

On an Alleged Argument for the Proper Binding Condition*

Klaus Abels
University of Connecticut

This paper investigates a paradigm of questions with raising adjectives: ✓*How likely to win is John?* vs. **How likely to be a riot is there?* It has been claimed that the paradigm argues for Fiengo's (1977) Proper Binding Condition. This paper shows that the argument cannot be maintained once the paradigm is extended to sentences like ✓*How likely is there to be a riot?* An explanation in terms of Relativized Minimality is offered.

1. Introduction

In this paper I investigate English sentences containing so called raising adjectives such as *likely* and *certain*. I focus on an asymmetry between the construction in example (1) with a lexical subject and that in example (2) with an expletive subject—the basic paradigm. The asymmetry shows up under question formation as in (3) and (4)—the asymmetric paradigm.

- (1) ✓ John is likely/certain to win the race.
- (2) ✓ There is likely/certain to be a riot.
- (3) ✓ How likely/certain to win the race is John?
- (4) * How likely/certain to be a riot is there?

The asymmetric paradigm has received some attention in the literature (Barss 1986; Boeckx to appear; Epstein 2001; Kroch and Joshi 1985; Lasnik and Saito 1992; Lasnik in press; Nomura 2001). There is another way of forming *how likely*-questions, discussed in detail only in Boeckx (to appear) and Nomura (2001), as in (5) and (6)—the symmetric paradigm.

- (5) ✓ How likely/certain is John to win the race?
- (6) ✓ How likely/certain is there to be a riot?

Classic accounts of the asymmetric paradigm rest basically on Fiengo's (1974, 1977) formulation of the Proper Binding Condition (PBC). The PBC ensures that every trace has a c-commanding antecedent at surface structure.¹

* Thanks are due Cedric Boeckx, Željko Bošković, Luisa Martí, Masashi Nomura, the audience at HUMIT, and the editors of this volume for comments, patience and support. Special thanks go to Howard Lasnik. All errors and shortcomings are to be blamed entirely on me.

¹ Fiengo uses precedence and command. For the replacement by c-command cf. Reinhart (1976).

- (7) Proper Binding Condition (PBC) (Fiengo 1977:45 #33)
 In surface structure S_{α} , if $[e]_{NP_n}$ is not properly bound by $[\dots]_{NP_n}$, then S_{α} is not grammatical.

The PBC was designed to rule out, among other things, downward movement, for example in passives and *wh*-questions (8). (8) is ruled out by the PBC, because the trace of the matrix subject is not c-commanded by its antecedent *who* at surface structure.

- (8) t_{who} wonders who John ate lunch.

The PBC-account of the asymmetric paradigm postulates that *likely* is lexically ambiguous between a raising and a control adjective. Accordingly, example (1), where the matrix subject is a true argument, is treated as ambiguous (9), but example (2), where the subject is expletive *there*, is unambiguous. Example (2) only has a raising analysis (10). The control structure is unavailable since expletives do not control PRO (Chomsky 1981:328).

- (9) John_i is likely ✓PRO_i/ ✓ t_{John} to win the race.
 (10) There is likely *PRO/ ✓ t_{there} to be a riot.

The asymmetric paradigm can be accounted for under these assumptions coupled with the PBC. On the raising analysis, (3) and (4) are both out by the PBC since the subject does not c-command its trace, as (11) and (12) show.

- (11) [How likely t_{John} to win] is John $t_{\text{how likely ...}}$? * by PBC
 (12) [How likely t_{there} to be a riot] is there $t_{\text{how likely ...}}$? * by PBC

However, a control structure is available for (3) but not for (4) ((13) and (14)).

- (13) [How likely PRO to win] is John $t_{\text{how likely ...}}$? ✓
 (14) [How likely PRO to be a riot] is there $t_{\text{how likely ...}}$? * bc *there* cannot control

As pointed out in Lasnik and Saito (1992), this account makes the correct prediction that idiom chunks pattern with expletives in relevant respects (examples (15) and (16)). Unlike expletives, idiom chunks are quasi-arguments, which can in principle control PRO as (17) from Chomsky (1981:328) shows. (16) is thus not ruled out by the inability of idiom chunks to control PRO *per se*; rather, the control analysis of (16) requires *advantage* to receive a θ -role from *likely/certain*. (16) is thus out as a violation of θ -theory on a par with (18). (Notice that the symmetric paradigm also extends to idiom chunks (19).)

