Howard Lasnik
Tim Stowell

Weakest Crossover

1. Introduction

1.1. Pronouns Construed as Bound Variables

A pronoun may be construed as a variable in contexts such as (1):

(1) a. Every man thinks he is lucky
b. Who dislikes his boss
c. Which man did you say dislikes his boss
d. No man should mistreat his friends

In (1a–d) the pronoun *he* or *his* is ambiguous between a referential construal, where it refers to a particular individual, and a bound variable construal, where it behaves like a variable bound by the *wh*-phrase or Quantifier Phrase (QP). We will assume that the bound variable construal is represented structurally by means of coindexing the pronoun with the *wh*-phrase or QP.[1]

Following Chomsky (1973), May (1977), and Huang (1982), we also assume that every *wh*-phrase and QP must undergo movement to an operator position in order to bind its trace as a variable at the level of Logical Form (LF). We assume further that *wh*-phrases appear in the Spec position of CP, whereas other quantifiers taking clausal scope are adjoined to IP (see Chomsky (1986a)). Thus, at LF the examples in (1a–d) are represented as (2a–d), respectively:

(2) a. $[_{IP}$ every man$_i$ $[_{IP}$ t_i thinks he$_i$ is lucky]]
b. $[_{CP}$ who$_i$ $[_{IP}$ t_i dislikes his$_i$ boss]]
c. $[_{CP}$ which man$_i$ [did you say $[_{IP}$ t_i dislikes his$_i$ boss]]]
d. $[_{IP}$ no man$_i$ $[_{IP}$ t_i should mistreat his$_i$ friends]]

We are grateful to Noam Chomsky, James Higginbotham, Kyle Johnson, Carlos Otero, and David Pesetsky for helpful discussion of this material, to Samuel D. Epstein, Mamoru Saito, Jun Abe, and Sung-Ho Ahn for valuable suggestions on an earlier version, and to two *Linguistic Inquiry* reviewers for extremely insightful commentary. In addition, we have benefited from comments and suggestions from audiences at the University of Massachusetts (Amherst), MIT, and the University of California (Irvine), and from participants in classes at the University of Connecticut and at UCLA.

[1] An alternative mode of representation, with somewhat different properties, is provided by linking theory, as in Higginbotham (1983; 1985). We will leave for another occasion investigation of how the phenomena considered here might be addressed within the linking framework.

Linguistic Inquiry, Volume 22, Number 4, Fall 1991
687–720

The availability of the bound variable construal is dependent on the structural position of the pronoun in the sentence. For instance, the pronoun *his* cannot be construed as a variable bound by the QP *no man* in (3a) or the *wh*-phrase *who* in (3b):

(3) a. *His$_i$ boss regrets that no man$_i$ is lucky
b. *His$_i$ mother wonders who$_i$ Jane saw t$_i$

This is because the scope of the *wh*-phrase or QP is restricted to the embedded clause, while the pronoun *his* occurs outside this domain in the matrix clause, as shown by the LF representations of (3) in (4):

(4) a. *[his$_i$ boss regrets [$_{CP}$ that [$_{IP}$ no man$_i$ [$_{IP}$ t$_i$ is lucky]]]]
b. *[his$_i$ mother wonders [$_{CP}$ who$_i$ [$_{IP}$ Jane saw t$_i$]]]

Assuming that the scope of a quantifier corresponds to its c-command domain at LF, the relevant restriction can be stated as the LF binding condition in (5):

(5) A pronoun P is construed as a variable bound by a quantifier Q only if Q binds P at LF.

In (5) we assume the usual definition of the binding relation:

(6) X binds Y iff (a) X and Y are coindexed, and (b) X c-commands Y.

Condition (5) is satisfied in (2), but not in (4), where the bound reading is consequently unavailable.

1.2. Strong Crossover

Although (5) is a necessary condition on the availability of the bound variable construal of a pronoun, it is not sufficient, as shown by the structures in (7), with the LF representations in (8):

(7) a. *Who$_i$ did you say he$_i$ made you visit t$_i$
b. *He$_i$ saw me visit nobody$_i$
(8) a. *[$_{CP}$ who$_i$ [did you say [$_{CP}$ he$_i$ made you visit t$_i$]]]
b. *[$_{IP}$ nobody$_i$ [$_{IP}$ he$_i$ saw me visit t$_i$]]

The quantifiers c-command the pronouns in (8) in accord with (5), but the bound variable construal of the pronouns is still excluded. These are instances of *strong crossover* (SCO), so named by Wasow (1972) (see Postal (1971)). In Chomsky (1981) (henceforth *LGB*), SCO is treated as a violation of Condition C of the binding theory:

(9) A. An anaphor is A-bound in its governing category.
B. A pronominal is A-free in its governing category.
C. An R-expression is A-free.

The *LGB* theory assumes that variables (*wh*-trace and QP-trace) are treated as R-expressions with respect to Condition C. In SCO configurations like (8a–b), Condition

C is violated if the pronoun is coindexed with the *wh*-phrase or QP and its trace, since the pronoun *he* A-binds the variable trace. Hence, the pronoun must have a distinct index to satisfy Condition C, thus excluding the bound construal.

A different account of SCO is offered by Chomsky (1982), who attributes it to the *LGB* definition of *variable* in (10):[2]

(10) α is a variable if and only if α is locally $\bar{A}$-bound and in an A-position.

(11) α is locally bound by β if and only if α is X-bound by β, and if γ Y-binds α, then either γ Y-binds β or $\gamma = \beta$.

Since the trace in (8a–b) is locally A-bound by the pronoun *he*, it is not locally $\bar{A}$-bound and therefore fails to satisfy (10). Therefore, the trace is not a variable. Furthermore, it is not well formed as any other type of empty category. It can't be a pronominal anaphor PRO, since it occurs in a governed object position. It can't be an anaphoric NP-trace, since it is not A-bound within its governing category. Finally, it can't be a null pronoun (small pro) since there are no agreement features to license it in this context. Thus, SCO configurations are excluded in this theory because the trace of the *wh*-phrase or QP does not correspond to any well-formed type of empty category. Therefore, the pronoun must have a distinct index in (8) so that the trace can be locally $\bar{A}$-bound and satisfy the definition of a variable.

1.3. Weak Crossover

In the examples discussed thus far, either the trace of the QP c-commands the coindexed pronoun and the bound construal of the pronoun is possible, as in (2), or the coindexed pronoun c-commands the trace of the quantifier in a SCO configuration, as in (8). Now consider the structures in (12), with the LF representations in (13), where neither the pronoun nor the trace c-commands the other:

(12) a. *The fact that he_i owned a gun implicated $everyone_i$
b. *Who_i does his_i boss dislike
c. *$Which\ man_i$ did you say his_i boss dislikes
d. *His_i friends should mistreat $no\ man_i$

(13) a. *$everyone_i$ [$_{IP}$[$_{NP}$ the fact that he_i owned a gun] [$_{VP}$ implicated t_i]]
b. *who_i [$_{IP}$ does [$_{NP}$ his_i boss] [$_{VP}$ dislike t_i]]
c. *$which\ man_i$ [$_{IP}$ did you say [$_{CP}$[$_{NP}$ his_i boss] [$_{VP}$ dislikes t_i]]]
d. *$no\ man_i$ [$_{IP}$[$_{NP}$ his_i friends] should [$_{VP}$ mistreat t_i]]

Here too the bound construal of the pronoun is excluded, despite the fact that Condition (5), Condition C of the binding theory, and the definition of a variable (10) are all satisfied. These structures exemplify *weak crossover* (WCO), so named by Wasow (1972) because the judgments concerning the unavailability of the bound construal of the pronoun are less robust than in the case of strong crossover.

[2] See *LGB* (pp. 185, 330). Chomsky (1982) attributes the analysis to Dominique Sportiche; see also Sportiche (1983).

The contrast between the WCO examples in (12) and the examples in (1) suggests the descriptive generalization (14):

(14) In a configuration where a pronoun P and a trace T are both bound by a quantifier Q, T must c-command P.

Most theories of WCO assume the descriptive correctness of (14) and seek to derive it from more general principles. We will discuss just one of these theories: that of Koopman and Sportiche (1983). They attribute WCO to the Bijection Principle, (15):

(15) There is a bijective correspondence between variables and Ā-positions. (That is, each operator must Ā-bind exactly one variable, and each variable must be Ā-bound by exactly one operator.)

Their account rests on the *LGB* definition of "variable" in (10), which (as they point out) does not distinguish between pronominal elements and empty categories. This means that in a WCO configuration, where no c-command relation holds between the trace T and the pronoun P, both P and T count as variables: since neither binds the other, they are both locally Ā-bound by the *wh*-phrase or QP. Consequently, each *wh*-phrase or QP binds two variables at LF and the Bijection Principle is violated. In contrast, each pronoun in (2) is locally A-bound by a *wh*-trace or QP-trace at LF, so it is not defined as a variable by (10), and no violation of (15) arises. Thus, the interaction of the definition of a variable (10) and the Bijection Principle (15) derives the descriptive generalization (14).[3]

In the WCO examples in (13) the offending pronoun is contained within an argument NP that c-commands the trace. But Stowell (to appear) notes that if the pronoun is contained within an adjunct phrase, no WCO effect arises even if (14) is violated, as in (16):

(16) a. Who$_i$ did you say [$_{CP}$ t$_i$ was a liar] [$_{PP}$ before you met him$_i$]
b. Who$_i$ did Jan say [$_{CP}$ she admired t$_i$] [$_{PP}$ in order to please him$_i$]

In response to these and related data, Stowell proposes a theory of WCO based on the descriptive generalization (17):

(17) In a configuration where a quantifier Q locally binds a pronoun P and a trace T, P may not be contained in an argument phrase XP that c-commands T.

In (16) the pronoun is contained within an adjunct phrase that c-commands the trace and (17) is satisfied, whereas the standard WCO examples like (13) involve pronouns within argument phrases that c-command the trace, where (17) is violated.

It is not our intent to propose or defend a particular theory of WCO in this article, and for our purposes the choice between (14) and (17) is immaterial. Therefore, we will attempt to abstract away from the empirical dispute between these conflicting descriptive

[3] Safir's (1984) Parallelism Constraint on Operator Binding (PCOB) is an alternative account of WCO. Our discussion is, as far as we can tell, neutral between these two approaches.

generalizations, and from the rival theories of WCO that they underlie. To this end, we will concentrate on configurations analogous to (13), where (14) and (17) converge in their empirical coverage.

1.4. Weakest Crossover

Quantifier binding is often equated with Ā-binding, since *wh*-phrases and other QPs normally participate in Ā-binding configurations at LF. Thus, (14) can be restated as (18), as in Koopman and Sportiche (1983) and Safir (1984):

(18) In a configuration where a pronoun P and a trace T are both Ā-bound by a category C, T must c-command P.

In a similar vein, (17) can be restated as (19):

(19) In a configuration where a category C Ā-binds a pronoun P and a trace T, P may not be contained in an argument phrase XP that c-commands T.

(18) and (19) incorporate the hypothesis that the principle(s) responsible for WCO effects apply in all configurations of Ā-binding; let us refer to this as the *Generalized WCO Hypothesis*.

