

# Chapter one

## On Control

---

*Norbert Hornstein*

### 1.1 Introduction

---

Recently, control has become a hot topic, largely for Minimalist reasons.<sup>1</sup> In particular, the current passion for reexamining the conceptual foundations of Universal Grammar (UG) has prompted grammarians to reconsider whether, and how much, the grammatical processes underlying control configurations differ from the movement operations that lie behind raising constructions. The idea that the grammatical processes responsible for these two types of structures are less different than generally believed is *not* novel.<sup>2</sup> What is different is the conceptual setting afforded by Minimalist commitments, most importantly the premium now placed on “simpler” theories that eschew both theory internal levels (like d-structure) or formatives (like PRO) and multiple ways of establishing grammatical dependencies (by either movement or construal). The venerable dualism between control and raising offers a tempting target for those impressed with the razor sharp Ockhamism of the Minimalist program.

Before reviewing the issues in detail, it is worth pausing to slightly expand on the cursory comments above as they provide some of the methodological motivations for reconsidering the distinction between raising and control. The main motivation comes from two sources.

First, the observation that control brings with it considerable grammatical baggage, for example a whole additional module of the grammar (the Control module whose function it is to determine the controllers of PRO and the interpretation that a particular control structure carries), a theory internal formative – PRO – with its own idiosyncratic distributional requirements (it occurs in the subject positions of non-finite clauses and this prompts otherwise conceptually and empirically problematic technology, for example null case, to track this fact), and a set of grammatical processes (construal rules) *added to the movement processes already assumed to be available*, whose function it is to establish dependencies quite similar to those that movement already affords.

Second, the observation that the whole distinction conceptually rests on assuming properties of UG that are currently considered problematic from an MP perspective. Two such problematic conditions are d-structure and that part of the theta criterion that stipulates that theta roles and chains be biuniquely

related (viz. be in a one to one relation). Both properties are methodologically suspect given minimalist commitments. Given Chomsky's (1993) objections to postulating levels like d-structure, it is natural to see if it is possible to eliminate d-structure and its various requirements entirely from the grammar. Moreover, from a purely conceptual point of view it is quite unclear why chains should be barred from bearing more than one theta role. Eliminating both d-structure and the biuniqueness condition on theta roles has obvious conceptual attractions. Interestingly, neither assumption fits comfortably with the assumption that control, like raising, is derived via movement. Thus rethinking the relation between movement and control immediately bears on these methodological issues and thereby on questions of optimal grammatical design.

This chapter reviews some of the recent hubbub surrounding these issues. Before proceeding, however, one last caveat is in order. Like all reviews, this one will suffer from the limited knowledge, perspective, interests and prejudices of the reviewer. The route taken here is but one of many. I hope that the trip proves interesting. However, I am sure that it is not the only worthwhile itinerary, nor the least contentious. Caveat lector!

## 1.2 Raising versus control

---

From the earliest work in Generative Grammar, Raising (1a) has been contrasted with Control (1b).

- (1)a. John seems/appears/is likely to like Mary
- b. John tried/hoped/expected to like Mary

The contrast has been drawn along several dimensions.

First, there are various empirical differences between the two kinds of constructions despite the obvious surface similarity. These include the following.

In (1a) John is understood as discharging a single thematic function. He is the *liker of Mary*. This is coded by assigning a single theta role to *John*, that of the external argument of the embedded verb *like*. In contrast to this, John in (1b) is perceived as bearing two roles. He is at once *Mary's liker* and also a *trier/hoper/expector*. More, theoretically, *John* is related to a pair of thematic positions, the external argument of the matrix predicate and that of the embedded one.

This difference in the thematic properties of raising versus control brings with it a host of others. Two are noteworthy.

First, the subject position of raising predicates can be filled with expletives and idioms (all the while retaining the idiomatic interpretation) while those of control predicates cannot be.

- (2)a. There seemed/appeared/was likely to be a man here
- b. \*There tried/hoped/expected to be a man here

- (3)a. The fat seemed/appeared/was likely to be in the fire  
 b. \*The fat tried/hoped/expected to be in the fire

Second, passivizing the embedded clause in raising constructions leaves the original meaning of the sentence largely intact. (4a,b) illustrate this property of “voice transparency”; the two sentences being virtual paraphrases of one another.

- (4)a. The doctor seemed to examine Mary  
 b. Mary seemed to be examined by the doctor

This contrasts sharply with what happens if we passivize in a control structure. (5a,b) are not voice transparent, not being even remote paraphrases.

- (5)a. The doctor tried to examine Mary  
 b. Mary tried to be examined by the doctor

These two facts can be directly tied to the thematic differences between raising and control verbs. One can explain the contrast between (2) and (3) by observing that expletives like *there* and idioms cannot bear theta roles. However, this is what they must do in examples like (2b) and (3b), on the further assumption that all theta roles of a predicate must be assigned. The trouble with (2b) and (3b) is that the theta roles that must be assigned cannot be assigned to any appropriate element given the resistance it is natural to assume that expletives and idioms have to bearing such roles. In (2a) and (3a) there is only one theta role to assign and as such no similar problem arises.

Analogous considerations account for (4) and (5). In control structures like those in (1b), the matrix subject fulfills two thematic requirements; that of the matrix verb and that of the embedded one. Thus, in (5a) the doctor is both *trier* and *examiner*, while Mary is merely the *examinee*. However, by passivizing the embedded clause, Mary becomes the matrix subject and thereby assumes two thematic functions, *trier* and *examinee*, while the doctor has just one thematic role, the *examiner*. This accounts for the clear difference in meaning between (5a,b). In contrast, the subject position of raising predicates is not associated with any thematic role. Consequently, similar manipulation of the embedded clause in (4) has no thematic repercussions. The thematic status of the doctor in both (4a,b) is that of *examiner* and that of Mary is that of *examinee*. As the thematic properties of the two constructions are undisturbed, the sentences remain essentially paraphrases.

These empirical differences between raising and control structures have generally been traced to an underlying difference in grammatical etiology. Specifically, the operations underlying raising configurations are distinct from those that generate control structures. The former are the province of movement rules. In contrast, control structures are either formed by applications of the non-movement rule Equi NP deletion (in the Standard Theory) or, in more recent Government Binding (GB) analyses, control is the result of construal rules that relate a phonetically null DP, “PRO,” to its antecedent. Structurally, then, raising structures differ from control configurations as in (6).

- (6)a. John<sub>1</sub> seems [t<sub>1</sub> to like Mary]  
 b. John<sub>1</sub> seems [PRO<sub>1</sub> to like Mary]

(6a) indicates that *John* has moved from the embedded subject position, leaving a co-indexed trace in its movement site. In (6b) *John* controls a base generated null expression *PRO*. The indicated indexations, therefore, arise from two different processes: movement for raising and construal for control. Similarly, the empty categories arise from two different processes: movement in the case of raising and lexical insertion in the case of control. The empirical differences noted above trace back to these different derivational histories that invoke different types of grammatical operations and employ different formatives.

It is worth considering the details. Specifically: what forces movement in (6a) and requires construal in (6b)? The answer, in a GB (and also an *Aspects*) style theory, is d-structure. The theoretical basis within GB for distinguishing the two constructions relies on contrasting traces and PROs: PROs head chains, traces do not; d-structure implements this difference. In fact, the classical distinction between raising and control follows seamlessly from the assumption that d-structure exists. Consider the reasoning.

D-structure has two distinctive properties: it is input to the transformational component and the locus of all thematic discharge; a representation of “pure GF- $\theta$ .”<sup>3</sup> Thus, prior to “displacement” operations (i.e. transformations) that rearrange phrase markers, words/morphemes are assembled into d-structure phrase markers by being lexically inserted into the available theta-positions. After lexical insertion, transformations apply to map d-structure phrase markers into others.

Given the requirements of d-structure, transformations cannot relate theta-positions (via movement) as all theta-positions have been filled by lexical insertion in forming the d-structure phrase marker. In particular, d-structure prohibits the movement in (7) from the embedded subject to the matrix subject position as the matrix subject is a theta position and so must be filled at d-structure. Consequently, it is unavailable for occupancy via movement.

- (7) John<sub>1</sub> tried [t<sub>1</sub> to like Mary]

In this way, d-structure thematic requirements make movement between  $\theta$ -positions impossible and thus prohibit control relations (which involve multiple theta-roles) from being the observed manifestations of movement operations.

Furthermore, if d-structure has *only*  $\theta$ -positions filled (in addition to all such positions being filled), then raising structures must be products of movement. In particular, a structure like (8), a non-movement version of a raising construction, is ill-formed at d-structure given that the matrix non-theta subject position is filled by *John*. Given d-structure thematic requirements, this position must be vacant at d-structure, that is it cannot be filled by lexical insertion. The only remaining option is to fill it at some later phase of the derivation, that is by movement.

- (8) John<sub>1</sub> seemed [PRO<sub>1</sub> to like Mary]

In sum, the classical vision of d-structure as the representation of pure GF- $\theta$ , that is the phrase marker where all and only thematic information is grammatically rendered, theoretically forces the empirical distinction between raising and control. Raising is required where d-structure thematic conditions prohibit the insertion of lexical material (e.g. the subject position of a raising verb as in (6a)) while it is prohibited where d-structure thematic conditions require the presence of lexical insertion (e.g. the subject positions of the control complement and matrix subjects in (6b)).

The Theta-Criterion further buttresses this view of d-structure, in particular the idea that all thematic information is discharged via lexical insertion. The relevant feature of the Theta-Criterion is the requirement that there be a biunique relation between  $\theta$ -roles and chains, in particular, that every chain bear *at most* one  $\theta$ -role. This effectively prohibits all movement from one  $\theta$ -position to another. But if movement into  $\theta$ -positions is forbidden, yet all  $\theta$ -roles must be discharged, then the only alternative is to fill each  $\theta$ -position via lexical insertion. The step from the Theta-Criterion to the postulation of PRO and construal rules that relate PROs to their antecedents is a short one.<sup>4</sup>

In sum, given the canonical view about d-structure and the Theta-Criterion, we theoretically derive the fact that raising is due to movement while control is due to construal. From this, the observed empirical differences follow. It's a nice story and it has had a lot of staying power. Nonetheless, there are problems given minimalist scruples. Here are three.

First, it requires the postulation of a theory internal formative PRO. PRO is similar to, but different from a trace. In a GB style theory, at LF, both traces and PROs are categories without phonetic contents. They are of the form "[<sub>NP</sub> e]<sub>I</sub>." PROs have the same structure at LF. However, whereas traces are the residues of movement, PROs are lexical expressions which receive their indices not in the course of the overt derivation but at LF via construal operations. That a lexical item (PRO) and a grammatical formative (trace) should be essentially identical at LF is quite unexpected. Indeed, identity of structure suggests that either both are lexical items or both are grammatical formatives.

Within a minimalist setting, things get worse. Nunes (1995) and Chomsky (1998) have noted that traces are theory internal constructs which are best avoided on conceptual grounds if possible. In place of traces, copies have been pressed into service.<sup>5</sup> This leaves the proposed structure of (6a) as (9).

(9) John seemed [John to like Mary]

However, (9) now raises a question for PRO. What is its structure? What kind of *lexical* item do we have here? It cannot simply be a null pronoun, as in many environments, for example (6b), it requires an antecedent. It might simply be a null reflexive, but if so why *must* it be phonetically null? Recall that minimalism resists the postulation of theory internal entities. Thus, the more the features of PRO are idiosyncratic (i.e. the more PRO is distinguished from more run of the mill lexical items) the less explanatory weight PRO has in a minimalist setting. Humdrum is best. However, within GB, PRO is hardly

run of the mill given that it brings in its train a whole module of grammar, the Control module, whose job it is to find its antecedent and provide an interpretation depending on its grammatical setting. We shall see that PRO continues to enjoy a special status in some recent reanalyses.

A second worry arises from the fact that control requires construal rules. Recall that Minimalism places a premium on simple theories. In the present context, it favors theories of UG that minimize rule types. A theory that has *both* movement rules and construal rules has two ways of establishing inter-nominal dependencies in the grammar. Ockham dictates that a theory with just one set of operations is preferred. On the assumption that movement is independently required (e.g. for feature checking), this suggests that construal processes should be eliminated, including those that underlie control.<sup>6</sup>

A third problem with the standard GB approach to control lies with its reliance on d-structure. As outlined above, the distinction between raising and control relies on the thematic requirements d-structure places on derivations. However, d-structure is a theory internal level (see Chomsky, 1993) and should be avoided if possible. If DS-centered accounts of control are suspect given standard minimalist assumptions then this casts doubt on GB approaches to control given their reliance on d-structure requirements for their explanatory punch.

There are other more technical problems with the GB approach to control, some of which we review below. However, even this brief discussion hopefully indicates that control theory as earlier conceived fits ill with the main methodological emphases of the Minimalist Program. The question is how radical a departure from traditional approaches is warranted empirically and is desirable theoretically. Various answers to this question have been offered and we review some below.

In sum, from a minimalist perspective, the standard approach to control presents various difficulties and motivates a search for alternatives to the standard GB accounts. In what follows, I will assume that these problems are sufficient reason to re-examine control phenomena.

### 1.3 Some basic properties of control

---

The theory of control must address two issues: (i) where PRO appears, and (ii) how it is interpreted. Let us consider these in turn.

#### 1.3.1 *The distribution of PRO*

PRO most conspicuously appears in the subject position of non-finite clauses. (10a–e) are thus fully acceptable, while (10f–i) are not.

- (10)a. John tried/hoped PRO to eat a bagel
- b. John tried/preferred PRO eating a bagel

- c. John thinks that [PRO eating/to eat a bagel] would be fun
- d. John saw Mary before PRO leaving the party
- e. John told Mary where PRO to eat a bagel
- f. \*John hoped that PRO eat a bagel
- g. \*John preferred for Bill to meet PRO
- h. \*John hated Bill meeting PRO
- i. \*John talked to PRO

What unifies the first four cases is that PRO sits in the subject position of a non-finite clause, in contrast to the four examples that follow.

It is reasonable to assume that in (10f–i) PRO is governed by various lexical heads. In (10f) it is the finite morpheme in Infl, in (10g,h) the verb *meet* and in (10i) the preposition *to*. If one further assumes that non-finite clauses do not contain lexical heads (at least of the relevant sort), then the presence of PRO in (10a–e) correlates with the absence of (head) government of these Spec IP positions.<sup>7</sup> Thus, the distribution of PRO conforms to the descriptive generalization (11).

- (11) PRO can only appear in ungoverned positions

Note that (11) does not prohibit generating PRO in a governed position. This is fortunate as PRO can be base generated in object position so long as it moves to an ungoverned position by s-structure (SS). (11) is a generalization that holds at s-structure or later.

- (12) John<sub>1</sub> tried [PRO<sub>1</sub> to be recognized t<sub>1</sub>]

As is well known, there were several attempts within GB to reduce (11) to more basic principles. This is not the place to review these efforts.<sup>8</sup> Suffice it to say, that these analyses do not currently enjoy much support. In section 1.4, we review two kinds of minimalist approaches to the distribution of PRO.

### 1.3.2 *The interpretation of PRO*

Control structures come in two varieties: local and long distance.

- (13a). John hopes [PRO to eat a bagel]
- b. John hopes that [ [PRO eating a bagel] will be fun]

In the earliest treatments, two distinct rules were involved in the derivations of (13a,b). Equi NP deletion applied to yield (13a) while Super Equi was involved in the derivation of (13b). Equi interacted closely with the Principle of Minimal Distance in determining the antecedent of PRO in (13a) while this principle did not regulate applications of Super Equi.<sup>9</sup> GB has honored essentially this same analysis by endorsing a distinction between structures of Obligatory versus Non-obligatory control (OC vs. NOC).<sup>10</sup>

OC and NOC differ in several important ways. Consider the following paradigm illustrating the interpretive properties of obligatory control structures.<sup>11</sup>

- (14)a. \*It was expected PRO to shave himself
- b. \*John thinks that it was expected PRO to shave himself
- c. \*John's campaign expects PRO to shave himself
- d. John expects PRO to win and Bill does too (= Bill win)
- e. \*John<sub>i</sub> told Mary<sub>j</sub> PRO<sub>i+j</sub> to leave together/each other
- f. The unfortunate expects PRO to get a medal
- g. Only Churchill remembers PRO giving the BST speech

(14a) shows that an obligatory control PRO requires an antecedent. (14b) indicates that this antecedent must be local and (14c) indicates that it must c-command the PRO.<sup>12</sup> (14d) shows that this PRO cannot have split antecedents.<sup>13</sup> PRO in (14f) only has the "de se" interpretation in that the unfortunate believes *of himself* that he will be a medal recipient. (14g) has the paraphrase (15a), not (15b). On this reading only Churchill could have this memory for Churchill was the sole person to give the speech. The two different readings follow on the assumption that obligatory control PRO must have a c-commanding antecedent. This requires "only Churchill" to be the binder. The unavailable reading has "Churchill" as the antecedent. This is possible in (15b) where the pronoun can have a non-c-commanding antecedent.

- (15)a. Only Churchill remembers himself giving the BST speech
- b. Only Churchill remembers that he gave the BST speech

PRO in non-obligatory control environments contrasts in every respect with the obligatory control cases.

- (16)a. It was believed that PRO shaving was important
- b. John<sub>i</sub> thinks that it is believed that PRO<sub>i</sub> shaving himself is important
- c. Clinton's<sub>i</sub> campaign believes that PRO<sub>i</sub> keeping his sex life under control is necessary for electoral success
- d. John thinks that PRO getting his résumé in order is crucial and Bill does too
- e. John<sub>i</sub> told Mary<sub>2</sub> that PRO<sub>1+2</sub> leaving together/each other was important to Bill
- f. The unfortunate believes that PRO getting a medal would be boring
- g. Only Churchill remembers that PRO giving the BST speech was momentous

(16a) indicates that non-obligatory control PRO does not require an antecedent. (16b) demonstrates that if it does have an antecedent it need not be local. (16c) shows that the antecedent need not c-command this PRO. (16d) contrasts with (14d) in permitting a strict reading of the elided VP, that is the reading in which it is John's resume which is at issue. (16e) can support split antecedents,

(16f) can have a non “de se” interpretation and (16f) is consistent with many people other than Churchill recalling that the BST speech was a big deal. Note that each non-obligatory control reading contrasts with those available in the obligatory control examples in (14). The cases in (14) and (16) contrast in one further interesting way; the former can be paraphrased with PRO replaced by a reflexive while the interpretive doubles of (16) replace PRO with pronouns. (17) illustrates this with the counterparts of (14c) and (16c).

- (17)a. \*John’s<sub>i</sub> campaign expects himself<sub>i</sub> to shave himself  
 b. Clinton’s<sub>i</sub> campaign believes that his<sub>i</sub> keeping his sex life under control is crucial for electoral success

In short, the differences in obligatory and non-obligatory control structures duplicate, where applicable, what one finds with locally bound anaphors versus pronouns. This makes sense if PRO is actually ambiguous – an anaphoric expression in obligatory control configurations and pronominal in NOC structures. The question facing the theoretician is to explain, first, why PRO should display these two sets of properties and why we find the OC properties in “Equi” configurations while we find the NOC cluster in “Super Equi” structures. In other words, why do the properties cluster as they do and why do they distribute as they do.

## 1.4 Two approaches to the distribution of PRO

---

### 1.4.1 *Null case*

Chomsky and Lasnik (1993) argue that the standard GB analysis of control in terms of government (see (11) above) cannot be reconciled with the last resort motivation for movement within Minimalism.<sup>14</sup> The empirical difficulty is illustrated in (18).

- (18)a. We never [PRO<sub>1</sub> expected to be found t<sub>1</sub>]  
 b. \*We never expected [PRO<sub>1</sub> to appear to t<sub>1</sub> [that Bill left]]

If movement is last resort and PRO must be ungoverned then the threat of being governed should suffice to license PRO’s movement in (18a). But if it suffices in (18a) why is it insufficient in (18b)? Chomsky and Lasnik argue that the two cases fall together if we assume that PRO has a case that must be checked.<sup>15</sup> To get this to work technically, they assume that PRO has its own (idiosyncratic) case, dubbed “null case,” that can only be checked by the T<sup>0</sup> of non-finite control clauses. PRO (and PRO alone) can and must check null case. Given these assumptions, (18b) violates greed in that PRO moves from one case position to another thereby violating last resort.

This same set of assumptions suffices to account for the facts in (10) repeated here.

- (10)a. John tried/hoped PRO to eat a bagel  
 b. John tried/preferred PRO eating a bagel  
 c. John thinks that [PRO eating/to eat a bagel] would be fun  
 d. John saw Mary before PRO leaving the party  
 e. John told Mary where PRO to eat a bagel  
 f. \*John hoped that PRO eat a bagel  
 g. \*John preferred for Bill to meet PRO  
 h. \*John hated Bill meeting PRO  
 i. \*John talked to PRO

In (10a–e), it is assumed that PRO is in the Spec of a T<sup>0</sup> able to check null case. Hence their acceptability. In contrast to this, PRO is in a nominative case position in (10f), and accusative position in (10g–i). These are not positions in which null case can be checked and so the structures underlying these sentences are ungrammatical.

The most fully worked out version of this null case approach to obligatory control is Martin (1996).<sup>16</sup> Following Stowell (1982), he argues that control infinitives differ from raising infinitives in that the former have tensed Infs. In effect, Martin provides motivation for the assumption that some infinitives can check case by assimilating them to the class of clauses that uncontroversially do so: finite clauses. Both finite IPs and control infinitives are +Tense and so it is not surprising that their respective Spec IPs are case positions. Given the further assumption that only PRO bears null case, only it can appear in the Spec IP position of tensed infinitives.

As mentioned, Martin follows Stowell in observing that some embedded clauses appear to have tense specifications. Stowell distinguishes between control clauses and raising clauses in that the latter always share the tense specifications of the clause they are embedded under while the former do not. The examples in (19) illustrate the point.