- (15) ✓ Advantage was likely/certain to be taken of John. (cf. (2))
 (16) * How likely/certain to be taken of John is advantage? (cf. (4))
 (17) ✓ Advantage was taken of John after PRO being taken of Sally.
 (18) * Advantage tried PRO to be taken of John.
 (19) ✓ How likely/certain is advantage to be taken of John? (cf. (6))

On an Alleged Argument for the Proper Binding Condition

Structures (11) and (12), which violate the PBC, are examples of remnant movement: an XP containing a trace is moved to a position where the trace contained in XP is no longer c-commanded by its antecedent (20).

(20) $[[_{XP} \dots t_{\alpha} \dots] \dots [\alpha [\dots t_{XP} \dots]]]$

In recent syntactic literature, remnant movement is often argued for or assumed to exist (e.g. Kayne 1998; Koopman and Szabolcsi 2000; Müller 1998 etc.) while the PBC is often argued to be wrong. To many, the PBC is also suspect since it is a representational condition. These researchers replace Fiengo's PBC by a derivational version which forces every step of movement to target a c-commanding position, without imposing a global output constraint (e.g. Epstein, Groat, Kawashima and Kitahara 1998: 100-106; Kitahara 1997: 70). The derivational formulation of the PBC still rules out downward movement as in (8) but allows remnant movement. I reserve the term 'PBC' in the remainder of this paper for Fiengo's version of the PBC (7) and argue against it. This paper is fully compatible with the derivational version.

On the assumption that remnant movement is allowed in principle, the account of the asymmetric paradigm in terms of the PBC is lost.² Since the asymmetric paradigm was taken by some to furnish a strong argument for the PBC (Kroch and Joshi 1985; Lasnik and Saito 1992), two questions arise: How strong is the argument for the PBC from these sentences? How can the asymmetric paradigm be accounted for in a framework without the PBC? In this paper I concentrate on the first question. First I review three arguments against the PBC-account of the asymmetric paradigm. Then I investigate how the symmetric paradigm fares with respect to the PBC and show that the symmetric paradigm cannot be reconciled with it. Finally I suggest a solution to the problem in terms of Relativized Minimality.

2. Problems for the PBC-Account from the Literature

The first problem (cf. Huang 1993) concerns the assumed ambiguity of example (1). Some predicates, like *certain*, are ambiguous between assigning a θ -role to their subject (21) and failing to do so (22).

- (21) John is certain that he will win the race.
(22) It is certain that John will win the race.

Somewhat surprisingly, this ambiguity is lost in example (1). On the control reading of (1) *John* should be theta marked by *certain* and coindexed with PRO, which gets its θ -role from *win*. This should give (1) a reading

² Barss (1986) and Lasnik (in press) follow a different path. For them remnant movement is allowed in principle, but sideward movement is not. The asymmetry then follows if expletives must be replaced at LF (Chomsky 1986 – Barss's route) or the features of the associate of the expletive must raise covertly (Chomsky 1995 – Lasnik's route). Idiom chunks are out in the asymmetric paradigm on the assumption that idioms have to be 'reassembled' at LF. As discussed in Boeckx (to appear), these theories cannot account for examples like (25) and (ft. 4 i).

synonymous with (21). Of course, the θ -role assigned by *certain* with an infinitival complement could be different from that in (21).

However, none of the authors who assume that *certain* and *likely* are ambiguous state what the second θ -role conferred to *John* by *likely* and *certain* is, or distinguish the two readings semantically. Example (1) fails to be perceived as ambiguous. This casts doubt on the independent motivation for the ambiguity analysis. Note that the assumption of two θ -roles was crucial in extending the PBC-account to idiom chunks (16).

The second problem is raised by Epstein (2001: 26), for whom (23) is acceptable.

- (23) How likely to appear/seem that Bush cheated, is it?/will it be?/become?

He takes this to be an immediate argument against the PBC-account of the asymmetric paradigm. He assumes that expletive *it* can never act as a controller. This assumption is not quite true, though. There are speakers who accept sentences in which expletive *it* controls PRO (24).³

- (24) It appeared that Bush was a liar after PRO seeming that he had known the truth all along.

The PBC-account now predicts a correlation between a speaker's acceptance of (23) and their acceptance of (24). This is so since (23) not only requires *it* to control PRO, but also to receive a θ -role from *likely* whereas (24) only requires control. According to the PBC-account all speakers who accept (23) should also accept (24). However, no such correlation holds. Crucially, there are speakers who accept (23) and yet reject (24). This indicates that the grammaticality or ungrammaticality of these sentences have different sources. Crucially, the status of (23) is independent of the properties of control.

