was, read)

- (29') a. (the (the, book))
 b. (read (read, (the (the, book))))
 c. (was (was, (read (read, (the (the, book))))))
 d. (was
 ((the (the, book)),
 (was (was, (read (read, (~~the (the, book)~~))))))

The last major operation of the computational system is *Spell-Out*, which strips away the phonological properties of lexical items from the semantic and formal features. It is then *Spell-Out* which is ultimately responsible for mapping syntactic objects of the type in (28') or (29') to PF and LF representations respectively.

Let us consider *Move* in more detail. *Move* is crucially restricted by two Economy principles, which enter into its definition, namely *Last Resort* and the *MLC*. In Chomsky's (1995) formulation, *Last Resort* states that movement of *A* to a target *K* is possible only if *A* enters into a feature-checking relation with *K*. Thus, given any complex syntactic object (*K* (*A*, *K*)) formed by movement, its two constituents *A* and *K* must be related by feature checking. This excludes ar-

bitrary applications of Move, as well as applications of Move that may be dictated by some further purpose, but that do not satisfy any needs of the pair (A, K). The overall effect of Last Resort is to effectively eliminate overgeneration in the movement component of grammar. Needless to say this represents one of the important departures of the minimalist approach from the Government and Binding (GB) approach of Chomsky (1981). GB grammars are allowed to overgenerate arbitrarily, and the overgeneration is then taken care of by filters at the interface. In a minimalist grammar, wild rule application is blocked by Last Resort in the course of the computation.

Note that, strictly speaking, Move applies to features, explaining why Last Resort crucially refers to it. A problem of empirical adequacy arises in connection with this latter assumption, to the extent that overt movement takes the shape of movement of entire phrases, rather than of (bundles of) features. According to Chomsky (1995) this problem is solved by taking into account the fact that (bundles of) features in isolation cannot be pronounced. Thus it is reasons of PF interpretation that force entire phrases to pied-pipe in overt syntax. In covert syntax, where PF features have been stripped away, we can assume that movement takes the pure form of Move F(eature).

Finally, the second crucial property of the operation Move, and the one that will concern us most directly here, is the MLC, as in (30):

(30) *Minimal Link Condition*

K attracts α only if there is no β , β closer to K than α , such that K attracts β

Consider for instance a typical context for a wh-island violation, as schematically represented in (30'):

(30') Q' ... [wh Q ... wh

Under the formulation of the MLC just given, in (30') movement of the lower *wh* to check its features against Q' is blocked, in that the *wh* in the Spec of Q can also move to Q' to do the same, and crucially is closer to it. Superraising effects can be accounted for in an analogous way. Consider (30''):

(30'') I seems [that it was told John [that IP]]

Under the MLC, movement of *John* to the matrix [Spec, IP] position to check the D-feature of I is blocked by the presence of the intermediate *it*, which can also move to the same position to check this feature. Chomsky (1995) also suggests that the formulation of the MLC is more natural if Move is thought of in terms of attraction. Thus incorporating Last Resort into this new conception, we can define Attract/ Move by saying that K attracts F if F is the closest feature that can enter into a checking relation with a sublabel of K.

In essence, the MLC exploits the generalization drawn by Rizzi's (1990) Relativized Minimality, itself anticipated by much work, including at least Culicover

and Wilkins (1984) and Aoun (1985). Indeed, the MLC derives two of the main consequences of Relativized Minimality, namely *wh*-islands and superraising. A systematic comparison of the two principles, however, reveals that they are not necessarily equivalent, when the wider class of weak islands is taken into account. Indeed, no explicit discussion of how to derive the latter, as opposed to just *wh*-islands, is provided by Chomsky (1995).

By contrast, a major difference between Relativized Minimality and the MLC explicitly noted by Chomsky (1995) concerns Travis's (1984) Head Movement Constraint (HMC). In essence, the MLC excludes those cases of the HMC where a head moves across another head attracted by the same target. However, unlike the HMC, the MLC does not necessarily exclude all cases of movement of a head across another head. Interestingly, the empirical adequacy of the HMC has been extensively discussed quite independently of the proposals of Chomsky (1995). A well-known *prima facie* problem for the HMC is represented by so-called long head movement, as studied notably by Rivero (1994). In essence, it appears that in a number of languages a lexical *V* moves to *C* under appropriate conditions, crossing an auxiliary in *I*. These cases are used by Roberts (1994) to argue in favor of a revision of Relativized Minimality that distinguishes between *A* and *A'* heads. As is evident, these same cases potentially favor the MLC, depending on whether some property of *C* can be shown to be involved, that attracts *V* but not *I*. Similar considerations apply to the revision of Relativized Minimality proposed by Baker and Hale (1990). It is then a purely empirical question whether the inability of the MLC to derive the HMC makes it inadequate in some respect, or conversely whether it provides a solution to long-standing problems with the HMC itself.