But there are constructions involving Ā-binding where WCO effects do not arise, despite the fact that both (18) and (19) are violated. Chomsky (1982) discusses one such case (the apparent absence of WCO effects in relative clauses); we examine his analysis in section 3. But the problem is more widespread; there are several other constructions that appear to freely violate (18) and (19). The relevant data are examined in section 2, but (20) provides a representative sample, involving *tough*-movement, parasitic gap, and topicalization constructions, respectively:

(20) a. Who$_i$ t$_i$ will be easy for us [to get [his$_i$ mother] to talk to e$_i$]
b. Who$_i$ did you stay with t$_i$ [before [his$_i$ wife] had spoken to e$_i$]
c. This book$_i$, I expect [its$_i$ author] to buy e$_i$

Each example in (20) is structurally parallel to the standard type of WCO configuration; there is an operator that locally Ā-binds both a pronoun and a trace (in violation of (18)) and the pronoun is contained in an NP that c-commands the trace (in violation of (19)).

We refer to these apparent violations of the Generalized WCO Hypothesis as instances of *weakest crossover*. There are two general strategies for responding to data of this sort. On the one hand, we could maintain the Generalized WCO Hypothesis, and reanalyze the binding relations in weakest crossover constructions so as to be consistent with (18) and (19), for example, by claiming that the pronouns in (20) are not truly bound by the quantifiers, at least at the level of representation relevant for WCO. Chomsky's (1982) analysis of relative clauses exemplifies this approach.

On the other hand, we could abandon the Generalized WCO Hypothesis, and revise the underlying theoretical principles in such a way that they predict the absence of WCO

effects in the environments of weakest crossover. This is the type of approach we advocate here; we suggest that the principles responsible for WCO apply only to variables that are locally bound by true quantifiers, and that the weakest crossover constructions in (20) do not count as such.[4] We propose that in the constructions exemplifying weakest crossover, where the operator is semantically nonquantificational, the trace is not a true variable, but rather a previously unrecognized type of empty category: a null R-expression, with binding properties analogous to those of names and definite descriptions (''epithets'') rather than to those of variables.

The rest of the article builds toward this analysis. In section 2 we discuss in more detail the three major environments of weakest crossover (parasitic gap constructions, *tough*-movement constructions, and topicalization constructions) and show that these fail to exhibit WCO effects even when the pronouns in question are clearly construed semantically as variables. Cleft, pseudocleft, and focus constructions raise additional problems for WCO theory, including our own account; we discuss these in appendix A.

In section 3 we review Chomsky's (1982) account of the presumed absence of WCO effects in restrictive relative clauses. We show that an account of this type, though potentially appealing, cannot handle all types of weakest crossover. In section 4 we construct our alternative account of weakest crossover phenomena, which is largely theory-neutral with respect to the rival theories of WCO alluded to in section 1.3. Our theory that certain Ā-bound traces function as null (nonvariable) R-expressions has implications for several issues relating to WCO and SCO, notably for the precise status of Condition C of the binding theory as it applies to variables.

2. Environments of Weakest Crossover

2.1. Parasitic Gaps

The first context of weakest crossover involves parasitic gap constructions, as in (21). Following standard practice, the licensing *wh*-trace is represented by *t*, and the parasitic gap is represented by *e*.

(21) a. Who$_i$ did Mary gossip about t$_i$ [despite your having vouched for e$_i$]
b. Which man$_i$ did you look at t$_i$ [after Mary had spoken to e$_i$]

The derivation of parasitic gaps is controversial. According to Chomsky (1982) and Koopman and Sportiche (1983), a parasitic gap is a base-generated empty category, which behaves just like true *wh*-trace with respect to the binding theory and the definition of a variable. Hence, a parasitic gap may not be A-bound by the licensing trace without creating a configuration that is formally equivalent to SCO. This explains why the licensing trace may not occur in a position that c-commands the parasitic gap (such as the matrix subject position) in examples otherwise comparable to (21):

[4] Sportiche (1983, 166) has already observed that the type of the operator is relevant to WCO. We discuss his proposals in footnote 11.

(22) a. *Who$_i$ t$_i$ gossiped about you [despite your having vouched for e$_i$]
b. *Which man$_i$ t$_i$ looked at you [after Mary had spoken to e$_i$]

Assuming that the adjunct phrase is external to VP, the VP-internal trace in (21) does not A-bind the parasitic gap, but the subject trace in (22) does, leading to the SCO-like effect.[5] Thus, the licensing trace must not c-command the parasitic gap, and vice versa. But then the *wh*-phrase must locally Ā-bind both its own trace and the parasitic gap, leading to a violation of the Bijection Principle, just as in WCO configurations. It is this violation to which Chomsky (1982) and Koopman and Sportiche (1983) attribute the somewhat marginal status of parasitic gap constructions. But this implies that parasitic gap constructions should have the same status as WCO violations, contrary to fact; we judge parasitic gap constructions like (21) to be considerably more acceptable than standard WCO configurations.

On the other hand, Contreras (1984) and Chomsky (1986a) argue that parasitic gaps are really traces derived by the movement of a null operator (NO) within the adjunct clause containing them. The chief argument for this theory, which we find persuasive, is that a parasitic gap may not occur in a Subjacency island within the adjunct clause, as first noted by Kayne (1983; 1984).[6] Assuming that Subjacency is a condition on movement, this suggests that a parasitic gap is really the trace of a NO that has undergone movement within the adjunct phrase.

The NO hypothesis raises complex issues about the precise nature of NO movement. In particular, the NO theory must provide a landing site for the NO and a principle to force its movement to this position. We will not go into these issues in any detail here. For the sake of concreteness, we will follow Chomsky (1986a, 65) in assuming that the NO must adjoin to the adjunct PP prior to S-Structure. At this point we do not wish to commit ourselves on the issue of what principle forces this movement. Perhaps, as Chomsky suggests, the NO must be 0-subjacent to the tail of the chain of the matrix *wh*-phrase in order to undergo a process of Chain Composition; alternatively, the NO may be an Ā-anaphor that must be bound by the matrix *wh*-phrase within a local binding domain, as in Aoun and Clark (1985). Either way, the parasitic gaps in (21) are locally Ā-bound by the NO rather than by the overt *wh*-phrase in the matrix clause. Thus, each operator (the overt *wh*-phrase and the NO) binds exactly one trace, and the existence of the parasitic gap does not lead to a violation of the Bijection Principle, so the marginality of parasitic gaps must be due to some other factor.

Now consider the examples in (23), where the adjunct clauses contain a bound pronoun in addition to the parasitic gap:

[5] As Barss (1986) has pointed out, examples such as (i) show that, regardless of where adjuncts are adjoined, the "real" trace need not c-command the parasitic gap:

(i) Which book$_i$ did John decide to tell his secretary [t$_i$ was unavailable] [before reading e$_i$]

See the discussion of (74) in section 4.4 below.

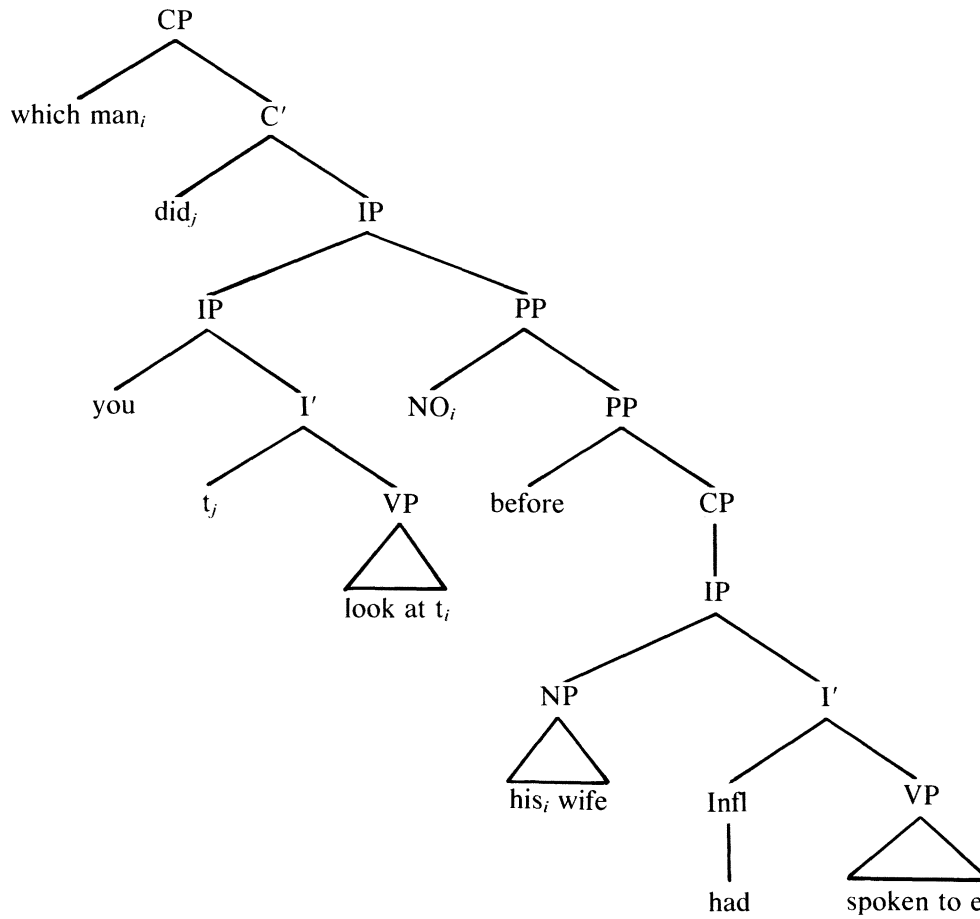
[6] Kayne assumed that parasitic gap constructions are not derived by movement, and therefore attributed these island effects to a condition on representations (his Connectedness condition on proper government).

(23) a. Who$_i$ did you gossip about t$_i$ [despite his$_i$ teacher's having vouched for e$_i$]
b. Which man$_i$ did you look at t$_i$ [before his$_i$ wife had spoken to e$_i$]
c. Who$_i$ did Leonard sleep with t$_i$ [in order for his$_i$ son to get hired by e$_i$]

These structures are problematic for both generalizations about WCO discussed above, regardless of whether a null operator analysis is assumed. Suppose first that there is no NO. Then the matrix *wh*-phrase locally Ā-binds three variables (the licensing *wh*-trace, the bound pronoun, and the parasitic gap) and the Bijection Principle and the generalization (18) are doubly violated. But these examples are no worse than their counterparts in (21).

Suppose instead that we adopt the null operator analysis. Then (23b) will have the S-Structure representation in (24):

(24)



In (24) the matrix *wh*-operator locally Ā-binds just one variable in an A-position (its trace). But the NO locally Ā-binds two variables: the parasitic gap and the pronoun coindexed with it. The same is true of the other weakest crossover examples in (23).

Although parasitic gaps are only marginally acceptable, especially when the adjuncts are tensed clauses with overt subjects, there is no WCO effect in these constructions; the bound reading of the pronoun is as freely available as the referential reading. So it seems that the NO is permitted to locally Ā-bind both a pronoun and a trace (the parasitic gap), contrary to (18); and the pronoun is free to occur within an NP that c-commands the parasitic gap, contrary to (19).