The third problem for the PBC-account comes from examples like (25). (These examples are discussed in Boeckx (to appear) and attributed there to Koji Sugisaki.) The judgments for (25) and (27) are contrastive. Many speakers do not find (27) fully acceptable, but all find a clear contrast between the examples in the direction indicated.⁴

³ An anonymous reviewer suggests that (i) is more or less grammatical and suggests on this basis that *Bush* might be controlling PRO in (24). This seems unlikely, though, since (i) is rejected by other informants who accept (24). The fact that (ii) is clearly ungrammatical speaks against the suggestion that *Bush* could be controlling the subject position of *seem*. Moreover, the *after*-clause in (24) is attached in the matrix clause, where it is not c-commanded by *Bush*, thus *Bush* is not even a potential controller. For the reviewer, the structure in (iii) might be interfering, where *seem like* is interpreted as *make it seem as though* – an interpretation that (24) clearly lacks.

- (i) Bush tried to seem that he had known the truth all along.
 (ii) * Bush seems that he had known the truth all along.
 (iii) ✓ Bush tried to seem like he had known the truth all along.

⁴ Notice that idiom chunks again pattern with expletives (example from Boeckx, to appear: #18).

- (i) *Who said that advantage was how likely to be taken of John? cf. (25)

On an Alleged Argument for the Proper Binding Condition

- (25) * Who said that there is how likely to be a riot?
- (26) ✓ Who said that John is how likely to win the race?⁵
- (27) ✓ Who said that it is how likely that there will be a riot?

Examples (26) and (27) show that *how likely* can remain in situ in a multiple question. In this respect *how likely* differs from *wh*-adjuncts like *why*, which cannot remain in situ. *How likely* is then not an adjunct.

Example (25) is just as bad as (4). It is reasonable to assume that both examples are ill-formed for the same reason. The PBC was originally formulated as a surface structure condition. However, since (25) does not involve overt *wh*-movement of *how likely to be a riot*, the PBC cannot be blamed for the ungrammaticality of the example.

The three problems for the PBC-account of the asymmetric paradigm I have reviewed are: (a) The proponents of the PBC-account have to assume that *likely* is ambiguous between a control and a raising predicate. No evidence for this assumption has been offered. (b) There are examples where non-controllers are acceptable as subjects of questions in the asymmetric paradigm. This is unexpected under the PBC-account. (c) The PBC-account incorrectly predicts that the asymmetry between arguments and expletives/idiom-chunks should disappear when *how likely/certain...* is not moved overtly.

3. A Problem of Undergeneration: The Symmetric Paradigm

In the previous section I reviewed three arguments that cast doubt on the PBC-account of the asymmetric paradigm. I now turn to the symmetric paradigm. How can examples (5) and (6), repeated here for convenience, be derived?

- (5) ✓ How likely is John to win the race?
- (6) ✓ How likely is there to be a riot?

Obviously *how likely* can either form a constituent with the infinitival as in the examples of the asymmetric paradigm, or it can be separated from the infinitival. We ask: Do the two constituency options reflect different underlying structures? Or are they derived from a single source? What is/are the underlying structure(s)?

Examples (28)-(34) immediately suggest one possible structure.

- (28) ✓ a likely scenario
- (29) ✓ a proposal likely to be accepted⁶
- (30) ✓ John became likely to win the race.
- (31) ✓ It became likely that John would win the race.

⁵ If covert *wh*-movement is full phrasal movement, example (25) can be handled by a stronger version of the PBC which makes it applicable at all levels of representation. I will not pursue this line here since I argue that even Fiengo's weaker version is too strong.

⁶ The different order between adjective and modified noun in the two examples follows the general pattern for English. Non-complement taking adjectives generally precede the head noun, but complement taking adjectives follow it: *a proud man* vs. *a man proud of his children*.

On an Alleged Argument for the Proper Binding Condition

I call the structure presented in (38) the *0-place hypothesis*. The reason for this terminological choice is the fact that *likely* is treated syntactically a non-complement-taking (0-place) adjective in (38) but a complement-taking one in (35).

We have seen that if structure (35) is the only available underlying structure, then the PBC-account faces an undergeneration problem, since the examples of the symmetric paradigm cannot be generated without violating the PBC. However, if structure (38) is also a possible underlying structure, then the PBC-account might be upheld for the asymmetric paradigm (pace section 2.) The following two subsections show that the 0-place hypothesis is wrong. This entails then that the PBC cannot be maintained even for these cases.

3.1 The categorial status of *how likely*

In order to keep the number of possibilities under consideration manageable, I first determine the categorial status of *likely/certain* in examples (1) through (6). In other words, I determine the category of YP in (38). The question is simply whether *likely* is an adverb or an adjective. This becomes an issue because in addition to its adjectival use ((28) through (31)), *likely* also has an adverbial use as example (40) shows.⁸ I argue that *likely* and *certain* are always adjectives in (1) through (6), though.