- (19)a. John decided/remembered [PRO to go to the party]  
 b. John believed [Mary to be the best player]

In (19a) the embedded event of going to the party takes place in the future with respect to John's decision (or recollection). In contrast, in (19b) Mary being the best player temporally coincides with John's belief. Stowell (1982) accounts for this by assuming that the embedded clause in (19a) has a T<sup>0</sup> specified for tense and so the embedded event can be temporally located independently of the main clause event. In contrast, the embedded T<sup>0</sup> of raising clauses is not specified for tense and so its temporal specification cannot be independent of the one in the main clause. Martin assumes that it is the presence of a temporal operator in control clauses that allows them to check (null) case.<sup>17</sup>

Martin (1996) provides further interesting evidence for the assumption that null case is a property of +tensed non-finite clauses. Here is one more bit. Martin claims that event denoting predicates cannot occur under raising predicates.<sup>18</sup>

- (20)a. \*Everyone believed Rebecca to win the game right then  
 b. \*The defendant seemed to the DA to conspire against the government at that exact time

Following Enç (1990), Martin assumes that event denoting predicates contain event-variables which must be bound by “tense or some other (temporal) operator in order to denote an individuated event” (Martin, 1996, p. 59). Without such an operator only a stative predicate is possible. The absence of a tensed  $T^0$  in raising predicates accounts for the oddity of the examples in (20).

With this in hand, Martin (following Boskovic, 1997) observes the following contrast between Romance *croire* and its English counterpart *believe*. The latter is a raising predicate with the exceptionally case marked subject of the embedded clause residing in its Spec VP.<sup>19</sup> For example, (21a) has the structure (21b).

- (21)a. John believes Mary to be tall  
 b.  $\text{John}_1 \text{ } [_{VP} \text{ Mary}_2 \text{ } [_{VP} \text{ t}_1 \text{ v } [\text{believes } [\text{t}_2 \text{ to be tall}]]]]$

The former, however, shows the characteristics of a control predicate, as the following French example indicates.

- (22) Je crois [PRO avoir fait une erreur]  
 I believe to have made a mistake

Interestingly, in Romance, the restriction to statives/habituals noted on embedded clausal complements of *believe* does not hold for *croire*. Contrast (23a) and (23b), its English counterpart.

- (23)a. Jean croit rêver  
 John believes to dream  
 b. John believes himself to dream

(23a) has the interpretation *John believes himself to be dreaming*. Note that this reading is unavailable for (23b). (23b) is unacceptable in English because *believe* is an ECM verb and so cannot take a control complement. What the data here appear to indicate are two things: first, that in French *croire* is not an ECM verb but a control verb and second, that when it is a control verb it allows eventive predicates. Martin concludes from this that control infinitives have tensed  $T^0$ s and that raising predicates do not.

One last interesting bit of corroborating evidence for this comes from a further observation. Boskovic observes the following interesting fact. There are cases of apparent ECM – like uses of *croire*, as in (24).<sup>20</sup>

- (24) Qui<sub>i</sub> Ana croyait-elle [t<sub>i</sub> plaire à Pierre]  
 “Who did Ana believe to please Pierre”

In these cases, moreover, the eventive reading noted in (23) is unavailable. Only a habitual interpretation is available.

- (25) \*Qui crois-tu rêver  
 “Who do you believe to be dreaming”

In sum, Martin provides interesting evidence that null case correlates with the presence of temporally active  $T^0$ s. In effect, he argues that null case can be checked by a  $T^0$  if and only if the  $T^0$  is +tense and –finite. If one further assumes that PRO can *only* bear null case, then a full account for the distribution of PRO follows.

### 1.4.2 Some problems for null case

Before discussing other approaches to the distribution of PRO, let me voice some skepticism regarding the proposed correlation between tense/eventive properties and the nature of  $T^0$  in raising and control structures. Consider first the claim regarding eventive predicates under raising versus control verbs.

The noted unacceptability of eventive predicates under raising verbs is most pronounced for the ECM predicate *believe* (see (20a)). There are perfectly acceptable raising constructions with embedded eventive predicates. So contrast the examples in (26) with (27).

- (26)a. Rebecca seemed to win the game right then  
 b. John appeared to take the wrong medicine  
 c. John is likely/certain/sure to eat a bagel
- (27)a. \*John believed Rebecca to win the game right then  
 b. \* John showed Bill to take the wrong medicine  
 c. \* John believed Bill to eat a bagel

This contrast suggests that the tense property of relevance relates to ECM verbs, rather than to raising predicates in general. But if so, then the conclusions Martin draws are too broad. The distinction of theoretical interest is raising versus control complements. Given this broad contrast, the cases in (26) and (27) should be entirely parallel.

Note, incidentally, that (26c) above has one other curious property. The complement clause is most naturally interpreted as in the future with respect to the tense of the matrix. Thus, it means, more or less, that it is *currently* likely/certain/sure that John *will* eat a bagel. This temporal ordering is unexpected given Stowell’s claim that raising predicates do not contain independent tense specification.

Martin (1996, pp. 80–105) observes similar facts and concludes from this that such cases involve control rather than raising. In fact, if one uses these temporal diagnostics, it appears that all the standard (non-ECM) raising predicates are actually ambiguous, with both a raising and a control structure. This assumption runs into problems. Consider how.

Assume that whenever one sees an eventive embedded predicate then the clause has a control structure. The sentences in (28) have embedded eventive predicates.

- (28)a. The shit appeared to hit the fan then  
 b. It seemed to start to rain exactly then  
 c. ?There appeared to enter several men at that very moment

(28a) means, roughly, *it appeared that pandemonium erupted then*. The presence of the punctual adverb reinforces this point. Given this, by assumption these sentences all have control structures, with the matrix subject controlling an embedded PRO. The problem is that were this so, we would expect idioms and expletives to be barred from the subject matrix position. (28) indicates that this is incorrect.

Observe that we also would expect to find no voice transparency in raising constructions where the embedded predicate had an eventive interpretation. This also appears to be incorrect.

- (29)a. The doctor seemed to then examine Mary  
 b. Mary seemed to then be examined by the doctor

In both (29a,b) the embedded clause is interpreted eventively. Nonetheless, voice transparency seems to hold. Recall that the absence of voice transparency and the limited distribution of expletives and idioms are the classical hallmarks of control. The absence of these thematic diagnostics in (28) and (29) suggests that they are standard raising, not control, constructions. If so, it can be argued that tense diagnostics do not correlate well with the standard thematic diagnostics that distinguish raising from control and this, in turn, argues against seeing some  $T^0$ s as assigners of null case.<sup>21</sup>

The data that Martin provides (following Boskovic, 1997) is also restricted to a contrast between ECM predicates in Romance and English. It does not contrast non-ECM raising predicates (such as *seem*) with control predicates. This is noteworthy for it is the latter contrast that is of obvious theoretical interest, not the former.

It is unclear what makes ECM predicates exceptionally able to assign case. However, it would not be remarkable if ECM verbs imposed semantic restrictions on their complements, selecting, for example, non-eventive clausal complements. Were this the case, then one would expect the correlations noted by Boskovic *even in the absence of case assignment in the embedded clause in control complements*. In other words, if Boskovic's observations pertain exclusively to ECM verbs then their peculiarities tell us nothing about the thematic contrasts between control and raising verbs. It is only if the properties that Boskovic noted are properties of raising predicates in general that his observations suggest that control predicates are null case markers.

A few last points. There are ECM verbs that do not display the diagnostics noted in the text. For example, *expect* takes eventive predicates and requires a temporal specification of the embedded event later than the matrix.

- (30) John expected Mary to leave the party

This implies that in such a case the embedded clause has a +tense  $T^0$ . This then further implies either that (i) the structure of (30) is similar to what one finds

with a *persuade* verb or (ii) that *Mary* is case marked by some element in  $C^0$  as is the case, arguably, for *want*.

There are problems for assumption (ii). Were this so, we would expect passivization to be blocked and *for* to appear overtly (in some configurations) parallel to what we witness with *want*. Both expectations are unrealized.

- (31)a. \*John was wanted to leave  
 b. John was expected to leave  
 c. John wants very much for Bill to leave  
 d. \*John expects strongly for Bill to leave

This leaves the (i)-option that what we have is a *persuade*-like configuration. However, this suggestion also has its problems. Recall the standard thematic differences between *persuade* and *expect* verbs. These are not suspended when we find the embedded clause with an eventive/future interpretation. For example, we find idioms and expletives with eventive/future readings (32a,b) and we find voice transparency (32c,d) are paraphrases.

- (32)a. John expected the shit to hit the fan at exactly 6  
 b. John expected there to erupt a riot  
 c. John expected the doctor to examine Mary then  
 d. John expected Mary to be examined by the doctor then

This is unexpected if these are control structures.

Finally, consider one last set of empirical difficulties.<sup>22</sup> Stowell (1982) observes that gerunds, in contrast to infinitives, vary their tense specifications according to the properties of the matrix verb.

- (33)a. Jenny remembered [PRO bringing the wine] (Stowell, 1982, (8b))  
 (J remembered a past event of bringing the wine)  
 b. Jim (yesterday) counted on [PRO watching a new movie (tonight)]  
 (J counted on a future event of movie watching)

In light of this, Stowell (1982) proposed that gerunds are generally marked –tense.

If correct, this constitutes a problem for the view that only +tense  $T^0$ s can assign null case and thereby license PRO given the presence of PRO in (33a,b).<sup>23</sup>

In fact, other gerunds provide potentially more serious problems. Pires (2001) identifies a class of TP defective gerunds which with respect to their tense properties are very similar to ECM infinitives in that their time interval must coincide with the event time of the matrix.

- (34) \*Bill last night avoided [PRO driving on the freeway this morning]

In such cases the argument against postulating a +tensed  $T^0$  on Stowell-like grounds is stronger still.

I have briefly reviewed some problems for the null case theory as proposed by Chomsky and Lasnik (1993) and most fully elaborated and developed in Martin (1996). Before moving on to consider another approach to the distribution of PRO (one that I personally prefer, I must confess), it is worth recalling the virtues of the null case account. It directly addresses one of the main characteristics of control phenomena; namely that it is *subjects* that are controlled. Once one gives up government as a core concept of grammar (and with it the generalization (11) that PRO only occurs in ungoverned positions) it becomes a challenge to identify the subject position of non-finite clauses in any sort of unified principled manner. Chomsky and Lasnik (1993) provide one way of doing this: via a case marking property unique to non-finite clauses. Martin's (1996) suggestion that null case is a property of some +tense T's is a reasonable attempt to provide a principled foundation for this idea and thereby account for the limited distribution of PRO.

### 1.4.3 (OC) PROs as traces

Consider now a second approach to the distribution of PRO.<sup>24</sup> It starts from one observation and one methodological qualm. The observation is that control, like raising, affects subjects of non-finite clauses. The qualm is that the theory of null case is stipulative and, hence, methodologically suspect. In fact, null case singles out PRO for special treatment in two distinct ways. First, it is the *only* lexical item able to check it or bear it. No other DP can bear null case, not even phonetically null expressions like WH-traces.<sup>25</sup>

- (35)a. John asked Bill \*Mary/PRO to eat a bagel  
 b. \*The man<sub>1</sub> (who) John asked Bill t<sub>1</sub> to eat a bagel

(35a) with *Mary* as the embedded subject is ungrammatical as *Mary* cannot check nor carry null case. PRO is fine in this position because it can. Note that (35b) is also unacceptable, presumably because t<sub>1</sub> cannot check/carry null case either. Thus, the only phonetically non-overt element able to bear null case is PRO.

Second, to my knowledge, it is the *only* lexical item whose case properties are grammatically specified. There are no other DPs, to my knowledge, that can bear but a single case.

Null case does not pattern with other structural cases in other ways. For example, it contrasts with WH-traces in not blocking *wanna* contraction. (36a) cannot be contracted as in (36c) while contraction for (36b) as in (36d) is fine.<sup>26</sup>

- (36)a. Who<sub>1</sub> do you t<sub>1</sub> want to visit Mary  
 b. I<sub>1</sub> want PRO<sub>1</sub> to visit Mary  
 c. \*Who do you wanna visit Mary  
 d. I wanna visit Mary

It thus seems that an accusative case marked non-phonetic WH-trace suffices to block sandhi effects like *wanna* contraction while a null case marked non-phonetic PRO does not.

In sum, whatever the other virtues null case may have, its peculiar characteristics are evident. With this as prologomena, consider a movement based approach to control.<sup>27</sup>

The core of such a theory is that OC PRO is identical to an NP-trace. It is the residue of overt A-movement. Thus, the overt structure of (37a) is (37b).<sup>28</sup>

- (37)a. John tried to win  
 b. [<sub>TP</sub> John [<sub>past</sub> [<sub>VP</sub> (John) [<sub>VP</sub> try [(John) to [(John) win]]]]]]]

Note that this is entirely analogous to what one finds in raising constructions with one caveat. In a raising construction, movement is from the embedded clause to a matrix non-theta position while in control structures, movement is via a matrix theta position. Note the trace in the spec of the matrix VP in (37b).

This proposal accounts for the distribution of PRO as follows. Assume that the subject positions of all non-finite clauses are not case marking positions (clearly the null hypothesis given the unacceptability of cases like (38a)). Then A-movement from this position is permitted. Further, as A-movement from case positions is prohibited (see 38b) and if (OC) PRO is the residue of A-movement, then we should never find PROs in case positions (see (10f-i), repeated here).

- (38)a. \*John hopes [Frank to leave]  
 b. \*John seems [t is nice]
- (10)a. John tried/hoped PRO to eat a bagel  
 b. John tried/preferred PRO eating a bagel  
 c. John thinks that [PRO eating/to eat a bagel] would be fun  
 d. John saw Mary before PRO leaving the party  
 e. John told Mary where PRO to eat a bagel  
 f. \*John hoped that PRO eat a bagel  
 g. \*John preferred for Bill to meet PRO  
 h. \*John hated Bill meeting PRO  
 i. \*John talked to PRO

(10f-i) are underivable as they all involve movement from a case position. (10a-e) are all acceptable as each involves movement from a non-finite clause via a matrix theta position ending in a matrix case position. If we assume that non-finite clauses do *not* assign case to their subjects, then movement from the embedded IP to the matrix IP will not be prohibited by last resort (i.e. Greed) as it is in (10f-i).

In addition, note that the null phonetic value of PRO (here just an NP-t) correlates with the fact that traces of A-movement are all phonetically null.<sup>29</sup> In other words, whatever it is that renders copies due to A-movement

phonetically null will extend to explain the null status of PRO. We need invoke no special null case properties and need invoke no special lexical item that is only able to carry such a case. This grouping of PRO with NP-t also accounts for why both behave similarly in *wanna* contraction contexts.

- (39)a. I seemta (seem+to) eat bagels every morning  
 b. I wanna (want+to) eat bagels every morning

Thus, the fact that (OC) PRO is typically a subject of non-finite clauses, follows from the basic premise that it is the residue of overt A-movement. This is merely a descriptive generalization. The proposal here is that control, like raising, is due to movement motivated by case concerns. It is not necessarily restricted to non-finite subject positions. If some non-subject position is not a case position, it can also be occupied by PRO. Lasnik (1995a,b), following a proposal by Munn, suggests that examples like (i) involve control.

- (i) John washed/shaved/dressed

Note that these do have the expected properties of control verbs. Thus, for example, *John* is related to two theta-roles in the case of these reflexive predicates. If one assumes that such verbs need not assign case to their objects, then one can treat these as cases of A-movement (and so, control). Note that were this a case of control, it would argue against the null case theory as here we would have a PRO unrelated to a  $T^0$ .

One other point. Were this a case of control, then we would need to explain why it is that sentences like (ii) are unacceptable.

- (ii) John<sub>1</sub> hoped Bill washed PRO<sub>1</sub>

This movement would not be blocked by case as the object position of *wash* need not be case marked. However, minimality (shortest Move) would be violated by moving over the intervening DP *Bill*. (40) lists the assumptions required to permit the sort of movement advocated here.

- (40)a. Theta roles are features  
 b. There is no upper bound on the number of theta features that a DP can have  
 c. Movement is Greedy  
 d. Greed is understood as “enlightened self interest”

Of these four assumptions, (40a,b) are the most contentious. They amount to rejection of the two basic ideas behind the standard theory of control: (i) that DS (or any analog) regulates the possibilities of movement and (ii) that there is any requirement limiting DPs to at most one theta-role. Both assumptions seem natural enough in a minimalist setting. Consider why.

Chomsky (1993) argues that grammar internal (non-interface) levels have no place in an optimal theory of grammar. DS, he observes, is such a level. As

such, it is methodologically suspect and should be eliminated. Chomsky (1993) provides some empirical arguments to this same conclusion. Assume that these considerations are decisive. Then one might expect the restrictions that DS placed on grammatical operations to also disappear. Recall, that it was the thematic restrictions that DS imposes that forced the distinction between raising and control in earlier theory. Without DS, however, we might expect the two types of operations to fall together, as proposed in a movement approach to control. Put more baldly, a movement theory of control is a natural consequence of the minimalist elimination of DS.

This conclusion has been resisted. Chomsky's proposed elimination of DS does not entail that the thematic restrictions found in earlier DS based theories play no role in the grammar. Chomsky (1995) achieves the functional equivalent of DS by assuming that theta-roles are *not* features. Coupled with the idea that all legitimate movement must check a feature, that is be greedy, the claim that theta-roles are not features results in a system where control cannot be reduced to movement as movement into theta positions is unmotivated and so blocked by Greed. (40a) amounts to a rejection of this way of re-introducing DS conditions once DS is abandoned. The assumption that theta-roles are features is aimed at allowing movement between theta positions.<sup>30</sup>

(40b) has a similar motivation. The prohibition against a DP's bearing at most one theta-role has no conceptual justification (in contrast to the idea that a DP bear at least one, which plausibly follows from a principle of Full Interpretation). The restriction, however, would follow were movement into theta positions prohibited. Given that control always involves the relation of a DP to at least two theta positions, a movement approach to control must countenance a DP's having more than one theta-role.

(40c,d) are standard minimalist assumptions. I here adopt the enlightened self-interest interpretation of Greed as I take theta roles to be features of predicates that DPs obtain by merging with them (via either pure merge or copy plus merge viz. Move). Theta features, then, are primarily properties of predicates and only derivationally properties of arguments. If so, checking a feature of one's target must suffice to license movement. Greed as enlightened self-interest permits this.

Given these four assumptions, then, a movement theory of control is viable. I have argued that they are conceptually natural (if pressed, I might say optimal) given a minimalist setting. I now want to consider empirical support for (40a,b) that arises outside the domain of control phenomena.

#### 1.4.4 *Movement into theta positions*

If control is movement then movement into theta positions must be possible. There is independent evidence that such movement obtains. I here review two cases in the literature.

Boskovic (1994, pp. 268 ff.) discusses the following examples from Chilean Spanish.<sup>31</sup>

(41)a. Marta le quiere gustar a Juan  
 Martha clitic wants to please to Juan  
 “Marta wants Juan to like her”

b. A Juan le quiere gustar Marta  
 “Juan wants to like Marta”

Three facts regarding examples like (41) are of present interest.

First, *a* case marks *Juan*. That *Juan* bears this case is a lexical property of the verb *gustar* which assigns a quirky/inherent case to its thematic external argument much as Belletti and Rizzi (1988) argue with respect to similar verbs in Italian.

Second, (41a) differs from (41b), as the glosses indicate, in that *Marta* has two thematic functions in the first sentence while *Juan* has two in the second. This suggests that both are control structures with the matrix subject controlling the external argument in the embedded clause. Third, note that in (41b) the matrix subject bears the inherent case marker *a*. Given that *a* is an inherent case assigned by the embedded verb *gustar*, this implies that *a Juan* has moved to the matrix subject position in (41b) from the embedded position in which *a* and its corresponding theta-role were assigned to *Juan*. Such movement is also observed in raising constructions with inherent case marked subjects (*al* is the case marker).<sup>32</sup>

(42) Al profesor le empezaron a gustar los estudiantes  
 “The professor began to like the students”

These three facts together implicate movement from one theta position to another. The *a* case on *Juan* in (41b) indicates movement from the domain of *gustar*, whence the inherent case marking was effected. However, in contrast to (42), *quiere* has an external thematic argument to assign and in (41b) *Juan* clearly bears it as the gloss indicates. The obvious implication is that *Juan* has moved from the thematic position of *gustar* to a thematic position of *quiere* and thereby obtained a second theta-role, as Boskovic (1994) concludes.<sup>33</sup>

Boskovic and Takahashi (1998) provides a second argument that movement into theta positions is possible. The phenomenon analyzed is long distance scrambling in Japanese.

As Boskovic and Takahashi note, long-distance scrambling – (43b) – is odd given minimalist assumptions that it seems unmotivated. The base form from which it is derived is well formed (43a).

(43)a. John-ga [Mary-ga sono hon-o katta to] ommotteiru (koto)  
 -nom -nom that book-acc bought that thinks fact  
 “John thinks that Mary bought that book”

b. Sono hon-o, John-ga [Mary-ga t<sub>i</sub> katta to] ommotteiru (koto)  
 “That book, John thinks that Mary bought”

The lack of semantic motivation for the movement is supported by the fact that unlike Wh-movement and topicalization, scrambling does not seem to establish

an operator-variable relation at LF.<sup>34</sup> In particular, it seems that the long scrambling *must* be undone at LF. (44) illustrates this.<sup>35</sup> Boskovic and Takahashi note that a (long) scrambled QP cannot take scope over a matrix QP subject.

- (44) Daremo<sub>1</sub>-ni dareka-ga [Mary-ga t<sub>1</sub> atta to] omotteiru (koto)  
 everyone-dat someone-nom Mary-nom met that thinks (fact)  
 “Everyone someone thinks that Mary met”

(44) must be interpreted with *dareka* (someone) scoping over *daremo* (everyone) despite the fact that the latter c-commands the former in overt syntax. This follows if the scrambled *daremo* must reconstruct at LF.

Where does the scrambled expression move to at LF? According to Boskovic and Takahashi, it moves to its theta position. This movement is obligatory for were it not to so move, it would violate full interpretation. In fact, as Boskovic and Takahashi argue, virtually all of the characteristics commonly associated with long scrambling follow if it is assumed that the scrambled expression is Merged in overt syntax in its scrambled position and then moved covertly to its theta position at LF. However, for this analysis to be viable, it must be possible to *move* to a theta position and this movement must have some sort of feature checking motivation given standard assumptions. Both desiderata are met if it is assumed that theta roles are sufficiently like features to license greedy movement, assumption (40a) above.<sup>36</sup>

To sum up, there is independent evidence for movement into theta positions. The existence of this movement indirectly enhances the proposal that control is the overt manifestation of movement via multiple theta positions. Why, after all, should UG have recourse to a specialized theory internal expression like PRO if movement can independently forge the requisite relations? The redundancy inherent in such an approach speaks against its adoption on purely methodological grounds.<sup>37</sup>

### 1.4.5 Conclusion

This section has discussed the two dominant current minimalist approaches to the distribution of PRO. Both attempt to respond to the fact that minimalist assumptions leave little room for the standard GB approach to this issue. The basic descriptive generalization is that PROs are subjects and are found in non-finite sentences.<sup>38</sup> How to account for this without theoretically alluding to government is the basic challenge that both approaches try to meet.