The really serious problem for the MLC is acknowledged by Chomsky (1995) when he states that the superraising violation, which falls under the MLC, is "far more severe" than the *wh*-island violation involving arguments, which should fall under the same principle. In fact, while Chomsky (1995) seems to take *wh*-extraction of arguments from a *wh*-island to be degraded, the classical study of Rizzi (1982) treats many instances of them in Italian as fully well-formed. Suppose they are, as we have assumed throughout. As already noted, Chomsky (1995) is fully explicit in construing the MLC as part of the definition of movement. If so, we are faced with the problem of a subset of movement cases that do not appear to conform to a part of the movement rule. As Chomsky (1995) is careful to point out, the minimalist framework leaves the problem completely open.

A second problem in locality theory that the minimalist approach of Chomsky (1995) leaves completely open is that of strong islands. An independent principle appears to be envisaged, presumably along the lines of Huang's (1982) Condition on Extraction Domains (CED). Still, it remains unclear whether the CED is to become a part of the definition of Move, like the MLC, or not. In either case, it also remains unclear why all types of movement are equally sensitive to strong islands and hence the CED, as opposed to weak islands and the MLC.

Finally, Chomsky's (1995) theory fails to make any predictions concerning a third classical problem in locality theory, which concerns a class of phenomena apparently irreducible to either the MLC or CED, accounted for by standard models within the principles and parameters framework in terms of some notion of head government. These are phenomena of the *that*-t class, typically involving asymmetries between object and subject. Chomsky's (1995) grammar excludes the notion of head government on what are essentially simplicity grounds. In particular, the (head, complement) relation and the (Specifier, head) relation are independently recognized by the grammar, in that they correspond to configurations created by Merge. Thus the only head government relation that does not correspond to a configuration created by Merge is the relation between a head and the Spec of its complement; the notion of head government then needs to be defined essentially for this case. Chomsky (1995), however, argues that this configuration systematically reduces to one of the other two. For instance, there is no Exceptional Case Marking of the subject of *believe*-type complements by the matrix V; rather, Exceptional Case Marking effects are due to the subject raising to the matrix (Spec, AgrO). Suppose we accept this argument against head government. This means of course that we can no longer have recourse to accounts of *that*-t phenomena based on this notion, such as Rizzi's (1990). But other branches of the theory make no predictions about them either.

Interestingly, the last three problems illustrated here for Chomsky's (1995) minimalist grammar are individuated as the fundamental problems in locality theory by Manzini (1992). In preminimalist terms, the first corresponds to the disjunction between argument and adjunct wh-movement with respect to antecedent government; the second to the disjunction between antecedent government, as relevant for weak islands, and the CED, as relevant for strong islands; and the third to the disjunction between antecedent government and head government. It seems fair to conclude that the minimalist approach makes no substantial progress, not only with respect to the first question, as explicitly recognized by Chomsky (1995), but also with respect to the other two. I shall concentrate here on the first one, leaving the other two for further study.

Let us consider (1) again in more detail. Suppose we have reached the stage in its derivation represented by (31):

(31) [do-Q'] [you wonder [what Q [to repair t_{what} how]]]

In order to derive (1) from (31), we need to move *how* to the checking domain of Q', hence in practice to the specifier of the higher C. This move in turn is legitimate under Last Resort since the Q feature is an attractor for the wh-feature. However, it equally clearly violates the MLC. In (31), movement of *what* to the specifier of the higher C is also licensed by Last Resort; for, the wh-feature of *what* is checked in the lower [Spec, CP] position, but it is not deleted, being +interpretable. Crucially, of the two wh-phrases, *what* is closer to Q' than *how*, where

closeness is straightforwardly defined in terms of c-command. Therefore, the availability of a closer attractee in the shape of *what* blocks movement of the farther attractee *how* to the attractor Q' under the MLC, as desired.

One question that Chomsky (1995) asks is why the outcome of (31) in (31'), with actual movement of *what* is equally ill-formed:

(31') *What do you wonder [_{t_{what}} to repair how]

The derivation is convergent. Indeed being +interpretable, the wh-feature on *how*, the only one which remains unchecked, does not require any checking. Chomsky (1995) therefore assumes that (31') fails because no well-formed meaning can be associated with it.

On the other hand, as already anticipated, the MLC does not just derive wh-island violations of the type in (1), but also semantically and/or pragmatically unrelated constraints on A-movement and head movement. The case that Chomsky (1995) chooses to illustrate this is superpassive/superraising, as in (32):

(32) *John seems [that it was told t_{John} [that IP]]

Consider the last stage of the derivation of (32), which corresponds to (30'') repeated below:

(30'') I seems [that it was told John [that IP]]

The I in (30'') is an attractor for DPs because of its D-feature. Movement of *John* to the checking domain of I, specifically to its Spec, therefore complies with Last Resort. However, movement of *it* to [Spec, IP] also complies with Last Resort, since its D-feature, though checked by the embedded I, is +interpretable and hence not deleted. Crucially, *it* is closer to the matrix I than *John* in straightforward c-command terms. Therefore *it* blocks the desired movement of *John* under the MLC. As before, the question can be asked why movement of *it* to [Spec, IP] yields an equally ill-formed result, as in (33):

(33) *It seems [that t_{it} was told John [that IP]]

In this case it is clear, however, that the derivation does not converge for reasons independent of the MLC. Thus even if we take *it* to be able to check Case features, it can do so only once, in the lower I position. After this, its Case feature, being –interpretable, is deleted. Therefore the Case feature associated with the higher I remains unchecked.

