REVIEW ARTICLE


Reviewed by Yoshiaki Kaneko, Ibaraki University*

1. INTRODUCTION. In this monograph, R(izzi) presents one of the most impressive extensions of government-binding (GB) (or, principles-and-parameters) theory. R starts with recharacterizing the notion of minimality, reformulates the Empty Category Principle (ECP), offers a restricted theory of indexing, and explores what consequences these proposals have for the theory of syntax.

In what follows, I will outline R's proposals and discuss some potential problems as well as possible extensions.

2. RELATIVIZED MINIMALITY.

2.1. The core intuition behind the notion of minimality is that a governor cannot govern inside the domain of another governor. In the configuration 1, for example, X cannot govern Y if Z is a closer potential governor for Y.

(1) ...X...Z...Y...

Chomsky 1986 formulates this intuitive idea in terms of the notion of minimality barrier.

(2) In the configuration...a...[γ...δ...β...], γ is a barrier for β if γ is the immediate projection of δ, a zero-level category distinct from β.1

An intervening head Z0 (=δ) triggers the minimality effect so that its immediate projection Z' (=γ) acts as a minimality barrier which blocks government of an element Y (=β) within Z' from an external governor X (=...
α). R calls an approach of this type Rigid Minimality, because a minimality barrier is assumed to block every kind of government relation.

R alternatively proposes Relativized Minimality, which is defined through the variable notion of ‘α-government,’ ranging over head government and antecedent government.

(3) Relativized Minimality: X α-governs Y if there is no Z such that
   (i) Z is a typical potential α-governor for Y,
   (ii) Z c-commands Y and does not c-command X.

The clause 3ii formalizes the notion of intervention hierarchically in terms of c-command.2

The two types of government used in R’s system are defined as follows:

(4) Head Government: X head-governs Y iff
   (i) X ∈ {A, N, P, V, Agr, T}
   (ii) X m-commands Y
   (iii) no barrier intervenes3
   (iv) Relativized Minimality is respected.

(5) Antecedent Government: X antecedent-governs Y iff
   (i) X and Y are coindexed
   (ii) X c-commands Y
   (iii) no barrier intervenes
   (iv) Relativized Minimality is respected.

The notion of typical potential α-governor is defined as follows:

(6) Z is a typical potential head governor for Y = Z is a head m-commanding Y.

(7) a. Z is a typical potential antecedent governor for Y, Y in an A-chain = Z is an A-specifier c-commanding Y.

2 R proposes another formulation of intervention in terms of m-command. He assumes that for functional heads the formulation in the text is used, while for lexical heads the formulation in terms of m-command is used. This means that specifiers of lexical XP’s are protected from external head government, while specifiers of functional XP’s are not. Notice that we could restate essentially the same distinction with the Chomsky’s notion of minimality barrier: the narrower definition for functional heads and the broader one for lexical heads. See note 1 above.

3 Following Cinque 1990, R assumes the following definition of barrier for government: XP is a barrier if it is not directly selected by an X° not distinct from [+V]. Here ‘selection’ means s-selection for lexical heads and c-selection for functional heads. Cinque proposes that C and I are not distinct from [+V]. If we follow this proposal, the XP’s selected by C° and I°, IP and VP, are not inherent barriers, and we can dispense with special devices assumed in Chomsky 1986 to void their barrierhood.
b. Z is a typical potential antecedent governor for Y, Y in an A'-chain = Z is an A'-specifier c-commanding Y.
c. Z is a typical potential antecedent governor for Y, Y in an X0-chain = Z is a head c-commanding Y.

The intuitive idea behind these definitions is that 'a typical potential α-governor for an element Y is a base-generated position that could bear the relevant kind of government relation to Y (p. 7).' As for antecedent government, three subcases are distinguished, depending on whether Y is a trace in an A-chain (NP-movement), in an A'-chain (wh-movement), or in an X0-chain (head-movement).