(40) ✓ There will likely be a riot.

Note first that the homonymy between the adjectival and the adverbial forms of *likely* is coincidental. *Certain* has two distinct morphological forms.

In examples with *certain* any ambiguity between the adjectival and the adverbial form disappears. In adjectival positions only the clearly adjectival form is licit as examples (41) and (42) indicate. In adverbial positions only the morphologically adverbial form is licit as (43) and (44) make clear. This is important, because it shows that *certain* is unambiguously adjectival.

- (41) ✓ He went down the path of certain failure.
(42) * He went down the path of certainly failure.
(43) ✓ John will certainly win the race.
(44) * John will certain win the race.

Recall that *certain* shares all the relevant properties of *likely*. In particular, *certain* gives rise to the asymmetric and the symmetric paradigms. According to the conclusion just reached, *certain* in (1) through (6) is unambiguously adjectival. *Likely* should be treated as an adjective on a par with *certain*.

The conclusion that these examples only tolerate adjectival predicates is corroborated by the ungrammaticality of (45) through (48), which involve the unambiguously adverbial *certainly*.

⁸ The homonymy between the adjectival and adverbial forms is probably not accidental, as a reviewer points out. It might be attributed to a phonological condition ruling out the putative regular form *likelily*.

- (45) * John is certainly to win the race.⁹ (cf. (1))
 (46) * There is certainly to be a riot. (cf. (2))
 (47) * How certainly is John to win the race? (cf. (5))
 (48) * How certainly is there to be a riot? (cf. (6))

The same conclusion is further supported by noticing that clearly adverbial *likely* does not tolerate being modified by *how* (49). For completeness the same fact is also illustrated for *certain* (50).

- (49) * How likely will John win the race?
 (50) * How certain(-ly) will John win the race?

I have argued that *likely* and *certain* are unambiguously adjectival in (1) through (6). The category of YP in (38) is adjective phrase (AP in (38')).

(38') [_{XP} [_{YP=AP} how likely] ... [_{IP/CP} ...]]

Boeckx (to appear) assumes that *likely* can be an adjunct. Since there is no nominal projection to which adjectival *likely* could adjoin, this entails that *likely* be treated as an adverb. The argument just presented thus immediately rules out Boeckx's version of the 0-place hypothesis.

3.2 Problems with the 0-place hypothesis

Recall that the 0-place hypothesis (38) was required to save the PBC-account of the asymmetric paradigm. The symmetric paradigm can only be distinguished from the asymmetric paradigm in terms of the PBC if (38) is a possible underlying structure. I consider two instantiations of structure (38). In the first, XP in (38) is identified with a small clause. The second structure I consider is one where AP and IP/CP are taken to be independent arguments of *be* – basically an analysis of *be* as a ditransitive. Neither of these hypotheses describes the properties of the construction adequately.

If *likely to win the race* could be analyzed as a small clause, we would expect it to be distributed like a small clause. This most basic prediction of the small clause analysis is not borne out. Example (51) shows that small clauses cannot be clefted. Examples (52) and (53) indicate that *likely to win the race* patterns with adjectives in being able to cleft.

- (51) * Bill a fool is what John considers.
 (52) ✓ Likely to win the race is what John considers Bill to be.
 (53) ✓ Slow is what John considers Bill to be.

This is of course not surprising, given that we have already established that *likely* can form an AP with the embedded clause (35). But if the small clause analysis were correct, it would be surprising if *likely* could not oc-

⁹ The example is bad as a synonym of *John is certain to win the race*. The example marginally has the adverbial construal of *certainly*, where it is synonymous with *Certainly, John is to win the race*. The same applies to example (46).

On an Alleged Argument for the Proper Binding Condition

cur in prototypical small clause environments – which indeed it cannot. The verb *consider* is a good test case. It allows small clause complements (54) even when headed by an adjective (55), and it disallows simple adjectival complements (56). Crucially *likely to...* is out under *consider*.

- (54) ✓ I consider [_{sc} John (to be) a fool].
- (55) ✓ I consider [_{sc} John (to be) smart].
- (56) * I consider sad.
- (57) * I consider likely (for John) to win the race.
- (58) * I consider likely for there to be a riot.

The next set of examples makes the same point. (59) and (60) show that while small clauses can function as the subject of *be*, the phrase *likely to...* cannot.