## 1.5 The interpretation of PRO

---

There are two broad approaches to the interpretive properties of control in the current literature. The first is a structural approach. Abstracting from many interesting differences, structuralists generally see the interpretive properties

of control as reflecting structural facts concerning the syntax of control configurations.<sup>39</sup> These properties, then, are analyzed as reflecting the internal workings of the grammar. The second approach is more eclectic. Eclectics see the interpretive properties of control as being due to the complex interaction of various modules, some of which are external to the syntax properly speaking. Chomsky sums up the view well:

the theory of control involves a number of different factors: structural configurations, intrinsic properties of verbs, other semantic and pragmatic considerations. (Chomsky, 1981, pp. 78–9)

This section concentrates on the first approach. The next reviews objections lodged against it from the more eclectically inclined. The main reason for this is methodological; all things being equal, structuralism is preferable to eclecticism. The reason is that all agree that grammatical structure is *part* of any adequate approach to control. What distinguishes structuralists from eclectics is whether this information exhausts what is needed. All things being equal then, structuralism is preferable if attainable. The main arguments in favor of eclecticism are the purported empirical inadequacies of structuralism. As such, starting from the former position and considering possible empirical problems that beset it should reveal the virtues and vices of both views.

As noted in section 1.3.2, not all PROs are interpreted in the same way. A contrast exists between PROs in Obligatory Control configurations (OC PRO) and those in non-obligatory structures (NOC PRO). The former have properties roughly equivalent to those of reflexives while the latter pattern interpretively largely like pronouns.<sup>40</sup> For illustration consider the two paradigms first outlined in section 1.3 and repeated here.

Recall that (14) and (16) differ completely. The OC PROs in (14) require local, c-commanding antecedents, require sloppy readings under ellipsis, forbid split antecedents, only have *de se* readings and require bound readings with *only* antecedents. The latter need no antecedents, need not be c-commanded locally by an antecedent if one is present, allow a strict reading under ellipsis, permit split antecedents, can have a non *de se* reading and can have a referential reading with *only* antecedents.

- (14)a. \*It was expected PRO to shave himself  
 b. \*John thinks that it was expected PRO to shave himself  
 c. \*John's campaign expects PRO to shave himself  
 d. John expects PRO to win and Bill does too (= Bill win)  
 e. \*John<sub>i</sub> told Mary<sub>j</sub> PRO<sub>i+j</sub> to leave together/each other  
 f. The unfortunate expects PRO to get a medal  
 g. Only Churchill remembers PRO giving the BST speech
- (16)a. It was believed that PRO shaving was important  
 b. John<sub>i</sub> thinks that it is believed that PRO<sub>i</sub> shaving himself is important  
 c. Clinton's<sub>i</sub> campaign believes that PRO<sub>i</sub> keeping his sex life under control is necessary for electoral success

- d. John thinks that PRO getting his resumé in order is crucial and Bill does too
- e. John<sub>1</sub> told Mary<sub>2</sub> that PRO<sub>1+2</sub> leaving together/each other was important to Bill
- f. The unfortunate believes that PRO getting a medal would be boring
- g. Only Churchill remembers that PRO giving the BST speech was momentous

There is an important fact worth bearing in mind in what follows. The data illustrated in (14) and (16) illustrate that OC PRO has a proper subset of the properties of NOC PRO. Thus, (16d) differs from (14d) in allowing a strict reading which the latter forbids. (16d) can support a sloppy reading. This holds for (most of) the other properties as well. This means that OC PRO and NOC PRO *cannot* be in free variation. In particular, positions inhabited by OC PRO must exclude NOC PRO. Were this not so, we would never be able to observe OC PRO as its signature properties would never emerge, being, as they are, a proper subset of NOC PRO. What this means theoretically is that we need some way of excluding NOC PRO when OC PRO is present. This has implications for any general theory of control and we return to it at the end.

### 1.5.1 *Some properties of a structural theory*

Assume for the nonce that OC PRO is the residue of NP movement. If so, the structure of a subject control verb is roughly as (45a) and an object control verb looks like (45b).

(45)a. DP<sub>1</sub> V [<sub>IP</sub> t<sub>1</sub> [<sub>I</sub> –finite] [<sub>VP</sub> t<sub>1</sub> V . . . ]]<sup>41</sup>

b. DP<sub>1</sub> V DP<sub>2</sub> [<sub>IP</sub> t<sub>2</sub> [–finite] [<sub>VP</sub> t<sub>2</sub> V . . . ]]<sup>42</sup>

Several features of these configurations are of interest in relation to the paradigm in (14).

First, the fact that OC configurations require local c-commanding antecedents follows directly from their being residues of movement. In fact, it is the same as the reason why A-traces require local, c-commanding, antecedents.

Second, the facts concerning sloppy readings under ellipsis, the prohibition against split antecedents, the required *de se* reading for (14f) and the bound reading in (14g) also follow given a movement analysis. The required sloppy reading parallels what one finds in raising constructions.<sup>43</sup>

(46) Bill was expected t to win and Harry was too

Split antecedents are barred on a movement analysis. For an OC PRO to have DP as its antecedent implies that PRO is the trace left by the movement of DP

from PRO's position. Split antecedents would require having two distinct DPs move from the very same position. This contemplated derivation is theoretically impossible and so split antecedents are ruled out.

It is worth noting that there is nothing that would rule out the possibility of OC PRO taking split antecedents were PRO a base generated expression any more than there is an explanation for why, given standard assumptions, locally bound reflexives cannot have split antecedents. Rather, the impossibility of split antecedents for reflexives and OC PRO is a stipulation motivated entirely on empirical grounds. The fact is tracked by an axiom explicitly forbidding split antecedents for reflexives and controlled PROs.<sup>44</sup>

Within a theory of control in which antecedents to PRO are functions of idiosyncratic lexical features of particular verbs the problem is perhaps worse. Why do control predicates designate only a single antecedent for PRO in OC configurations? Why not two (or more) antecedents? Were this possible then split antecedence would follow naturally. Thus, the impossibility of split antecedents and the fact that control predicates "choose" but a single controller are closely related facts. Both are immediately accounted for on a movement theory of OC.

The *de se* interpretation found in (14f) also follows. As a result of movement via multiple theta positions one ends up with a chain with multiple theta-roles. The natural semantic interpretation of such a syntactic object is of one expression, a chain, "saturating" several distinct argument positions. Semantically, this yields a complex monadic predicate, roughly of the form (47).

$$(47) \text{ DP}[\lambda x (\text{P}_{x,x})]$$

Such complex monadic predicates contrast with configurations of binding in forcing *de se* readings.<sup>45</sup> Thus, as movement via multiple thematic positions leads to the formation of complex monadic predicates and given that these only permit *de se* readings, it follows that if OC PRO is a residue of movement that it will require a *de se* interpretation. It is, once again, worth noting that this interpretive fact, though consistent with a base generation of a PRO approach, does not follow from it.

The facts concerning (14g) also follow. On the movement theory, the PRO is actually the residue of overt movement and so must have a c-commanding antecedent. *Only Churchill* is a possible antecedent. *Churchill* alone is not. Thus, the relevant structure must be one in which *only Churchill* is the antecedent of PRO. The logical form of (14g) is (48).

$$(48) \text{ Only Churchill } \lambda x (x \text{ remembers } x \text{ giving the BST speech})$$

In sum, if OC is formed via movement, then the interpretive properties illustrated in (14) follow rather directly.

One technical point is worthy of note before considering further facts friendly to the movement analysis. The above account relies on there being an LF chain relating controller to OC PRO at LF. This structural property is compatible

with many different views of movement. Thus, Hornstein (2000), Manzini and Roussou (2000), Martin (1996), and O'Neill (1995) implement a movement analysis in rather different ways. Nonetheless, one common feature of all four is that at LF PRO and its controller are links of a single common chain. This suffices to derive the facts above.<sup>46</sup>

### 1.5.2 *The Minimal Distance Principle*

There are other characteristics of OC PRO constructions that follow naturally if movement is involved. Rosenbaum (1967, 1970) observed that the controller/PRO relation generally obeys the Minimal Distance Principle (MDP). Thus examples like (49) must be controlled by the object, not the subject.

- (49) John<sub>1</sub> persuaded Mary<sub>2</sub> PRO<sub>\*1/2</sub> to go home

The vast majority of control structures with matrix transitive verbs like *persuade* require object control.<sup>47</sup>

Furthermore, when verbs can optionally alter their argument structures Rosenbaum (1967) noted shift in the potential controllers.

- (50)a. John<sub>1</sub> asked/begged/got Mary<sub>2</sub> PRO<sub>\*1/2</sub> to leave  
 b. John<sub>1</sub> asked/begged/got PRO<sub>1</sub> to leave

In (50a), *John*, the subject, cannot control PRO. In (50b), when the object is not generated (at least in overt syntax), the subject can be (and must be) the controller.

These sorts of cases can be accounted for in terms of the MDP. Assume for the moment that the MDP is a descriptively adequate generalization. Why does it hold? Note that it follows on a movement theory of control if one assumes that movement is governed by minimality, a standard assumption. To see this, consider what the derivation of (49) would have to be like were *John* the antecedent of PRO.

- (51)a. John [<sub>VP</sub> John persuaded Mary [<sub>IP</sub> John to [John go home]]]

The copies of *John* mark the history of derivation. Note that in moving from the embedded Spec IP to the matrix Spec VP *John* crosses the intervening DP *Mary*. This move violates minimality and is so barred. The only derivation not prohibited by minimality is one in which the DP in Spec IP raises to the next highest potential DP position, in this case the object. The derivation is illustrated in (51b).

- (51)b. John [<sub>VP</sub> John persuaded Mary [<sub>IP</sub> Mary to [Mary go home]]]

So, if OC PRO is the residue of A-movement, the MDP follows.

### 1.5.3 Adjunct control and movement

Rosenbaum (1970) extends the MDP to cases of adjunct control.

(52) John saw Mary after/before/while PRO eating a bagel

In (52) only *John* can control the PRO in adjunct position. Control from the object position is forbidden. This falls under the MDP on the assumption that *John* but not *Mary* c-commands the adjunct.

Unfortunately, this assumption is not obviously correct. For example, it is possible for objects to bind pronouns found within adjuncts, as Orson Wells taught us.

(53) John will drink no wine<sub>i</sub> before it<sub>i</sub> is ready for drinking

If we assume that to be interpreted as a bound variable a pronoun must be c-commanded by its antecedent, this implies that *it* in (53) is c-commanded by *no wine* at least at LF. But then objects should be able to control into adjuncts, contrary to fact.

Note that one might get around this difficulty if one assumed that the generalization described by the MDP holds for overt syntax. Were this so, then one could argue that at LF, the object might c-command the adjunct but in overt syntax it does not. The problem, given a minimalist grammatical architecture, is how to state this generalization. Hornstein (2000) offers a solution along the following lines.<sup>48</sup>

Minimalism has generally taken movement to be a complex operation made up of two simpler ones, copy and merge. Assume that this is so. Nunes (1995) observes that if this is correct, then it should be possible to copy an expression from one sub-tree and merge it into another.<sup>49</sup> Call this sequence of operations “sideways movement.” It is possible to analyze adjunct control as involving this sort of operation. In particular let us assume the following:<sup>50</sup>

- (54)a. Sideways movement is a species of movement and it is possible
- b. Adjuncts headed by *after*, *before*, *while*, etc. are adjoined to VP (or higher)

(54a) makes two claims: that sideways movement is a grammatically viable operation *and* that it falls under the same general restrictions that govern more conventional kinds of movement. (54b) says where the adjuncts of interest are merged. It assumes, contra Larson (1988) for example, that the adjuncts in (52) are basically adjoined to VP and are not generated within the VP shell. Assuming (54), consider a derivation of adjunct control structure like (55).

(55) John saw Mary after PRO eating lunch

The derivation proceeds as follows. First construct the adjunct; merge *eating* and *lunch* then merge *John* to *eating lunch* then merge *after*. This yields (56).

(56) after [John eating lunch]

Next construct the matrix. Merge *saw* and *Mary*. This yields a derivation with two unmerged sub-trees as in (57).

(57) [saw Mary] [after [John eating lunch]]

At this point, we have exhausted the numeration but the external argument position of *saw* is still unfilled. To fill it we copy *John* and merge it into the Spec VP position, as in (58).

(58) [John [saw Mary]] [after [John eating lunch]]

We then merge the two sub-trees and finish the derivation in the conventional manner yielding (59).<sup>51</sup>

(59) [John T<sup>0</sup> [[John [saw Mary]] [after [John eating lunch]]]]

There are several things to note about this derivation. First, *John's* sideways movement is what allows its case to be checked. Were it not to move from the adjunct, the derivation would crash as *John* would have an unchecked case. Second, the movement only occurs after *Mary* has merged into the complement of *saw*. Why could *John* not move into this position? Were it to do so then we could have object control so we need to discover what prevents this. I assume that economy restricts sideways movement and so one cannot move if the derivation can proceed without movement. In this case, we can merge *Mary* into this position so movement is blocked. This then forces *John* to move only after *Mary* has merged into the object position. This is why the controller of a PRO inside an adjunct cannot be an object.<sup>52</sup>

Thus, the spirit of the MDP can be captured in cases of adjunct control as well, if we see movement as the composite of Copy and Merge and thereby permit sideways movement.

This is actually a very positive result. The reason is that adjunct control displays all of the diagnostic properties of OC PRO.

PRO headed adjuncts require local, c-commanding antecedents.

- (60)a. \*John<sub>i</sub> said [that Mary left after PRO<sub>i</sub> dressing himself]<sup>53</sup>
- b. \*John's<sub>i</sub> picture appeared after PRO<sub>i</sub> shaving himself
- c. \*[that Bill left] seemed true before PRO<sub>arb</sub> noticing

The PRO in these adjuncts do not tolerate split antecedents:

- (61)a. \*John<sub>i</sub> said that Mary<sub>j</sub> left after PRO<sub>i+j</sub> washing themselves
- b. \*John<sub>i</sub> told Mary<sub>j</sub> a story after PRO<sub>i+j</sub> washing themselves

PRO headed adjuncts only have sloppy readings under ellipsis.

(62) John left before PRO singing and Bill did too

Thus, (62) only has the reading paraphrased in (63a). It cannot be understood as (63b).

- (63)a. . . . and Bill left before Bill sang  
 b. . . . and Bill left before John sang

In “Churchill” sentences – (64a) – they cannot take “Churchill” as antecedent. In other words, (64b) is not an adequate paraphrase of (64a).

- (64)a. Only Churchill left after PRO giving the speech  
 b. Only Churchill left after Churchill gave the speech

And within adjuncts of the appropriate type (e.g. purposives), they display the obligatory *de se* interpretation.<sup>54</sup>

(65) The unfortunate sent his report (in order to) PRO to get a medal

All in all then, cases of adjunct control display both MDP effects as well as the standard interpretive properties associated with OC structures. This is what we expect if they are indeed formed by movement.

It is quite interesting that cases of adjunct control can be analyzed in terms of movement. The reason is that such control is unlikely to be reducible to the thematic requirements of an embedding control predicate. In other words, the control one finds in such cases is *not* a function of the properties of the matrix predicate. Hence the controller cannot be some designated argument of the control predicate as has been proposed for cases of complement control. Nonetheless, these cases of control display all the diagnostic properties of OC. This suggests that, in at least some cases of obligatory control, the controller is structurally specified. Moreover, if the properties of OC PROs within adjuncts are structurally determined then it would be odd to treat cases of complement control (where the very same properties appear) as derived by entirely different operations and in entirely different ways.

This section has shown how movement accounts for some central features of obligatory control configurations. The account covers standard cases of complement control as well as adjunct control. The general characteristics of OC PRO make sense once OC PRO is seen as the residue of overt A-movement. In the next section, we consider some possible empirical problems for this sort of approach.

## 1.6 Problems for movement

---

Several authors have noted empirical difficulties for the movement analysis of control. This section reviews empirical challenges to the approach.

### 1.6.1 Promise: *and markedness*

One prominent argument rests on denying the general validity of Rosenbaum's MDP.<sup>55</sup> The form of the argument is as follows: movement is subject to minimality. Minimality is an inviolable condition on movement. Thus, we should find no cases of subject control in transitive control predicates. However, it is well known that there are a class of predicates that appear to violate the MDP, for example verbs like *promise* require subject control.

(66) John<sub>1</sub> promised Mary<sub>2</sub> PRO<sub>1/\*2</sub> to leave

Thus, a movement based approach to control cannot be correct as it is incompatible with violations of minimality.

There is something curious about this argument. It appears to concede Rosenbaum's observation that *in general* one finds object control in transitive control predicates. However, the existence of a semantically coherent class of exceptions (the *promise* class which includes *vow* and *commit* among others), is taken to indicate that the reduction of the MDP to minimality cannot be correct. Here is Landau's version of the argument.

Hornstein (1999) is aware of the exceptions to the MDP, but rather than attributing them to some hidden (double object) structure, he proposes to view the MDP as [a, sic] markedness condition. (Landau, 1999, pp. 231–2)

The problem is that the MDP is not a primitive in Hornstein's system: "the MDP reduces to the MLC [Minimal Link Condition/i.e. Minimality NH]." This reduction is taken to be a strong argument in favor of the whole approach. But then any properties of the MDP should follow from properties of the MLC. In particular, if the MDP is a markedness condition, so should the MLC be. This is clearly not the case, however; the MLC is exceptionless, and its violations are sharply ungrammatical, whereas the "marked" violations of the MDP (i.e. subject control) are perfect.

There are several replies that one can make to this argument. First, let us be clear what it would mean to say that the minimality should be understood as a "markedness condition." One way to interpret this is that verbs that fail to respect minimality are marked in the sense that environments that appear to allow it will necessarily be hard to acquire. What "hard" means is that its acquisition will be data-driven with the language acquisition device requiring considerable evidence before it abandons the view that the witnessed operations are actually permitted in the observed environment. That is all that "marked" means.<sup>56</sup> In this sense, then, a movement theory treats control verbs like *promise* as marked in that they should be harder to acquire than the normal run of transitive control verbs like *persuade*. As Hornstein (1999) observed (and Landau, 1999 recognizes), there is acquisition data from C. Chomsky (1969) indicating that it is indeed the case that *promise* predicates are acquired late.<sup>57</sup>

Note, the MDP and the minimality approach to control provides the beginnings of an *explanation* for *promise*'s odd acquisition profile. Contrary to Landau (1999), it makes little difference whether the MDP is itself part of UG or is reduced to more fundamental properties of UG. In either case, if interpreted according to the logic of markedness, the relevant UG condition simply states that derivations that deviate from the condition, be it the MDP or minimality, will require substantial data to acquire. This, to repeat, seems to be true for *promise* verbs.

Moreover, any view that "regularizes" the properties of *promise*, for example by claiming that it falls under a broader generalization in terms of which its behavior is grammatically impeccable, *cannot* account for why it is acquired late. If indeed there is nothing grammatically (be it syntactic or semantic) untoward about *promise* then it should be learned as smoothly as any other transitive control predicate.<sup>58</sup>

This does not yet settle the matter however. It is still behooves us to find out in what way *promise* is marked. There are several ways of stating this. Most baldly, one can state that *promise* type verbs are exceptional in allowing their objects to be ignored for purposes of minimality. This is not pretty, but it is not impossible to state.<sup>59</sup>

A second more interesting possibility is that *promise* has a null preposition in its structure similar to the one found in other *promise*-like verbs such as *vow* and *commit*.

(67) I<sub>1</sub> vowed/committed \*(to) Bill PRO<sub>1</sub> to leave<sup>60</sup>

In (67), there is no obvious problem for a movement analysis in that the object of the preposition does not c-command the position PRO occupies, at least in overt syntax. In this respect, the examples in (67) would be analogous to those one finds with the raising constructions in (68).

(68) John seemed/appeared to Mary t to be nice

As is well known, English is somewhat exceptional in allowing movement across the indirect object in a raising construction (see Chomsky, 1995). Even the movement in English might seem problematic given that the indirect object is known to induce principle C effects in its complement. In this regard, (67) and (68) are similar.

- (69)a. \*John seemed to them<sub>1</sub> to like the men<sub>1</sub>  
 b. \*John vowed to them<sub>1</sub> to hit the men<sub>1</sub>

Let us assume that *promise* is similar to these other control verbs in having an indirect object in overt syntax and that this preposition becomes null (perhaps by incorporating into the verb) in the course of the derivation.<sup>61</sup> In many other languages *promise* is an overt indirect object and in its nominal form the *to* surfaces, so this may not be an unreasonable assumption.

## (70) John's promise \*of/to Mary

If this were so, then the prepositional complement structure of *promise* would be opaque to the child and its late acquisition would be expected.

Observe that this is just one way of implementing the idea that *promise* is odd. It retains the view that it is marked and proposes a specification of how it is marked. However, to repeat, the movement analysis could survive if this specific proposal were incorrect and it was simply asserted as a lexical fact (similar to what one has with ECM verbs) that *promise* allows movement across its object.

One further point is worth making. Landau (1999) suggests that minimality violations are not found in other kinds of A-movement. This is not quite accurate. The problems with raising across the prepositional arguments of *seem* and *appear* have been touched on above. These are quite parallel to what we find with *vow* and *commit*. There are also raising analogs of *promise* in English.<sup>62</sup>

- (71)a. John<sub>1</sub> strikes Bill as t<sub>1</sub> dumb  
 b. It strikes Bill that John is dumb  
 c. It<sub>1</sub> strikes Bill as t<sub>1</sub> likely that Bill is dumb  
 d. John<sub>1</sub> strikes Bill as t<sub>1</sub> likely t<sub>1</sub> to be dumb

The raising verb *strikes* seems to violate minimality in allowing movement across its object in (71a,c,d). Here too, it is plausible to think that the apparent object is actually an indirect object. Whether this is so or not, it appears that contrary to what Landau (1999) asserts, there are apparent cases of A-raising that violate minimality.