2.2. Let us compare the two formulations of minimality (i.e. Rigid and Relativized versions of minimality) in terms of the following configurations.

(8) a. \[ ... X_0^0 ... [z'... Z_0^0 ... Y ...] \] head-gov.
b. \[ ... X_i^0 ... [z'... Z_0^0 ... Y_i ...] \] ante.-gov.
c. \[ ... X_P(A') ... [z'... Z_P(A'-spec) ... Y_i ...] \] ante.-gov.
d. \[ ... X_P(A) ... [z'... Z_P(A-spec) ... Y_i ...] \] ante.-gov.
e. \[ ... X_P(A') ... [z'... Z_0^0 ... Y_i ...] \] ante.-gov.
f. \[ ... X_P(A) ... [z'... Z_0^0 ... Y_i ...] \] ante.-gov.

The predictions of the two theories of minimality are indicated in the following table as to whether the intervening Z(P) triggers the minimality effect on the government relation involved in each case of 8a–f.

(9) Relativized: yes yes yes yes no no
Rigid: yes yes no no yes yes

The two theories are empirically equivalent for 8a and 8b. For 8c and 8d,

---

4 This implies that an element in an adjoined position never counts as a typical potential α-governor. It is predicted, for example, that an operator adjoined to IP or VP by QR never triggers the minimality effect. For this prediction, see pp. 20–21. Notice that it is also predicted that QR itself can be sensitive to the minimality effect triggered by A' specifiers. I will not discuss this issue here.
Relativized Minimality is more restrictive than Rigid Minimality, while, for 8e and 8f, the situation is reversed.

Let us first consider cases of 8a and 8b.

(10) *John tried [\text{CP} C^0 [\text{Bill to win}]] (cf. 8a)
\[
X^0 \quad \underbrace{\text{Z}^0} \quad Y
\]

(11) *Have they could t left? (cf. 8b)
\[
X^0 \quad \underbrace{\text{Z}^0} \quad Y
\]

Under either theory, the intervening heads trigger the minimality effect. Consequently, in 10, \textit{Bill} cannot be head-governed by \textit{tried}, so it cannot be Case-marked.\(^5\)\(^6\) In 11, which is a case of the Head Movement Constraint violation (cf. Chomsky 1986; Baker 1988), the trace cannot be antecedent-governed by \textit{have}. Hence the ECP is violated.\(^7\)

Let us turn to 8c and 8d. First, consider an instance of 8c (adjunct extraction out of \textit{wh}-islands).

(12) *How do you wonder [\text{which problem [ \text{PRO} to solve} \ t \ t']]
\[
\underbrace{\text{XP}} \quad \text{ZP} \quad Y
\]

Under Relativized Minimality, the ECP is violated, because \(t\) is neither antecedent-governed by \textit{how} nor \(\theta\)-governed. In contrast, the ECP is not violated in 13, because there is no intervening A’-specifier.

(13) How do you think [t that [Bill solved it t]]

Next, consider a case of 8d (Super Raising).

(14) *John seems that it is likely [t to win]
\[
\underbrace{\text{XP}} \quad \text{ZP} \quad Y
\]

In 14, the trace must be antecedent-governed because it is not \(\theta\)-governed. Under Relativized Minimality, it is not antecedent-governed, and hence the ECP is violated. In 15, where there is no intervening A-specifier for the A-chain (\textit{John, t}), the ECP is satisfied.

(15) John does not seem [t to be here]

R further argues that the so called 'inner island' phenomenon (cf. Ross 1984) should be treated as another instance of 8d.

---

\(^5\) Chomsky 1986 assumes that a featureless head does not yield the minimality effect. For this assumption, see section 3 below.

\(^6\) As R notes, this approach implies that we can account for the impossibility of head government from an external head governor in 10 without regarding CP as a barrier 'by inheritance.' For the notion of barrier by inheritance, see Chomsky 1986.