As I have already mentioned, despite its obvious success with the cases just reviewed, Chomsky's (1995) theory faces empirical problems. To summarize, there are essentially two of them and both are far from trivial. The first one is recognized by Chomsky (1995) himself, who acknowledges that the MLC makes no progress on the complex issue why superpassive/superraising violations of the type in (32) are much stronger than wh-island violations with arguments,

of the type in (2), though they are both supposed to fall under exactly the same constraint, namely the MLC. This reduces to the by now familiar problem of providing a minimalist theory of the two wh-extraction classes: for violations of wh-islands by wh-adjuncts appear to be otherwise comparable in severity to super-passive/superraising violations.

The second problem is not recognized by Chomsky (1995), but clearly arises as soon as we attempt to extend the MLC from wh-islands to all weak islands, including, for instance, the inner and factive islands reviewed above. Thus consider the simple inner island violation in (9). The relevant stage of the derivation is as in (34). For simplicity I have indicated *not* in what appears to be its base position, rather than moving it along with *do* to the C position; both outcomes are indeed known to be in principle possible in English:

(34) [do-Q] [you [not believe [that they fired him why]]]

The question is whether movement of *why* to [Spec, Q] can be blocked by *not* under the MLC. The answer is positive only if the negation can be shown to have the feature attracted by Q. If so, since it is closer to Q than the wh-phrase, movement of the latter is blocked. Neg, however, clearly does not have the wh-feature, since there is no context where it can independently be shown that Neg checks a Q feature. Thus inner islands cannot be derived by the MLC.

3. THE MINIMAL LINK CONDITION AND ITS PROBLEMS: FROM WH-ISLANDS TO WEAK ISLANDS

In trying to solve the problem in (34), it is worth casting a look back to Rizzi's (1990) Relativized Minimality, which does allow for the weak island generalization to be drawn. Within a GB framework, there are principles which insure that Spec-head configurations share features, for example, the WH-Criterion of Rizzi (1991), the NEG-Criterion of Haegeman and Zanuttini (1991), and so on. However, Relativized Minimality runs on another property, namely that wh-phrases, Neg phrases, and the like all are in A'-position. Let us try to reconstruct this double articulation of Rizzi's (1990) theory into Chomsky's (1995). To begin with, we must look for some feature that attractees of Q, Neg, and so on, share. For instance, they share the property of being weak DPs, or indefinites; syntactically their D appears to be empty, with wh, Neg, and other features realized in a lower functional position within the DP.

Let us provisionally settle for this property and consider (34) again in this light. We can suppose, as is standardly done for Relativized Minimality, that an abstract Neg-phrase sits in the checking domain of *not*. If we take Q to attract an indefinite D, the abstract Neg phrase associated with *not*, which by hypothesis is an indefi-

nite D, is now in a position to block movement of the farther away indefinite D represented by the *wh*-phrase. As for the fact that Neg phrases can never be seen to check Q features, or for that matter *wh*-phrases can never be seen to check Neg features, we can invoke a constraint independently postulated by Chomsky (1995), namely (35):

(35) Mismatch of features cancels the derivation

Unfortunately, it is easy to see that this solution does not work. The reason appears to be directly related to the choice of indefiniteness as the common property of all attractees interacting under the MLC. Indeed, there is any number of indefinites in grammar that are not attracted by *wh*-, Neg, and the like; these can normally intervene on any *wh*/ Neg-extraction path, without blocking it. Thus *a man* does not ordinarily block *wh*-movement, and so on. This may suggest that a property other than indefiniteness must be sought. However, there is another, more sweeping conclusion that can also be pursued, namely that potential attractees do not really interact with one another. Rather it is potential attractors that interact, as independently suggested by Manzini (1993a).