In short, the existence of a limited number of counter examples to the MDP do not, I believe, invalidate either Rosenbaum's initial generalization or attempts to reduce the MDP to more basic properties of UG. In fact, the marked nature of *promise*, displayed in its tardy acquisition profile, argues in favor of a theory that treats *promise* as an exception and speaks against regularizing its grammatical status.

### 1.6.2 Control shift<sup>63</sup>

There is a second phenomenon that suggests that what one sees in *promise*-like cases is sensitive to semantic conditions. Consider cases like (72).

- (72)a. John<sub>1</sub> asked/persuaded/begged/petitioned Mary<sub>2</sub> [PRO<sub>1/2</sub> to be allowed to leave early]  
 b. ?John<sub>1</sub> promised Mary<sub>2</sub> PRO<sub>1/2</sub> to be allowed to leave

As Landau (1999, pp. 212–14) notes, the examples above vary in acceptability from speaker to speaker and are subject to a variety of pragmatic influences in determining acceptability.<sup>64</sup> To my ear, the results are best when the embedded context is actually of the form *be allowed . . .* At any rate, the contexts where

this is acceptable are rather subtle.<sup>65</sup> What implications does this have for the notion that OC is parasitic on movement? So far as I can gather, none whatsoever. Let me explain.

The dominant view concerning these cases of control is that they are due to “special mechanisms” (this is Landau’s term). Various proposals exist but they all seem to involve some specialized mechanism designed to obtain the relevant control facts. There is little reason to think that such a special mechanism could not also be grafted onto the movement theory.

Moreover, these phenomena are problematic for a movement account only if they involve obligatory control, for only then does a movement approach to PRO propose that movement is involved in the derivation of the control structure.<sup>66</sup> Non-obligatory control involves the binding of a null pronominal expression (see below).

So the question arises whether in control shift contexts we have a case of OC or NOC. The data suggest that the latter is the case. Consider the following data.

- (73)a. John was asked/begged PRO to be allowed to leave early
- b. John’s mother asked/begged Mary PRO to be allowed to shave himself before dinner
- c. John petitioned/begged/asked Mary PRO to be allowed to leave early and Frank did too (OK with John’s leaving early)
- d. John asked/begged Mary PRO to be allowed to shave each other
- e. The unfortunate petitioned congress PRO to be allowed to get a medal

In each of these cases, it appears that the PRO allows NOC readings. In (73a) it has no antecedent. In (73b) the antecedent can be the non-c-commanding *John*. (73c) allows a strict reading under ellipsis. (73d) permits split antecedents and (73e) allows a non *de se* reading. As these are standard characteristics of NOC structures, it leads to the conclusion that control shift involves a change from an OC to a non-OC structure.

If this is correct, it leaves several questions open. For example, it does not address the issue of why in these cases we get a shift from OC to NOC. One possible reason for this comes from considering sentences like (74).

- (74)a. John asked/begged/petitioned Mary that Peter be allowed to leave
- b. \*John asked/begged/petitioned Mary<sub>1</sub> that she<sub>1</sub> be allowed to leave

The indicated binding of a pronoun in cases like (74) leads to unacceptability. If so, we have an explanation for why the analogous obligatory control reading should be unavailable. Perhaps the unavailability of the required reading allows the emergence of the other NOC structure. We note below that NOC only arises if OC is blocked. (74b) indicates that identifying the embedded subject with the matrix object in environments analogous to those in control shift leads to unacceptability. This plausibly allows the NOC structure to emerge and permit the otherwise blocked binding relations. Why, however, this holds largely with *be allowed to* and not more widely remains a mystery.



here agreement is indirect with PRO first agreeing with the embedded T-Agr and this complex then agreeing with the matrix F<sup>0</sup>. This is the configuration of PC. The important feature of the analysis is that finite T<sup>0</sup> must move to C thereby inducing a blocking effect on agreement and thus licensing the PC interpretation.

#### 1.6.4 *Some problems for Landau (1999)*

It is unclear (at least to me) how the blocking effect in (76b) is meant to work technically. The idea is reported as follows (Landau, 1999, pp. 80 ff.). Following Chomsky (1999, pp. 4 ff.), the F feature in (76a) is a probe with DP and PRO as goals. PRO is unspecified for the number feature in non-finite clauses. The matrix F has number features. AGREE is a “maximizing operation, matching as many features of the probe it can with the goal. Since semantic number is one of F’s features, PRO inherits it” (Landau, 1999, pp. 81). (76b) involves a different set of AGREE operations. First AGREE holds between the lower T-Agr and PRO.<sup>69</sup> Neither is specified for semantic number. Then the T-Agr raises to C. It then enters a second AGREE operation with the matrix F as probe. It is assumed that this raising blocks F from agreeing with PRO directly.

This proposal is not without some problems, I believe. First, it is assumed that both EC and PC constructions are CPs. This would appear to put PRO outside the reach of the matrix probe on a phase based theory by the Phase Impenetrability Condition (PIC).<sup>70</sup> The proposal must assume that this C is porous unless T-Agr adjoins to it. It is theoretically unclear why adjunction should so function. What is clear is that this assumption is empirically vital. Note that were this not assumed then F would be able to reach PRO in both structures in (76), nullifying the difference between the two.

Second, it is unclear how the probe F can “see” PRO in (76a) given that the matrix DP intervenes. It is generally assumed that AGREE is sensitive to minimality in that the Probe cannot see through the first accessible goal where feature Match applies (see Chomsky, 1998, p. 38, (40iii)). The minimality restriction is important to the probe/goal/AGREE based theory of feature checking. Chomsky, for example, accounts for the absence of nominative agreement in Icelandic constructions such as (77) on this basis (1998, p. 47, (51ii)). He notes that the nominative/Agr probe in the matrix cannot target the embedded object as goal as the *John* intervenes.

(77) me(DAT) seem (pl) [t<sub>me</sub> [John(DAT) to like horses(pl, NOM)]]

The structures proposed in (76) should prevent F from targeting PRO if minimality were in effect. The matrix DP intervenes between F and PRO in (76a) and F and T-Agr in (76b) and so should block any AGREE relation between F and these expressions. It appears, therefore, that minimality is irrelevant in these cases, and so the AGREE operation assumed here must be different from standard instances of AGREE. In other words, the AGREE operation underlying control is a special case of AGREE with properties specially suited

to the phenomenon at hand. Clearly, this negatively affects the explanatory force of the proposed analysis.

Third, it is unclear how AGREE is understood in this proposal. In Chomsky (1998, 1999) it is taken to be an operation that compares features of two expressions and checks those that need checking. The match is either under strict identity or non-distinctness. In the current proposal, it is neither of these operations. Rather it is both a checking and a copying rule. The probe checks to see if the goal's features match and *if* the feature sets do not conflict then the probe can transfer to the goal those that it has, but which the goal does not. This way, PRO acquires semantic features contextually (Chomsky, 1999, pp. 76–7 and esp. (99d)). Note that this is a different conception of AGREE in that the operation involves both checking and copying.

Furthermore, the copy operation appears to violate the Inclusiveness Condition. Landau (1999, p. 76) accepts that semantic number is “determined by the lexical entry” as opposed to purely formal phi-features that are appended to expressions prior to their entering the derivation. What of PRO? Landau proposes that PRO acquires its *semantic* number contextually. As noted, this violates Inclusiveness and so needs strong independent motivation in a minimalist context. Moreover, PRO is unique among lexical items in having its number specification contextually specified. Other lexical items, Landau (1999, p. 76) assumes, are inherently specified for semantic number “in the lexical entry.” This, then, makes PRO special in two ways: its null case is special and its semantic number features are special. As noted at the outset, the more the PRO's properties are special or unique the less its properties are explained.

Fourth, it is not clear how the mechanism proposed actually codes the basic control property. In fact, it is unclear whether the proposed AGREE operations suffice to determine the antecedent of PRO. Consider why. What AGREE does is check features. It is clear that having the same features as some other expression cannot be sufficient to establish antecedence or control. One might reply that it is not *having* the same features that is crucial but having one's features established by being checked against a common probe that results in the control relation. This would suffice to establish control between DP and PRO in (76a) as they are both taken as goals of the same F probe (issues of minimality aside). However, why should DP be taken as controller of PRO in (76b)? All they have in common is that they share some features as they have agreed with two heads that themselves agree. It seems that what is being assumed here is that if B agrees with A, C agrees with A, and D agrees with C then B and D agree as well. In (76b) this reasoning is as follows: If DP agrees with F and T-Agr agrees with F and PRO agrees with T-Agr then PRO agrees with DP. The assumption seems to be that DP controls PRO *in virtue of* PRO's agreement with DP. However, it is not clear how it is that PRO does so agree. Let me explain.

AGREE establishes a relation between two expressions. To see if the chain of specific agreements above implies control requires seeing what kind of relation AGREE is. The minimal assumption here must be that it is transitive. However, this is not enough. Landau must also be assuming that AGREE is

symmetric. However, this is hardly self-evident and seems counter to the view advocated by Chomsky (1998, 1999) where AGREE seems to be an asymmetric relation. It is not that A and B agree but that the Probe A agrees with the goal A and this introduces an asymmetry into the relation: so Probes AGREE with goals not vice versa as Probes c-command and seek out goals and have their features checked by goals rather than the reverse. If this is correct, that is if AGREE is not symmetric, then we cannot establish a control relation through the series of AGREE operations noted above even if we assume that AGREE is transitive. The reason is that we cannot deduce that DP agrees with PRO from the fact that the other agreement operations took place. But if we cannot do that we cannot derive the fact from the proposed set of operations that in (76b) PRO is controlled by DP as they have not been related by AGREE. Only if we assume that AGREE is symmetric (and thereby cancel the inherent asymmetry in a probe/goal system) and transitive can we get DP and PRO in any sort of relation able to undergird control.

It should be observed that similar difficulties underlie the standard case of EC as outlined above. In (76a), F agrees with both PRO and DP. But from this it does not *follow* that DP AGREES with PRO or vice versa. From aRc and aRb it does not follow that cRb or bRc even if we assume R is transitive. But if not, how does AGREE code the antecedence relation?<sup>71</sup>

There is another conceptual point I would like to briefly make. Say that AGREE is symmetric and transitive. Why should multiple agreement code control? In particular why should the fact that two DPs (one a *PRO*) both agreeing with the same functional head or both agreeing with different functional heads that agree with each other lead to control? Note, the question is not *could this be the case?* We can stipulate anything we want. The question is why it *should* be so. Why should multiple agreement operations set up control relations especially when the agreement is quite indirect? Clearly this is not so for other cases of multiple agreement. There is no conceptual reason that I can see why it should be so here. Let me put this another way. Landau (1999) seems to tacitly hold the following in order to get the relevant antecedence relation from AGREE: A and B agree iff A is the antecedent of B or B is the antecedent of A. In other words, it is not enough to observe that antecedence implies agreement, the converse must hold as well. There is little reason to think that anything like this is well motivated either empirically or conceptually.

Let me say this one more way. On a movement theory, control is the reflex of chain membership. In particular, it is the reflex of a chain's spanning multiple thematic positions. Control holds because of the COPY operation and is defined over identical instances of the same expression. On Landau's (1999) proposal, control is a reflex of AGREE but it is not clear why AGREE should have the power to establish antecedence, especially when it is as round about as in (76b) where the DPs only agree with each other in virtue of agreeing with a common functional head (in EC) or distinct heads (in PC).<sup>72</sup>

There are further problems with the proposed analysis. I briefly mention a few of them here. Section 1.4.2 reviewed some problems for theories of null case that used the standard diagnostics for the tense of non-finite clauses following the original work of Stowell (1982). These problems are compounded

in Landau's proposal. He assumes that both EC and PC control structures have PRO subjects. However, this defeats at most a principled version of the null case theory, one like Martin (1996) that tries to assimilate null case to nominative by having it be the property of +tense non-finite clauses. Landau must reject this version of the null case theory and leave it simply as a stipulation that control T<sup>0</sup>s assign case regardless of their feature composition. He codes this by assuming that control infinitives all contain anaphoric Agr projections regardless of their tense specifications. This use of Agr works but has no apparent motivation beyond coding the fact that both environments can host PRO. It is presumably this anaphoric Agr feature that licenses PRO via some sort of null case. This makes evident the stipulative nature of the null case theory and leaves us without any principled explanation for the distribution of PRO. In effect, Landau's theory appears to sacrifice a principled account for the distribution of PRO in order to accommodate the PC property of some control configurations.

Note incidentally, that this approach requires some version of the EPP. The reason is that in (83a) the PRO must raise to get outside the domain of the lower T-Agr and get close enough to the higher matrix F projection. There is currently considerable skepticism regarding the status of the EPP.<sup>73</sup> This proposal appears to require it.

There is also a problem with the generalization that Landau proposes. It ties PC to +tensed non-finite clauses. However, we have evidence that gerunds too can have PC readings.

(78) John prefers meeting/to meet at 6

If this is correct, then these gerunds must have +tensed T<sup>0</sup>s and have movement of this T<sup>0</sup> to C. Section 1.4.2 reviewed evidence that gerunds did not have +tense T<sup>0</sup>s yet these can display PC. Recall that Stowell (1982) argued against treating gerunds as +tense. Moreover, there is very little evidence that gerunds ever have a CP layer. There are no gerundive indirect questions for example, nor do they come with overt complementizers. If they are bare TPs, a common assumption (see Pires, 2001 for discussion), then they should never be able to host PROs at all, nor should they be able to manifest PC readings, contrary to fact.

There is one more curious feature of Landau's analysis. He argues that there is movement of T-Agr when the latter is tensed, but not when it is untensed. What is curious is that the latter set of cases manifest a strong aspectual dependency between the matrix and the embedded clause. Thus, for example, the event referred to in the embedded clause in (79) is cotemporaneous with respect to the matrix event. In (79a) John's finishing drinking the wine terminates an event of drinking while in (79b) the starting point initiates the drinking and in (79c) the trying spans the drinking.

- (79)a. John finished PRO drinking the wine  
 b. John started PRO drinking the wine  
 c. John tried drinking the wine

How should one code this dependency? Note that it cannot be via selection given the intervening C between the matrix verb and the embedded T-AGREE. One way of finessing this problem is to raise T-Agr to C (if there is one). However, by assumption, this T-Agr does not raise or it would permit PC readings. Thus there are dependencies that would naturally be coded by raising to C that cannot be so handled on this analysis. On the other hand, Landau does not say what one gains by raising +tensed T-Agr to C. It is against the spirit of the minimalist program to simply assume that such raising occurs without any other interface requirement being furthered. One can force such movement by placing a feature in C. However, this does not explain *why* there should be such movement and is thus, on methodological grounds, without explanatory foundation.

One last point. The movement required to track the PC data must be overt given the motivation behind AGREE as an operation able to dispense with LF like operations. There is, to my knowledge, little evidence that such overt T to C movement takes place, at least in English. Thus, in embedded questions, there is little evidence of T to C movement in overt syntax for full finite clauses, hence the absence of Aux inversion in embedded clauses. One could treat this movement as covert. However, here too there is a problem. It is unclear that covert movement suffices to block the AGREE operation licensing partial control from applying *prior* to this movement. But then the structure of EC should be met and the structural distinction between EC and PC will be idle.

The above remarks have concentrated on technical aspects of Landau (1999). They could surely be finessed. However, I would conclude from the above that Landau (1999) is more a description of partial control than an explanation of it. The question still lingers, however, whether the existence of partial control is incompatible with a movement theory or whether the latter can be supplemented to accommodate the EC/PC distinction.

### 1.6.5 *Movement and partial control*

If there is no element like PRO how is partial control to be accommodated? There are several brute form approaches that could be pressed into service which, though not particularly elegant, are no worse than the other approaches in the literature. The least interesting (yet adequate) approach would be to treat PC as the result of an optionally applicable meaning postulate licensed by certain matrix verbs when taking control complements. If Landau's (1999) description is correct, the relevant class of verbs are those that take non-finite +tense complements.

What would the meaning postulate look like? It would have roughly the content of (80).

- (80) If "DP Vs [<sub>TP</sub> to VP]" then "DP Vs [<sub>TP</sub> DP and some contextually specified others to VP]"

With regard to an example like (81a), (80) would license the inference (81b).

- (81)a. John wants to meet at 6  
 b. John wants John and some contextually specified others to meet at 6  
 c. John wants [John to meet at 6]

The PC interpretation is then provided by (81b) or, redundantly, the conjunction of (81c) and (81b). (81c) alone is what the syntax licenses. It licenses the interpretation that John wants himself to meet at 6. As it is a property of verbs like *meet* that they require that the class of “meeters” be semantically plural, (80) applies, licensing the inference (81b). But now John wants both himself and some contextually specified other(s) to meet at 6 and this suffices to give the predicate the requisite semantic plurality and the indicated partial reading. In cases involving verbs like *meet* the application of (80) is forced, otherwise unacceptability results. In other cases, it may or may not apply (recall (80) is optional) yielding either a PC or EC reading as the case may be.<sup>74</sup>

This is clearly a *very* uninteresting account of the PC phenomenon as it is specifically crafted to cover the partial reading data. Despite this, the approach has more or less the correct empirical requirements for the purpose at hand. For example, if we assume that anaphors such as *themselves* and *each other* require syntactic licensing (a standard assumption), the meaning postulate approach implies that structures that license PC readings will not be able to license reciprocals and plural reflexives. This is correct (as Landau (1999) observes). Note that (81a) only has the syntactic phrase marker (81c) in which *John* (and only *John*) is the subject of the embedded clause. Being singular, it cannot license a plural anaphor (cf. (82)) and so the sentence is unacceptable.

- (82)a. \*John wants to meet each other  
 b. \*John wants to wash themselves

A second virtue of this approach relates to the fact that PC is a property of control *complements*. It is absent from adjuncts.

- (83)a. \*John saw Mary after/without meeting/gathering at 6  
 b. \*John saw Mary early (in order) PRO to meet/gather at Max’s at 6

This too is expected if PC is due to a meaning postulate tied to a higher predicate. Predicates exercise their lexical powers over their arguments, not over adjuncts. As such, we might expect to find PC absent within control adjuncts.<sup>75</sup>

(83b) offers a possible contrast between this approach and Landau’s (1999) proposal. The latter, recall, ties partial readings to the presence of +tense non-finite inflections. The purpose clause in (83b) shows some of the hallmarks of being +tense. For example, we can have contrasting adverbs in the main and the adjunct clauses.

- (84) John saw Mary yesterday (in order) to leave early tomorrow

If this is a diagnostic of +tense non-finite inflections, then the purpose clause must be tensed (see Landau, 1999, p. 71). But then we would expect to see PC

in such structures, contrary to fact. Observe that these purpose adjuncts have all of the characteristic properties of obligatory (adjunct) control.<sup>76</sup> For example, it is subject oriented, (85a); requires a prominent antecedent, (85b); and a local one ((85c) is only acceptable with the purpose clause modifying the matrix); requires sloppy readings under ellipsis, (85d); forbids split antecedents, requires a *de se* reading, (85a); and gets the bound reading with *only* DP antecedents, (85f).

- (85)a. John<sub>1</sub> saw Mary<sub>2</sub> in order PRO<sub>1/2</sub> to get a medal
- b. John<sub>1</sub>'s mother saw Mary in order PRO<sub>\*1</sub> to get a medal
- c. Bill<sub>1</sub> said John saw Mary in order PRO<sub>1</sub> to get a medal
- d. John<sub>1</sub> saw Mary in order PRO<sub>1</sub> to get a medal and Bill did too
- e. \*John<sub>1</sub> saw Mary<sub>2</sub> in order PRO<sub>1+2</sub> to shave each other
- f. Only John saw Mary in order PRO to get a medal

Thus, PROs in these contexts act like OC PROs, lodge in non-finite sentences that appear +tense, and also forbid PC readings. Were PC a property of +tense non-finite inflections we would expect PC to arise here as well, contrary to fact. This is compatible with a meaning postulate approach on the assumption that PC is due to lexical powers exercised by predicates over their embedded complements.

Consider one further interesting case.

- (86) John wants/prefers to talk about himself all together at dinner tonight

(86) can have a PC reading, as *all together* indicates. If so, what binds the reflexive? If the embedded subject has a plural or unspecified number specification then why does the reflexive appear in the singular? It is generally the case that reflexives agree in syntactic number with their antecedent.

- (87)a. The committee congratulated itself/\*themselves
- b. The pants were folded over themselves/\*itself

If so, this should indicate that the syntactic antecedent of *himself* in (86) has singular formal number features. In short, it appears that in PC cases, there is evidence both that the PRO is semantically plural and that it is syntactically singular. This is a problem for Landau's (1999) proposal.

However, the meaning postulate approach predicts that a singular reflexive should be licensed here as, *in the syntax*, the embedded subject is a copy of *John*.

- (88) John wants/prefers John to talk about himself . . .

On Landau's (1999) account, the embedded subject is not singular but is either plural or unspecified for number. Neither assumption is consistent with the fact that a *singular* reflexive is licensed here.<sup>77</sup>

I conclude that it is *possible* to combine a movement account with some other mechanism to obtain the PC facts. The meaning postulate approach, though very crude, has the desired properties: it is tied to specific lexical predicates and so allows a distinction between EC and PC complements, being a meaning postulate it is not expected to interact with the syntax and so not license plural anaphors, and being a lexical property it is not expected to affect adjuncts.

This section has sketched an approach to partial control consistent with the movement theory of control. This suffices to demonstrate that the simple existence of partial control does not *in and of itself* argue against the possibility of a movement approach to obligatory control. This said, the existence of PC readings does raise a challenge for the approach to control phenomena outlined here. One of the charms of the movement approach is that it derives the basic distributive and interpretive facts concerning control without further assumptions. Once it is seen as an instance of movement, all the details follow. Partial control phenomena do *not* follow from the movement theory. What I have argued is that they are compatible with movement, not that they follow from it. Thus the EC versus PC distinction as well as why PC holds at all remains as a puzzle for the minimalist ambition of entirely eliminating any specific mention of control within UG.<sup>78</sup>

### 1.6.6 Control in nominals

Finally, let us turn to what (in my view) is the greatest challenge to the movement theory of control. We have, to this point, restricted attention to control within verbal complements. However, there is also quite a bit of (at least) apparent control inside nominals, as in (89).

(89) John's attempt/plan/desire PRO to leave

The question is how this is to be handled given a movement theory.