\(^7\) For the moment, let us assume the following definition of the ECP: A nonpro-nominal empty category must be (a) \(\theta\)-governed, or (b) antecedent-governed.
(16) a. It is in this way that I think that John fixed the car t
   b. *It is in this way that I don't think that John fixed the car t

In 16b, the adverbial PP cannot be clefted across the intermediate clause containing negation. In this structure, R claims, ZP = the negative operator, XP = in this way, and Y = the trace. Therefore, the clefted PP cannot antecedent-govern its trace, and the ECP is violated.  

Rigid Minimality, on the other hand, can predict no minimality effect in these cases, because the intervening elements, which problem in 12, it in 14, and the negative operator in 16, are not head categories. Thus, for cases of 8c and 8d, we can regard Relativized Minimality as superior to Rigid Minimality.

Let us now turn to 8e and 8f. First, consider a case of 8e.

(17) *How did John announce [NP a plan [CP t' [PRO to fix
   XP the car t]]]
   \(Z^0\)

Under Rigid Minimality, the intervening head plan triggers the minimality effect, so t' is neither antecedent-governed nor \(\theta\)-governed. Hence, the ECP violation results.  

Relativized Minimality, in contrast, predicts no minimality effect in this case. A head never qualifies as a potential antecedent governor for a trace in A'-chain. In R's system, however, since CP complements selected by nouns are generally intrinsic barriers (cf. note 3 above), t' is neither antecedent-governed nor \(\theta\)-governed. Thus, for this case, we cannot choose between the two theories.

Finally, consider a case of 8f.

---

8 This analysis is based on the assumption that negative elements occupy some A'-specifier position at the relevant level of representation. This assumption, however, is controversial. For this problem, see pp. 17-24. For relevant discussions, see Chomsky 1989; Pollock 1989; Laka 1990. Furthermore, even if negative elements occupy some A'-specifier position, negative operators such as not and never might be heads on their own as well as A'-specifiers. If so, some devices would be necessary to eliminate this potential ambiguity.

9 In fact, Chomsky 1986 accounts for adjunct extraction from wh-islands and Super Raising in terms of barrier by inheritance, instead of minimality barrier. For example, the CP intervening between how and its trace in 12 is a barrier by inheritance in his system, blocking the required antecedent-government relation. For problems with this account, see Cinque 1990.

10 Chomsky 1986 suggests that the that-trace phenomena can be treated as instances of 8e. For this topic, see section 3 below.
(18) John seems [t to be intelligent]

\[ XP \rightarrow \quad Z^0 \quad Y \]

This is an example of (local) Raising. Under Relativized Minimality no minimality effect is predicted, as discussed above (cf. 15). Under Rigid Minimality, on the other hand, the intervening head V^0 should trigger the minimality effect. Hence, Rigid Minimality must say something special to void the undesirable minimality barrier.\(^\text{11}\) Thus, for this case, Relativized Minimality can be regarded as superior to Rigid Minimality.

2.3. We have compared the two theories in terms of the configurations 8a–f. The results of the comparison would lead us to the conclusion that Relativized Minimality is more natural than Rigid Minimality, and, as a whole, is on the right track. Relativized Minimality, however, raises a conceptual question: Why should the principle of minimality be relativized as it is in R’s formulation? Can’t there be another way of relativization? For example, Baker and Hale 1990 claim that antecedent government by a head should be relativized as to whether the head is lexical or functional. If their claim is correct, we may expect head government also to be relativized as to whether the head governor is lexical or functional. However, if head government is relativized in such a way, we cannot account for why the functional head C^0 blocks head government of Bill and PRO by the lexical head tried in the following sentences (cf. 10 above).

(19) a. *John tried [\( C^0 \) [Bill to win]]

b. John tried [\( C^0 \) [PRO to win]]

Thus, we need some principle to define the notion of ‘possible relativization.’

We might here raise another interesting question: Are other notions of locality relativized? As an illustration, let us consider the notion of local binding.