In Manzini's (1993a; 1993b) model, in particular, the basic operation of grammar is Form Dependency, which in the most elementary case takes as its input two lexical items and forms an ordered pair out of them, as in (36):

(36) *Form Dependency*
 $a, b \rightarrow (a, b)$

Movement/binding paths are built by successive applications of the rule. For the sake of exemplification, we can consider *wh*-island violations with extraction from argument positions, since there is a clear complementation path connecting the *wh*-phrase to its trace. Thus in (8) the dependency connecting the operator and its variable takes the form in (37):

(37) *Quanti—Q—sai—t_{sai}—Q—abbiano—assegnato—t_{quanti}*

It can clearly be seen in (37) that while the intermediate attractor Q intervenes on the path of the dependency (and so would a Neg or similar head), the attractee *wh*-phrase in its Spec is not even included, as is generally the case for left branches. The weak island generalization suggested by Manzini (1993a) on the basis of such representations is that Op-vbl dependencies are excluded if they cross another Op, as in (38), where Op subsumes heads such as Q and Neg and vbl represents a variable (i.e., again an indefinite), which ends up in the scope of both:

(38) *Weak Islands Generalization*
 $*(Op, \dots, Op, \dots, vbl)$

I do not intend to pursue here the Form Dependency approach any further, since it only has empirical import in domains, such as strong islands and head-government, as well as constituency, which are not discussed here. I shall, however, pursue the generalization in (38), based not on the intervention of a potential attractee but of a potential attractor on a movement path, within the conventional minimalist framework.

My starting point is once again (31) and (34). Obviously enough, it is not sufficient to theorize that Q prevents Q' from attracting *how* on the basis of the fact that both are attractors for the same *wh*-feature. Indeed this does not predict that an intervening Neg acts as a block for a *wh*-dependency. Suppose, however, we have recourse to the idea already introduced above, that the appropriate formulation of minimality is sensitive to the fact that all of the interacting operators attract an indefinite D. Both of the examples in (31) and (34), repeated below, are finally captured:

(31) [*do-Q'*] [*you wonder [what Q [to repair t_{what} how]]]*

(34) [*do-Q*] [*you [not believe [that they fired him why]]]*

In (31) the indefinite *what* moves to the closest attractor Q; however, we can assume that movement of the indefinite *how* to the attractor Q' is blocked by the intervening attractor Q. Similarly in (31) the indefinite *why* cannot move to the Q attractor because of the intervening Neg attractor. In other words, suppose we have reached the stage of the derivation in (31) or (34) with the Q' and Q attractors respectively active. All that we need to block the derivation is a principle to the effect that Q'/Q can attract an indefinite D only down to the next attractor of indefinite Ds, in practice Q in (31) and Neg in (34). In either case the lower *wh*-phrase is not included in the attraction scope of the higher operator, and the derivation crashes.

The formulation of the minimality constraint that we arrive at takes then the form in (39):

(39) *Minimality*

Given an attractor feature F and an attractee feature A_F, F attracts A_F only down to the next attractor F' for A_F

Let us apply (39) to the extraction of *wh*-phrases across embedded interrogatives. In (31), as already anticipated, Q' attracts an indefinite D, but only down to the next attractor Q. This means that *how* cannot be attracted. On the other hand, it is possible to construe *what* as its attractee. But if *what* is attracted, we can assume that the resulting structure does not have a meaningful interpretation, exactly as does Chomsky (1995).

Though Chomsky (1995) himself leaves the question completely open, it is interesting to speculate on what meaning principle could have this effect. One candidate is the clause of the Bijection Principle of Koopman and Sportiche

(1982), which prevents a variable from being bound by more than one operator. There are no known exceptions to it, though the converse clause, namely that each operator binds one and only one variable, is potentially undermined, for instance, by parasitic gap. It is natural to assume that movement of the same *wh*-phrase to two different [Spec,Q] positions results in the binding of its trace by two different Q operators. But no interpretation seems to be associated with such a configuration, resulting in the unacceptability of (31'), indeed on semantic rather than syntactic grounds, in keeping with Chomsky's (1995) insight.

One immediate advantage of an attractor-based formulation of minimality over attractee-based ones is pointed out already by Manzini (1993a). This is that it does not need to rely on the postulation of empty attractees. Thus consider the simple case of *wh*-extraction out of an interrogative sentence introduced by *if*, as in (19) above. The relevant stage of the derivation is represented in (40):

(40) [do-Q'] [you wonder [[if-Q] [they fired him why]]]

Rizzi's (1990) Relativized Minimality works in (40) only if there is an otherwise unmotivated empty *wh*-phrase in [Spec, Q]; exactly the same appears to be true under the MLC, though Chomsky (1995) does not recognize this explicitly. Obviously enough, the Minimality in (39) does not require any *wh*-phrase other than *why* in order to work; for, it is based on *if* closing off the domain of attraction of Q'. As Manzini (1993a) argues in detail, there are a considerable number of contexts in which the no empty attractees theorem has empirical import, as will also become evident from the discussion to follow.

The next major case to be considered is *wh*-extraction from an inner island, as illustrated in (34). The MLC cannot block the derivation; Relativized Minimality can, but apart from any other consideration, it needs to postulate an otherwise unmotivated empty *Neg*-phrase in [Spec, NegP]. Under the present Minimality the derivation is blocked, because the only indefinite whose features could check the matrix Q feature (i.e., the *wh*-phrase) is not within the immediate scope of Q, but rather of *Neg*.