Parasitic gaps also occur in argument NPs, with marginal acceptability, in structures such as (25) and (26):

(25) a. Which man$_i$ is [everyone who asks the boss to talk about e$_i$] usually interested in t$_i$
b. Who$_i$ did [Mabel's stories about e$_i$] annoy t$_i$
(26) a. Which man$_i$ is [everyone who asks his$_i$ wife to talk about e$_i$] usually interested in t$_i$
b. Who$_i$ did [his$_i$ mother's stories about e$_i$] annoy t$_i$

These constructions raise the same problem for WCO theory, regardless of whether a NO analysis is assumed, since the bound reading of the pronoun is freely available in the examples in (26), which are no worse than their counterparts in (25).

If there is no NO, the Bijection Principle and the generalization (18) are doubly violated, since the overt *wh*-phrase locally Ā-binds three variables: the pronoun, the parasitic gap, and the licensing trace. These structures also doubly violate the generalization (19), since the bound pronoun is contained within an argument NP that c-commands the parasitic gap, and it is also contained within a larger NP that c-commands the licensing trace. If there is a NO, it locally Ā-binds both the pronoun and the parasitic gap, and (18) is violated. And the bound pronoun occurs in an NP that c-commands the parasitic gap, so (19) is also violated. Thus, both types of parasitic gap construction seem to be immune to WCO effects, contrary to the expectations of the generalizations in (18) and (19). They constitute our first case of weakest crossover.

2.2. Tough-*Movement and* Too/Enough-*Movement*

WCO effects also fail to appear in *tough*-movement constructions. Here, essentially following Chomsky (1977; 1981), we assume that a NO moves to the Comp position of the infinitival complement of the *tough* adjective:

(27) a. John$_i$ should be easy for you$_j$ [NO$_i$ [PRO$_j$ to love e$_i$]]
b. John$_i$ will be hard [NO$_i$ [PRO to persuade Mary$_j$ [PRO$_j$ to vouch for e$_i$]]]

Once again the NO seems not to induce a WCO effect:

(28) a. John$_i$ should be easy for [his$_i$ wife] [NO$_i$ [PRO to love e$_i$]]
b. John$_i$ was hard [NO$_i$ [PRO to persuade his$_i$ boss [PRO to vouch for e$_i$]]]

In (28a) the NP containing the bound pronoun might, as indicated, be the object of *for* in a matrix PP, rather than the subject of a *for*-infinitive; see Chomsky (1973), Fiengo

and Lasnik (1974). If so, it is locally A-bound by the subject of the main clause and does not pose a problem for WCO theory. But in (28b) the NP containing the bound pronoun is clearly within the infinitive, and the NO locally Ā-binds both it and the trace. But there is no WCO effect.

The same is true of the NO in complements of *too* and *enough*, which also involve NO movement (see Chomsky (1977)):

(29) a. This book$_i$ was too obscene [NO$_i$ [PRO to have [its$_i$ author] publicize e$_i$]]
b. John$_i$ isn't old enough for us [NO$_i$ [PRO to ask [his$_i$ wife] to give up e$_i$]]

Here too the NO locally Ā-binds both its trace and the pronoun *his* or *its*, in violation of (18) and (19).

In these constructions the antecedent of the NO is a referring NP in an argument position. Therefore, one might be tempted to dismiss these examples as irrelevant to the theory of WCO, by claiming that the "bound" pronoun is really functioning as a referential expression that simply happens to corefer with the antecedent of the NO. But since the NO must be coindexed with its referential antecedent by hypothesis, this claim will only eliminate the WCO violation if it is assumed that coreferential NPs need not be coindexed. Reinhart (1981) proposes a theory of coreference that incorporates this latter assumption. This theory still needs to account for the usual noncoreference effects that arise in structures where a pronoun c-commands an R-expression or an R-expression c-commands a pronoun in its governing category, as in (30):

(30) a. He said that Mary likes John
b. John believes him to be clever

In such environments Reinhart's theory relies on the interaction of various syntactic and pragmatic principles, the details of which need not concern us here. But even Reinhart's theory assumes that coindexation is a precondition for bound variable construal; thus, whereas pronouns may "accidentally" corefer with referential NPs, this type of coreference is not an option with (nonreferential) QPs.

But no WCO effect arises in *tough*-movement and *too/enough*-movement constructions, even when the matrix subject is a variable bound by a *wh*-phrase or quantifier, and the pronoun is interpreted semantically as a bound variable, as in (31):

(31) a. Which boy$_i$ [t$_i$ will be easy [NO$_i$ [PRO to persuade his$_i$ boss to vouch for e$_i$]]]
b. Nobody$_i$ should be impossible [NO$_i$ [PRO to ask his$_i$ wife to talk about e$_i$]]
c. Every book$_i$ will be easy for you [NO$_i$ [PRO to persuade its$_i$ author to publicize e$_i$]]
d. No child$_i$ is old enough for us [NO$_i$ [PRO to ask [his$_i$ mother] to give up e$_i$]]

Since these QP antecedents do not themselves refer, the possibility of "accidental" coreference is excluded. Indeed, these are the same QPs that trigger WCO effects in

the standard contexts discussed in section 1. But the pronouns in (31) can be construed as bound variables, and on this reading they must be syntactically bound by the NOs and their QP antecedents. Despite the violation of (18) and (19), there is no WCO effect.

The bound reading is also possible with sloppy identity under VP-deletion, which also requires syntactic binding of the pronoun in the antecedent of the VP gap:[7]

(32) a. John$_i$ will be easy [NO$_i$ [PRO to get his$_i$ wife to vouch for e$_i$]], and Bill will (be) [e] too
b. Darren$_i$ was not old enough for us [NO$_i$ [PRO to ask his$_i$ mother to give up e$_i$]] and neither was Ricky [e]

Thus, *tough*-movement and *too/enough*-movement constructions, like parasitic gap constructions, are problematic for the theory of WCO.

2.3. *Topicalization*

A third construction where expected WCO effects are absent is topicalization, where, according to Chomsky (1977; 1981), the Topic binds a null *wh*-phrase that has moved to Comp:

(33) a. John$_i$ [NO$_i$ [I believe his$_i$ mother loves e$_i$]]
b. This book$_i$ [NO$_i$ [I would never ask its$_i$ author to read e$_i$]]

It is harder to exclude an account based on coreference (as distinct from variable binding) in this case, since *wh*-phrases and other QPs do not occur as Topics. However, we can appeal to the "sloppy identity" test (that is, the bound variable reading under VP-deletion) in examples like (34):

(34) This book$_i$ [NO$_i$ [I would never ask its$_i$ author to read e$_i$]], but that book$_j$ [NO$_j$ [I would ____]]

Assuming that the bound variable reading of the pronoun is a reliable diagnostic of true syntactic binding, it seems that topicalization raises the same problem for the Generalized WCO Hypothesis, as embodied in (18) and (19).

2.4. *Weakest Crossover*

The three constructions discussed in this section all fail to exhibit WCO effects in environments where the Generalized WCO Hypothesis predicts that they should occur. In the next section we examine a fourth construction where WCO effects are either absent or obscured: relative clauses. We review Chomsky's (1982) analysis of the binding relations in relative clauses, which enables him to explain away the apparent violations of the generalization (18) and thereby preserve the Generalized WCO Hypothesis. We

[7] See Lasnik (1976), Sag (1976), and Reinhart (1983) for arguments that sloppy identity involves variable binding.

then show that this sort of approach cannot be extended to the full set of weakest crossover environments.

3. An LF′ Approach to Weakest Crossover

3.1. The LF′ Theory

Chomsky (1982) indicates that pronouns in potential WCO contexts within relative clauses can be construed as coreferential with the head noun—that is, with no WCO effect:

(35) a. the man$_i$ [who$_i$ [his$_i$ mother loves t$_i$]]
b. the book$_i$ [which$_i$ [its$_i$ author read t$_i$]]

Although each *wh*-phrase in (35) appears to locally bind both a pronoun and a trace, Chomsky suggests that this is not a true counterexample to the Bijection Principle. He constructs a novel analysis of the binding relations in relative clauses to accommodate these facts, based on the idea that the pronouns in (35) are not coindexed with the moved *wh*-phrases at the level of LF, where the Bijection Principle applies. To account for the intuition of coreference, he posits a distinct level of representation (LF′) where the pronoun and the *wh*-phrase *are* coindexed, and a rule of Predication that brings about this reindexation.

Like Higginbotham (1980) and Safir (1986), we disagree with Chomsky's empirical claim that WCO effects are fully absent in restrictive relatives, even in examples like (35). However, with appositive relative clauses, we (and the literature) are in full agreement with Chomsky's judgment that there is no WCO effect:

(36) a. Gerald$_i$, who$_i$ his$_i$ mother loves t$_i$, is a nice guy
b. This book$_i$, which$_i$ its$_i$ author wrote t$_i$ last week, is a hit

Thus, appositive relatives also exhibit "weakest crossover."

Although Chomsky's theory of relative clause reindexing at LF′ was designed with restrictive relatives in mind, we can transpose it from restrictives to appositives and evaluate it in that context without regard to the status of the judgments. In fact, Safir (1986) has pursued exactly this approach to appositives with some success. Hence, it is worth looking more closely at the mechanical details of the LF′ theory, in order to determine whether it might be extended to account for the cases of weakest crossover discussed in the previous section.

Chomsky's LF′ theory holds that the *wh*-phrase in a relative clause is free to bear any index at D-Structure, S-Structure, and LF, so that it need not be coindexed with the head NP (or head noun). Ultimately, the *wh*-phrase and its trace must be coindexed with the head NP, but this is achieved by a late rule of Predication, which derives LF′ from LF. Predication can be thought of as a rule that equates the index of the head NP with that of the *wh*-phrase and its trace. Thus, although the indexing in (36) might be necessary at LF′, the indexing at S-Structure and LF might look quite different, so that

the LF′ representation (36b) could correspond to the LF representation (37a) or (37b), where the head NP *this book* and the relative *wh*-phrase *which* bear distinct indices:

(37) a. This book$_i$, which$_j$ its$_j$ author wrote t$_j$ last week, is a hit
b. This book$_i$, which$_j$ its$_i$ author wrote t$_j$ last week, is a hit

In order for a pronoun in a relative clause to be construed as a variable bound by the head NP, it must be bound by the head NP at LF′. Since LF′ is derived from LF by the rule of Predication, this implies that the bound pronoun in question must be bound at LF either by the head NP or by the *wh*-phrase. If *its* were bound by the *wh*-phrase at LF, as in (37a), the Bijection Principle would be violated and a WCO effect would be expected, since *wh* would locally bind both the pronoun and the trace. But if *its* is coindexed with the head NP at LF, as in (37b), then the Bijection Principle is satisfied, since *wh* locally binds a unique variable, namely, its trace. Thus, Chomsky's LF′ theory attributes the absence of WCO effects in relative clauses like (36b) to the availability of the indexing in (37b) at LF, where the Bijection Principle applies.[8]

3.2. *Limitations of the LF′ Approach*

3.2.1. How the Theory Would Be Extended. For the sake of discussion, let us suppose that the LF′ analysis of binding relations in relative clauses is correct. The crucial aspect of the LF′ theory is its claim that WCO effects disappear when a pronoun that appears to be (locally) bound by an operator is really bound by some external NP with a distinct index, such that the index of the external NP is later equated with the index of the operator by the Predication rule deriving LF′. A successful extension of the theory to *tough*-movement, topicalization, and parasitic gap constructions will depend on the existence of an external NP in each case that is plausibly linked to the NO by Predication.