Relevant to this question are the strong crossover phenomena.

\(^{11}\) Chomsky 1986 tries to solve this problem by introducing the new notion of extended chain. Notice that if we allowed the intervening V^0 to trigger the minimality effect, we could rule out cases of Super Raising such as 14 as instances of 8f. Cases of \textit{wh}-islands (cf. 12) and inner islands (cf. 16b) could be treated in a parallel way, if the intervening I^0 and/or V^0 were allowed to trigger the minimality effect. Such a treatment, however, would require further special devices in order to permit grammatical cases such as 13 and 16a.
(20) *Who₁ does he₁ think that Mary loves t₁?
Roughly speaking, there are two types of approaches to these phenomena. Under one approach, 20 is ruled out as a violation of the following condition (or definition).

(21) A variable must be locally A'-bound.
In 20, the trace is a variable by definition (a variable is a Case-marked empty category). However, since it is locally A-bound by he, it is not locally A'-bound. Thus 20 violates 21 (cf. Chomsky 1982; Koopman and Sportiche 1982). Notice that under this account A'-binding is blocked by the intervening A-binder. In other words, local binding is not relativized for the distinction of A-/A'-binding.

The other approach is based on the Binding Principle C. Since the trace, an A'-bound empty category, is a variable, it counts as an R-expression. However, it is A-bound by he and violates the Binding Principle C. Under this approach, A'-binding itself is not blocked by the intervening A-binder, that is, local binding is relativized for the distinction of A-/A'-binding.

I have no decisive evidence for either approach, but it should be noticed that the strong crossover phenomena may have a non-trivial implication on whether local binding is relativized.

3. A CONJUNCTIVE FORMULATION OF THE ECP.
3.1. The classical formulation of the ECP (Chomsky 1981) is intended to capture the subject-object asymmetry with respect to the that-trace phenomena.

(22) ECP I: A nonpronominal empty category must be (i) lexically governed, or (ii) antecedent-governed.
(23) a. Who do you think [t' (that) Bill saw t]  
b. Who do you think [t' (*that) t saw Bill]
Traces of extracted subjects, which are not lexically governed, require antecedent government by t', which is blocked by the intervention of the complementizer that. In contrast, traces of objects, which are always lexically governed by the heads, do not require antecedent government.

The original ECP, however, cannot capture the complement-non-complement asymmetry with respect to extraction out of wh-islands (cf. Huang 1982).

(24) a. ??Which problem do you wonder how John could solve t t  
b. *Which student do you wonder how t could solve the problem t
c. *How do you wonder which problem John could solve t t
The adjunct trace in 24c is lexically governed by the verb within VP. However, it still requires antecedent government as the subject trace in 24b.

If we replace lexical government with θ-government (lexical government with θ-marking), this complement-noncomplement asymmetry can be captured (cf. Stowell 1981; Chomsky 1986).

(25) ECP II: A nonpronominal empty category must be (i) θ-governed, or (ii) antecedent-governed.

The subject trace in 24b is neither antecedent-governed nor θ-governed, because the subject is not governed by the verb, and the adjunct trace in 24c is neither antecedent-governed nor θ-governed by the verb, because the adjunct is not θ-marked by the verb. Thus, the ECP is violated in 24b–c. In contrast, the object trace in 24a is θ-governed by the verb, satisfying the ECP.12

R argues, however, that the ECP II is conceptually as well as empirically inadequate. The conceptual problem is that the ECP II is defined by combining the two clauses disjunctively. A disjunctive definition always accompanies the question of why the disjunctively listed independent notions constitute a natural class in some relevant respect.13 The empirical problem is that I0, which is not a lexical head, seems to license a VP trace.

(26) a. I asked John to go home, and [VP go home] I think [t’ that [he did t]]
   b. ?... and [VP win the race] I wonder whether he did t

These examples suggest that the class of proper governors should include some functional heads in addition to lexical heads.