If the Minimality in (39) is correct, we furthermore expect it to predict not only weak island violations produced by *wh*-extractions but also data of the type in (25)–(26), concerning abstract *Neg*-movement. At the relevant stage of the derivation, (25)–(26) present substructures of the type in (41) and (42) respectively:

(41) [*non* [*so* [[*se-Q*] *veda nessuno*]]]

(42) Q [*sai* [*che* [*non* [*veda nessuno*]]]]]

Consider first (41). Abstract movement of (the relevant features of) *nessuno* to the checking domain of *non* is clearly blocked under (39) by the fact that the *Neg*-phrase is not in the immediate scope of *Neg*, but rather of Q. In contrast, in (42) abstract movement of *nessuno* to Q is blocked by the fact that it is not in the immediate scope of Q, but rather of *non*. No empty Specs are required by this

theory. To be more precise, attraction of *nessuno* cannot be effected literally by the Q feature in (41), since apart from everything else, the Q feature is satisfied by merger of *if*. Rather, it must be the case that a weak feature, call it M, is associated with Q as well as with Neg, and both act as attractors for N-words in abstract syntax. In this way we also explain why both (41) and (42) are well formed, though with different interpretations. When *nessuno* is in the immediate scope of Q it is attracted by its M feature; while it is attracted by the M feature of Neg when it is in its immediate scope. As for the nature of the M feature, this appears to relate to the intensional properties of both Q and Neg: the other major element that attracts N-words in Italian is hypothetical *se* ('if'). The label M was chosen to suggest Modality, given that the same set of elements act as licensors for the subjunctive, as argued for Italian by Manzini (1996).

One major set of data is still unaccounted for, namely wh-extraction from factive islands, as for instance in (14). The position at which the property of factivity is realized and wh-dependencies are blocked is identified with the Spec of the factive C by Melvold (1991) and with the C itself by Hegarty (1992) and Manzini (1993a). However, a factive appears to be a Topic in the informational structure, its content being pragmatically presupposed. Thus the factive itself and/or its C head appear to have absolutely nothing in common with operators of the type of Q or Neg, which correspond if anything to a Focus. Not surprisingly, the empirical arguments used in favor of the identification of the factivity barrier with C lack any real force. Thus according to Manzini (1993a), in Greek the property of factivity appears to be realized as a specialized C, namely *pou*, as opposed to the ordinary declarative *oti*. However, Roussou (1994), from which the relevant data are taken, points out that *pou* creates a strong rather than a weak island. Conversely, *oti* can be seen to create weak islands in contexts comparable to English or Italian factives.

One possible way out of this predicament is to invoke the idea that a factive sentence, not unlike any other Topic, is only licensed by the presence of a Focus. Suppose then the verb that governs the factive sentence intrinsically bears some F specification. This needs to be checked against some F head. All that we require in order to block wh-extraction out of a factive configuration is that the relevant wh-phrase is in the immediate scope of this F head and not of Q. For the sake of the present discussion, I shall simply indicate this F head as immediately superordinate to the factive verb, and thus taking its complement in its immediate scope. If so, at the next to last step of the derivation (14) has the structure indicated in (43):

(43) [do-Q] [you [F regret [that they fired him why]]]

Given (43), the Minimality in (39) applies trivially as in all of the other cases examined: *why* is in the immediate scope of F rather than Q, and the latter cannot attract it.

Several empirical ramifications of the analysis provided so far are worth mentioning briefly. First, we predict that overt Focusing blocks extraction of the relevant class of *wh*-phrases. This appears to be correct, as in the contrast between (44) and (45); thus there are Focus islands, comparable to *wh*-islands, Neg islands, and so on:

(44) *Why do you think [that only him they would fire t_{why}]

(45) Who do you believe [that only for this reason they would hire t_{who}]

Second, if the relevant class of *wh*-extractions is sensitive to factive and Focus islands, so should be Neg-movement. According to Manzini (1993a) this is indeed the case, though examples of the type of (46) are well-formed:

(46) *Non mi dispiace che veda nessuno*
I do not regret that s/he see anybody

Assuming that there is an F (factivity or Focus) operator in the position immediately superordinate to the matrix V, we would expect (46) to be excluded at least on the negative reading of *nessuno*, on the basis of the fact that the Neg-phrase is not in the immediate scope of the Neg attractor. The acceptability of (46) nevertheless is not problematic, since it is strictly dependent on the availability of a nonfactive reading. Indeed, in the scope of Neg, the property of factivity can itself be lost. It is much easier to show that Neg-movement is blocked by an overt Focus, as is indeed the case in (47):

(47) **Non penso che solo in questa fabbrica abbiano licenziato nessuno*
I do not believe that only at this plant they fired anybody

The last prediction of the framework is that Focus movement should be blocked by Q, Neg, and F. Interestingly, it turns out that not all F-phrases are sensitive to weak islands, but only the counterparts to bare *wh*-phrases, as exemplified in Italian (48)–(49) with *wh*-island contexts and (50)–(51) with inner island ones:

(48) *Tre lavori, mi chiedo a chi abbiano già assegnato*
Three jobs, I wonder to whom they have already given out

(49) **Tre, mi chiedo a chi ne abbiano già assegnati*
Three I wonder to whom they have already given out of them

(50) *Tre lavori, non credo che abbiano già assegnato*
Three jobs I do not believe that they already gave out

(51) **Tre, non credo che ne abbiano già assegnati*
Three I do not believe that they already gave out of them

Thus we come back full circle to the problem we started with in section 1, and that I have not yet examined after the reformulation of the Minimality in (39),

namely the problem of the two wh-extraction or indeed F-extraction classes, to which I turn next.