In the case of parasitic gap constructions such as those in (23), the NO would be coindexed with the licensing *wh*-phrase at LF′ and would be free to bear a distinct index at LF. The pronoun would be coindexed with the licensing *wh*-phrase and its trace rather than with the NO and the parasitic gap. Thus, at S-Structure and LF (23b) would be indexed as in (38):

(38) Which man$_i$ did you look at t$_i$ [NO$_j$ [before [his$_i$ wife had spoken to e$_j$]]]

Then no WCO effect would be expected, since the NO would bind just one variable (the parasitic gap). The pronoun *his* would still be construed as a variable, by virtue of being bound by the matrix *wh*-operator at S-Structure and LF (and by the NO at LF′).[9]

[8] Safir (1986) adopts Chomsky's theory, but revises it to apply only in the case of appositives. To force WCO effects in restrictive relatives, Safir assumes that the head NP must bind a restrictive relative *wh*-pronoun at LF. He proposes that appositives are not attached to the syntactic structure of the sentence until the level of LF′, at which point Predication applies. The PCOB is inapplicable at LF′ by hypothesis, on analogy with Chomsky's Bijection Principle–based account.

[9] This account would have to assume that the licensing trace c-commands the adjunct phrase containing the pronoun and the parasitic gap; otherwise, a WCO violation would arise with respect to the matrix *wh*-phrase in (38), which would locally Ā-bind the licensing *wh*-trace and the pronoun.

The LF′ analysis can be applied analogously to the other contexts of weakest crossover. In a *tough*-movement or *too/enough*-movement construction, the NO in the clausal complement would be coindexed with the controlling matrix subject at LF′, while bearing a distinct index at LF. In a topicalization structure, the Topic NP would not be coindexed with the null *wh*-operator until LF′. In each case, the pronoun would be coindexed with the external antecedent rather than with the NO at LF, thereby avoiding a WCO violation.

3.2.2. Tough-*Movement and Reconstruction.* However plausible this extension of the LF′ approach might appear a priori, a more detailed examination of the theory's consequences reveals that it cannot succeed in at least two (and possibly three) of the configurations at issue. Interestingly, a different type of difficulty arises in each case. We will begin with adjectival complement structures, but we must first make a brief detour into the theory of reconstruction and its relation to apparent violations of the binding theory.

Williams (1986) and Barss (1986) note that an anaphor in a moved *wh*-phrase may be bound by the subject of any clause intervening between the *wh*-phrase and its trace:

(39) a. Which picture of himself$_i$/herself$_j$ did John$_i$ say t′ that Mary$_j$ bought t
b. Which of each others'$_{i/j}$ friends did [Joe and Frank]$_i$ think t′ that [Sue and Sally]$_j$ visited t

Anaphors must normally be bound by an NP within their governing category. The data in (39) can be accounted for if it is assumed that the binding theory applies at a level (presumably LF) where all or part of the *wh*-phrase containing the anaphor may be reconstructed to any position in the chain headed by the *wh*-phrase—that is, either to position *t* or to position *t′* in (39).

Chomsky (1986a), citing Kearney (1983), observes that a matrix *wh*-phrase does not exhibit reconstruction effects with respect to the position occupied by the parasitic gap:

(40) a. Which books about himself$_i$ did John$_i$ file t before Mary$_j$ had read e
b. *Which books about herself$_j$ did John$_i$ file t before Mary$_j$ had read e

This suggests that the parasitic gap is not part of the chain headed by the matrix *wh*-phrase, at least at the level where reconstruction applies. This is consistent with our extension of the LF′ theory, which maintains that the parasitic gap is not even coindexed with the matrix *wh*-phrase at S-Structure, let alone part of the same chain.

But in other environments of weakest crossover, this is not the case. In particular, *tough*-movement constructions license reconstruction, as shown by (41):

(41) a. One picture of himself$_i$ should be easy for John$_i$ to draw t
b. Twenty pictures of himself$_i$ would be too much to expect John$_i$ to draw t in one day

The fact that reconstruction is possible in (41) suggests that the matrix subject NP in these constructions serves as the head of the chain containing the trace of the NO in

these constructions. But this is only possible if the matrix subject is coindexed with the NO and its trace when reconstruction applies.

This potentially contradicts our extension of the LF′ analysis of weakest crossover, which holds that the NO and the matrix subject may bear distinct indices prior to LF′. To accommodate (41), one would have to assume that the matrix subject has two options available at S-Structure and LF: it can be coindexed with the NO and function as the head of the extended Ā-chain (in which case reconstruction is possible); or it can bear a distinct index until LF′ (in which case it has the potential to license weakest crossover). This would predict that reconstruction effects and weakest crossover should be incompatible. But (42) shows that this is not the case:

(42) [Which picture of herself$_i$]$_j$ t$_j$′ was easy for John [NO$_j$ [PRO to persuade its$_j$ previous owner to let Mary$_i$ buy t$_j$]]

Here the anaphor behaves as though the NP containing it has been reconstructed to the position of t_j. This should be possible only if the matrix subject is coindexed with the NO and its trace, since the matrix subject is itself the trace of the *wh*-phrase containing the anaphor. But then there should be a WCO effect, since the Bijection Principle would be violated in (42) if the NO and the matrix subject are coindexed. Judgments are delicate here, but we do not detect a WCO effect in (42).

3.2.3. Topicalization without Null Operators. The LF′ analysis depends on the existence of an external antecedent for the pronoun to be coindexed with. In most weakest crossover constructions, this aspect of the analysis is unproblematic. In (appositive) relative clauses, the head NP is a plausible candidate. In *tough*-movement constructions and parasitic gap constructions, the existence of the NO is solidly motivated, and the matrix subject or *wh*-phrase is a plausible external binder for the pronoun.

But in topicalization constructions, it is not obvious that there is a null *wh*-type operator involved. Lasnik and Saito (in preparation) argue that under certain circumstances topicalization simply involves adjunction to IP, as in Baltin (1982), with no *wh*-operator involved. In *tough*-movement constructions, the existence of the NO is motivated by the fact that the trace behaves as though it were Ā-bound, and the NO is the only possible Ā-binder for the trace. But this argument does not carry over to topicalization, since the topicalized NP is itself in an Ā-position, and it can function as the Ā-binder of the trace.[10]

If Lasnik and Saito's analysis is correct, then the LF′ theory cannot account for the absence of WCO effects with topicalization, as in (34), repeated here without the NO as (43):

(43) [$_{IP}$ This book$_i$ [$_{IP}$ I would never ask its$_i$ author to read e$_i$]], but [$_{IP}$ that book$_j$ [$_{IP}$ I would ____]]

The Extended WCO Hypothesis would predict a WCO effect here, since the Topic is

[10] See Baltin (1982) and Lasnik and Saito (in preparation) for arguments that topicalization can (and sometimes must) involve IP adjunction of the Topic rather than movement of a NO.

itself the potential local Ā-binder of the pronoun, and there is no external NP with a distinct index to serve as an independent antecedent of the pronoun. Because the LF′ theory attributes the absence of WCO effects to the existence of an external binder, it cannot be extended to account for weakest crossover in topicalization without abandoning the IP-adjunction analysis.

3.2.4. Parasitic Gap Licensing. The LF′ theory encounters a different sort of problem with parasitic gaps, which is even more devastating. As Chomsky (1982) observes, parasitic gaps must be licensed at the level of S-Structure; LF licensing is insufficient. Thus, (44a) is well formed, but (44b–c) are not—despite the fact that these examples involve Ā-binding by quantifiers at LF, as in (45):

(44) a. What$_i$ did you say John filed t$_i$ without reading e$_i$
b. *Who$_j$ did you say t$_j$ filed what$_i$ without reading e$_i$
c. *John filed no book$_i$ without reading e$_i$

(45) a. [What$_i$ [who$_j$ [you said t$_j$ filed t$_i$ without reading e$_i$]]]
b. [No book$_i$ [John filed t$_i$ without reading e$_i$]]

Evidently the NO in a parasitic gap construction must be licensed by an Ā-binding configuration at S-Structure. Furthermore, the *wh*-operator licensing the NO and the parasitic gap must bind the NO. This is shown by (44b), where the matrix clause contains a *wh*-phrase c-commanding the NO at S-Structure. (44b) is bad precisely because the c-commanding *wh*-phrase is not coindexed with the NO and parasitic gap.

In order to capture the S-Structure licensing condition, we must assume that NO (and the parasitic gap) must be Ā-bound by the *wh*-phrase at S-Structure. This is incompatible with the LF′ theory of weakest crossover, which crucially assumes that the two operators are not coindexed prior to LF′. The only way that the LF′ theory could allow for weakest crossover parasitic gap structures to satisfy both the S-Structure licensing condition and the Bijection Principle would be to assume that the two operators are coindexed at S-Structure, then contraindexed at LF by some new mechanism, and then, finally, coindexed again at LF′ via Predication. This hardly seems plausible.

We submit that our extension of Chomsky's LF′ theory of weakest crossover, though initially plausible, is not correct. More generally, we believe that any approach to weakest crossover that relies on a mechanism of reindexation (that is, equation of formerly distinct indices) is doomed to a similar fate. Hence, we will attempt to construct a solution based on the assumption that the offending pronouns are coindexed with the operators and their traces throughout the derivation—including the level where the principle responsible for WCO effects applies. This is the topic of the next section.

Our analysis relies neither on the existence of an external binder for the pronoun nor on the assumption that there is a level of LF′ distinct from LF. Instead, we relate the occurrence of WCO effects to an intrinsic semantic property of the quantifier that locally Ā-binds the pronoun and the trace. Although our theory is based primarily on the weakest crossover constructions discussed in section 2, it extends naturally to ac-

count for the behavior of relative clauses vis-à-vis WCO. If our theory is right, it undermines the primary motivation for the level of LF′, which would no longer be needed to account for the distribution of WCO effects.

4. Two Kinds of Operators

4.1. True Quantifiers and Weak Crossover

So far, we have discussed four contexts of weakest crossover. As we have shown, the unexpected absence of WCO effects in these constructions cannot be adequately explained away in terms of "accidental" coreference or equation of indices at LF′. Assuming, then, that a bound pronoun is coindexed throughout the derivation with the operator that binds it—surely the null hypothesis—we must seek another explanation for the absence of WCO effects.

As a point of departure, consider again the contrast between simple structures of quantification in (46) and *tough*-movement constructions with quantified antecedents in (47):

(46) a. *Mary asked me [who$_i$ [PRO to persuade his$_i$ mother$_j$ [PRO$_j$ to vouch for t$_i$]]]

b. *[Nobody$_i$ [his$_i$ mother vouched for t$_i$]]

(47) a. Who$_i$ t$_i$ should be easy [NO$_i$ [PRO to persuade his$_i$ mother$_j$ [PRO$_j$ to vouch for e$_i$]]]

b. Nobody$_i$ should be easy [NO$_i$ [PRO to persuade his$_i$ mother$_j$ [PRO$_j$ to vouch for e$_i$]]]

These examples all exhibit the same structural relations between the pronoun and the trace of the operator (and between the operator and each of them), so these factors cannot be responsible for the contrast between WCO in (46) and weakest crossover in (47). The nature of the ultimate antecedent of the pronoun is also irrelevant, since the same quantifiers that give rise to WCO in (46) fail to do so in (47).