Thus, R reformulates the ECP in terms of the notions of formal licensing (=FL) and identification (=ID).

(27) ECP III: A nonpronominal empty category must be
   (i ) properly head-governed (FL), and
   (ii ) antecedent-governed or θ-governed (ID).

The ECP III is formulated conjunctively, and traces must satisfy both of the principles of FL and ID. ‘Proper head government’ in 27i means

---

12 The marginality of 24a is due to the Subjacency condition violation, which is milder than the ECP violation.
13 For other conjunctive formurations of the ECP, see Stowell 1986; Aoun et. al. 1987; Chomsky 1986.
'government by $X^0$ within $X'$.' The class of head governors includes A, N, P, V, and I (that is, Agr and T).\textsuperscript{14}

3.2. Chomsky 1986, assuming the ECP II above, suggests that the \textit{that}-trace phenomena can be accounted for in terms of the notion of minimality barrier.

(28) *Who do you think $[\text{CP } t' [C' \text{ that } [t \text{ left}]]$] The subject trace, which is not $\theta$-governed, must be antecedent-governed. In 28, however, the intervening overt complementizer triggers the minimality effect, and the minimality barrier $C'$ blocks the antecedent government relation between $t'$ and $t$.

R argues against this proposal on the following grounds. First, no principled explanation is offered as to why an overt $C^0$ blocks antecedent government while a nonovert $C^0$ does not.

(29) Who do you think $[t' 0 [t \text{ left}]]$ ($0=\text{a null } C^0$) Second, sentences such as 30 suggest that an overt $C^0$ does not block antecedent government.

(30) How do you think $[t' \text{ that } [\text{Bill solved the problem } t]]$ Since the adjunct trace must be antecedent-governed, if \textit{that} blocked antecedent government, 30 should violate the ECP.

Thus R rejects the standard idea that the \textit{that}-trace effect comes from the failure of antecedent government of the subject trace. In fact, Relativized Minimality excludes the possibility of the standard idea as a theorem, because a head never blocks antecedent government in A'-chains (recall the discussion about 12 and 13 in section 2).

R argues instead that the \textit{that}-trace effect is due to the failure of proper head government of the subject trace. The relevant structure of 28 is as follows (cf. Pollock 1989; Chomsky 1989):

\textsuperscript{14} R assumes that Agr$^0$ is a governor only if it has features, that is, in tensed clauses, and it only governs the element(s) it agrees with. T$^0$, on the other hand, is always a governor.
In this structure, the subject trace $t$ is antecedent-governed by $t'$, and satisfies the ID clause of the ECP. Although it can be head-governed by $Agr^0$, however, since it is outside $Agr'$, it cannot be properly head-governed by $Agr^0$. Furthermore, it is not properly head-governed by $C^0$, which is not a head governor. Hence, the subject trace violates the FL clause of the ECP.

Thus, in R's system, subject traces in the specifier position of IP (or $Agr_P$ in a more articulated structure) in general cannot satisfy the conjunctively formulated ECP, and therefore subjects cannot be extracted unless something special happens to them. He mentions the three strategies to permit subject variables: the insertion of subject resumptive pronouns (ex. Swedish), the extraction of VP-adjoined post-verbal subjects (ex. Null Subject Languages such as Italian), and the conversion of $C^0$ into a proper governor. In what follows, let us consider the last strategy, which is crucial to account for sentences like 29.

According to R's proposal, $C^0$, which is generally inert for government, can be turned into a governor through spec-head agreement in the domain of Comp. He proposes that in English a tensed complementizer can be expanded as that or Agr, assuming that Agr can be an independent head or be assigned to another head as a feature (or a set of features).