- (59) ✓ [Mandela free] would pose a bigger threat than [Mandela behind bars].
(Aarts 1992, cf. also Safir 1983; Stowell 1983, 1991)
- (60) * Likely (for Ian) to win would be worse than likely (for him) to lose.

The examples showing that the small clause analysis is inadequate could be increased ad libitum. I conclude from this brief discussion that the small clause analysis is wrong.

The second option to consider is that *be* has a ditransitive construal: *Be* would take a 0-place adjective as its first complement and a clause as its second argument. Just like the small clause analysis, this analysis runs into grave problems. If the position where *likely* goes were an argument position of *be*, then other (semantically similar and by hypothesis 0-place) adjectives should go into the same slot. Examples (61) through (64) illustrate that this expectation is not borne out.

- (61) * There is possible to be a riot.
- (62) * John is possible to win the race.
- (63) * There is proud to have been a riot.
- (64) ✓ John is proud to have won the race.

If there is no such ditransitive *be* and (35) is the only available structure, then the ungrammaticality of (61) through (63) is immediately explained.

Possible takes CP complements with arbitrary PRO in subject position (65) but does not tolerate raising infinitivals ((66) and (67)). This selectional restriction against a raising complement rules out (61) and (62).

- (65) ✓ It seems possible to win the race.
- (66) * John seems possible to win the race.
- (67) * There seems possible to be a riot.

Similarly for *proud*. The PRO subject of the complement of *proud* is controlled by the subject of *proud*. Raising is not allowed with *proud* ((68) vs. (69)). Hence the ungrammaticality of the expletive subject in (63).

- (68) ✓ John seems t_{John} proud PRO to be here.
 (69) * There seems proud to be a riot.

The restrictive hypothesis according to which only structure (35) is available explains (61) through (64) immediately. On the alternative hypothesis where *be* is taken to be ditransitive, the selectional restrictions of the adjective should not play a role since the adjectives do not take a complement at all. On this hypothesis *be* must allow raising infinitivals as its second complement (cf. (2) and (6)). And it should do so with any (semantically suitable) adjective – counter to fact.

Neither of the two considered versions of the 0-place hypothesis (38) gives satisfactory results. The 0-place hypothesis suffers from severe overgeneration in both of its guises. Given the results of sections 3.1 and 3.2 it is hard to see what other form it might take. I then conclude this subsection with an argument from poverty of imagination: since no conceivable version of the 0-place hypothesis has the right properties, it must be wrong, and (38) is not an available structure.

I argued at the beginning of section 3 that under the complement-taking hypothesis (35), the sentences of the symmetric paradigm cannot be derived without violating the PBC (cf. (37)). Given that the complement-taking hypothesis is correct and exclusive, the symmetric paradigm is a case of PBC-violating remnant movement. Hence the PBC is wrong. The asymmetric paradigm, which used to be taken as one of the strongest arguments for the PBC, turns out not to furnish an argument at all once the full paradigm is looked at. This confirms the hunch we had at the end of section 2.

4. Towards an Account

In this section I explore a solution to the puzzles raised by the data investigated here. I follow the intuition pursued in Boeckx (to appear), who suggests an account for the facts in terms of Relativized Minimality (Rizzi 1990 and for illuminating recent discussion Starke 2001). I will repeat here the core data that an account has to cover.

- | | | | |
|----------|---|---|---------------------|
| (1) | ✓ | John is likely to win the race. | Basic paradigm |
| (2) | ✓ | There is likely to be a riot. | |
| (15) | ✓ | Advantage is likely to be taken of John. | |
| (3) | ✓ | How likely to win the race is John? | Asymmetric paradigm |
| (4) | * | How likely to be a riot is there? | |
| (16) | * | How likely to be taken of John is advantage? | |
| (26) | ✓ | Who said that John was how likely to win? | |
| (25) | * | Who said that there is how likely to be a riot? | |
| (ft. 4i) | * | Who said that advantage was how likely to be taken of John? | |
| (5) | ✓ | How likely is John to win the race? | Symmetric paradigm |
| (6) | ✓ | How likely is there to be a riot? | |
| (19) | ✓ | How likely is advantage to be taken of John? | |

On an Alleged Argument for the Proper Binding Condition

The examples of the basic paradigm are unproblematic. The embedded subject can raise out of its clause in the usual fashion. The asymmetric paradigm indicates that expletive *there* and idiom chunks cannot move past *how*, but regular argument DPs can. As argued in Boeckx (to appear), this kind of pattern is reminiscent of weak islands. Weak islands block the extraction of certain items but not others. This suggests a treatment in terms of Relativized Minimality.