The most obvious difference between (46) and (47) is the null versus overt status of the operator that locally binds the pronoun and trace. But this difference cannot provide the basis for a generalized account of weakest crossover, since overt *wh*-phrases in questions trigger WCO effects, whereas their overt counterparts in appositive relative clauses do not. The contrast between appositives and restrictives shows this even more strikingly; in appositive relatives the *wh*-phrase is obligatorily overt and there is no WCO effect, whereas in restrictive relatives the *wh*-phrase is optionally null but there is still a mild WCO effect, at least for some speakers (including us). Moreover, there are certain types of parasitic gap constructions with NOs that do trigger WCO effects; we will look at these shortly.

The only factor that correlates almost perfectly with the distribution of WCO effects is the intrinsic quantificational status of the local Ā-binder of the pronoun. A WCO effect seems to occur just when the pronoun and trace are locally Ā-bound by a true QP (or

by a trace of a true QP). If the local Ā-binder is either a referential NP (topicalization) or an operator bound by an external antecedent (appositive relatives, *tough*-movement constructions, and parasitic gap constructions), then there is no WCO effect.[11]

Before going any further, we should be more explicit about what we mean by a "true Quantifier Phrase." For our purposes, a true QP is composed of a quantifier Q and a nominal term T defining a range R that Q quantifies over, such that R is a possibly nonsingleton set. For instance, in the true QP *which man*, Q is *which*, T is *man*, and R is a set of two or more men; in the true QP *everyone*, Q is *every*, T is the morpheme *one*, and R is a set of three or more human individuals. (The precise size of the sets in R is subject to various grammatical and pragmatic principles that need not concern us here.)

In contrast, the local Ā-binders in the weakest crossover examples in (47) have a radically different status. First consider the status of the NO in a typical *tough*-movement structure:

(48) John$_i$ was easy [NO$_i$ [PRO to please t$_i$]]

The NO$_i$ in (48) does not range over any nonsingleton set; its reference is fixed to that of its antecedent—in this case, John.[12] The same is true of topicalization structures, regardless of which analysis of these structures is assumed:

(49) a. [John$_i$ [we like t$_i$]]
b. [John$_i$ [NO$_i$ [we like t$_i$]]]

In the IP-adjunction analysis, the local Ā-binder is not a true QP, since it is simply a referring NP, containing no Q. In the NO analysis of topicalization, the NO contains no Q with any semantic content, and its range R is equivalent to the denotation of its antecedent, the referring NP.

The status of the NO in parasitic gap constructions is not immediately transparent. We want to say that the NO is a logically inert operator, just like the NO in *tough*-movement structures. But there is another possibility that must at least be considered.

[11] The idea that the intrinsic semantic status of the operator is relevant for the binding properties of its trace is not unprecedented. Kayne's (1984) analysis of subject clitics in French complex inversion assumes that A-positions and Ā-positions are sometimes distinguished in terms of the intrinsic properties of the element occupying the positions.

A more direct precursor of our analysis is Sportiche's (1983) treatment of certain instances of what we have called weakest crossover, that is, the absence of WCO effects in relative clauses and clefts. Sportiche captures this by restating the Bijection Principle so as to be sensitive to whether the Ā-binder is an operator or not. Although he does not provide a pretheoretical definition of the notion "operator," it serves a role in his analysis that is analogous to that served by our term "true quantifier." His analysis differs from the one that we will offer in section 4.2, however, in not distinguishing between the locally Ā-bound trace of an "operator" and that of a "nonoperator"; both are treated as syntactic variables. (The difference is attributed instead to the fact that the Bijection Principle only applies to operators.)

[12] The situation is a bit more complicated in *tough*-movement structures where the matrix subject is a variable bound by a true QP, as in (i):

(i) Who$_i$ t$_i$ was easy [NO$_i$ [PRO to criticize e$_i$]]

The situation here is essentially parallel to that obtaining with parasitic gaps, which we discuss shortly. The basic idea is that the trace of the NO functions in this case as a covariable, or dependent variable.

Hypothetically, one might think of the NO as having the same logical status as its antecedent (the matrix *wh*-phrase). On this view, the NO would have the same Q and R as its antecedent, and by this logic it might count as a true QP, binding its own variable (the parasitic gap). In other words, what is at issue here is whether parasitic gap constructions contain a single true QP at LF, or two identical true QPs.

We can dispose of this alternative view by considering the meanings of these constructions. Compare the NOs binding parasitic gaps in (50) with the overt *wh*-phrases in (51), involving pairs of identical *wh*-phrases:

(50) a. Which man$_i$ does everyone [NO$_i$ [who meets e$_i$]] admire t$_i$
b. Who$_i$ did he hire t$_i$ [NO$_i$ [without talking to e$_i$]]
(51) a. Which man$_i$ [t$_i$ wonders [which man$_j$ his$_i$ mother vouched for t$_j$]]
b. Who$_i$ [t$_i$ expects [his$_i$ mother to visit who$_j$]]

All these examples involve pairs of identical operators with the same values for Q, T, and R. But each *wh*-phrase in (51) quantifies independently over R; the value assigned to the trace of one *wh*-phrase need not covary with the value assigned to the trace of the other. In contrast, the value assigned to the parasitic gaps in (50) must covary with the value assigned to the licensing trace; thus, the only possible type of answer to either question in (50) is one where the parasitic gap and the licensing trace have the same value. Thus, the interpretation of parasitic gap constructions supports the view that the NO is logically equivalent to the vacuous NO in a *tough*-movement construction such as (52) rather than to the independent *wh*-operators in (51):

(52) Who$_i$ t$_i$ was easy [NO$_i$ [PRO to please e$_i$]]

Thus, the absence of WCO effects with parasitic gap constructions conforms to our generalization, as does the presence of WCO effects in structures with two identical but independent *wh*-phrases parallel to (51):

(53) a. *Which man$_i$ [t$_i$ wonders which man$_j$ his$_j$ mother vouched for t$_j$]
b. *Who$_i$ [t$_i$ expects [his$_j$ mother to visit who$_j$]]

Interestingly, there is a dialect of English that allows parasitic gaps in certain types of NPs to have a multiple question interpretation analogous to (51), as discussed in Clark (1983) and Stowell (1985). Thus, (54a–b) are ambiguous in this dialect; they can be interpreted either as "normal" parasitic gap constructions where the parasitic gap covaries with the licensing trace, or as multiple questions, anticipating (55a–b) as possible answers:

(54) a. Who did [Mary's stories about e] amuse t
b. Who did you give [your pictures of e] to t
(55) a. Mary's stories about Jenny amused Jim
b. I gave my pictures of Sally to Sue

The structural conditions governing the availability of this reading, discussed in Stowell (1985), need not concern us here. On the "normal" reading, these parasitic gaps behave just like their counterparts in (50); the NO is semantically vacuous. But on the multiple question reading, the NO has essentially the same logical status as the overt wh_j in (51), and it is thus appropriate to analyze the NO as a true QP on this interpretation.[13]

As our analysis predicts, there is no WCO effect on the "normal" parasitic gap interpretation; the pronoun and the parasitic gap are locally bound by a NO that is not a true QP:

(56) a. Who$_i$ did [his$_i$ mother's stories about e$_i$] amuse t$_i$
b. Who$_i$ did you give [his$_i$ friends' pictures of e$_i$] to t$_i$

But there is a WCO effect on the multiple question interpretation; thus, *his*$_i$ cannot be construed as a bound variable in (57a–b) if these questions anticipate answers like (58a–b):

(57) a. *Who$_j$ did [NO$_i$ [his$_i$ mother's stories about e$_i$]] amuse t$_j$
b. *Who$_j$ did you give [NO$_i$ [his$_i$ friends' pictures of e$_i$]] to t$_j$
(58) a. John$_i$'s mother's stories about him$_i$ amused Bill
b. I gave Sam$_i$'s friends' pictures of him$_i$ to Bob

Thus, the distribution of WCO effects with parasitic gaps correlates precisely with the logical status of the operator that locally binds the parasitic gap; the NO induces a WCO effect only if it is interpreted as a true QP.

Next consider appositive and restrictive relative clauses like those in (35a) and (36a), repeated here:

(35) a. %the man$_i$ [who$_i$ [his$_i$ mother loves t$_i$]]
(36) a. Gerald$_i$, [who$_i$ [his$_i$ mother loves t$_i$]]

First consider the appositive in (36a). Clearly, the *wh*-phrase that locally binds the pronoun here is not a true quantifier. In this respect we agree with Jackendoff (1977), who treats the appositive *wh*-phrase as a pronoun coreferential with the head NP. Thus, the absence of WCO effects accords with our expectations.

The situation is rather different with the restrictive relative clause in (35a). We have annotated this example with a % to indicate the dispute in the literature over its WCO status, but we will assume that there is a WCO effect here, as noted previously. The situation is clearer with respect to restrictive relatives with QP heads:

(59) a. *[every man$_i$ [that his$_i$ mother rejected t$_i$]]
b. *[no house$_i$ [that its$_i$ owner sold t$_i$]]

Why are there WCO effects in restrictive relative clauses? The relation between the *wh*-

[13] As Stowell notes, the availability of this reading supports the NO hypothesis, since without the NO there would only be one *wh*-type operator, and it would be difficult to provide a logical representation of the multiple question reading.

phrase and the head NP (or QP) is not one of coreference, since the head of a restrictive relative clause has no intrinsic reference (as is obvious in (59); see Higginbotham (1980)). But this in itself does not predict a WCO effect, since one might expect the restrictive *wh*-operator to be analogous to the NO in a parasitic gap or *tough*-movement construction. After all, the restrictive *wh*-operator is arguably bound by an external antecedent (the head NP or QP) and it has sometimes been suggested that it is a vacuous operator with a logical translation equivalent to "such that" (see Chomsky (1986a, 85)). Thus, the occurrence of WCO effects with restrictive relatives is, a priori, potentially problematic for our account of weakest crossover.

But we are not convinced that the *wh*-operator in a restrictive relative clause is logically vacuous. On the contrary, we believe that it is a true QP in the sense defined above, even when it is phonetically null, as in (59). Recall that the head of a restrictive relative clause does not itself refer; rather, the reference of the entire NP is determined conjointly by the head noun and the relative clause. More precisely, the reference of the full NP is the intersection of the sets defined by the head noun and the relative clause. Thus, the reference of *the men who John saw* is the intersection of the set of men (denoted by the head) with the set of human individuals that John saw (denoted by the relative clause). In this sense, the restrictive *wh*-operator may be a true QP ranging independently over a possibly nonsingleton set distinct from that of the head, even though the full NP will have a unique referent if the head NP is definite and singular. Viewed in these terms, the occurrence of WCO effects with restrictive relatives is consistent with our proposal.[14]

To summarize, as suggested by Sportiche (1983), there is a logically based distinction between the structures where WCO effects arise and those where they do not: WCO effects arise only in contexts where a pronoun is locally Ā-bound at LF by a true quantifier ranging over a possibly nonsingleton set. This implies that WCO does not follow exclusively from the structural configuration of Ā-binding, contrary to the Generalized WCO Hypothesis; the logical status of the operator in the Ā-position must be taken into account.

4.2. Null Epithets

So far we have offered a descriptive generalization about the distribution of WCO effects, but we have provided no real theory about why this generalization holds. We believe that this issue can be fruitfully addressed by reconsidering the syntactic binding types of the traces involved in these constructions.