The problem for a movement account is quite complex as the data are very intricate.<sup>79</sup> The aim here is *not* to handle all of the data that have been put forward. Rather it is to suggest one way of construing the facts so that their properties are compatible with a movement account. The form of the argument that I present is not very novel. The suggestion is that what we find in nominals is not obligatory control. This suggestion goes back at least to Williams (1980) who argued that PRO within nominals did not require syntactic antecedents and so were different in kind from verbal cases of control. Thus, alongside (89) we find (90) where there is no overt syntactic controller.

(90) any attempt/plan/desire PRO to leave

Moreover, the contexts in (90) are ones that support arbitrary readings, as Williams (1980) also observed.

(91) Any attempt/plan/desire to conceal oneself

In verbal settings, obligatory control contexts do not permit the arbitrary interpretation of PRO.

- (92) \*It was attempted/planned/desired to conceal oneself

There are other indicators that what one has in such contexts is not obligatory control. For example, nominals can support split antecedents, in contrast to their verbal counterparts.

- (93)a. John approved Bill's attempt/plan PRO to sneak each other/ themselves into the party  
 b. John distrusted Bill's desire PRO to promote each other/ themselves  
 c. \*John said that Bill attempted to sneak each other/ themselves into the party  
 d. \*John said that Bill desired to promote each other/ themselves

Note too that in the case of (93a,b) one of the antecedents is non-local in the sense of being across an intervening subject. This is forbidden in the verbal cases in (93c,d) where *John* cannot be an antecedent of PRO.

It is also possible to get strict readings under ellipsis.

- (94)a. John's plan to sneak himself into the convention was not as clever as Mary's  
 b. John's attempt to sneak himself into the convention failed though Mary's succeeded

In both (94a,b) it is possible to read the ellided portion strictly, for example (94a) can be understood as Mary's plan for John to sneak himself in.

There is one further contrast of interest. It seems that verbs and their nominal counterparts differ with respect to partial control. Verbs like *try/attempt* resist PC interpretations. However, their nominal counterparts are quite felicitous with the PC reading.

- (95)a. \*John said that the chair attempted to meet together at 6  
 b. John criticized the chair's attempt to meet together at 6

Thus, on a range of diagnostics for OC it appears that verbs and their nominal counterparts differ. This suggests that Williams (1980) was largely correct in assuming that what we have in nominals is non-obligatory control.<sup>80</sup>

Consider what this means if correct: the movement theory is not put at risk by the control facts within nominals. The reason is that the movement theory of control only applies to obligatory control structures. Non-obligatory control is *not* due to movement. Rather, as we will see below, it rests on the assumption that NOC's properties are due to the presence of a null pronominal (akin to *pro*) in NOC contexts. If this is correct, then we do not expect to see the marks of OC in nominal structures.<sup>81</sup>

This is, of course, a negative result. It does not say what is going on within nominals. Rather it says that whatever it is, it is not the province of the theory of movement. For present purposes this is enough. If correct, it insulates the movement theory of control from the nominal data.

However, it still leaves the question of how to explain the fact that in examples like (89) there is a very strong preference for understanding John to be both the attempter/planner and the leaver. Why so? What follows are some very speculative remarks aimed at this question.

DPs with genitive “subjects” have one general interpretive feature. The genitive is interpreted as in some contextually specified way related to the rest of the DP.<sup>82</sup> This holds even where there is no apparent control going on at all.

- (96)a. John’s book
- b. John’s claim that Frank kissed Mary
- c. John’s rumor that Frank kissed Mary

In each of these cases, *John* is interpreted as related *in some way* to the book, claim or rumor. These specifications can be very lax. For example, (96a) allows John to be the owner, author, subject, caretaker, etc. For (96b,c) more obvious relations are salient. Thus, one natural way for a claim to be John’s is that he made it. But another can be that he defends it, moots it, discovers it, publicizes it, or publishes it. Similarly for rumors. Thus, (96b,c) can be used with the understanding that John is the claimer or originator of the rumor but they need not be so understood. What is required is that there be a contextually specified relation between the genitive and the following nominal.

With this in mind, consider another instance of a similar restriction within sentences.

- (97)a. As for John, Mary likes him a lot
- b. Concerning John, Mary said that he was great
- c. As for the Canadiens, I like Toe Blake
- d. Concerning Toe Blake, the Habs knew how to win

Sentences with left positioned “hanging” topics like those in (97) require that the topic be related in some way to the proposition that follows. One rather easy way to accomplish this is by having a pronoun interpreted as anchored to the topic. Thus, the most salient reading of (97a,b) is the one in which *John* is the antecedent of *him/he*. This sort of binding, however, is not required. The relation can be quite a bit looser, as illustrated by (97c,d). Here, the sentences are felicitous if one knows, for example, that Toe Blake coached the great Montreal hockey teams of the late 1950s and early 1960s. The relation can be quite loose. But a relationship there must be. And if one is salient, it is often the one adopted.

With this in mind, consider cases like (89), repeated here, one more time.

- (89) John’s attempt/plan/desire PRO to leave

*John* needs to be related to the nominal that follows in some way. Clearly, one very salient way to relate to the following is to be the leaver and the attempter/planner/desirer. However, given the right context, it seems to me that this is not at all forced. So, for example, John need not be the planner in (89) if he is related to the plan in some other way, say as the backer of the plan, or expositor. Nor need he be the leaver. So for example, imagine a plan John is backing in which he gets hidden by being buried in a pit. Then, one can use a phrase like (98) to refer to this plan where *him* refers to *John*.

(98) John's plan to bury him in the pit just won't work

Here John is not the planner, nor is he a burier under the intended interpretation. However, *John* is related in the roundabout way described and so all is well. I believe that similar ingenuity leads to acceptable readings in the other cases as well. If this is characteristic of control within nominals, then it suggests that what has been interpreted as control, is not due to mechanisms of control at all, but is the reflex of a restriction imposed by the aboutness relation that genitive's impose on DP phrases.

Two more points before leaving this very complex topic. First, once the problem is finding an aboutness relation then many factors may come into play. For example, consider the sentence in (99).

(99) John's attempt to leave was amusing

One plausible paraphrase of (99) is (100).

(100) John attempted to leave and it was amusing

Note that under the paraphrase (100) of (99) John is necessarily understood as both attempter and leaver. Note that the paraphrase involves verbal control, not control within nominals. (100) is not the only paraphrase of (99) but it is a plausible one and maybe even a salient one. To the degree it is salient, to that degree *John* will be understood as "controlling" the PRO. Not because it does so in (99) but because it does so in (100) and (100) is thought to be *in the context* a good paraphrase of (99). In this way, the properties of verbal control can "leak" into those of nominal control. However, it is worth bearing in mind that there might well be cases in which this sort of paraphrase is not particularly felicitous and in those cases the indicated (control) inference will be blocked.<sup>83</sup>

Second, the discussion here is much too cursory to be compelling. However, my aim has been more modest. Though the facts regarding control in nominals are very complex and though I have barely scratched the surface, I hope to have shown that the control facts as generally described, at least in some central cases, are more elusive than generally thought. Furthermore, there are reasons for thinking that the control seen here is not obligatory control at all. If this is correct, it cannot be understood as undermining the movement approach to OC.

### 1.6.7 Conclusion

This section has reviewed a variety of objections to a movement theory of OC. I have argued that none of the proposed objections are fatal. This said, the facts discussed are complex and the conclusions reached here should be considered very tentative. What I would like to emphasize, however, is that given the theoretical virtues of the movement theory, it behooves us to have compelling empirical reasons against it before we abandon it. The objections reviewed above, though provocative, do not yet meet this higher standard, in my view.

## 1.7 Non-obligatory control (NOC): a short note

---

For the movement theory, NOC<sup>84</sup> is the elsewhere case. What this means to say is the following: if the movement theory is on the right track, then OC and its attendant properties receive a principled grammatical account in terms of movement via multiple theta positions. When such movement can occur, it must. When it cannot occur, NOC results. Several questions arise assuming this story to be correct.

First, are NOC structures all of a piece? In other words, do they all embody identical structures or are there various kinds of NOC configurations?

Second, what underlies the NOC properties noted in (16) above? Is there some structural reason for the fact that NOC configurations do not require local *c*-commanding antecedents, license strict readings under ellipsis, etc.?

Third, why is it that OC and NOC structures are in complementary distribution? More exactly, why is it that NOC cannot occur precisely where OC does? Note that we empirically know that this is *not* possible. Were it so we would have little evidence for the existence of OC given that the properties of OC configurations are a proper subset of those displayed by NOC structures. However, we would like a theoretical reason for why it is that NOC is possible just in case OC is not.

Let us now consider NOC PRO in light of these questions. There is one particularly interesting fact about NOC structures, at least in a core set of cases. They occur within islands. Thus, the examples in (16), repeated here, all find NOC PRO inside a subject of a finite clause.

- (16)a. It was believed that PRO shaving was important  
 b. John<sub>i</sub> thinks that it is believed that PRO<sub>i</sub> shaving himself is important  
 c. Clinton's<sub>i</sub> campaign believes that PRO<sub>i</sub> keeping his sex life under control is necessary for electoral success  
 d. John thinks that PRO getting his résumé in order is crucial and Bill does too  
 e. John<sub>i</sub> told Mary<sub>2</sub> that PRO<sub>1+2</sub> leaving each other was important to Bill  
 f. The unfortunate believes that PRO getting a medal would be boring  
 g. Only Churchill remembers that PRO giving the BST speech was momentous

What makes this observation interesting is that if OC is the reflex of movement, then we would expect OC to be exempt from subjects as these are islands.<sup>85</sup> This seems to be correct.<sup>86</sup>

If movement prohibits relating the PROs in (16) to their antecedents what grammatically underpins them? The only realistic option is some sort of binding operation. Thus, the relation between *John* and *Mary* and *PRO* in (16e) must be some sort of binding or co-reference.<sup>87</sup> And if this is so, then it is likely that *PRO* is (more or less) some sort of pronominal in these configurations. To state things baldly, NOC *PRO* is (roughly) akin to *pro*, a null pronoun. It is interesting to note the sentences in (16b–g) all have paraphrases with overt pronouns in place of *PRO*.

- (101)b. John<sub>i</sub> thinks that it was believed that his/him<sub>i</sub> shaving himself was important
- c. Clinton's<sub>i</sub> campaign believes that his/him<sub>i</sub> keeping his sex life under control is necessary for electoral success
  - d. John thinks that his/him getting his résumé in order is crucial and Bill does too
  - e. John<sub>1</sub> told Mary<sub>2</sub> that their/them<sub>1+2</sub> leaving each other was important to Bill
  - f. The unfortunate believes that his/him getting a medal would be boring
  - g. Only Churchill remembers that his/him giving the BST speech was momentous

Were NOC *PRO* (more or less) equivalent to *pro* then the pattern observed in (16) is what we would expect to find given the facts in (101).<sup>88</sup> It is also what we would expect theoretically. Pronouns are not subject to the locality and prominence conditions characteristic of A-traces. Thus if NOC *PRO*s are akin to pronouns while OC *PRO*s are A-traces we would expect to find the attested differences in their distribution and interpretation.

This said, it is not actually crucial for the movement theory of control that NOC *PRO* be *pro*. What is crucial is that it *not* be the residue of movement. I mention this, for some have argued that what we find in cases of NOC *PRO* are logophors rather than pronouns.<sup>89</sup> Landau (1999, pp. 138–9; (66)–(68)) cites the following data from Kuno (1975).

- (102)a. John said about Mary that it would be easy \*(for her) to prepare herself for the exam
- b. John sued Mary for divorce because it was no longer possible \*(for her) to support him

Kuno analyzes such cases as involving logophors, rather than mere pronouns, to account for the contrast with overt pronouns.<sup>90</sup>

As said, the movement theory is compatible with the idea that some NOC *PRO*s are logophors rather than pronouns. However, whether this is a significant difference awaits further specification of the basic licensing properties of

logophors: Do they require c-commanding antecedents? Do they support split antecedents? Do they allow strict readings under ellipsis? Can they be interpreted as bound variables? I do not know. One element that is commonly assumed to be a logophor, Japanese *zibun*, is interpretable as a bound pronoun but is not able to support split antecedents – (103a) – nor license a strict reading under ellipsis – (103b) only has a sloppy reading.<sup>91</sup>

- (103)a. Takahashi<sub>1</sub>-ga Mariko<sub>2</sub>-ni [Kenji<sub>3</sub>-ga zibun<sub>1/3/\*1+2/\*1+3</sub>-o  
                   -nom           -dat           -nom                    SELF-acc  
 suisenshita-to] tsugeta  
 recommended reported  
 “Takahashi reported to Mariko that Kenji recommended self”
- b. Takahashi<sub>1</sub>-ga zibun-o home, Kenji<sub>2</sub>-mo soo-shita  
                   -nom self   praise           -too so did  
 “Takahashi praised himself and Kenji did too”

If *zibun* is a typical logophor, it cannot be the case that all NOC PROs are logophors given the data presented in (16) above.

Note too that whatever NOC PROs are they can be bound by quantificational antecedents and still display the same range of properties.<sup>92</sup> For example, split antecedents are permitted with NOC PROs. Contrast the examples in (104). (104a) requires OC PRO and prohibits split antecedents while (104b) has an NOC PRO and permits them. Note that the relevant reading is one in which the value of *someone* varies with that of the antecedent.

- (104)a. \*Everyone<sub>1</sub> persuaded someone<sub>2</sub> PRO<sub>1+2</sub> to wash each other  
 b. Everyone<sub>1</sub> persuaded someone<sub>2</sub> that PRO<sub>1+2</sub> washing each other  
 would amuse Mary

What is descriptively important, then, is that whatever logophors are, their properties track those of pronouns with respect to the features noted in (16). If they do not, then *at least some* NOC PROs cannot be logophors.<sup>93</sup> If, however, logophors are essentially pronouns with other added complexities, then the story provided here, the one that treats NOC PRO as essentially *pro*, can be modified without much difficulty.<sup>94</sup>

Consider now the second issue; the trading relation between OC and NOC PRO. Why are OC and NOC PRO in complementary distribution? In particular, why does NOC PRO not occur wherever OC PRO does?

There are several possible approaches to this question. One can simply stipulate that *pro* must be licensed by specific features of T<sup>0</sup> and that these features are absent, for some reason, in OC configurations. However, this tack leaves us in ignorance as to why NOC configurations congregate within islands. Why, for example, isn't there a verb just like *hope* but which displays the NOC properties in (16) rather than the OC features of (14)?

Consider another approach to this complementarity in terms of economy. Hornstein (2000) proposes that UG prefers movement operations to construal

processes. Slightly more precisely, if an expression E can be related to a position P via movement, then it must be associated via movement. Only if movement cannot establish the liason can construal operate to link E to P. If we assume that something like this obtains within UG then the presence of OC removes the possibility of NOC as the former are more “economical” than the latter. Note, if this is correct, then we will expect to find NOCs within islands as these are configurations from which movement is barred.<sup>95</sup> Furthermore, where movement is licit we will expect that NOC is absent, the reason being that as movement (and so OC) suffices for derivational convergence, the more costly non-movement derivation, that brings NOC interpretations in its train, will be barred by economy.

This kind of story brings with it many additional questions. For example, why is construal more costly than movement? Is there any additional evidence that construal or binding operations are “last resort” in the relevant sense? How should one code the trading relation between movement and construal? All these are legitimate concerns, though we do not address them here.<sup>96</sup> However, what is of interest now is that so interpreting OC and NOC makes available a possible answer to a pressing question: why OC and NOC are in complementary distribution.

So, we answer the three questions above as follows. NOC has the properties in (16) because NOC PRO is *pro*, a null pronominal. NOC might involve logophors as well as pronouns without affecting the basics of the movement approach to OC. And, OC and NOC are in complementary distribution because movement is preferred to construal/binding. As such, we will find NOC where movement cannot apply, within islands, and expect to find NOC absent where movement is possible.<sup>97</sup> This cut seems roughly correct.<sup>98</sup>

## 1.8 Extensions

---

The earlier sections have defended a movement approach to control. It is movement-based in two senses. First, OC PRO is identified with A-traces, the residues of overt A-movement. Second, NOC *pro* is parasitic on A-movement in that it is available only if A-movement is prohibited. This section reviews analyses that develop the movement theory and thereby add to its empirical vivacity.

### 1.8.1 *Backwards control*

Polinsky and Potsdam (2000) analyze a case of “backwards” control in Tsez, a language of the Caucasus.<sup>99</sup> Backwards control describes cases in which the controlled PRO (asymmetrically) c-commands its antecedent. This sort of control configuration is unexpected in standard approaches to control as it should lead, among other difficulties, to a principle C violation once the control relation is established. However, as Polinsky and Potsdam observe, backwards

control can be accounted for in a movement based theory in a rather straightforward manner. Let us consider the facts and the argument.

Polinsky and Potsdam argue that backwards control (BC) occurs in clauses like (105).

- (105) PRO<sub>1/\*2</sub> [kidba<sub>1</sub> ziya bisra] yoqsi  
 girl-erg cow-abs feed-inf began  
 "The girl began to feed the cow"

Both *-oqa* (begin) and *-ica* (continue) occur in such syntactic configurations. Observe that the PRO in (105) necessarily takes *kidba* as its antecedent. The evidence for the proposed structure is the following.

First, these verbs are thematic in that they impose selectional restrictions on their subjects. They cannot host idiomatic subjects. The subject of the idiom in (106a) cannot move to the matrix clause in (106b).

- (106)a. T'ont'oha buq bac'xo  
 darkness-erg sun-abs eat-pres  
 (lit.: Darkness ate the sun)  
 "The sun eclipsed"
- b. \*T'ont'oha buq bac'a baq  
 darkness-erg sun-abs eat-inf begin-fut  
 "The sun will begin to eclipse"

Moreover, the subject positions of these BC verbs impose animacy and volitionality restrictions on potential subjects, hence the unacceptability of (107b) and (108b).<sup>100</sup>

- (107)a. Haca nesis xot'o zek'si  
 door-erg his foot-abs hit-past-evid  
 "The door hit his foot"
- b. \*Haca nesis xot'o zek'a yaq  
 door-erg his foot-abs hit-inf begin-fut  
 "The door will begin to hit his foot"
- (108)a. Kidber hazab bukay-n  
 girl-dat suffering see-past-nonevid  
 "The girl suffered"
- b. \*Kidber hazab bukada y-oq-si  
 girl-dat suffering see-inf II-begin-past-evid  
 "The girl began to suffer"

Second, the case marking on the overt subject is always that which is found on subjects in the embedded clause. For example, *bic'zi boqa* (understand.inf) takes a dative subject regardless of whether it is embedded under *-oqa*.<sup>101</sup>

- (109) kidber hisab bic'zi boqa y-oq-si<sup>102</sup>  
 girl-dat math-abs understand-inf II-begin-past-evid  
 "The girl began to understand math"

Third, In Tsez, scrambling is rather free both to the left and the right. However, it is also clause bounded. In particular, scrambling out of an infinitive is not permitted. With this in mind, scrambling can be used as a diagnostic of sentence structure. In *-oqa* constructions, the overt subject cannot scramble with matrix clause elements.

- (110)a. hul [kidba ziya bisra] y-oq-si  
 yesterday girl cow feed began  
 "Yesterday, the girl began to feed the cow"
- b. \*kidba hul [ziya bisra] y-oq-si  
 Girl yesterday cow feed began

Moreover, it is possible to scramble the whole embedded clause and when one does the subject cannot be left behind but must scramble with the rest of the clause.

- (111)a. hul [kidba ziya bisra] y-oq-si  
 yesterday girl cow feed began  
 "Yesterday, the girl began to feed the cow"
- b. hul y-oq-si [kidba ziya bisra]  
 Yesterday began girl cow feed
- c. \*hul kidba y-oq-si [ziya bisra]  
 Yesterday girl began cow feed

This is what we expect if the overt subject *kidba* is part of the complement clause.

Event quantification data further support the proposed BC phrase structure. Consider (112a). It is ambiguous with *uyrax* (four times) modifying the embedded verb (four feedings, (112b)) or the matrix verb (four beginnings, (112c)). (112d), in contrast, only has the reading in which the embedded clause is modified (four feedings). This is what we would expect if the overt subject were in the embedded clause.

- (112)a. Uyrax kidba ziya bisra y-oq-si  
 four times girl cow feed began  
 "The girl began to feed the cow four times"
- b. Uyrax [kidba ziya bisra] y-oq-si
- c. [Uyrax kidba ziya bisra] y-oq-si

- d. [kidba uyrax      ziya bisra] y-oq-si  
 girl    four times cow feed    begin

Polinsky and Potsdam present other evidence all pointing to the same conclusions; namely that in Tsez BC constructions the subject position of the matrix is obligatorily null, thematic, and obligatorily bound by the embedded overt subject. They account for these facts by proposing that BC constructions involve covert movement of the embedded subject to the matrix theta position at LF. Thus, examples like (113a) have LF structures like (113b).

- (113)a. EC [kidba ziya bisra] yoqsi  
 EC girl    cow feed    begin
- b. Kidba [kidba ziya bisra] yoqsi

At LF, *kidba* moves to the matrix theta position of *-oqa* and thereby assumes the matrix theta role.

Polinsky and Potsdam provide independent evidence for the proposed LF A-movement. Tsez reflexives are clause bound (114a). However, a reflexive in the matrix of a BC construction can be bound by a DP in a lower clause (114b). This makes perfect sense if the lower DP raises to the matrix clause at LF and from there binds the reflexive.

- (114)a. Babir<sub>1</sub> nesa nesir<sub>1/2</sub> etin      [uza<sub>2</sub> yutku roda]  
 Father refl                      wanted    boy house make  
 “The father wanted for himself that the boy build the house”
- b. Nesa nesir<sub>1</sub> oqsi    [yesi zek’a<sub>1</sub> yutku roda]  
 refl                      begin this man    house make  
 “Then the man began for himself to build a house”

Polinsky and Potsdam offer further elaborations of the proposal sketched here and they discuss various technical issues related to its implementation. However, their main point is twofold. First, that the standard theories of control that involve PRO and binding cannot easily account for BC constructions. In fact, in most versions of the standard approach, BC should simply be impossible. Second, they show that it is possible to explain the properties of BC phenomena if one adopts a movement approach to control. In fact, Polinsky and Potsdam note that if theta-roles are syntactic features (as they must be in a movement theory) then “the theory leads us to expect that theta-role features could . . . be weak and thus checked covertly. In the case of Control, a weak theta-role would yield a Backward Obligatory Control configuration in which the lower argument is overt and the higher one is unpronounced” (p. 2).<sup>103</sup>

In sum, Polinsky and Potsdam (2000) offers a plausible account of backward control phenomena in terms of a movement approach to control. The properties of these constructions resist explanation in terms of more standard PRO based accounts.