4. THE TWO EXTRACTION CLASSES

The informal generalization arrived at in section 1 on the basis of the data concerning wh-extractions is that bare wh-phrases are sensitive to weak islands, while pied-piped ones are not. This is essentially the conclusion arrived at by Manzini (1990, 1992), and it receives decisive support from reconstruction phenomena. Indeed, if reconstruction of non-wh-material takes place in an overtly pied-piped wh-phrase, along the lines of Chomsky (1995), only a bare wh-word is left in scope position in abstract syntax; if so, we expect reconstruction to display sensitivity to weak islands. This is indeed the case.

Following Chomsky (1995), we can assume that the first step in the reconstruction process is raising of the wh-word to take scope over the wh-phrase, as indicated for Italian (7) in (52). We can then assume that in the operator position, everything but the operator deletes; all nonoperator material is realized in the trace position, producing a structure like (53), interpreted roughly as in (54):

(52) [quanti] [_t_{quanti} posti] . . . [t]

(53) [quanti] . . . [_t_{quanti} posti]

(54) [quanti x] . . . [x posti]

Following Chomsky (1995), the nonoperator material is, of course, not literally reconstructed in the trace position. Rather, we can think of the trace as a copy of the wh-phrase. The operator head is then deleted *in situ*, while the nonoperator material is deleted in the fronted position. Chomsky (1995) furthermore assumes that a nonreconstructed second structure exists side by side with (53) or (54). For (7), such a structure corresponds simply to (52).

The basic empirical evidence in favor of reconstruction involves the interaction of wh-movement with anaphora, for instance, reflexivization. In Chomsky's (1995) terms, both (55) and its reconstructed counterpart (56) are well formed in abstract syntax:

(55) John wondered [which picture of himself] Bill saw [t]

(56) John wondered [which] Bill saw [_t_{which} picture of himself]

Under natural locality assumptions about reflexivization we predict that in (55) the reflexive will be coreferential with *John*, while in (56) it will be coreferential with *Bill*. Indeed the sentence that corresponds to both of these structures appears to be ambiguous precisely in the way predicted.

Suppose then we force reconstruction in an example involving overt pied-piping by forcing the *in situ* reading of a reflexive, as in (57). This is perfectly possible in the absence of islands, as in (57a), but impossible if reconstruction crosses a wh-island, as in (57b). By comparison, (58) shows that if the reading of the reflexive forces absence of reconstruction, then the presence of a wh-island becomes irrelevant:

- (57) a. *Mi chiedo che ritratto di sè tu pensi che Gianni ti darà*
I wonder which portrait of himself you think that John shall give you
b. **Mi chiedo che ritratto di sè tu sappia a chi Gianni darà*
I wonder which portrait of himself you know to whom John shall give
- (58) a. *Gianni si chiede che ritratto di sè tu pensi che ti darò*
John wonders which portrait of himself you think that I shall give to you
b. *Gianni si chiede che ritratto di sè tu sappia a chi darò*
John wonders which portrait of himself you know to whom I shall give

In (58), reflexivization forces pied-piping, whereby *Gianni* and *sè* appear in a local configuration; the wh-island therefore does not create a violation. In (57), by contrast, reflexivization forces reconstruction, whereby *Gianni* and *sè* are again in a local configuration; hence the wh-island does create a violation. In the absence of wh-islands, on the other hand, both construals are equally possible.

A cross-check of the correlation between reconstruction and sensitivity to weak islands can be obtained by looking at bound pronouns. If the correlation holds, we predict that a pronoun can be reconstructed inside the scope of a quantifier in the absence of weak islands, as in (59), but not across it, as in (60). The judgments in (59)–(60) seem once again consistent with the predictions; what is more, the reading of (60) in which the quantifier does not bind the pronoun is perfectly acceptable, again as expected:

- (59) *Mi chiedo quale dei suoi amici tu pensi che ciascuna ragazza sceglierà*
I wonder which of her friends you think each girl will choose
- (60) **Mi chiedo a quale dei suoi amici tu sappia che cosa ciascuna ragazza darà*
I wonder to which of her friends you know what each girl will give