The *LGB* binding theory in (9) recognizes three major types of NPs: anaphors, which are subject to Condition A; pronominals, which are subject to Condition B; and R-expressions, which are subject to Condition C. R-expressions are subdivided into *names* and *variables*. The technical term *name* is somewhat misleading, since it includes not only proper names but also definite descriptions such as *the man*.

[14] Significantly, Montalbetti's (1984) Overt Pronoun Constraint distinguishes between restrictive and appositive relative clauses in just the way that WCO does. See appendix B.

Following common practice, we will use the term *epithet* for definite descriptions that function quasi-pronominally, that is, with linguistic antecedents. Although epithets such as *the bastard* are commonly cited in the literature, almost any simple definite NP may so function; see Stowell (1979):[15]

(60) a. I went to visit John Smith$_i$ last week, but the man$_i$ was too busy to talk to me
b. All of Bill$_i$'s friends love the guy$_i$

The intrinsic syntactic binding type of an NP does not fully determine its referential properties, since these are affected by the binding relations that the NP bears to other NPs. For example, the pronouns in (61a) are coindexed with no other NPs and therefore may refer directly, whereas the pronouns in (61b–c), which are bound by variables, are themselves construed as variables and do not refer:

(61) a. She$_i$ shot him$_j$
b. Every man$_i$ t$_i$ thinks he$_i$ is lucky
c. No president$_i$ t$_i$ expects people to hate him$_i$

Thus, an element that belongs to a single binding type (pronoun) may belong to two different types of referential categories.

The same is true of epithets. The examples in (60) involve epithets that corefer with referential NPs, but, as noted by Hornstein and Weinberg (1988), epithets may also function as nonreferential bound variables in contexts like those in (62):

(62) a. [Some barber in every midwestern city$_i$] hates the place$_i$
b. Which assailant$_i$ did Mary escape from t$_i$ before the bastard$_i$ had a chance to steal her money

In (62) the epithets function logically as bound variables, just like their pronominal counterparts (compare the "inversely linked" structures of May (1977)).

Although pronouns and epithets may both function logically as bound variables, they retain their distinctive behavior with respect to binding conditions. Thus, an epithet typically corefers with an NP that precedes it in the discourse, but that NP may not occur in an A-position that c-commands the epithet:

(63) a. *John Smith$_i$ denied that the man$_i$ was too busy to see me
b. *Bill$_i$ thinks Mary told me to visit the guy$_i$

In this respect, epithets differ from true pronouns:

(64) a. John Smith$_i$ denied that he$_i$ was too busy to see me
b. Bill$_i$ thinks Mary told me to visit him$_i$

[15] We share the misgivings of Higginbotham (1985) about the term *epithet* for this usage of definite NPs, since there is no sense in which this quasi-pronominal usage requires that the NP be interpreted as an epithet for (or comment about) the antecedent. As Higginbotham notes, even demonstratives such as *that man* may behave quasi-pronominally in this way. However, we defer to the tradition in the literature on the use of the term.

This follows if epithets, unlike pronouns, must obey Condition C.[16]

Bound pronouns and epithets both differ from QP-trace and *wh*-trace in not being subject to WCO, as shown by (65)–(66):

(65) a. I went to visit John Smith$_i$ last week, but his$_i$ secretary told me he$_i$/the man$_i$ was unavailable
b. All of Bill$_i$'s friends say his$_i$ mother loves him$_i$/the guy$_i$

(66) a. [Some tenant in every apartment building$_i$] has asked its$_i$ owner to repaint the place$_i$
b. Which assailant$_i$ did Mary escape from t$_i$ before his$_i$ partner joined up with the bastard$_i$

Now consider the status of a trace that is locally Ā-bound by an element that is not a true QP, in the sense described in the previous section. The *LGB* definition in (10) does not distinguish between the trace of a non-QP and the trace of a true QP; both are classified as syntactic variables. The fact that both types of traces exhibit SCO effects is captured in *LGB* in terms of Condition C:

(67) a. *Sam$_i$ was easy to tell him$_i$ [PRO$_i$ to make Mary visit e$_i$]
b. *Who$_i$ did you talk to t$_i$ [after he$_i$ asked you to stay with e$_i$]
c. *John$_i$, I think he$_i$ told Mary to visit e$_i$
d. *John$_i$, who$_i$ I think he$_i$ said Mary likes e$_i$, . . .

But the trace of a non-QP, unlike *wh*-trace and QP-trace, does not exhibit WCO effects. Thus, the trace of a non-QP behaves just like an epithet: it exhibits SCO effects but not WCO effects. We suggest that this similarity is not accidental, and that this type of trace represents a previously unrecognized syntactic type of empty category: a null epithet. If the principle responsible for WCO applies only to true variables, then the absence of WCO in weakest crossover contexts follows from the fact that the trace of a non-QP is an epithet rather than a variable.

Suppose that this is correct. We would then distinguish between two types of locally Ā-bound traces: null variables and null epithets. They differ from each other in that null variables are locally bound by a true quantifier and exhibit WCO effects, whereas null epithets are locally bound by an element that is not a true quantifier. Certainly a null epithet may be construed semantically as a variable if it is Ā-bound by a true QP, but this is also possible for a pronoun or overt epithet that is Ā-bound by a quantifier, as we have shown.[17]

[16] On epithets and Condition C, see Lasnik (1989), which further develops proposals of Lasnik (1976). Note that the acceptability of (62) indicates that although the epithets are bound by the quantifiers at LF, they are not bound for the purposes of Condition C. This might follow either from a difference between Ā-binding and A-binding, or, alternatively, from limiting Condition C to S-Structure.

[17] Cinque (1984; 1990) also argues that the empty categories involved in parasitic gap and *tough*-constructions are not variables. However, he treats them not as names or epithets but rather as null pronominals (pro). We discuss Cinque's theory in appendix B.

4.3. *A Typology of R-Expressions and Variables*

If epithets and variables represent distinct binding types, we would expect a broader typology of null versus overt binding types than that envisaged by Chomsky (1982), to reflect this distinction. Chomsky's typology is based on the binary features [±pronominal] and [±anaphoric]. If an empty category (EC) is [+a, −p], it is an anaphor subject to Condition A and has the distribution of NP-trace.[18] If an EC is [−a, +p], it is a pronominal subject to Condition B and has the distribution of pro. If an EC is [+a, +p], it is a pronominal anaphor subject to Conditions A and B, and has the distribution of PRO. If an EC is [−a, −p], it is subject neither to Condition A nor to Condition B; instead, it observes the "elsewhere" condition of the binding theory, Condition C.[19]

Chomsky (1982) considers eliminating Condition C from binding theory, deriving SCO effects from the definition of a variable in (10). But this leaves a residue of Condition C effects unaccounted for, especially with respect to epithets; see Higginbotham (1985), Lasnik and Uriagereka (1988), and Lasnik (1989). For this reason, we will continue to assume that R-expressions are subject to Condition C, although we will return to this issue shortly.

Chomsky's feature system does not distinguish between R-expressions and variables; both are [−a, −p]. The distinction between them is assumed to follow from the null versus overt distinction; variables are simply null R-expressions. But our proposal that null variables and null epithets must be distinguished from each other suggests that the null versus overt status of an R-expression does not invariably determine its binding type. If null elements with the features [−a, −p] subdivide into R-expressions and variables, we might expect the same to be true of overt elements, deriving the typology in (68):

(68)

	Overt categories	*Null categories*
R-expressions	Names, definite NPs	Null epithets
Variables	(Overt variables)	*Wh*-trace, QP-trace

For null elements, the distinction between variables and epithets requires a modification of the definition of a variable such as (69):

[18] The null anaphor NP-trace differs from overt anaphors in one important respect: it is not the head of an A-chain like an overt reflexive or reciprocal, but occupies a subsidiary position on an A-chain. This difference could be eliminated at LF if reflexive pronouns undergo movement to Infl as in Lebeaux (1983) and Chomsky (1986b). The reflexive and its trace could then be analyzed along the lines suggested by Bouchard (1982) for reflexive clitics in Romance, whereby the reflexive is a subject clitic and its antecedent forms a chain with the direct object position. Then the position occupied by the overt reflexive would be directly analogous to an NP-trace in a passive or raising construction.

[19] The status of null categories with respect to intrinsic feature content is unclear. Chomsky (1982) suggests that all null categories are assigned features on the basis of the syntactic environments in which they occur and that these features can be changed in the course of a derivation; this is the so-called functional theory of ECs. Brody (1984) has challenged the proposal that features may be changed in the course of a derivation and has argued that ECs are inserted in a structure with freely assigned but permanent values for [±a, ±p]. Yet another theory of ECs holds that they have intrinsic features. As far as we can tell, this dispute has no direct bearing on our proposals, so we will ignore it here. For further discussion, see Epstein (1984) and Lasnik and Uriagereka (1988).

(69) X is a variable iff:
  a. X is in an A-position, and
  b. X is locally Ā-bound by a true QP.

For null epithets, we might adopt a definition such as (70):

(70) X is an R-expression iff:
  a. X is in an A-position, and
  b. X is locally Ā-bound by a non-QP.

The disadvantage of (70) is that, unlike (69), it does not generalize to overt R-expressions, since overt epithets and names do not have to be Ā-bound by anything. If we wish to avoid a licensing condition that applies only to null epithets, we might eliminate (70) and simply treat nonvariable R-expressions as a default case, that is, as [−a, −p] elements that fail to satisfy the definition of a variable in (69).

If a [−a, −p] EC that fails to satisfy (69) is a well-formed EC (namely, a null epithet), then SCO will follow not from failure of licensing of the trace (as in Chomsky (1982)) but from Condition C. Consider a classic SCO context like (71):

(71) *Who$_i$ did he$_i$ tell Mary to visit t$_i$

Assume that the trace t_i cannot be a null anaphor or pronoun for the usual reasons; it also cannot be a variable since it is not locally Ā-bound and therefore does not satisfy (69). Suppose it is permitted to be a null epithet as the default case. Then it falls under Condition C, like other epithets, and (71) is immediately excluded.[20]

A potential problem with defining null epithets in purely negative terms (that is, as a default value for nonvariable [−a, −p] elements) is the simple fact that null epithets, unlike overt epithets, must be Ā-bound at S-Structure. Thus, we must account for the contrast between (72a) and (72b):

(72) a. I tried to visit John$_i$ last week, but I was unable to persuade the guy$_i$ to see me
  b. *I tried to visit John$_i$ last week, but I was unable to persuade e$_i$ to see me

In order to exclude (72b), we must ensure that a null epithet cannot simply be base-generated in any context where an overt epithet can occur. It is possible that the ECP is responsible for this. Suppose that an EC in a governed argument position must have its features identified, in the sense of Rizzi (1986). Since there are no agreement features that locally license the EC, its features must be identified under antecedent government by an XP. If the EC is an epithet, its antecedent may not A-bind it without violating Condition C. Therefore, the epithet must be Ā-bound by its antecedent governor. Thus, the interaction of the ECP, Condition C, and the definition of a variable (69) may be sufficient to derive the effects of the licensing condition (70).

[20] Note that (71) is not excluded by the Bijection Principle, at least under the theory of Koopman and Sportiche (1983), since the bound pronoun is locally Ā-bound by the *wh*-phrase, but the EC is not.