Polinsky and Potsdam's analysis of backwards control in Tsez extends to other cases in Brazilian Portuguese (BP) described in Farrell (1995). Farrell (1995) discusses the "periphrastic causative construction" illustrated in (115).<sup>104</sup>

- (115)a. A mulher fez o nenê dormir  
 The woman made the baby to sleep  
 "The woman put the baby to sleep"
- b. Eu mandei o sapateiro concertar esse sapato  
 I had the cobbler fix these shoes

Farrell (1995) presents evidence that these constructions have the thematic structure of control complements but the surface syntax of ECM constructions. Thus, in overt syntax, in a GB style analysis, (126a) would have the structure (116) and so display backwards control.

- (116) A mulher [fez [PRO<sub>i</sub> [o nenê<sub>i</sub> dormir]]

Farrell (1995) offers several kinds of evidence for the conclusion that *fazer* (make) and *mandar* (have) have empty thematic objects. Let us review some of these.

First, some evidence that these verbs have a *persuade*-like thematic structure. They do not display voice transparency, that is (117a) is not a paraphrase of (117b).

- (117)a. Eu mandei/fiz o médico examinar a minha filha  
 I had/made the doctor to examine to my daughter  
 "I had/made the Dr examine my daughter"
- b. Eu mandei/fiz a minha filha examinanda pelo médico  
 I had/made to my daughter examined by the doctor

Expletives are barred from these constructions.

- (118)a. Pro<sub>expl</sub> é óbvio que eu sou forte  
 it is obvious that I am strong
- b. \*Aquilo faria Pro<sub>expl</sub> ser óbvio que eu sou forte  
 That would make it obvious that I am strong

They also impose selectional restrictions on the subject of the embedded clause. Thus, *mandar* does not allow inanimate DPs.

- (119) \*Ele mandou a minha temperatura aumentar  
 He had to my temperature to-rise  
 "He had my temperature rise"

These facts all follow if the subject of the embedded clause is also the thematic complement of the matrix *fazer/mandar*.

However, there is also evidence that the thematic complement is not a syntactic object of the matrix in overt syntax. For example, in contrast to standard object control verbs, these cannot be passivized. Compare (120a) and (120b).

- (120)a. Os alunos foram forçados a estudarem mais  
The students were forced to study more
- b. \*O nenê foi feito dormir  
The baby was made to sleep

These two constructions also differ with respect to being able to use a subject pronoun for the thematic object. Compare (121a,b).

- (121)a. A professor mandou/fez eu apagar o quadro  
The teacher had/made me erase the board
- b. \*A professor proibiude eu apagar o quadro  
The teacher prohibited me from erasing the board

(121b) prohibits the *eu* form as expected given that this is not allowed in object positions quite generally (122b). However, it is allowed in (121a), suggesting that in overt syntax *eu* is in the embedded subject position, a position that permits *eu* (122a).

- (122)a. Eu falei com ela  
I spoke with her
- b. \*Ela viu eu  
She saw me

The above illustrates the arguments that Farrell (1995) presents to support the conclusion that thematically *fazer/mandar* are *persuade*-like though in their overt syntax they are more like *expect*.<sup>105</sup> Farrell (1995) treats these constructions as instances of backwards control. The movement analysis of control can treat these analogously to the Tsez cases via a process of movement of the embedded subject to the internal theta position of the matrix verb at LF. It is less clear how these data are to be accommodated given a PRO based approach to control.

### 1.8.2 Brazilian Portuguese null subjects

Consider now a second kind of control, one that appears in sentences that look a lot like finite clauses.<sup>106</sup>

Brazilian Portuguese (BP) is in the process of losing pro-drop. It is still possible to find *pro* headed matrix clauses but only if the subject is non-referential. Thus we observe the contrast in (123).

- (123)a. Esta chovendo  
is raining  
“it is raining”
- b. Mataram o presidente  
killed the president  
“Someone killed the president”
- c. \*Comprou um carro novo  
bought-3sg a car new  
“He bought a new car”

As shown, weather and impersonal constructions permit pro-drop. However, referential use of *pro* is no longer available in matrix clauses.

Apparent cases of referential null subjects, however, are found in embedded clauses.

- (124) O Joao<sub>i</sub> disse [que ec<sub>i</sub> comprou um carro novo]  
John<sub>i</sub> said that he<sub>i</sub> bought a car new

How is the asymmetry between main and embedded clauses to be accounted for? Ferreira (2000) and Rodrigues (2000) propose that what we find in (124) is not *pro* but control. In particular, they assume, that in cases like (124) the embedded clause need not case mark the subject; this optional case marking being related to the fact that the verbal paradigm is undergoing simplification, which in turn is related to the loss of referential null subjects. In support of this proposal they observe that the empty category in embedded subject position in (124) displays the diagnostic properties of OC. Thus, this empty category requires a local c-commanding antecedent (125a–c), forbids split antecedents (125d), requires a sloppy reading under ellipsis (125e), has to have the bound reading with *only DPs* (125f), and requires *de se* readings (125g).

- (125)a. \*pro<sub>expl</sub> parece que ec tinha telefonado  
“It seems that he had telephoned”
- b. \*o Joao<sub>i</sub> disse [que a Maria acha [que ec<sub>i</sub> e esperto]]  
“John said that Mary thinks that he is smart”
- c. \*A Mae do Joao<sub>i</sub> acha [que ec<sub>i</sub> e esperto]  
“John’s mother thinks that he is smart”
- d. \*O Joao<sub>i</sub> disse [que Maria<sub>j</sub> acha [que ec<sub>i+j</sub> sao espertos]]  
“John said that Maria thinks that they are smart”
- e. O Joao acha que ec vai ganhar a corrida e a Maria tambem  
“John thinks that he will win the race and Mary thinks that she will too”

- f. So o João acha que *ec* vai ganhar a corrida  
 “Only John thinks that he himself will win the race”
- g. O Reagan esta convencido de que *ec* foi um dos nelhores presidentes dos EUA  
 “Only Reagan is convinced that he himself was one of the best presidents of the USA”

Further evidence that the empty categories in (125) are not null pronouns comes from the fact that they cannot act as resumptive pronouns, in contrast to *pro* in other Romance languages. Consider the contrast between (126a) and (126b).

- (126)a. Ese el tipo que<sub>1</sub> Maria conoce a la mujer [con quien *ec*<sub>1</sub> se casó]  
 “That is the guy that Maria knows the woman he married”
- b. Esse é o rapaz que<sub>1</sub> a Maria conhece a garota que *ele*<sub>1</sub>/*\*ec*<sub>1</sub> gosta  
 “This is the guy that Maria knows the girl that he likes”

The *ec* in (126a) is a null pronoun that can act as a resumptive expression. Overt pronouns in BP can play a similar role as (126b) indicates. However, there is no null pronoun in BP as there is in Spanish.

In sum, the empty expression in these constructions looks very much like an OC PRO.

Both Ferreira (2000) and Rodrigues (2000) propose that these empty categories are residues of overt movement. The possibility of such movement is related to the idea that the agreement system in finite clauses is undergoing radical simplification and this is related to the progressive loss of pro-drop in BP. As Ferreira and Rodrigues note, if we assume that in BP the embedded clauses optionally bear case features, we can permit movement from the embedded clause just as we do in standard control complements. We can further relate this quirk in BP to Chomsky’s (1999) idea that only a full complement of phi-features must assign case.<sup>107</sup>

There are two more interesting set of facts. First, it appears that BP allows raising from finite clauses.

- (127)a. O João parece que *ec* comprou um carro novo  
 John seems that he bought a car new
- b. O João disse que *ec* comprou um carro novo  
 John said that he bought a car new

As Ferrira (2000) observes, the acceptability of cases like (127a) is not surprising given the proposal that the *ec* in (127b) is due to movement.

Lastly, it appears that one sees similar phenomena inside adjuncts. Most interestingly, the same requirement found in English adjunct control structures (viz. that the subject be the antecedent) extends to these cases.

- (128)a. O João<sub>1</sub> foi embora depois que ec<sub>1</sub> brigou com a Maria  
 John<sub>1</sub> left after he quarreled with Mary
- b. \*O João conheceu a Maria<sub>1</sub> depois que ec<sub>1</sub> ficou rica  
 John knew Mary after that she became rich-fem

Rodrigues (2000) notes that these adjunct cases display the diagnostics characteristic of OC. (129a) indicates that the *ec* requires an antecedent. (129b) shows that the antecedent must be local and (129c) that it c-command the *ec*. (129d) indicates that split antecedents are prohibited, (129e) that only sloppy readings are allowed, and (129f) that bound readings are required with *so* (= only).

- (129)a. \*ec<sub>expl</sub> chove toda vez que ec fala com o Paulo  
 It rains every time that she speaks with Paulo
- b. A Ana<sub>1</sub> disse que a Maria<sub>2</sub> olha para o chão toda  
 Ana said that Maria looks at the ground every  
 vez que ec<sub>1/2</sub> fala com Paulo  
 time that she speaks with Paulo
- c. [A Mãe da Maria<sub>1</sub>]<sub>2</sub> olha para o chão toda vez  
 The mother of Mary looks at the ground every time  
 que ec<sub>1/2</sub> fala com Paulo  
 that she speaks with Paulo
- d. \*O Luca<sub>1</sub> disse que a Ana<sub>2</sub> chorou pra caramba depois que ec<sub>1+2</sub>  
 Luca said that Ana cried a lot after that they  
 deixaram o Brasil  
 left Brazil
- e. A Ana<sub>1</sub> voltou o Rio depois que ec<sub>1</sub> ficou grávida e  
 Ana returned to Rio after that she got pregnant and  
 o Luca também  
 Luca too  
 (= after Luca got pregnant, NOT after Ana did)
- f. Só o Maluf<sub>1</sub> ficou triste depois que ec<sub>1</sub> perdeu as eleições  
 (= Only M is an x such that x got upset after x lost the elections)

In sum, Ferreira (2000) and Rodrigues (2000) show that certain empty categories in finite complements and adjuncts behave just like OC PRO. They show how to extend the analysis of control as movement in non-finite clauses to control in these “finite” configurations. These facts are compatible with a non-movement approach to control. However, it is interesting that these BP clauses tolerate both control and super-raising, suggesting that both raising and control are reflexes of the same kind of operation, as a movement

approach to control would lead us to expect. Second, instances of both complement control and adjunct control are possible from these “finite” configurations supporting the idea that both are generated by the same grammatical operations, viz. movement on the present proposal.

### 1.8.3 PRO gate effects

Consider the following contrast noted in Higginbotham (1980).<sup>108</sup>

- (130)a. \*Mary’s/his<sub>1</sub>/him<sub>1</sub> kissing his<sub>1</sub> mother made everyone<sub>1</sub> late
- b. \*Who did Mary’s/his<sub>1</sub>/him<sub>1</sub> kissing his<sub>1</sub> mother upset t<sub>1</sub>
- c. PRO<sub>1</sub> kissing his<sub>1</sub> mother made everyone<sub>1</sub> late
- d. Who did PRO<sub>1</sub> kissing his<sub>1</sub> mother upset t<sub>1</sub>

(130a,b) exemplify weak cross over (WCO) effects which Higginbotham analyzed as barring the binding of a pronoun by an antecedent to the right of the pronoun. In (130a) the antecedent is *everyone* (or its LF trace); in (130b) it is the WH-t left by movement of *Who*. What is interesting is that (130c,d) seem to allow the binding prohibited in (130a,b). The PRO in the gerund acts as a “gate” permitting the indicated binding. Higginbotham dubbed this amelioration effect with respect to WCO PRO gate effects. A natural question that arises is why PROs act as gates. Kiguchi (2000) provides an explanation in terms of the movement theory of control. The account proceeds roughly as follows.

Kiguchi (2000) observes that A-movement circumvents the WCO condition. (131) illustrates this.

- (131)a. \*Who<sub>1</sub> did it seem to his<sub>1</sub> mother t<sub>1</sub> liked Bill
- b. Who<sub>1</sub> t<sub>1</sub> seemed to his<sub>1</sub> mother t<sub>1</sub> to like Bill

A’-moving *who* in (131a) does not allow pronominal binding as it violates WCO. However, raising *who* to the matrix subject and then A’-moving to Spec CP does not similarly induce a WCO violation, as the acceptability of (131b) shows. Kiguchi’s (2000) idea is to generalize this to cases like (130c,d). In particular, if OC PRO is simply an A-trace, then OC PRO should obviate WCO effects. If the PROs in (130c,d) are OC PROs then we can account for their gate-like status.

There is evidence that these PROs are indeed OC PROs. Thus, such PROs require antecedents.

- (132)a. \*PRO shaving himself impressed Mary
- b. \*PRO shaving himself made it seem cold outside

Second, they resist split antecedents.

- (133) \*John<sub>1</sub> said that PRO<sub>1+2</sub> shaving themselves upset everyone<sub>2</sub>

Third, they require *de se* readings and only permit bound readings in *only* sentences. Contrast the sentences in (134) with those in (135) where we have an overt pronoun. (135a) is ambiguous with both a *de se* and non-*de se* reading. (134a) only has the *de se* interpretation. (135b) only has a strict reading with the pronoun coreferential with Churchill. This is the reading missing in (134b) where the PRO is interpreted as bound by *only Churchill*.

- (134)a. PRO receiving the medal unnerved the unfortunate  
 b. PRO giving the speech upset only Churchill.

- (135)a. His/him receiving the medal unnerved the unfortunate  
 b. His/him giving the speech upset only Churchill.

Last of all, there are locality and prominence conditions on the antecedent. Thus, the antecedent of PRO must be the most prominent of a series of DPs and cannot be buried inside another DP.

- (136) \*PRO<sub>1</sub> shaving himself made Mary believe John<sub>1</sub>

- (137) \*PRO<sub>1</sub> shaving himself upset John's<sub>1</sub> mother

Kiguchi (2000) shows that these data all follow if we assume that the PRO in PRO gate constructions are all the residue of sideways A-movement, the same movement operative in adjunct control.

In sum, the relevant PROs display the diagnostics of obligatory control, which is what we would expect if they were the residues of A-movement, in this case, sideways A-movement. As such, we can account for their gate-like properties along the lines sketched above.

This approach to PRO gate effects makes a second interesting prediction. Those PROs that cannot be the residue of movement should not act as gates and should not ameliorate WCO violations. This too seems to be correct. Consider sentences like (138).

- (138)a. \*Who<sub>1</sub> did the fact that PRO<sub>1</sub> cooking his<sub>1</sub> lunch took all afternoon annoy t<sub>1</sub>  
 b. \*The fact that PRO<sub>1</sub> cooking his<sub>1</sub> mother lunch took 30 minutes kept no one<sub>1</sub> in the kitchen

The PROs here are OC PROs as they occur within complex NP islands. As such, they cannot be the residues of movement but must be null pronominals like *pro*. As such, they should not aid in ameliorating WCO effects.

Moreover, we expect these PROs to function like pronouns in other ways. As (139) indicates, this expectation is born out.

- (139)a. The fact that PRO cooking for oneself takes time annoys John  
 b. The fact that PRO<sub>1</sub> cooking herself lunch took 30 minutes made John angry at Mary<sub>1</sub>

- c. The fact that PRO<sub>1</sub> cooking himself lunch took 30 minutes made John's mother angry
- d. The fact that PRO<sub>1+2</sub> cooking themselves lunch took 30 minutes made John<sub>1</sub> angry at Mary<sub>2</sub>
- e. The fact that PRO<sub>1</sub> receiving the medal took 30 minutes annoyed the unfortunate<sub>1</sub>
- f. The fact that PRO<sub>1</sub> giving the speech took 30 minutes annoyed only [Churchill]<sub>1</sub>

These PROs do not require proximate prominent antecedents (139a–c). They allow split antecedents (139d), non-*de se* readings (139e) and strict readings with *only* (139f). These are the hallmarks of a pronoun. It is thus not surprising that like other pronouns they are subject to WCO conditions and do not function like OC PROs in ameliorating WCO effects.

In sum, Kiguchi (2000) provides an independent argument that OC PROs are residues of A-movement and that NOC PROs are pronoun like. The former act like gates with respect to WCO just like A-traces do in more familiar raising constructions. The latter, being pronominal, do not. This account of PRO gate effects constitutes independent evidence for the supposition that OC PROs are formed by movement.

## 1.9 Conclusion

---

Recently there has been a resurgence of interest in control phenomena prompted by the theoretical turn to minimalism. Like all actions in some theoretical direction, this one has prompted a reaction against movement based approaches. This review has had three aims: to re-advertise the virtues of a movement approach, to parry some of the principal arguments against its viability and to present some more recent evidence that supports it. It is for the reader to judge how successful this effort has been. However, before closing, let me reiterate two points made at the outset. First, there are many versions of the movement approach. I am most familiar with one of these and so I have illustrated the virtues of the movement theory using it. However, much of what was advanced here in one technical garb could have been equally well promoted exploiting other details of implementation. Second, it is worth bearing in mind that these movement based analyses are of relatively recent vintage. We should thus expect empirical problems to beset them. What I find interesting is not that they indeed have empirical weaknesses, but that there are not more of them.<sup>109</sup> Movement theories can only get off the ground by rejecting long-held core assumptions about the principles that regulate grammatical structure and derivations. It must reject the theta criterion and the idea that thematic roles can only be saturated by merge/lexical insertion (i.e. cannot be saturated by movement). It is surprising, I believe, that setting these assumptions aside does not result in theoretical chaos. If nothing else, putting these assumptions to one side has allowed us to explore our earlier theoretical

assumptions more closely. It has also opened up, or, more accurately perhaps, revived, novel ways of thinking about control phenomena. Even if these should prove to be incorrect, I hope that they will serve as useful foils for future more theoretically principled and empirically richer proposals.

## Notes

- 1 Various people have commented on earlier drafts of this chapter. I would like to thank David Lightfoot, Juan Uriagereka, Acrisio Pires, Klea Grohmann, and Martin Hackl for their kind indulgence. I would also like to thank Randy Hendrick for detailed comments on an earlier draft. This work was supported by NSF grant SBR-9817569.
- 2 Bowers (1973) is the earliest proposal that I know of in the generative literature for treating raising and control as essentially the same. Others have made similar proposals since, including Bresnan (1982).
- 3 This concept goes back to the earliest models of generative grammar. The thematic properties of d-structure are roughly identical to those enjoyed by kernel sentences in a *Syntactic Structures* style theory. Kernel sentences were input to transformational processes and were the locus of (what we now call) theta-roles. *Aspects* substitutes the base for kernel sentences. The base is a pre-transformational phrase marker generated by phrase structure rules. Like kernel sentences, it is the input to the transformational component and the locus of thematic information. This role for the base has been retained in some form in all subsequent theories.
- 4 Having both d-structure and the theta criterion as parts of UG is redundant and so, undesirable. Note that there is little to methodologically recommend the idea that DPs (or chains) can bear but single theta-roles. There is nothing odd about the idea that a single expression should saturate several variable positions and hence bear several theta-roles. Thus structures like (i) are perfectly coherent and plausibly represent a situation on which one expression, *John* bears two theta-roles, viz. the washer and washee.

(i) John  $\lambda x$  ( x wash x)

If so, on its own, the biuniqueness assumption stipulated in the theta criterion is methodologically as interesting as all stipulations are. This does not imply, however, that the stipulation is empirically without merit, only that it has no conceptually independent motivation.

- 5 Move is resolved into two more basic operations, Copy and Merge. In place of traces, therefore, we find copies. There are other ways of implementing the idea that traces be replaced by copies (e.g. via a re-merge operation) but which technical implementation is adopted is of no relevance here.
- 6 I do not mean to say that movement *is* independently motivated. Only that a theory with both kinds of processes is to be avoided *ceteris paribus*. I assume that the preferred reduction is construal to movement but the methodological point remains even if the other tack is taken.
- 7 One needs to further assume that in examples like (10e) the +WH  $C^0$  does not govern across IP.
- 8 See Bouchard (1984) where there is an attempted account of (11) in terms of case theory and Chomsky (1981) where the distribution of PRO is accounted for in binding theoretic terms on the assumption that PRO is at once +pronominal and

+anaphoric. For a brief review of these issues against a minimalist backdrop see Hornstein (1997).

- 9 See Rosenbaum (1970).
- 10 See Williams (1980), Koster (1984), Lebeaux (1984–5) and Manzini (1983).
- 11 This paradigm derives from Williams (1980), Lebeaux (1984–5), Higginbotham (1992) and Fodor (1975).
- 12 At least in the standard cases. We return to this below in the discussion of sideways movement.
- 13 The split antecedent diagnostic has been challenged by Landau (1999, p. 136). He claims that OC PRO can support split antecedents “in most environments.” I believe that this is simply incorrect. There is a very clear contrast between cases like (14e) and (16e). The latter easily permits the binding of a reciprocal while the former strongly resists this.

The counterexample offered by Landau are sentences like (i).

- (i) John<sub>1</sub> persuaded/suggested to Mary<sub>2</sub> [PRO<sub>1+2</sub> to get themselves a new car]

However, these sorts of examples show very little. First, many native speakers of English find examples like (i) simply unacceptable. Second, even for those who accept these marginally, they reject the sentences when the anaphor is replaced with a reciprocal.

- (ii)a. \*John persuaded Mary to get each other a car  
 b. \*Did you suggest to Mary to get each other a new car?

Third, for some of these same speakers, examples like (iii) where the reflexive is replaced by a pronoun are also marginally acceptable. Thus, there may not be a complementary distribution between anaphors and pronouns in this context.

- (iii) ?\*John<sub>1</sub> persuaded Mary<sub>2</sub> [PRO<sub>1+2</sub> to get them<sub>1+2</sub> a new car]

Fourth, myriad other examples are uniformly rejected.

- (iv)a. \*John persuaded Mary PRO to wash/shave themselves/each other  
 b. \*John persuaded Mary PRO to drive each other/themselves around

Fifth, for those who accept examples like (i), there is no clear contrast between them and (v), these latter not being control contexts at all.