Expansions in 32 are optional, and $C$ can be left unexpanded. That and an unexpanded $C$ are inert for government, while Agr as the head of

\[ (i) \quad \ldots \ T^0 [_{VP [_{VP \ldots} \text{manner Adv}]] \ldots \]

In this structure, their traces can be properly head-governed by $T^0$ across the transparent VP segment (p. 50). For the treatment of other types of adjuncts, see pp. 46–51.
Comp belongs to the class of governors. As a general property of Agr, Agr must be licensed by coindexation with its specifier. Thus, when Agr is the head of Comp, it must be coindexed with the specifier to be licensed. It follows that the specifier position of Comp must be occupied either by a wh-operator as in 33a or by a trace as in 33b (=29).

\[(33) \quad \text{a. Who}_1 \text{ Agr}_1 [t_i \text{ Infl left}]\]
\[(33) \quad \text{b. Who}_1 \text{ do you think } [t_i \text{ Agr}_i [t_i \text{ Infl left}]]\]

In these structures, the subject traces are properly head-governed by Agr in Comp, and hence the Formal Licensing condition of the ECP is satisfied.\(^{16, 17}\)

3.3. We have seen how the conjunctively formulated ECP III, along with Relativized Minimality, accounts for the that-trace phenomena. However, as R notices, the ID clause of the ECP III has some conceptual and empirical problems which lead him to remove it from the ECP, as will be discussed in the next section.

Then, in this subsection, let us concentrate our attention on the FL clause: traces must be properly head-governed.

One of the problematic cases for this clause is \(X^0\)-movement: traces created by \(X^0\)-movement in general are not properly head-governed (p. 41, note 8). In 34, for example, the trace of *did* is not properly head-governed.

\[(34) \quad [\text{IP Did } [\text{IP Mary t [see John]]}]\]

R suggests that the head-government requirement holds only for empty XP's, so that this requirement is not relevant to empty heads. It is not clear why this must be the case, but this might be related to the fact that a head can qualify as an antecedent governor as well as a head governor.

We also might raise a question as to the level at which the FL clause should apply. R assumes that the FL clause applies at S-structure for traces created in the syntax (p. 39). This assumption is based on his observation that the antecedent-government requirement can be satisfied through reconstruction, which is assumed to be an LF process, while the head-government requirement cannot.

---

\(^{16}\) The intermediate trace in 33b is properly head-governed by the matrix verb and antecedent-governed by *who.*

\(^{17}\) R further makes a detailed discussion on other problems deeply related to the complementizer system such as the relative complementizer, the doubly-filled Comp phenomena, and so on. I will not discuss these topics here.
In this connection, Kikuchi and Takahashi 1990 point out that if the VP-internal subject hypothesis is assumed, the subject trace in VP is not properly head-governed at S-structure in the VP-preposing construction, and hence the FL clause should be violated.

(35) I asked John to go home, and [ti go home] I think that he, did t
To solve this problem, they propose that the subject trace is properly head-governed by the head verb go, arguing that when X is a lexical category, X' can be ignored for the purpose of proper head government. We could, however, interpret this example as indicating that the FL clause can be satisfied through reconstruction. Thus, further investigations should be made as to the appropriate level(s) at which this clause should apply.

The definition of head government also raises a fundamental conceptual question: Why is C excluded from the class of (proper) head governors? We find essentially the same problem in the notion of θ-governor in Chomsky 1986. Anyway, this question must be answered in a principled way in the future.

4. Referential Indices.

4.1. We have seen that the ECP is formulated conjunctively as follows:

(36)(=27) ECP III: A nonpronominal empty category must be
    (i) properly head-governed (FL), and
    (ii) antecedent-governed or θ-governed (ID).
However R argues that this formulation still raises a number of questions, both conceptual and empirical.

One of the conceptual problems is that there is a significant redundancy between the head-government requirement and the θ-government requirement. Another problem is that the ID clause is formulated disjunctively.

Empirically, the ECP III cannot capture the argument-adjunct asymmetry in wh-movement from wh-islands. The disjunctive definitions in the ECP II and the ID clause of 36 are intended to capture the following complement-noncomplement asymmetry (see section 3.1 above).