On the basis of the evidence that precedes we can then tentatively conclude that there is a correspondence between pied-piping, or lack of reconstruction, and insensitivity to weak islands, as well as a correspondence between reconstruction, or lack of pied-piping, and sensitivity to weak islands. Obenauer (1994) provides further support for the same conclusion through a detailed investigation of French (L. Rizzi, personal communication). As illustrated in (3)–(4), adjuncts, exactly like arguments, are not uniformly sensitive to wh-islands. Again, in order to predict the absence of a weak island violation in (4), we only need to assume that its abstract representation involves no reconstruction of the pied-piped wh-phrase. In

the case of apparently bare *wh*-phrases, such as *what* or *who*, which tendentially show no sensitivity to weak islands, we can assume that a lexical restrictor is present in the form of the feature $\pm$ human. This analysis also predicts the contrast illustrated in (5)–(6) between Italian *che* and *che cosa*, which both correspond to *what*, except that *cosa* overtly acts as a restrictor to *che*.

At this point, all that remains is connecting the generalization I have arrived at concerning the two *wh*-extraction classes with the theory of weak islands provided by Minimality. The fundamental parallelism just argued for between lack of pied-piping and reconstruction actually complicates the problem, rather than simplifying it. For Minimality is conceived of as applying in the course of the derivation, but its interaction with reconstruction seems to require that it applies to representations obtained under Chomsky's (1995) model of grammar only after the Spell-Out point. A solution that immediately suggests itself is that a representational, rather than a derivational model of Minimality, and perhaps of grammar in general, is called for, in the manner of Rizzi (1990) or Brody (1995). I shall show, however, that a derivational solution is equally possible.

The fundamental insight of Obenauer (1984) and Cinque (1990) is that the empty category associated with a pied-piped nonreconstructed *wh*-phrase is a pronoun. Thus we can assume that the *wh*-phrase is inserted directly in [Spec, CP], while a zero D correlate is independently merged in θ -/Case position, as in (61):

(61) what [do-Q] [you wonder [how to [repair D] t_{how}]]

The fact that the *wh*-phrase does not move at all in the derivation in (61) obviously explains why it is not sensitive to weak islands. On the other hand, the equally obvious problem with such a derivation is that, though not sensitive to weak islands, pied-piped *wh*-phrases are sensitive to other diagnostics for movement, notably strong islands. In order to account for strong island effects, it is therefore necessary to assume that D covertly moves to the checking domain of *wh*-, as in (62). Indeed, we independently know that such a relation between *wh*- and D is necessary in order for the sentence to be interpretable at all:

(62) what [D-[do-Q]] [you wonder [how to [repair t_D] t_{how}]]

As for what triggers the movement in (62), one possibility is that the *wh*-phrase, like any QP, must be closed off by a D. Since the movement of D to the domain of *wh*- is not triggered by a Q feature, the intervention of an attractor of the same class (i.e., Q, Neg, or F) on the path of movement is irrelevant for Minimality; once again, therefore, weak island effects are not observed. In short, when *wh*-phrases move, they necessarily give rise to weak island effects. When they do not give rise to weak island effects, it is because the *wh*-phrase is merged *in situ* and abstract movement connects it with its θ -/Case position; since this movement is not triggered by Q, it is insensitive to the intervention of attractors of the F-class, hence to weak islands.

We are now in a position to consider why the nonmovement strategy in (61)–(62) is possible only with wh-phrases which include a lexical restriction. Chomsky's (1995) theory assigns to pied-piped nonreconstructed (52) the interpretation in (63):

(63) [*Quanti* x [*x posti*]] . . . x

Following Chomsky (1995), we assume that the head of the wh-phrase is an operator; raising it within its own phrase, as in (52), gives rise to a structure where it binds a variable among the arguments of N. If the nonoperator part of the wh-phrase is not reconstructed, we can assume that for each value assigned to the x variable, the wh-phrase itself takes on the properties of a referential expression. Thus its link with the θ -/Case position within VP is a bound pronoun one. This interpretation is of course supported by the pronoun-like D in (61)–(62).

Consider by contrast a bare wh-phrase. The wh-operator cannot bind a variable within the wh-phrase; therefore, its link with the θ -/Case position must correspond to a true operator–variable one. Suppose we make the natural assumption that such an interpretation cannot be supported by a construal of the wh-dependency as a bound pronoun one. If so, the absence of a lexical restriction implies a true movement derivation, hence, indirectly, sensitivity to weak islands. Similarly, it is easy to see that reconstruction can only apply on the basis of a true movement derivation. Thus if a wh-phrase with a lexical restriction is inserted in the position where it surfaces, as it can, weak islands are circumvented, but reconstruction is impossible. On the other hand, if the same wh-phrase undergoes true movement, it is possible to reconstruct it, but the derivation will display sensitivity to weak islands.