In fact the class of elements that selectively block extraction from the complement of *likely* extends beyond *how*. Many degree modifiers give rise to the same pattern in more or less strong form. This is illustrated in (70).

Unfortunately, there is a considerable degree of speaker variation and squishiness involved in (70). Additional research is clearly needed. Nonetheless, (70) is suggestive of the broader generalization (71).

- (70) Good
- | | | |
|----|---|---|
| a. | ↓ | There is likely to be a riot. |
| b. | ↓ | There is very likely to be a riot. |
| c. | ↓ | There is somewhat likely to be a riot. |
| d. | ↓ | There is extremely likely to be a riot. |
| e. | ▼ | Who said that there is how likely to be a riot. |
- Bad

- (71) A non-argument cannot pass a degree-modifier.¹⁰

Following Starke's (2001) discussion of Relativized Minimality, we can summarize it in two slogans: An item in class α cannot cross another item in the same class (72); an item in class α can cross items in other classes β (73).

- (72) * α_2 α_1 α_2
 (73) ✓ α_2 β_1 α_2

Boeckx (to appear) argues that *how* qua *wh*-word is a non-specific indefinite. He further argues that the relevant dependency between the matrix subject position and the embedded subject position in the good examples of the asymmetric paradigm involves specific or definite DPs ((3) and (26)), whereas that dependency involves non-specific, non-referential, indefinites in the bad examples ((4), (16), (25)).¹¹ The asymmetric paradigm then immediately reduces to (72) in the bad and to (73) in the good cases.

Although I fully agree with the logic of Boeckx's solution for the asymmetric paradigm and although I cannot develop a classification of elements to replace Boeckx's with (but cf. ft. 12), I would like to point out a problem that appears to be fairly grave. Boeckx assumes that all *wh*-elements are non-specific indefinites. This should preclude any *wh*-word from escaping

¹⁰ I am skirting the fascinating question of what happens in different kinds of comparatives and in superlatives here. I hope to be able to take up that question in future work.

¹¹ Technically Boeckx does not assume that expletive *there* raises. The relevant relation for him is actually the agreement relation between the matrix verb and the non-specific, non-referential, indefinite associate of the expletive. This detail is irrelevant for our discussion, however.

out of the sentential complement of *how likely*. This prediction, however, is not borne out as (74) indicates.

(74) ✓ Who is how likely to win the race?

We are then left with generalization (71).¹² I assume that it reduces to Relativized Minimality along the lines suggested in Boeckx (to appear) and ft. (12).

We still have not explained why expletives and idiom chunks become good in the examples of the symmetric paradigm. Suppose that a derivation for the examples of the symmetric paradigm could be provided in which the idiom chunk and the expletive do not cross the degree modifier. The first idea that comes to mind is to extrapose the embedded infinitival prior to raising.

(75) Derivation

- a. [how [likely [there to be a riot]]]
- b. [[there to be a riot] ... [how [likely t_{ip}]]]
- c. [is [[there to be a riot] ... [how [likely t_{ip}]]]]
- d. [there [is [[t_{there} to be a riot] ... [how [likely t_{ip}]]]]]
- e. [[how [likely t_{ip}]] [C° [there [is [[t_{there} to be a riot] ... $t_{how\ likely\ t}$]]]]]

The crucial step is from (c) to (d). Extracting *there* directly from under *how* would violate Relativized Minimality. The prior extraposition ((a) to (b)) removes *there* from the c-command domain of *how*, making raising from the embedded to the matrix subject position possible. Note again that it is irrelevant here whether the extraposing movement is leftward or rightward.

Nomura (2001), commenting on an earlier version of this paper, discusses (75) and argues that it is implausible on two counts: (i) Why doesn't the extraposition step freeze the extraposed element turning it into an island? (ii) Why is extraposition possible just in case *wh*-movement of *how likely t* happens later on? I.e. why couldn't example (27) involve string vacuous extraposition and become grammatical? Alternatively, assuming leftward extraposition, what rules out this version of (27): **Who said that there is to be a riot how likely?*

The first problem does not seem particularly grave. The examples usually used to illustrate that extraposition creates islands are CPs and PPs respectively. Both CP and PP are cyclic domains (see Riemsdijk 1978 for PPs). Although Chomsky (1999, 2000) considers only vP and CP to be *phases*, we might extend this treatment to PP as well. Cyclic domains require movement to the edge (specifier, Comp) to allow extraction, but non-cyclic domains do not.

¹² Generalization (71) might be made to follow from Starke's (2001) theory of locality. If Starke is right, a *wh*-phrase constitutes an intervener for a particular kind of movement which Starke calls θ -movement. Starke's (2001: 7.7) discussion suggests that idiom chunks (and presumably expletives) can only undergo θ -movement. If this speculation is on the right track, then it follows that *how* blocks extraction of idiom chunks or expletives.