(37) a. *Who do you wonder whether t can help us?
   b. ??Who do you wonder whether we can help t?
   c. *How do you wonder whether we can help Bill t?
However, R argues that the complement-noncomplement asymmetry is not a true generalization and even subjects can be extracted from wh-
islands if their traces are properly head-governed as in 38a. Adjuncts, in contrast, can never be extracted from \textit{wh}-islands. The ECP III cannot capture this argument-adjunct asymmetry.

(38) a. ?*Who do you wonder whether we believe \textit{t} can help us?
b. ??Who do you wonder whether we believe \textit{we} can help \textit{t}?  
c. *How do you wonder whether we believe \textit{we} can help Bill \textit{t}?

The notion of \(\theta\)-government is also inadequate in that it cannot account for the fact that not every complement can be extracted from \textit{wh}-islands.

(39) John weighed (a) apples/(b) 200 lbs  
(40) What did John weigh \textit{t}?  
(41) ?What did John wonder how to weigh \textit{t}?

\textit{What} in 40 is ambiguous: it can be either a patient as apples in 39a or a measure phrase as 200 lbs in 39b. When extracted from a \textit{wh}-island, however, it can only be interpreted as a patient. A measure phrase, though selected and \(\theta\)-marked by \textit{weigh}, cannot be extracted from \textit{wh}-islands like ordinary adjuncts.

4.2. In order to avoid these problems, R claims that the ECP should be reduced to the principle of FL.

(42) ECP (Final): a nonpronominal empty category must be properly head-governed.

He further claims that the disjunctive ID clause has no independent theoretical status and its essential effect is accounted for by restricting the use of index.

Essentially along the lines of the classical theory in Chomsky 1965, he proposes that the use of referential index is subject to the following principle:

(43) A referential index must be licensed by a referential \(\theta\)-role. A referential index is only legitimate on an element associated to a referential \(\theta\)-role such as agent, theme, patient, goal, and so on. Nonreferential \(\theta\)-roles such as manner, measure, quasi-\(\theta\)-role assigned to an idiomactic element, and so on cannot license referential indices. For example, a measure phrase selected by \textit{weigh} can bear no referential index.

The binding relation is defined as follows:

(44) \textit{X} binds \textit{Y} iff (i) \textit{X} c-commands \textit{Y}, and (ii) \textit{X} and \textit{Y} have the same referential index.

It follows from 43 and 44 that binding relations are restricted to elements
Let us consider how the restricted indexing theory accounts for the cases problematic for the ECP III.

(38)  a. *Who_{i} do you wonder whether we believe t_{i} can help us?  
    b. ??Who_{i} do you wonder whether we believe we can help t_{i}?  
    c. *How do you wonder whether we believe we can help Bill t?  

In 38c, the adjunct operator how is not assigned any referential index at D-structure under the principle 43, because it receives no referential $\theta$-role. Thus, it cannot be connected to the trace through binding. Consequently, to satisfy the principle of Full Interpretation, the operator how must be connected to the variable through some other device: antecedent government. However, the required antecedent government is blocked under Relativized Minimality. Hence, the structure receives no interpretation. In 38a–b, although antecedent government is impossible, the referential operators can be connected to the traces through binding, because the referential indices are licensed by the referential $\theta$-roles these operators receive at D-structure. A parallel explanation holds for the absence of ambiguity in 41.

In general, referential arguments can always make use of binding in addition to antecedent government, while other elements can have access only to antecedent government. This is the essential effect of the ID clause: adjunct traces, in contrast to object traces, must be antecedent-governed.\footnote{Antecedent government is redefined as follows: X antecedent-governs Y iff (i) X and Y are nondistinct (in category, feature content, indices if indexed, etc.), (ii) X c-commands Y, (iii) no barrier intervenes, and (iv) Relativized Minimality is respected. Notice that the nondistinctness requirement rules out the possibility of extended chains in Chomsky 1986 as 'crazy chains.'}

A crucial difference between the two modes of connection is that binding is an unbounded relation, while antecedent government is an intrinsically local relation. In other words, referential arguments can always take part in long A'-dependencies through binding, within the limits posed by the Subjacency, while other elements can never use this option. The argument-adjunct asymmetry in extraction from wh-islands is subsumed under this general referential-nonreferential asymmetry.