Finally, I noticed in the initial discussion that it is not only pied-piped and/or reconstructed wh-phrases, or Focus phrases, that display sensitivity to weak islands, but in general, all indefinites that undergo covert movement, such as the Italian polarity item *nessuno*. Again our analysis provides us with an account for the behavior of these elements as a single class. Quite simply, covert movement cannot help but correspond to true movement, albeit of features. In particular, since polarity items move under the attraction of an F-class operator, their movement will be blocked by closer F-class attractors. This yields the desired weak island effects on interpretation illustrated in Italian by the contrast between (25) and (27).

REFERENCES

- Aoun, J. (1985) *A Grammar of anaphora*. The MIT Press, Cambridge, MA.
 Aoun, J., and A. Li (1993). *Syntax of scope*. The MIT Press, Cambridge, MA.

- Baker, M., and K. Hale (1990). Relativized Minimality and pronoun incorporation. *Linguistic Inquiry* 21, 289–297.
- Brody, M. (1995). *Lexico-Logical Form*. The MIT Press, Cambridge, MA.
- Chomsky, N. (1964). *Current Issues in Linguistic Theory*. Mouton, The Hague.
- Chomsky, N. (1973). Conditions on Transformations, in S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*. Holt, Rinehart and Winston, New York.
- Chomsky, N. (1981). *Lectures on Government and Binding*, Foris, Dordrecht.
- Chomsky, N. (1986). *Barriers*. The MIT Press, Cambridge, MA.
- Chomsky, N. (1995). *The Minimalist Program*. The MIT Press, Cambridge, MA.
- Cinque, G. (1990). *Types of A'-dependencies*. The MIT Press, Cambridge, MA.
- Culicover, P. W., and W. Wilkins (1984). *Locality in Linguistic Theory*. Academic Press, New York.
- Haegeman, L., and R. Zanuttini (1991). Negative heads and the NEG Criterion. *The Linguistic Review* 8, 233–251.
- Hegarty, M. (1992). *Familiar complements and their complementizers*. Unpublished manuscript, University of Pennsylvania, Philadelphia.
- Huang, J. (1982). *Logical relations in Chinese and the theory of grammar*. Ph.D. dissertation. MIT, Cambridge, MA.
- Kiss, K. 1993. Wh-movement and specificity. *Natural language and Linguistic Theory* 11, 85–120.
- Koopman, H., and D. Sportiche (1982). Variables and the Bijection Principle. *The Linguistic Review* 2, 185–213.
- Kroch, A. (1989). Amount quantification, referentiality and long wh-movement. Unpublished manuscript, University of Pennsylvania, Philadelphia.
- Manzini, M. R. (1990). Locality theory: Towards a unification, W. Chao and G. Horrocks (eds.), *Levels of representation*. De Gruyter, Berlin.
- Manzini, M. R. (1992). *Locality: A theory and some of its empirical consequences*. The MIT Press, Cambridge, MA.
- Manzini, M. R. (1993a). Syntactic dependencies and their properties: Weak islands, in A. Alexiadou and T. A. Hall (eds.), *Studies on Universal Grammar and typological variation*. John Benjamins, Amsterdam.
- Manzini, M. R. (1993b). Form Dependency and the theory of phrase structure. To appear in *Proceedings of the Tilburg Conference on Rightward movement*.
- Manzini, M. R. (1996). Il congiuntivo. In P. Benincà et al. (eds.), *Italiano e dialetti nel tempo*, Bulzoni, Roma.
- Melvold, J. (1991). Factivity and definiteness. *MIT Working Papers in Linguistics* 15, 97–117.
- Obenauer, H. (1984). On the identification of empty categories. *The Linguistic Review* 4.2.
- Obenauer, H. (1994). *Aspects de la syntaxe A-barre* Thèse de Doctorat d'État. Université de Paris VIII.
- Pesetsky, D. (1987). Wh-in-situ: Movement and unselective binding. In E. Reuland and A. ter Meulen (eds.), *The representation of (in)definiteness*. The MIT Press, Cambridge, MA.
- Rivero, M.-L. (1994). Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory*, 12, 63–120.

- Rizzi, L. (1982). *Issues in Italian syntax*. Dordrecht.
- Rizzi, L. (1990). *Relativized minimality*. The MIT Press, Cambridge, MA.
- Rizzi, L. (1991). Residual verb second and the wh-criterion. *University of Geneva Technical Reports*.
- Roberts, I. (1994). Two types of head movement in Romance, in N. Hornstein and D. Lightfoot (eds.), *Verb movement*. Cambridge University Press, Cambridge.
- Ross, J. R. (1984). Inner islands. *Berkeley Linguistic Society* 10, 258–265.
- Roussou, A. (1994). *The syntax of complementizers*. Ph. D. dissertation, University College London.
- Szabolcsi, A., and F. Zwarts (1990). Semantic properties of composed functions and the distribution of wh-phrases. *Proceedings of the 7th Amsterdam Colloquium*, ITLI.
- Travis, L. (1984). *Parameters and effects of word order variation*. Ph.D. dissertation, MIT, Cambridge, MA.