(74), (3), and (26) are allowed since definites and the *wh*-subject in (74) can move by other kinds of movement (A-movement and Starke's SQ-movement). The asymmetric paradigm and generalization (71) can be derived this way. Unfortunately, space constraints prevent me from introducing Starke's approach to locality in any detail here and to develop this suggestion.

The examples of the symmetric paradigm (76v) are instances of the allowed (interlacing) configuration (78), whereas the in-situ examples of the asymmetric paradigm are cases of the disallowed configuration (77)/(72).

The similarity between the *how likely* paradigm and Richards' cases is fairly abstract. It is the difference between interlacing, order preserving chain configurations and non-interlacing, crossing chain configurations. Richards examples involve movement of multiple elements to different specifier positions of the same head, whereas in the present case the expletive moves to SpecIP and *how likely t* moves to SpecCP. Despite this superficial difference, the abstract similarity is rather striking: order preservation is good, re-ordering is bad.

The remaining examples of the asymmetric paradigm involving movement ((3), (4), and (16)) have not yet been accounted for. (3) is of course trivial. The rest are equally straightforward. The relevant structures differ minimally from those of (76v). They are given below as (76v')

(76v') [[[[how[likely t_{ip}]]X^o...[[there to be a riot]... $t_{how\ likely\ t}$]] [C^o[there[is t_{xp}]]]]

Crucially, the moved phrase is not *howP* but the bigger XP. The movement of XP cannot affect the relation between the expletive and *howP* in this case, since no new c-command relation between the two is established. These examples of the asymmetric paradigm are then examples of the banned configuration (77)/(72). They are out exactly on a par with the in-situ ones.

In this section I have suggested a way of deriving the complete paradigm without the PBC, but relying instead crucially on Relativized Minimality.

5. Conclusion

In this paper I have argued that the asymmetric paradigm, according to Lasnik and Saito (1992) one of the strongest cases for the PBC, should not and cannot be reduced to the PBC because of the symmetric paradigm. The argument for the PBC vanishes. This is a good result, since the PBC has fallen from grace anyway.

The result was obtained by showing (contra Boeckx to appear) that an unspoken assumption of most of the literature is correct and that *likely* is unambiguously a complement-taking adjective in the relevant contexts. This is crucial for the argument from the symmetric paradigm.

Finally I proposed a way to reduce the full paradigm to Relativized Minimality. A first attempt came up against two theoretical problems: it involved instance of optional movement (IP extraposition)—notoriously difficult to account for under minimalist assumptions. It also involved making this movement (IP extraposition) contingent on a later step of *wh*-moving *how likely t_{ip}* —this seemed to lead to a problem of computational look ahead. To avoid these problems, I postulated that extraposition is in fact obligatory. The nature of the driving force was left open. This treatment in turn suggested that the asymmetric paradigm instantiates order preservation. It is a particularly interesting case of order preservation, since it does not involve multiple specifiers of a single head, but specifiers of different heads (*there* is in SpecIP,

On an Alleged Argument for the Proper Binding Condition

howP in SpecCP). This indicates that Richards' (1997) way of deriving the order preserving generalization (via tucking in) is too narrow since *there* would have to tuck in non-locally to the lower head I° . A more general version along the lines of (77) seems to be required (cf. Starke 2001). (77), crucially, is a representational condition. It is hard to see how (77) could be recast in derivational terms.

Tucking in violates Chomsky's *extension condition* (cf. discussion in Richards 1997). Given the extension condition, there is hope that grammatical relations like c-command and conditions of the grammar (the derivational version of the PBC for example) might become reducible to cyclicity (cf. Epstein et al. 1998; Epstein 1999, 2001 for discussion). If acyclic operations exist, then this hope vanishes. Every non-tree extending operation is a problem from this perspective, whether it be acyclic head movement (Chomsky 1995), the non-tree extending operation of tucking-in, or acyclic merger of adjuncts (invoked in Boeckx's (to appear), account of the asymmetric paradigm). All violate the extension condition. All make the full derivation of grammatical relations and conditions of the grammar impossible. All must crucially invoke representations on which to define relations and conditions.

We seem to be left with two options: Either the grammar allows acyclic operations (non-local tucking in) and Relativized Minimality can be formulated as a derivational constraint on the operation attract, or Relativized Minimality is a condition on representations. If the conclusions about acyclic operations reached in the preceding paragraph are correct, then either option makes the grammar in effect representational.