Thus far, we have not commented on an additional prediction: that the final quadrant of (68), the category of overt variables, should be instantiated. In fact, Koopman and Sportiche (1983) and Engdahl (1985) argue that such elements exist in languages such as Vata and Swedish. They show that in certain syntactic contexts, null variables are excluded on the basis of independent principles (such as the ECP). In these contexts, overt variables (homophonous with pronouns) occur instead. Abstracting away from ECP-type effects, these overt variables have the distributional properties of *wh*-trace; they clearly differ from resumptive pronouns of the sort that occur in left dislocation constructions in English. In particular, they pattern with *wh*-trace with respect to tests such as across-the-board extraction from coordinate structures and WCO effects. Since the properties of overt variables are peripheral to our main concern, we will not pursue this interesting issue.

4.4. The Domain of Condition C

Chomsky (1986b) suggests that the domain within which an R-expression must be A-free is not identical for R-expressions and variables. Whereas an R-expression must be A-free in the entire root clause, a variable need only be A-free within the c-command domain of the operator that binds it. *Tough*-movement constructions and purpose clauses are often cited in support of this distinction, but we will argue against this analysis.

Our view is based on the interaction of two properties of parasitic gap constructions that we have already discussed. Consider the status of parasitic gaps with respect to Condition C. As Chomsky (1982) observes, the subject-object licensing asymmetry illustrated in (73) can be explained rather straightforwardly under the assumption that some principle prevents the licensing trace from c-commanding the parasitic gap:

(73) a. Who$_i$ did you stay with t_i [without your talking to e_i]
b. *Who$_i$ t_i stayed with you [without your talking to e_i]

That c-command, as opposed to some sort of Case-agreement condition, is the relevant factor here is supported by Barss's (1986) observation that subject traces are, in fact, able to license parasitic gaps if the anti-c-command condition is respected:

(74) a. Which book$_i$ did you discover t_i [before Sam had read e_i]
b. *Which book$_i$ t_i was unavailable [before John had read e_i]
c. Which book$_i$ did John decide to tell his secretary [t_i was unavailable] [before reading e_i]

This anti-c-command requirement is, of course, strongly suggestive of a Condition C effect. Within the context of Chomsky's (1982) theory of parasitic gaps, which does not involve a NO, this is consistent with the status of the parasitic gap as a variable; the parasitic gap must be A-free within the domain of the operator binding it (namely, the matrix *wh*-phrase). On the other hand, the anti-c-command requirement is surprising from the perspective of the null operator analysis, since the NO ought to restrict the

applicability of Condition C to its own c-command domain, in particular to the adjunct clause in (73b) or (74b).

Since the NO analysis is well founded on the basis of the Subjacency effects discussed by Chomsky (1986a), we conclude that (73) and (74) argue either against the assumption that parasitic gaps are variables or against the version of Condition C in Chomsky (1986a) that provides for the "sheltering" of a variable by an operator. On the other hand, if parasitic gaps are really epithets rather than variables, then Chomsky's original (1982) account of the anti-c-command condition can be retained, even if Chomsky's (1986b) Condition C is adopted. As an epithet, the parasitic gap must be A-free in the root clause, even if it is locally Ā-bound by the NO within the adjunct clause.

There is one further consideration relevant to this conclusion. Contreras (1984) has argued against the anti-c-command condition, citing apparent Condition C effects in structures parallel to (73a) and (74a):

(75) ??Mary discovered them$_i$ [before John had read the papers$_i$]

Though we agree with Contreras that coreference is difficult in (75), we do not believe that this represents a true Condition C effect. Epithets normally observe Condition C, as shown by (76a), but they show no Condition C effects in structures parallel to (73a), (74a), and (75), as in (76b). Moreover, the position of the antecedent of an epithet shows the same subject-object asymmetry as the position of the licensing trace in a parasitic gap construction, as shown by the contrast between (76b) and (76c):

(76) a. *John$_i$ was told [that the bastard$_i$ was unwelcome]
b. Mary kicked John$_i$ [before the bastard$_i$ had a chance to grab her]
c. *John$_i$ kicked Mary [before the bastard$_i$ grabbed her]

Whatever is responsible for the noncoreference effect in (75), the possibility of coreference in (76b) casts doubt on Contreras's claim that the object necessarily c-commands the adjunct, and thus provides some support for Chomsky's original (1982) interpretation of the subject-object asymmetry in (73)–(74) as a Condition C effect.[21] If parasitic gaps are really null epithets rather than variables, it is possible to retain this analysis within the null operator theory.

Our analysis of parasitic gaps has obvious implications for the other constructions where null epithets occur. Specifically, if the trace of a NO always has the status of an epithet rather than a variable, then we ought to expect comparable Condition C effects in *tough*-movement constructions, topicalization constructions, and appositive relatives (as opposed to interrogative constructions and restrictive relatives, where the trace is defined as a variable by (69) and WCO effects show up).

In *tough*-movement constructions, this claim appears to be trivially falsified, since the local binder of the NO (the matrix subject) is in a c-commanding A-position. However, we believe that this can be explained by reconsidering the relationship between

[21] See Saito (in press) for further arguments, and for detailed analysis of apparent counterexamples.

Condition C and other binding relations. Note that in *tough*-movement constructions, the matrix subject is not just *permitted* to bind the NO and its trace; it is *required* to bind them, by the rule of Predication (or possibly control) that interprets these constructions. We suggest that in structures where Predication requires that an A-position X must bind another category Y, Condition C does not affect this binding relation. As suggested in section 4.3, Condition C is a kind of "elsewhere" condition, checking binding relations between pairs of A-positions that are not independently determined by other principles. More precisely, we suggest the principle (77):[22]

(77) If an A-position X A-binds a category Y as a result of Predication (or control), then Condition C does not apply to A-binding of Y (or its chain) by X.

Since Predication requires that the subject of a *tough* adjective must A-bind the NO in the *tough* complement, Condition C does not apply to A-binding relations between the subject (X) and the trace of the NO (the tail of the chain of Y). This draws the desired distinction between *tough* constructions and parasitic gap constructions. Although the NO in a parasitic gap construction must be $\bar{A}$-bound by a *wh*-phrase in the matrix Comp position, no principle directly requires that the NO be bound by any particular A-position; the matrix operator need not originate in any particular A-position as far as the NO is concerned. Therefore, no particular A-position is identified by (77) as being immune to Condition C effects as a binder, so it is still predicted that the licensing trace may not A-bind any R-expression, including the parasitic gap. Notice that if this approach can be maintained in general, then the original version of Condition C, as in Chomsky (1981), can be reinstated. The exemption for variables granted in Chomsky (1986b) is now seen to be incorrect in certain instances, and to be unnecessary, following from the "elsewhere" nature of Condition C, in the remaining instances.

5. Conclusion

The binding behavior of NO-trace suggests the existence of a previously unrecognized type of empty category: a null, nonvariable R-expression, corresponding to a null version of an epithet. Like epithets and other nonvariable R-expressions, this new type of empty category must be A-free in the root sentence, apart from situations where coindexation is forced by a principle. This accounts, among other things, for the fact that a parasitic gap may not be c-commanded by the licensing trace. Furthermore, like epithets and proper names, the null epithet does not exhibit WCO effects. This indicates that whatever principle is responsible for WCO effects must apply only to true syntactic variables, that is, to categories that are locally $\bar{A}$-bound by true quantifiers.

[22] Examples such as (i) seem to necessitate granting the Condition C exemption for control as well as for Predication:

(i) John$_i$ is easy for himself$_i$ [NO$_i$ [PRO$_i$ to talk to t$_i$]]

This example is perfect even though PRO binds the trace (and even with both inside the domain of the NO).

Appendix A: Clefts, Pseudoclefts, and Focus

Cleft, pseudocleft, and focus constructions exhibit a rather confusing mix of WCO data. We will briefly consider the implications of our theory for each of these constructions in turn. Clefts and pseudoclefts superficially resemble our weakest crossover constructions in lacking WCO effects in some cases:

(78) a. It was John$_i$ [who$_i$ his$_i$ mother was talking about e$_i$]
b. It was this book$_i$ [(wh$_i$) that I got its$_i$ author to read e$_i$]
(79) a. ?[Who$_i$ his$_i$ mother was talking about e$_i$] was John$_i$
b. ?[What$_i$ I got its$_i$ author to talk about e$_i$] was this book$_i$

Assuming with Chomsky (1977) that *wh*-movement applies in the embedded CP in both constructions, there is a potential WCO violation involving the bound pronoun. Our theory leads us to expect WCO effects here, given the logical function of the *wh*-operators. The cleft and pseudocleft clauses both have an interpretation resembling that of a restrictive relative clause; thus, (78a) and (79a) can be paraphrased as 'John is the (only) individual that his mother was talking about' or 'The (only) individual that his mother was talking about was John'. The trace should be a true variable, since the covert quantifier *only* ranges over a larger domain of individuals: John (and no other individual in the domain of discourse) is such that his mother was talking about him. But little if any WCO effect is discernible; *his* can be coreferential with the satellite NP *John*, which is presumably coindexed with *wh* and its trace.

A possible explanation for the absence of WCO effects here is LF reconstruction, since clefts and pseudoclefts, like topicalization structures, exhibit reconstruction effects involving Conditions A and C:

(80) a. [Who$_i$ John$_i$ saw t$_i$] was himself$_i$
b. It was himself$_i$ [(wh$_i$) that John$_i$ saw t$_i$]
c. *[Who$_i$ he$_i$ saw t$_i$] was John$_i$
d. *It was John$_i$ [(wh$_i$) that he$_i$ saw t$_i$]
Compare:
e. Himself$_i$, John$_i$ likes t$_i$
versus
f. *John$_i$, he$_i$ likes t$_i$

In (80b) *wh*-trace should violate Condition C as an instance of SCO, and the satellite reflexive should violate Condition A in its derived position. These apparent violations might be explained in terms of a reconstruction operation optionally reconstructing the satellite NP into the position of the trace. If reconstruction fails to apply, (80a–d) are all excluded as SCO violations. If reconstruction does apply, then (80c–d) are still excluded, since Condition C is violated by the reconstructed name. On the other hand, (80a–b) are permitted, since the reconstructed anaphor satisfies Condition A. If the

principle responsible for WCO also applies to the output of (optional) reconstruction, then that principle would see a name rather than a trace in (78) and (79), and no WCO would be expected.

Given a reconstruction account of the amelioration of WCO in (78)–(79), we would expect WCO effects to reappear if the satellite constituent is a QP, since the QP would still have to undergo QR from its reconstructed position, which would recreate the WCO configuration. With QPs, this is difficult to test, since QPs are generally impossible as satellite constituents in clefts. But with pseudoclefts, QPs are reasonably acceptable, and a WCO effect is indeed discernible:

(81) a. ?What$_i$ I read t$_i$ to its$_i$ author was nothing$_i$
'I read nothing to its author'
b. *What$_i$ its$_i$ author read t$_i$ to me was nothing$_i$
'Nothing was read to me by its author'

The opposite kind of problem arises with focus constructions, where WCO effects unexpectedly show up. As discussed by Chomsky (1976), these constructions display WCO. In (82) the focused constituent is identifiable by focal stress:

(82) a. His$_i$ mother SHOT John$_i$
b. ??His$_i$ mother shot JOHN$_i$
c. His$_i$ mother bought a PICTURE of John$_i$

Chomsky suggests that the focused object *JOHN* in (82b) undergoes movement to a focus position, thus creating a WCO configuration.