Finally, let us consider a case of A-movements.

(45)  a. It seems that someone told Bill that …
b. It seems that Bill was told that ...

c. *Bill seems that it was told that ...

As 45c indicates, the A-dependency between Bill and the trace must be strictly local, in spite of the availability of binding.

R attributes this strict locality requirement on A-dependencies to the $\theta$-criterion. If the $\theta$-criterion is defined in terms of chains, chain formation is obligatory for $\lambda$-movements because $\theta$-roles and Cases are transmitted through $\lambda$-chains. And if the notion of chain is defined in terms of antecedent government, 45c is ruled out as a violation of the $\theta$-criterion, because the intervening it blocks the antecedent-government relation between Bill and the trace under Relativized Minimality.\footnote{Cinque 1990 suggests that if we assume that NP-traces are not referentially autonomous and cannot bear any indices, the obligatory chain formation in $\lambda$-movements is attributed to the unavailability of binding in $\lambda$-movements. This alternative approach, however, might raise a question for $\lambda'$-movements: Can operators be regarded as referentially autonomous?}

4.3. We have seen how the ECP is reduced to the proper head government requirement and how the restricted theory of indexing accounts for the argument-adjunct asymmetry.

Though the notion of referentiality plays a crucial role in R's system, this notion is not easy to define. For example, R suggests the possibility of locative and temporal elements as referential arguments (p. 91), but it is not clear what kind of referential $\theta$-roles these elements receive. Furthermore, Cinque 1990 argues that in addition to the notion of referential $\theta$-role, intrinsic properties of arguments play a significant role in defining this notion. Thus it is still not fully clear how to define the notion of referentiality.\footnote{Cinque also claims that the notion of Discourse-linking in Pesetsky 1987 can be subsumed under the notion of referentiality.}

Another problem is that although disjunctive statements are entirely eliminated from the ECP, a new disjunctive statement might have to be introduced in LF, that is, the notion of well-formed operator-variable structure in LF (let us call it LF-binding) might have to be stated disjunctively as follows:

\begin{equation}
(46) \text{An operator } X \text{ LF-binds a variable } Y \text{ iff}
\begin{align*}
(i) & \quad X \text{ binds } Y \text{ through referential coindexing, or} \\
(ii) & \quad X \text{ governs } Y \text{ through a chain of antecedent-government relations.}
\end{align*}
\end{equation}
A possible solution for this problem might be to assume that indices are assigned as in the standard theory of indexing, and only indices assigned to referential arguments are identified as referential indices. The binding relation is defined in terms of referential indices as in 44. The antecedent-government relation is defined in terms of indices regardless of referentiality. Given these assumptions, we might be able to define the well-formed operator-variable structure in terms of coindexing and c-command without disjunctive statements.

5. CONCLUDING REMARKS. We have summarized R’s main proposals and discussed some questions and potential problems. In spite of these questions and problems, this monograph is full of important proposals and suggestions which deserve serious considerations, and offers a promising theoretical framework for the analysis of symmetries and asymmetries observed in various extraction phenomena. There is no doubt that R has made a significant contribution to the theory of syntax.

REFERENCES

AOUN, JOSEPH; NORBERT HORNSTEIN; DAVID LIGHTFOOT; and AMY WEINBERG. 1987. Two types of locality. LI 18.537-77.


KIKUCHI, AKIRA, and DAIKO TAKAHASHI. 1990. Agreement and small clauses. ms. Tohoku University.

KOOPMAN, HILDA, and DOMINIQUE SPORTICHE. 1982. Variables and the bijection principle. TLR 2.139-60.