- (v) John<sub>1</sub> expected Mary<sub>2</sub> to get themselves<sub>1+2</sub> a new car

It is worth noting that none of this affects the NOC examples in (16) where the analogous examples are all perfectly fine. For example, contrast (iv) and (vi).

- (vi) John persuaded Mary that [PRO washing/shaving each other/themselves] would amuse Sam

Cases like (vi) are uniformly acceptable.

I conclude that the claim that OC PRO resists split antecedents is well grounded and that OC and NOC PRO contrast significantly along this dimension.

- 14 Minimalist reasoning also casts suspicion on government as a primitive grammatical relation. To the degree that government cannot be exploited as a descriptive predicate, the standard GB analysis fails. See Chomsky (1993, 1995) for discussion.
- 15 This, in effect, returns us to a version of Bouchard’s proposal that reduced the distribution of PRO to case theory. The details of the Chomsky and Lasnik proposal, however, are rather different.

- 16 See also Landau (1999) and Wurmbrand (1998) for further refinements.  
 17 For some criticism of Stowell's assumption that raising clauses have different tense properties from control clauses see Brugger (1997). Martin (1996) also notes that some raising clauses have independent tense specifications.  
 18 These examples are from Martin (1996, p. 59, n. 82). Event denoting predicates are meant to contrast with stative predicates like (i).

(i) John believes Mary to be tall

I do not fully agree with Martin's specification of the facts presented here. See the following note for a brief discussion.

- 19 I assume that case is checked in the outer spec of *v* rather than AgrO, though nothing hangs on this. I also assume for expository purposes that the raising is overt and that further V-movement provides the requisite linear order.  
 20 ECM only appears to be acceptable with subjects that are WH-traces, not full DPs; why is unclear.  
 21 Observe that this does not argue against treating some T<sup>0</sup>s as tensed and others as untensed, as suggested in Stowell (1982). Rather, it argues against correlating this property of T<sup>0</sup> with any null case assignment properties. For a sophisticated consideration of related issues and possible answers to the objections raised here see Martin (1996, ch. 2).  
 22 This section is based on the discussion in Pires (2001). See this paper for a detailed discussion of the points briefly made here. For related observations based on Japanese control configurations, see Aoshima (2001).  
 23 Virtually all current theories of control would require a PRO in such configurations, including those like Wurmbrand (1998) that treat some control structures as PRO-less. The reason is that these gerunds alternate with those that can have overt accusative subjects.

(i) John remembered Bill bringing the wine

- 24 This section focuses on OC PRO. NOC PRO is discussed in section 7. The details of this proposal exploit the ideas outlined in Hornstein (1999, 2000). However, there are other proposals that adhere to the logic outlined here though they differ in technical implementation. See, for example, Manzini and Roussou (2000) and O'Neill (1995). Both these proposals involve the assumption that the same grammatical operations underlie control and raising, though they differ as to how this difference is to be characterized.  
 25 Note that *who* is deletable in (26b) so that the problem is not with some case that the WH element must bear.  
 26 These data are discussed in Boeckx (2000) and Hornstein (2000).  
 27 Before proceeding, however, one point is worth making: these problems for null case do *not* imply that this approach to the distribution of PRO is wrong. Perhaps it is empirically superior to the movement based account outlined below. However, it is clear that the null case theory has *methodological* problems which should prompt a minimalist to hope to do better.  
 28 I assume here the copy theory of movement. A deleted copy is represented in brackets, e.g. (John). I also assume that the predicate internal subject hypothesis is correct. Thus, all subjects are theta marked within the immediate projection of a thematic head. For present purposes, I adopt the idea that the EPP holds for all clauses, including non-finite clauses.  
 29 Why this is so is an interesting question. See Nunes (1995) for a possible answer.

- 30 It is not clear what makes a feature a feature. What is relevant here is the assumption that theta-roles are feature-like in the following way: checking one suffices to allow greedy movement.
- 31 His discussion is based on earlier work by González (1988, 1990).
- 32 These cases parallel certain examples in Tsez, a caucasian language, discussed below in section 1.8. These data also contrast in an interesting manner with what occurs in Icelandic control structures. Here raising and control structures contrast in that one cannot get DPs that bear the quirky case of an embedded predicate surfacing as the subject of a control predicate though this is common in raising constructions (Thráinsson, 1986). It remains unclear why it is that Icelandic differs in this regard from Tsez and Spanish. However, Icelandic control clauses are quite problematic for most current minimalist accounts. For example, it appears that control clauses bear regular case, not simply null case (Sigursson, 1991). Moreover, control clauses contrast with raising clauses in requiring V to I(nflection) movement for some reason (see Sigursson, 1989; Hornstein, 1990). It is unclear how (or whether) these differences relate to the facts noted above. However, it is clear that the Icelandic facts and the Tsez, Spanish facts are pulling in opposite directions.
- 33 Boskovic provides numerous other examples that lead to this same conclusion, viz. that movement between thematic positions is possible. He (1994, p. 273) notes a rather interesting case from Italian first discussed in Burzio (1986). It concerns sentences like (i).
- (i) (expletive) *ne vorrebbero arrivare molti all festa prima di Mario*  
of them would want to arrive many to the party before Mario  
“Many of them would want to arrive at the party before Mario”

As noted, *molti* is clearly located in the embedded clause in overt syntax. However, it is also interpreted as the matrix clause subject (in conjunction with the cliticized *ne*). Burzio notes that these examples are a problem for theta theory as the matrix subject theta-role is unsaturated in both DS and SS. Note that these structures pose no serious problem for a movement theory if we assume that movement through theta positions is licit at LF. For other cases that can be analyzed similarly, cases of “backwards control,” see the discussion in section 1.8 below.

- 34 Saito (1992) provides a lot of evidence establishing this fact.
- 35 Boskovic and Takahashi review other evidence pointing to this conclusion. The interested reader is referred to their paper and the references cited there.
- 36 Boskovic and Takahashi (1998) rely on the assumption that long scrambling must be base generation as movement is greedy while Merge need not be. This relies on an assumption in a vintage form of minimalism in which Move differed from Merge in being subject to Greed (see Chomsky, 1995). However, recent versions of minimalism have assumed that Merge too is subject to Greed (see Chomsky, 1998). If so, the theory internal motivation for having the long scrambled expression merged into its surface position is weakened, though the empirical reasons offered still stand.
- 37 Which, of course, does not mean that it is false. Boskovic (1994) offers a way of doing this. However, what is clear is that adopting this sort of view requires a methodological compromise that, *ceteris paribus*, is best avoided.
- 38 However, recall that this is only the *typical* case. If inherent reflexives like *wash*, *dress*, *comb*, *shave*, etc. also involve movement, then this constitutes a further problem for a case theoretic approach to control that ties null case to the properties of certain T<sup>0</sup>s.

- 39 This is the standard view within generative grammar starting with Rosenbaum (1970). It assumes that the core properties of control can be largely traced to the structural properties of control configurations.
- 40 As observed in section 3, the distinction between OC and NOC PRO goes back to the earliest days of generative grammar where the rule of Equi NP Deletion was contrasted with Super Equi. It survived in various forms into the EST era. For discussion, see Chomsky and Lasnik (1977), Fodor (1975), Koster (1984), Lebeaux (1984–5), Manzini (1983), and Williams (1980). Though not every author endorsed exactly the same distinction, the opposition between the two cases of control has been widely recognized.
- For present purposes, I do not distinguish between types of pronouns: logophoric versus bound versus referential, etc. There may well be further differences that this crude distinction glosses over. For the nonce, we abstract away from these differences.
- 41 This derivation assumes that the EPP holds in non-finite clauses. This assumption has been challenged in Castillo, Drury, and Grohmann (1999) Epstein and Seely (2000), and Hornstein (2000). There are good theory internal reasons to suppose that a movement approach to OC fits poorly with the idea that the EPP holds in non-finite clauses. However, for present purposes, I abstract away from these issues. See the work noted above for discussion.
- 42 (46) has an underlying small clause structure. I ignore these details here.
- 43 Furthermore, the strict reading is not derivable if we adopt the approach to NOC PRO outlined in section 1.7 below. In particular, to get a strict reading requires having a pronoun-like expression in cases like (14d). So the LF must be like (i) for ellipsis to be licensed (see Merchant, 1999 for discussion of the semantic equivalence conditions required for ellipsis).

- (i) John<sub>1</sub> wants pro<sub>1</sub> to win and Bill<sub>2</sub> wants pro<sub>1</sub> to win

However, we assume below that *pro* can be used in control structures only if movement is barred. This is not the case in these configurations. As such, the “pronominal” PRO required here to license the strict reading is prohibited. As such, only the sloppy reading is available. The semantic form of the sentence is (ii).

- (ii) John  $\lambda x\{x \text{ wants } x \text{ to win}\}$  and Bill  $\lambda x\{x \text{ wants } x \text{ to win}\}$

This yields the sloppy reading. Ellipsis is possible because the two VPs mutually entail one another.

- 44 See Lasnik and Uriagereka (1988, ch. 5) for a review of Higginbotham’s original discussion of this. See Hornstein (2000, ch. 5) and Lidz and Idsardi (1997) for proposals that local anaphors are also the residues of overt A-movement.
- 45 See Salmon (1986) for discussion. He points out that there is an important semantic difference between an expression saturating two argument positions and two distinct expressions that are in a binding relation but with each saturating its own argument position. The former has a *de se* reading that the latter lacks. (ia) illustrates a case in which a pronoun bound by a quantificational antecedent allows a non *de se* reading. This contrasts with (ib) where only a *de se* reading is possible.

- (i)a. No unfortunate<sub>1</sub> expected that he<sub>1</sub> would receive a medal  
 b. No unfortunate expected PRO to receive a medal

The contrast shows that binding is insufficient to force a *de se* reading. See Hornstein (2000, ch. 2) for further discussion of this point.

- 46 Martin (1996) offers a hybrid theory in which PRO is base generated but there is a kind of clitic raising that results in a chain being formed between the PRO and its antecedent at LF. Roussou and Manzini (1997) develop a theory based on feature attraction rather than movement but they too end up with a chain at LF. O'Neill (1995) and Hornstein (1999) implement the movement theory using a standard overt A-movement analysis. They too end up with a single chain at LF mediating the controller and controllee.
- 47 We return to some exceptions to this generalization in a moment.
- 48 This solution is based on earlier work by Nunes (1995) on parasitic gaps where sideways movement operations are advocated.
- 49 See Nunes (1995) where parasitic gaps are analyzed in this way. The idea is also explored by Bobaljik and Brown (1997) and Uriagereka (1998) for head movement.
- 50 Several other assumptions are required for the following proposal to be viable. As the details have been discussed in Hornstein (2000), I do not review them here. However, one important additional assumption is (i)
- (i) The Extension Condition holds for all grammatical operations including the merger of adjuncts
- (i) is needed to account for CED effects. Note that the present proposal allows apparent movement from an adjunct. However, this is mere appearance. At the time that movement occurs, what will *become* the adjunct has not yet adjoined. Once adjoined it triggers CED effects. Assumption (i) is required to derive this result. (i) treats adjunction like any other kind of merge operation in requiring it to obey extension. Thus, adjuncts cannot be counter cyclically merged. This differs from Chomsky (1993) where adjunction is treated as an exception to the Extension Condition. For discussion, see Hornstein (2000).
- 51 The derivation abstracts away from whether accusative case is checked by overt movement.
- 52 What drives this conclusion then is an economy assumption. In Hornstein (2000) I assumed that the economy assumption was similar to the one underlying the preference for Merge over Move as in Chomsky (1995). However, this may be incorrect. There is evidence that sideways movement may be inherently more "expensive" than movement within a single sub-tree. In other words, structural descriptions involving two sub-trees are less preferable than those involving but one. If so, sideways movement will be delayed as long as possible. This too will derive the fact that adjuncts are controlled by subjects, not objects (at least in transitive constructions).
- 53 This sentence is acceptable with the indicated reading if the adjunct is interpreted as modifying the matrix verb. This is, however, an irrelevant reading for the point being made. Note that if one preposes the adjunct, the preferred reading is modification of the matrix verb.
- 54 These cases were pointed out to me by John Nissenbaum (private communication).
- 55 See Culicover and Jackendoff (2001) and Landau (1999).
- 56 I suspect that this is not how Landau understands the notion "marked." I suspect that for him it means something like "OK, but not perfectly acceptable." This is a non-technical use of the term which has no theoretical standing so far as I can tell. Marked constructions can be perfectly acceptable (e.g. Verb raising in Romance) and unmarked constructions rather unacceptable (e.g. center-embedded sentences). There exists no theoretical relationship between being marginally acceptable and grammatically marked, so far as I know.

- 57 Gary Milsark (p.c.) believes that the data in C. Chomsky (1969) contains an even more powerful argument for the markedness of *promise* than the one noted here. Chomsky (1969) claimed late acquisition for the *promise* structures. However, Milsark observes that her data make a different argument: the acquisition data are all over the place with some 9-year-olds never getting it right and some 6- and 7-year-olds doing just fine. It appears that some speakers never quite get the hang of sentences like “Fred promised Bill to leave.” If this is correct, as Milsark notes, it provides an even more direct argument for the markedness of *promise*. It appears that it really is very difficult to finally converge on subject control readings in sentences like this, so difficult that some never do manage. This is precisely what the MDP (and the movement based account that subsumes it) would lead one to expect.
- 58 Let me beat this horse dead with an example. Farkas (1988) proposes that control is sensitive to the RESP(ONSIBILITY) relation in that controllers are intentional initiators of situations that they bring into effect. This is intended to cover both *persuade* and *promise* verbs. Say that this is so. We now have a problem. Why is it that the two verbs are learned at different rates? Why is *promise* more marked? It must be because it and *persuade* are not the same with respect to the RESP relation in some way. Indeed Farkas’ proposal is intended to be understood in terms of markedness. But then we are back where we started. How do we code this? In the verb type. But whether the verb is marked with respect to RESP or minimality seems of little moment logically.
- 59 In fact, Lakoff (1970) explores various ways of lexically marking exceptions to grammatical processes.
- 60 Note that (i) *without* the preposition displays object control.

(i) John committed Bill PRO to leave early

This suggests that what allows subject control in the *commit to* construction is the presence of the preposition.

- 61 Kitihara (1996) has a proposal to this effect. There are various ways to implement this. One might argue that *to* incorporates into *promise* or that it is null to begin with. What is important is that where overt movement applies a preposition is present so minimality is irrelevant and at LF the presence of the preposition can be ignored. This can be implemented in a derivational theory in various ways.
- 62 I owe this point to discussion with Jacek Witkós.
- 63 For a good review of the various approaches to this effect see Landau (1999, ch. 5.3).
- 64 Some might consider the subtlety of this idiolectally highly variable data sufficient grounds for placing such cases outside the domain of core grammar. Those inclined to such a conclusion can skip the rest of this section.
- 65 Chomsky (1980) notes that changing *be allowed to* to *get permission for* disrupts the possibility of this control shift. Compare (72) with (i).

- (i)a. John<sub>1</sub> asked Mary<sub>2</sub> PRO<sub>7<sup>1</sup>/2</sub> to get permission to leave  
 b. John<sub>1</sub> promised Mary<sub>2</sub> PRO<sub>1/7<sup>2</sup></sub> to get permission to leave

The cases in (i) are very infelicitous with the shifted readings, if acceptable at all.

- 66 I owe the following to discussion with Juan Uriagereka.
- 67 What of the second kind of partial control reading, the one involving *promise* in (72b). I sadly have little of interest to say about it. It appears, however, to be even more sensitive to minor perturbations than the *ask* cases are. See Landau (1999, pp. 212 ff.) for comments. He notes that Comrie (1984) rejects cases like (ib) and

Chomsky (1980) finds that substituting *get permission* for *be allowed to* in (ia) leads to only the non-shifted reading. It contrasts with the *ask* cases in several interesting ways, however.

First, it does not invert control so much as expand it. Thus, the examples in (i) allow both subject and object control.

- (i)a. ?John promised Mary PRO to be allowed to go to the movies
- b. ?John promised Mary PRO to be healthy at game time

This contrasts with the *ask* cases where the object control disappeared under control shift.

In other respects, however, it patterns like *ask/beg/petition*. The PRO does not require an antecedent.

- (ii) It was promised to be allowed to wash before dinner

It allows split antecedents.

- (iii) John promised Mary PRO to be allowed to wash themselves before dinner

It seems to permit a strict reading under ellipsis.

- (iv) John promised Frank to be allowed to rest before the talk and Mary too (Frank promised Mary that Frank would be allowed to rest)

These data support the idea that *promise* also involves NOC under control shift. However, why it contrasts with *ask* remains unclear.

- 68 His diagnostic properties are largely similar to those outlined above but there are some differences regarding the status of split antecedents. For current purposes I assume that PC and EC differ only on this single dimension made evident in the capacity of the PRO to license collective predicates like *meet*.
- 69 Actually, if this is an AGREE operation, then it is likely that the Probe is T-Agr and it AGREES with the goal PRO prior to movement of PRO to Spec TP. This would be the standard AGREE configuration.
- 70 Landau (1999, p. 18) observes that he must revise the PIC to allow PRO to be visible but the embedded T-Agr to be outside the purview of the matrix probe. There are no apparent ill effects of doing this, he notes, but it is unclear why it should be that the Specifier of IP should be “visible” while the head of IP should not be, given that the label of the IP projection is identical to the head on bare phrase structure assumptions.
- 71 There is one more curious fact. The proposal must be assuming that AGREE is not a checking operation but an assignment operation. Consider why. Say that all AGREE did was see if features matched up to non-distinctness. Then in (76a) the DP could match F on all phi-features while PRO could match F only on person and gender features being unspecified for number. This would then allow PRO to be unspecified for number in (76a) and so, it would seem, permit a PC reading. So the assumption must be that this is not possible, that in such a case AGREE implies identity not non-distinctness. Now consider (76b). If AGREE were identity then the T-Agr must have the same phi-features as F and PRO must as well. But then F, PRO and DP must all have the same number specification. But then PC should not hold. So, AGREE cannot be feature checking.

That means that the current technology requires that AGREE involve assigning features, not merely checking them. This, you will recall, appears at odds with inclusiveness if we interpret this to mean that the semantic properties of an

expression cannot be changed derivationally. In sum, AGREE cannot be feature checking and must involve feature assignment.

- 72 Recall that control is a much tighter relation than co-reference. The relation between a controller and a PRO in OC structures is quite different from the relation that holds between an antecedent to a bound or co-referential pronoun. Noting this is sufficient to establish that two expressions sharing identical features is not sufficient to establish the relation that underlies control in a way to distinguish it from pronominal binding.
- 73 See Castillo, Grohmann, and Drury (1999), Epstein and Seely (2000), Groat (1999), and Hornstein (2000) for discussion.
- 74 Such an approach to PC assumes that verbs like *meet* do not require *grammatically* plural subjects. This implies that sentences like (i) are grammatical though semantically deviant.

(i) John met

The restriction on *meet* is that there be understood to be a plurality of meeters, not that there be a plural subject. This seems correct; anyhow given the acceptability of sentences like (ii).

- (ii)a. The committee met at 6  
b. John met with the men at 6

- 75 Note that the gerundive structure of the adjunct is not the relevant factor given the possibility of PC with gerunds.

(i) John prefers meeting/gathering at 6

- 76 Landau (1999, p. 134) accounts for the absence of PC in these configurations by noting that the adjunct is an island and so not subject to AGREE. This would then block an AGREE relation between the T-Agr element in the adjunct and the higher T. The problem with this analysis, however, is that it does not explain two features of the indicated structures. First, they display all the other diagnostic properties of OC, as indicated. Second, it requires assuming that logophors (these occupy the subject position of infinitives within islands) cannot have PC-like readings. But this is incorrect.

(i) John said that PRO meeting/gathering together at 6 would be fine with him

The PRO in (i) can sustain a PC-like reading. If so, why cannot it do so in the adjunct case in the text? The movement theory answer is that the PRO in the purpose clause is an OC PRO and so patterns as such, while the PRO in (i) is not an OC PRO. See below for further discussion of PC readings and NOC interpretations.

- 77 The idea that PC structures are the result of a null *tachi*-like element that appears in Japanese PC structures has been suggested by Boeckx (p.c.). He suggests that *tachi* is an adjunct. He also assumes that adjuncts can be merged non-cyclically. In fact, given that *tachi* is phonetically null, it is possible that it gets merged *only* at LF. Let us also assume, following Boskovic (2001) that adjunction to elements in theta positions is prohibited. This implies that *tachi* can only be found on copies in non-theta positions. This would then yield a structure like (i) for (97).

(i) John wants [[[John]+tachi] to [John talk about himself] all together . . .

Note that the antecedent for *himself* is plausibly the lowest copy of *John* and it is singular. The matrix *John+tachi* is not the relevant binder.

This implementation is even consistent with the idea that adjunction is not acyclic. If we assume that Boskovic is correct in that adjunction to elements in thematic positions is blocked, we can adjoin *tachi* to *John* after copying it and before merging it to Spec TP. We could also distinguish EC constructions from PC constructions by supposing that the former have no TP projections while the latter do. Given that Spec TP is non-thematic, adjunction of *tachi* will be possible in the first case but not the latter.

This story might still have some problems. For example, it does not say why purpose clauses cannot have PC readings despite being tensed. However, the ideas behind this implementation are consistent with a movement analysis of control and would serve as well.

78 The facts as reported in Landau (1999) require some further elaboration, I believe. It is not correct that all semantically plural predicates permit a PC reading under verbs like *want* or *decide*. For example, the following seem to resist PC readings.

- (i) \*John wants/decided to PRO be similar/be touching/to look alike/to sing alike/be a group/to be mutually supporting

These are all predicates that are fine with overt plural subjects. The cases in (i) contrast with Landau's examples in that the embedded verbs all resist additional comitative *with* adjuncts.

- (ii) \*John is similar/is a group/sings alike with Bill

This suggests that the relevant generalization behind PC is that certain verbs can select embedded comitatives. How this is to be executed, however, is unclear at present. For further interesting discussion of PC and the subtleties that surround it see Pires (2001).