We have come full circle: the representational PBC has been replaced by a representational version of Relativized Minimality.

References

- Aarts, Bas (1992). *Small Clauses in English. The Nonverbal Types*. Berlin: Mouton.
- Barss, Andrew (1986). Chains and Anaphoric Dependence: On Reconstruction and Its Implications. Doctoral dissertation. MIT. Distributed by MITWPL.
- Boeckx, Cedric (to appear). Agree or Attract? A Relativized Minimality Solution to a Proper Binding Condition Puzzle. *Natural Language and Linguistic Theory*.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (1986). *Knowledge of Language*. New York: Praeger.
- Chomsky, Noam (1995). Categories and Transformations. In, *The Minimalist Program*, Cambridge, MA: MIT Press, pp. 219-394.
- Chomsky, Noam (1999). Derivation by Phase. *MIT Occasional Papers in Linguistics* 18. Distributed by MITWPL.
- Chomsky, Noam (2000). Minimalist Inquiries: The Framework. In Roger Martin, David Michaels and Juan Uriagereka (eds.), *Step by Step: Essays on Minimalism in Honor of Howard Lasnik*, Cambridge, MA: MIT Press, pp. 89-155.
- Epstein, Samuel David, Erich M. Groat, Ruriko Kawashima and Hisatsugu Kitahara (1998). *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.
- Epstein, Samuel David (1999). Un-Principled Syntax and the Derivation of Syntactic Relations. In Samuel David Epstein and Norbert Hornstein (eds.), *Working Minimalism*, Cambridge, MA: MIT Press, pp. 317-345.

Abels

- Epstein, Samuel David (2001). Deriving the Proper Binding Condition. Paper presented at *TiLT*, Utrecht, NL.
- Fiengo, Robert (1974). Semantic Conditions on Surface Structure. Doctoral dissertation. MIT. Distributed by MITWPL.
- Fiengo, Robert (1977). On Trace Theory. *Linguistic Inquiry* 8: 35-62.
- Huang, C.-T. James (1982). Logical Relations in Chinese and the Theory of Grammar. Doctoral dissertation. MIT. Distributed by MITWPL.
- Huang, C.-T. James (1993). Reconstruction and the Structure of Vp: Some Theoretical Consequences. *Linguistic Inquiry* 24: 103-138.
- Kayne, Richard (1998). Overt Vs. Covert Movement. *Syntax* 1: 128-191.
- Kitahara, Hisatsugu (1997). *Elementary Operations and Optimal Derivations*. Cambridge, MA: MIT Press.
- Koopman, Hilda and Anna Szabolcsi (2000). *Verbal Complexes*. Cambridge, MA: MIT Press.
- Kroch, Anthony and Aravind Krishna Joshi (1985). The Linguistic Relevance of Tree Adjoining Grammar. Ms. Philadelphia, Pennsylvania
- Lasnik, Howard and Mamoru Saito (1992). *Move A: Conditions on Its Application and Output*. Cambridge, MA: MIT Press.
- Lasnik, Howard (in press). Feature Movement of Agreement at a Distance? In Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers and Hans Martin Gärtner (eds.), *Remnant Movement, F-Movement, and the T-Model*, Amsterdam: John Benjamins.
- Müller, Gereon (1998). *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Dordrecht: Kluwer Academic Publishers.
- Nomura, Masashi (2001). Extraposition or Scattered Deletion? Ms. Storrs, CT
- Reinhart, Tanya (1976). The Syntactic Domain of Anaphora. Doctoral dissertation. MIT. Distributed by MITWPL.
- Richards, Norvin W. III (1997). What Moves Where When in Which Language? Ph.D. dissertation. MIT. Distributed by MITWPL.
- Riemsdijk, Henk van (1978). *A Case Study in Syntactic Markedness*. Dordrecht: Foris.
- Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- Safir, Kenneth (1983). On Small Clauses as Constituents. *Linguistic Inquiry* 14: 730-735.
- Starke, Michal (2001). Move Dissolves into Merge: A Theory of Locality. Doctoral dissertation. University of Geneva.
- Stowell, Timothy (1983). Subjects across Categories. *The Linguistic Review* 2: 285-312.
- Stowell, Timothy (1991). Small Clause Restructuring. In Robert Freidin (ed.), *Principles and Parameters of Comparative Grammar*, Cambridge, MA: MIT Press, pp. 182-218.

Department of Linguistics
University of Connecticut U-1145
337 Mansfield Road
Storrs, CT 06269-1145
USA

klaus.abels@uconn.edu
<http://www.sp.uconn.edu/~kla98002>