From the perspective of our theory, (82b) is surprising. First, the focused NP does not seem to be a true quantifier, so its trace should be a null epithet in terms of our typology. Second, if reconstruction can circumvent the effects of WCO with clefts and pseudoclefts, it is not obvious why this is not also possible here. To explain the WCO effect in (82b), we must (a) find a true QP to bind the trace of Focus as a variable and (b) prevent the focused NP from reconstructing. Both problems are solved if the focused NP is assumed to contain a covert operator *only*, which carries the semantic import of focusing. Being nonvacuous, such an operator would bind a true variable, rather than a null epithet, and would be unable to reconstruct for the same reason that reconstruction is blocked in (81).

Appendix B: The Overt Pronoun Constraint and Ā-Bound Pro

There is another phenomenon, similar to WCO, that also seems to be sensitive to our distinction between variables and null epithets. This is Montalbetti's (1984) Overt Pronoun Constraint (OPC), which, in effect, prohibits an overt pronoun from being locally bound by a variable when a null pronoun could occur instead, as with overt subject pronouns in pro-drop languages. As Montalbetti observes, parasitic gaps do not behave

like true variables with respect to the OPC. He presents a Spanish example involving a parasitic gap locally binding an overt subject pronoun:

(83) A quiénes$_i$ contrató t_i el director [sin persuadir e_i de que ellos$_i$ viajen a Lima]
'Who$_i$ did the director hire t_i without persuading e_i that they$_i$ should go to Lima?'

Montalbetti points out that the same is true of the trace in *tough*-movement constructions. In each case it is possible for the null epithet to locally bind an overt subject pronoun. Likewise, it seems that restrictive relatives like (84) display strong OPC effects, whereas appositives like (85) display only very weak effects:

(84) *el hombre$_i$ a quien$_i$ dijiste t_i que él$_i$ debìa irse
the man to whom (you) told that he should leave
'the man who you told that he should leave'

(85) ?Juan$_i$, a quien$_i$ dijiste t_i que él$_i$ debìa irse
Juan to whom (you) told that he should leave
'Juan, whom you told that he should leave'

There is an obvious similarity between OPC effects and WCO effects; both involve a restriction against pronouns functioning as bound variables. The fact that both types of effects disappear in weakest crossover contexts is surely no coincidence. It is unclear whether WCO and OPC effects can be traced to a common principle. Nevertheless, our distinction between null epithets and (null) variables provides a means for stating the relevant partial generalization: only true variables are subject to the OPC.

Cinque (1984; 1990) also uses OPC effects and various other diagnostics to distinguish between true variables and what we have called null epithets. Cinque differs from us in that he treats the empty category as a null pronominal ("small pro"), licensed under Ā-binding. As he observes, the absence of WCO effects is then immediate: pronouns, unlike variables, do not display WCO. Although this proposal is attractive, there are certain difficulties with it.

First, as Cinque notes, the existence of apparent SCO effects with, for example, parasitic gaps is unexpected, since pronouns no more show SCO than they do WCO. Cinque suggests that such apparent SCO effects are actually the result of the licensing requirement on pro. When not governed by (rich enough) personal Agr, pro "can be identified only via Ā-binding" (p. 150). Identification will fail, Cinque goes on to argue, if there is an intervening A-binder, as there would inevitably be in a configuration resembling SCO.

It is not clear to us that this is so. If the requirement is, as Cinque presents it throughout, one of simple Ā-binding, an intervening A-binder should be irrelevant. On the other hand, the requirement could be restated to demand *local* Ā-binding, with the desired result. But as far as we can tell, there is no independent justification for such a requirement, and conceptually, it is unclear why an intervening element with the same

index as the null pronominal should prevent it from being identified, given that any number of intervening elements with different indices are permitted.

A second potential problem with Cinque's suggested approach to weakest crossover is that it seems overly inclusive. Crucially for Cinque, "long movement," (that is, configurations in apparent violation of island constraints) must pattern exactly like parasitic gap and *tough*-movement constructions, and he provides evidence that in certain respects this is the case. However, he concedes that the evidence that long movement escapes WCO is not entirely clear, and he speculates about why this should be so.

But WCO clearly does arise in many long movement configurations, in contradiction to Cinque's proposal. (87), involving "long movement," seems no less a WCO violation than (86), with its standard successive cyclic movement:

(86) ?*Who$_i$ does his$_i$ mother think that you like t$_i$
(87) ?*Who$_i$ does his$_i$ mother wonder whether you like t$_i$

That it is not just the island violation that is responsible for the seriously degraded status of (87) is evident from the fact that if *his* is not construed as covarying with *t*, the example is improved.

It should also be noted that there are configurations of weakest crossover that cannot, according to Cinque's analyses, involve base-generated pro, but must involve trace. Cinque acknowledges that this is so, observing that his analysis "requires a different treatment for the absence of weak crossover effects in Topicalization and appositive relatives . . ." (p. 199, fn. 59). He does not hint at what this different treatment might be, so it is difficult to evaluate his overall analysis of weakest crossover, or to compare his account with ours, except to note that we provide a unified analysis of the phenomenon, whereas Cinque will presumably have a bifurcated one.

References

Aoun, J. and R. Clark (1985) "On Non-overt Operators," in *Southern California Occasional Papers in Linguistics 10*, Department of Linguistics, University of Southern California, Los Angeles.

Baltin, M. (1982) "A Landing Site Theory of Movement Rules," *Linguistic Inquiry* 13, 1–38.

Barss, A. (1986) *Chains and Anaphoric Dependence*, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Bouchard, D. (1982) *On the Content of Empty Categories*, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Brody, M. (1984) "On Contextual Definitions and the Role of Chains," *Linguistic Inquiry* 15, 355–380.

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. (1976) "Conditions on Rules of Grammar," *Linguistic Analysis* 2, 303–351.

Chomsky, N. (1977) "On *Wh*-Movement," in P. Culicover, T. Wasow, and A. Akmajian, eds., *Formal Syntax*, Academic Press, New York.

Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, N. (1982) *Some Concepts and Consequences of the Theory of Government and Binding*, MIT Press, Cambridge, Massachusetts.

Chomsky, N. (1986a) *Barriers*, MIT Press, Cambridge, Massachusetts.
Chomsky, N. (1986b) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.
Cinque, G. (1984) "A-bar Bound pro vs. Variable," ms., Università di Venezia.
Cinque, G. (1990) *Types of A'-Dependencies*, MIT Press, Cambridge, Massachusetts.
Clark, R. (1983) "Parasitic Gaps and Split Anaphora," ms., UCLA, Los Angeles, California.
Contreras, H. (1984) "A Note on Parasitic Gaps," *Linguistic Inquiry* 15, 698–701.
Engdahl, E. (1985) "Parasitic Gaps, Resumptive Pronouns, and Subject Extractions," *Linguistics* 23, 3–44.
Epstein, S. D. (1984) "A Note on Functional Determination and Strong Crossover," *The Linguistic Review* 3, 299–305.
Fiengo, R. and H. Lasnik (1974) "Complement Object Deletion," *Linguistic Inquiry* 5, 535–572.
Higginbotham, J. (1980) "Pronouns and Bound Variables," *Linguistic Inquiry* 11, 679–708.
Higginbotham, J. (1983) "Logical Form, Binding, and Nominals," *Linguistic Inquiry* 14, 395–420.
Higginbotham, J. (1985) "On Semantics," *Linguistic Inquiry* 16, 547–593.
Hornstein, N. and A. Weinberg (1988) "Logical Form: Its Existence and Its Properties," in A. Cardinaletti, G. Cinque, and G. Giusti, eds., *Constituent Structure: Papers from the 1987 GLOW Conference*, Università di Venezia.
Huang, C.-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Jackendoff, R. (1977) *$\bar{X}$ Syntax: A Study of Phrase Structure*, MIT Press, Cambridge, Massachusetts.
Kayne, R. (1983) "Connectedness," *Linguistic Inquiry* 14, 223–249.
Kayne, R. (1984) *Connectedness and Binary Branching*, Foris, Dordrecht.
Kearney, K. (1983) "Governing Categories," ms., University of Connecticut, Storrs.
Koopman, H. and D. Sportiche (1983) "Variables and the Bijection Principle," *The Linguistic Review* 2, 139–160.
Lasnik, H. (1976) "Remarks on Coreference," *Linguistic Analysis* 2, 1–22.
Lasnik, H. (1989) *Essays on Anaphora*, Kluwer, Dordrecht.
Lasnik, H. and M. Saito (in preparation) *Move Alpha*, MIT Press, Cambridge, Massachusetts.
Lasnik, H. and J. Uriagereka (1988) *A Course in GB Syntax*, MIT Press, Cambridge, Massachusetts.
Lebeaux, D. (1983) "A Distributional Difference between Reciprocals and Reflexives," *Linguistic Inquiry* 14, 723–730.
May, R. (1977) *The Grammar of Quantification*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Montalbetti, M. (1984) *After Binding*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Postal, P. (1971) *Cross-Over Phenomena*, Holt, Rinehart and Winston, New York.
Reinhart, T. (1981) "Pragmatics and Linguistics," *Philosophica* 27 (I).
Reinhart, T. (1983) *Anaphora and Semantic Interpretation*, University of Chicago Press, Chicago, Illinois.
Rizzi, L. (1986) "Null Objects in Italian and the Theory of *pro*," *Linguistic Inquiry* 17, 501–557.
Safir, K. (1984) "Multiple Variable Binding," *Linguistic Inquiry* 15, 603–638.
Safir, K. (1986) "Relative Clauses in a Theory of Binding and Levels," *Linguistic Inquiry* 17, 663–689.
Sag, I. (1976) *Deletion and Logical Form*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Saito, M. (in press) "Extraposition and Parasitic Gaps," in C. Georgopoulos and R. Ishihara, eds., *Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda*, Kluwer, Dordrecht.
Sportiche, D. (1983) *Structural Invariance and Symmetry in Syntax*, Doctoral dissertation, MIT, Cambridge, Massachusetts.

Stowell, T. (1979) "Restricting Relatives," in E. Battistella, ed., *Proceedings of the Ninth Annual Meeting of NELS*, GLSA, University of Massachusetts, Amherst.
Stowell, T. (1985) "Licensing Conditions on Null Operators," in J. Goldberg et al., eds., *Proceedings of the West Coast Conference on Formal Linguistics*, Stanford Linguistics Association, Stanford University, Stanford, California.
Stowell, T. (to appear) "Adjuncts, Arguments, and Crossover," *Natural Language and Linguistic Theory*.
Wasow, T. (1972) *Anaphoric Relations in English*, Doctoral dissertation, MIT, Cambridge, Massachusetts.
Williams, E. (1986) "A Reassignment of the Functions of LF," *Linguistic Inquiry* 17, 265–299.

(Lasnik)
Department of Linguistics
University of Connecticut
Storrs, Connecticut 06269-1145
LASNIK@UCONNVM.BITNET

(Stowell)
Department of Linguistics
UCLA
405 Hilgard Avenue
Los Angeles, California 90024-1543
IML7TIM@UCLAMVS.BITNET