79 For a very good review, from which I have freely, (though selectively) borrowed here, see Culicover and Jackendoff (2000). I will not be able to deal with all the examples that this paper presents. Nor am I at all confident that the movement theory can be extended to cover all the complexities of the control facts in nominals. What I am more confident of is that there is too little theoretical structure to the proposal offered by Jackendoff and Culicover. What they suggest is that control is basically to be coded in terms of theta theory. But that is about *all* that they propose. Basically, the theory of control consists of a list of controllers coded by thematic role. For example, *try/attempt* is agent control, *tell* is addressee control, *persuade* is patient control, *promise* is source control, etc. (see Culicover and Jackendoff 2000, pp. 38 ff.). Thus, the theory of control is essentially a list. Perhaps they are correct. However, I believe it is worth resisting their conclusion for the time being for two reasons. First, if they are right, then it will be relatively easy to work out the details of the theory. Putting lists together is not all that difficult. Second, the grammatical relation between nominals and their verbal counterparts is not always well understood theoretically. Until the relation between the two is better understood it seems premature to conclude that the nominal tail should wag the verbal dog. Of course, it may well be that the nominal side is the dog and the verbal one the tail and, if so, perhaps the present movement approach to control is misconceived.

80 Randy Hendrick (private communication) notes that were OC a function of movement (as proposed here) then the absence of OC inside DPs leads us to expect the absence of raising as well. As is well-known, there is no raising within DP.

- (i)a. \*John's<sub>1</sub> appearance [t<sub>1</sub> to leave]
- b. John appeared to leave

Hornstein (2000) suggests a way of getting control (but not raising) within DPs via sideways movement. The data reviewed here suggests that this view is incorrect. What then blocks sideways movement? One possibility is to assume that the external arguments of DPs are not like those in verbs. The latter are licensed by *v*. It is plausible that Nouns do not have *v* projections and so do not have external thematic arguments. A similar claim is made by Williams (1994, p. 52). If so, the sideways movement postulated in Hornstein (2000) would be barred by Greed and the absence of OC inside DPs would follow.

- 81 Juan Uriagereka and İlhan Cagri (private communication) note that the distinction between raising and control within nominatives might not be as clean cut as generally assumed. Thus there are odd holes in the paradigm of OC structures.

- (i)a. \*John's trial (from *try*) PRO to leave
- b. \*John's start PRO to leave
- c. \*John's finish PRO leaving

And there are some cases of apparent raising that seem remarkably acceptable.

- (ii) (?) John's likelihood PRO to break a leg rose considerably with the advent of the icy conditions on the slopes

These data suggest that the sharp contrast assumed to hold between raising and control in nominals might be somewhat more blurred than usually thought.

- 82 See, for example, Keenan (1987).

- 83 This sort of reasoning might also extend to "object" control cases.

- (i)a. John's order to Frank to wash himself/each other/ourselves
- b. John's pressure on Frank to have himself/themselves/ourselves admitted

The cases involving *himself* strongly favor the object reading. This is not required, however, as the licensing of plural anaphors and *ourselves* indicates. Reasoning analogous to the one on the text could hold for the cases in (i). For example in (ii).

- (ii) I hated John's order to Frank to leave

This has as one reasonable paraphrase (iii).

- (iii) John ordered Frank to leave and I hated it.

Here, *Frank* is necessarily understood as the leaver. If (ii) is so understood, then we will have the appearance of object control.

- 84 NOC is the counterpart of what used to be called Super-Equi.

- 85 They are islands, at least, if the movement is upwards, as would be the case here, rather than sideways, as would be the case with adjunct control.

- 86 Landau (1999, p. 134) offers a similar account for NOC. He proposes that AGREE cannot penetrate islands and so OC cannot be found within them. However, contrary to what is proposed here, Landau (1999) treats adjuncts as islands for A-movement. As noted above, this leaves the fact that they display the full range of OC properties unaccounted for. Nonetheless, Landau (1999) is parallel to the suggestion in Hornstein (1999, 2000) in that it treats the absence of OC in certain constructions as due to their island properties.

- 87 Randy Hendrick (private communication) points out that there are cases that look like control with overt pronouns:

- (i) John confessed that he/\*Mary was lazy

Here it appears that we need the indicated coreference. I would assume that the indicated coreference is a consequence of the lexical meaning of *confess* and not really a kind of control similar to what is found in cases of OC discussed above.

- 88 One possible problem with treating this PRO as *pro* is that it runs foul of the generalization that *pro* does not exist in English which is not a pro-drop language. However, it is possible to interpret the pro-drop parameter in a slightly different way: not as banning *pro* from English, but as prohibiting *pro* from finite subject or case positions. *Pros* which are finite clause subjects can have referential functions, while *pros* in NOC configurations cannot be used referentially. Hornstein (2000) proposes that referential pronouns are lexical elements, in contrast with bound or expletive pronouns. The latter are treated as grammatical formatives while the former are assumed to be in the lexicon. If we assume that all DPs coming from the lexicon require case, then only pro-drop languages will allow *pro* to be referential. In particular, NOC *pro* will not have a referential or deictic use. This seems more or less correct empirically. Thus, (i) cannot be used in a context where NOC PRO is contextually understood deictically, as say John, in the relevant context.

- (i) PRO<sub>1</sub> washing would amuse Mary<sub>2</sub>

- 89 See Landau (1999, pp. 137 ff.) and works therein cited.  
90 Landau cites other cases as well. However, these are not clearly all of a piece. For example, his example (139, (69)) improves dramatically when the tenses are “sequenced.”

- (i) John’s wife thought that PRO to indulge himself in drinking \*is/was inappropriate

Note further that the whole sentence improves again when we drop *in drinking*. It is unclear why this should make any difference, but it does. I similarly find the putative unacceptability of examples like (ii) (see p. 140) rather doubtful.

- (ii) Mary discussed with friends of John the possibility of him/\*PRO divorcing her

Analogous cases with reflexives are fine.

- (iii) Mary discussed with friends of John the possibility of him/PRO dressing himself in a tux for dinner

I should add that both cases improve for me if *John* bears genitive case. But this is true in general as I find phrases like *friend of John* awkward.

- (iv) Mary discussed with friends of John’s the possibility of him/PRO dressing himself in a tux for dinner

- 91 See Motomura (2001) for discussion.  
92 One must, of course, control for binding restrictions such as being in the scope of the quantificational antecedent.  
93 One more point. It might be that logophors are more like referential pronouns than bound pronouns. Hornstein (2000) argues that deictic pronouns are very different from “grammatical” pronouns like epithets and bound pronouns. If logophors are deictic, then they are not last resort expressions of the kind discussed in Hornstein (1999, 2000). It is interesting to observe that the examples in (102) are not particularly felicitous with quantificational antecedents for the overt pronouns.

- (102)a. ?\*John said about nobody<sub>1</sub> that it would be easy for him<sub>1</sub> to pass the exam  
 b. \*John sued every woman<sub>1</sub> for divorce because it was no longer possible for her<sub>1</sub> to support him

This suggests that logophors are not bindable and so may well be deictic pronouns of sorts, at least in the English examples provided.

- 94 Of course issues will remain. For example, why NOC PROs are logophors and not simple pronouns. See Sells (1987) for a fuller discussion of logophors. His discussion indicates that logophors (or what have been taken to be such) show quite a bit of cross linguistic variation. Thus, different “roles,” in Sell’s sense, serve as logophoric antecedents in different languages. As such, it is not clear what empirical consequences follow from saying that some expression is a logophor. This leaves the question of whether NOC PRO is (ever) a logophor or a pronoun somewhat up in the air.
- 95 Landau (1999, p. 134) makes an analogous suggestion concerning AGREE.
- 96 See Hornstein (1999, 2000) for discussion of this in the context of both control theory and binding theory. See Aoun, Choueri, and Hornstein (2000) for a discussion of movement as last resort in the context of resumptive pronoun constructions. See also Hornstein (2000) for a general discussion of pronouns along these lines.
- 97 There are other issues concerning interpretation within control clauses that this chapter ignores. For example, we say nothing about arbitrary control except to say that it is a species of NOC. Note that it is compatible with the views presented here that arbitrary control is a species of OC with an implicit controller (see Epstein, 1984). However, I am skeptical that this is always so given examples like (i).
- (i)a. John thinks that shaving oneself is important to Bill  
 b. John thinks that shaving oneself would amuse Bill
- It is not clear where the implicit controller would reside in these cases and yet an arbitrary reading is available.
- 98 Landau (1999) can be read as making a similar distinction. For him OC PRO is anaphoric and requires Agreement. This, in conjunction with the assumption that Agree does not see into islands forces non-anaphoric PRO within islands. What Landau (1999) does not address is why non-anaphoric PRO cannot occur where the reflexive can. One can answer this by assuming that anaphors must be used where they can be used. One complication within Landau’s (1999) theory is that OC PRO is not an anaphor in the sense of being subject to binding principle A. The requirement that anaphors be used in preference to pronouns is typically cast against the background of the binding theory.
- 99 Polinsky and Potsdam (2000) note that this phenomenon has also been documented in other languages, including Brazilian Portuguese (Farrell, 1995) (see below for discussion), Japanese (Kuroda, 1965, 1978; Harada, 1973).
- 100 Polinsky and Potsdam note that these effects are similar to Perlmutter’s (1970) account of *begin* in English.
- 101 This is similar to Boskovic (1994) argument concerning dative subjects in *want* constructions in Chilean Spanish discussed above.
- 102 This sentence is somewhat degraded due to the volitionality requirement placed on the subject of *-oqa*.
- 103 At the end of the paper, Polinsky and Potsdam suggest that the correct parameterization is not on the strength of theta features but on EPP features and

the pro-drop parameter. However, this suggestion leaves the main thrust of their proposal intact; movement offers an account of BC while standard approaches do not.

104 All the data are taken from Farrell (1995).

105 For further data see Farrell (1995).

106 This section is based on work by Ferreira (2000), Modesto (2000) and Rodrigues (2000).

107 This construction has one further interesting property. It appears to allow partial control readings.

(i)a. A Maria disse que ec se encontrou na rua  
Mary said that ec SE meet-3 SG in-the street

b. A Maria disse que se encontrarem na rua  
Mary said that SE meet-3 PL in the street

It appears that all speakers accept a PC reading for (ia). Some also accept it for (ib), some do not. What is interesting is that the availability of the PC reading seems unaffected by the number features on the verb. Note that in these cases these features are overt. Recall that Landau (1999) ties PC readings to inflections *unspecified* for number. This suggests that where number is overtly specified that partial control readings should be affected. However, as indicated in (i) this seems to be incorrect.

These constructions involve further quirky properties that go beyond the scope of this review. However, it seems clear that studying the properties of these sorts of “control” might well shed light on PC given the overt number morphology available.

108 This section is based on Kiguchi (2000) and Kiguchi and Hornstein (2001).

109 There remain several problems that have not been adequately addressed here. We very briefly noted (note 32) the facts concerning quirky case in raising and control structures in Icelandic. They still await complete analyses in a movement based account. Second, Randy Hendrick (private communication) reminds me of an argument in Lasnik and Saito (1992) that goes as follows. They observe the following contrasts:

- (i)a. How likely to win is John  
b. \*How likely to be men at home are there  
c. \*How likely to be taken of John is advantage

They note that if *likely* could get a control as well as a raising analysis then we could analyze cases like (ia) as control. This would then make the relevant cut. As expletives and idioms are barred from control clauses, only (ia) could have a control structure. The raising would be barred by the requirement that traces be properly governed (the Proper Binding Condition).

Whatever the current status of the Proper Binding Condition, the asymmetry noted in (i) is unexpected on the present analysis unless we explain the differences in (ia) versus (ib,c) in some other way, without adverting to control. Consider the following.

Cases like (ia) have one further distinctive property. It appears that the subject is topic-like in being old information. Thus, (ia) is very odd discourse initially. It is felicitous only if we are talking about John’s chances of finishing the race and someone then asks (ia). Note, further, that indefinites are not acceptable here, though definites are. This is what we would expect if the subject had to be a discourse topic.

- (ii)a. \*How likely to finish the race is/a man someone/are many people
- b. How likely to finish the race is that man/the man/John/?everyone

Last of all, sentences like (ia) have distinctive intonations, unlike those found in conventional questions. There is a rising at *is* and falling at *John* in (ia). In fact, these questions are unacceptable in embedded contexts where normal questions are required (contrast (iiia) and (iiib)).

- (iii)a. ?\*I wonder how likely to finish the race he is
- b. I wonder how likely he is to finish the race

What might this be telling us? Say that for some reason the subject in these constructions is actually a topic (maybe even in topic position and hence not easily embeddable under *wonder*). Then, we would need something that was topic-like here. On the (reasonable) assumption that expletives and idioms cannot be topics (old information) then the cases in (ib) and (ic) would be ruled unacceptable regardless of the control versus raising distinction.

There is some independent reason for thinking that this might be the case. Note that (iva) and (ivb) are about as voice transparent as the raising constructions in (v) (setting aside topicality issues).

- (iv)a. How likely to examine Mary is the Dr
- b. How likely to be examined by the Dr is Mary
- (v)a. The Dr is very likely to examine Mary
- b. Mary is very likely to be examined by the Dr

This is what we would expect if both (iv) and (v) were raising constructions.

There is surely more to say about these (and other) cases. However, this seems like a fine place to stop.

## References

- Aoshima, S. (2001). Mono-clausality in Japanese Obligatory Control Constructions. Unpublished manuscript. University of Maryland, College Park.
- Aoun, J., Choueiri, L. and Hornstein, N. (2001). Resumption, movement and derivational economy. *Linguistic Inquiry*.
- Belletti, A. and Rizzi, L. (1988). Psych-verbs and  $\theta$ -theory. *Natural Language and Linguistic Theory*, 6, 291–352.
- Bobaljik, J. and Brown, S. (1997). Interarboreal operations: head movement and the extension requirement. *Linguistic Inquiry*, 31, 357–66.
- Boeckx, C. (2000). A note on contraction. *Linguistic Inquiry*, 31, 357–66.
- Boskovic, Z. (1994). D-structure, theta-criterion, and movement into theta-positions. *Linguistic Analysis*, 24, 247–86.
- Boskovic, Z. (1997). *The syntax of nonfinite complementation: An economic approach*. Cambridge, MA: MIT Press.
- Boskovic, Z. (2001). Floating quantifiers and theta role assignment, *Proceedings of NELS*, 31.
- Boskovic, Z. and Takahashi, D. (1998). Scrambling and last resort. *Linguistic Inquiry*, 29, 347–66.
- Bouchard, D. (1984). *On the content of empty categories*. Dordrecht: Foris.
- Bowers, J. (1973). *Grammatical relations*, PhD dissertation, MIT, Cambridge (reprinted as *The Theory of Grammatical Relations* (1981). New York: Garland).

- Bresnan, J. (1982). Control and complementation. In J. Bresnan (ed.), *The mental representation of grammatical relations*. Cambridge, MA: MIT Press.
- Brugger, G. (1997). Expletive auxiliaries. Unpublished manuscript. UCLA.
- Burzio, L. (1986). *Italian syntax*. Dordrecht: Reidel.
- Castillo, J. C., Drury, J. E., and Grohmann, K. K. (1999). Merge over move and the extended projection principle. *University of Maryland Working Papers in Linguistics*. Vol. 8, 63–103.
- Chomsky, C. (1969). *The acquisition of syntax in children from 5 to 10*. Cambridge, MA: MIT Press.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1980). On binding. *Linguistic Inquiry*, 11, 1–46.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale and S. J. Keyser (eds.), *The View from Building 20. Essays in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, N. (1998). Minimalist issues: the framework. *MIT Occasional Papers in Linguistics*. Vol. 15, Cambridge, MA. MITWPL. To appear in R. Martin, D. Michaels, and J. Uriagereka (eds.), *Step by step*, Cambridge, MA: MIT Press.
- Chomsky, N. (1999). *Derivation by phase*. MS, Cambridge, MA: MIT.
- Chomsky, N. and Lasnik, H. (1977). Filters and control. *Linguistic Inquiry*, 8(3).
- Chomsky, N. and Lasnik, H. (1993). The theory of principles and parameters. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann (eds.), *Syntax: An International Handbook of Contemporary Research*. Berlin and New York: Walter de Gruyter. Reprinted in Chomsky (1995), *The minimalist program*. Cambridge, MA: MIT Press. pp. 13–127.
- Comrie, B. (1984). Subject and object control: syntax, semantics and pragmatics. *BLS*, 10, 450–64.
- Culicover, P. and Jackendoff, R. (2000). *Control is mostly semantic*. MS. Ohio State and Brandeis University.
- Culicover, P. and Jackendoff, R. (2001). Control is not movement. *Linguistic Inquiry*, 32, 493–511.
- Enç, Muvet (1990). On the absence of the present tense morpheme in English. University of Wisconsin, Madison. Unpublished manuscript.
- Epstein, S. D. (1984). Quantifier-pro and the LF representation of PRO-arb. *Linguistic Inquiry*, 15.
- Epstein, S. D. and Seely, T. D. (2000). *SPEC-ifying the GF subject: eliminating A-Chains and the EPP within a derivational model*. MS, University of Michigan.
- Epstein, S. D., Groat, E. M., Kawashima, R., and Kitahara, H. (1998). *A derivational approach to syntactic relations*. Oxford: Oxford University Press.
- Farkas, D. F. (1988). On Obligatory Control. *Linguistics and Philosophy*, 11, 27–58.
- Farrell, P. (1995). Backward control in Brazilian Portuguese. In J. M. Fuller, H. Han, and D. Parkinson (eds.), *Proceedings of ESCOL'94*. Ithaca, NY: Department of Linguistics and Modern Languages, Cornell University. (pp. 116–27).
- Ferreira, M. (2000). *Hyperraising and null subjects in Brazilian Portuguese*. MS. Univesidade Estadual do Campinas, Campinas.
- Fodor, J. (1975). *The language of thought*. New York: Thomas Y. Crowell.
- González, N. (1988). *Object raising in Spanish*. New York: Garland.
- González, N. (1990). Unusual inversion in Chilean Spanish. In P. Postal and B. Joseph (eds.), *Studies in Relational Grammar*. Chicago: University of Chicago Press.

- Groat, E. (1999). Raising the case of expletives. In Samuel D. Epstein, and Norbert Hornstein (eds.), *Working minimalism*. (pp. 27–44). Cambridge, MA: MIT Press.
- Harada, S.-I. (1973). Counter-equi NP deletion. *Research Institute of Logopedics and Phoniatrics Annual Bulletin*, 7, 113–47.
- Higginbotham, J. (1980). Pronouns and bound variables. *Linguistic Inquiry*, 11, 679–708.
- Higginbotham, J. (1992). Reference and control. In R. Larson, S. Iatridou, U. Lahiri, and J. Higginbotham (eds.), *Control and grammar* (pp. 79–108). Dordrecht: Kluwer.
- Hornstein, N. (1990). Verb raising in Icelandic infinitives. *NELS*, 20, 215–229.
- Hornstein, N. (1997). Control in GB and Minimalism. *Glott International*, 2(8) (reprinted in Cheng and R. Sybesma (eds.), *The first glott international state-of-the-article book* (2000). Berlin: Mouton de Gruyter.
- Hornstein, N. (1999). Movement and control. *Linguistic Inquiry*, 30, 69–96.
- Hornstein, N. (2000). *Move! a minimalist theory of construal*. Oxford: Blackwell.
- Huang, C.-T. J. (1983). A note on the binding theory. *Linguistic Inquiry*, 14, 554–61.
- Keenan, E. (1987). A semantic definition of “indefinite NP”, in Eric Reuland and Alice ter Meulen (eds.), *The representation of (in)definiteness*. Cambridge, MA: MIT Press, (pp. 286–317).
- Keenan, E. (1989). A semantic definition of indefinite NP. In E. Reuland and A. G. B. ter Meulen (eds.), *The representation of (in)definiteness*. Cambridge, MA: MIT Press.
- Kiguchi, H. (2000). PRO-gate and movement, Generals Paper, University of Maryland, College Park.
- Kiguchi, H. and Hornstein, N. (2001). *PRO-gate and (sideways) movement*. MS, University of Maryland, College Park.
- Kitihara, H. (1997). *Elementary operations*. Cambridge, MA: MIT Press.
- Koster, J. (1984). On binding and control. *Linguistic Inquiry*, 15, 417–59.
- Kuno, S. (1975). Three perspectives in the functional approach to syntax. *Papers from the Parasession on Functionalism: Chicago Linguistics Society*, April 17, 276–336.
- Kuroda, S. Y. (1965). Generative studies in the Japanese language. Doctoral dissertation. Cambridge, MA: MIT.
- Kuroda, S. Y. (1978). Case marking, canonical sentence patterns, and counter-equi in Japanese. In John Hinds and Irwin Howard (eds.), *Problems in Japanese syntax and semantics*. Tokyo: Kaitakusha. (pp. 30–51).
- Lakoff, G. (1970). *Irregularity in syntax*. New York: Holt, Reinhart and Winston.
- Landau, I. (1999). *Elements of control*. PhD Dissertation, MIT, Cambridge.
- Larson, R. (1988). On the double object construction. *Linguistic Inquiry*, 19, 335–91.
- Lasnik, H. (1995a). Last resort. In S. Haraguchi and M. Funaki (eds.), *Minimalism and linguistic theory* (pp. 1–32). Tokyo: Hituzi Syobo.
- Lasnik, H. (1995b). Last resort and attract F. In *Proceedings of FLSM 6* (pp. 62–81). Bloomington, IN: Indiana University Linguistics Club.
- Lasnik, H. and Saito, M. (1992). *Move a: Conditions on its application and output*. Cambridge, MA: MIT Press.
- Lasnik, H. and Uriagereka, J. (1988). *A course in GB syntax*. Cambridge, MA: MIT Press.
- Lebeaux, D. (1984–5). Locality and anaphoric binding. *The Linguistic Review*, 343–63.
- Lidz, J. and Isardi, W. J. (1997). Chains and phonological form, in A. Dimitraidis, H. Lee, C. Moisset, and A. Williams (eds.), *Proceedings of the 22nd annual Penn Linguistics Colloquium. University of Pennsylvania Working Papers in Linguistics*, 8, 109–25.
- Manzini, M. R. (1983). On control and control theory. *Linguistic Inquiry*, 14, 421–46.
- Manzini, M. R. and Roussou, A. (2000). A minimalist theory of A-movement and control. *Lingua*, 409–47.
- Martin, R. (1996). *A minimalist theory of PRO and control*. PhD dissertation, University of Connecticut, Storrs.

- Modesto, M. (2000). Null subjects and rich agreement. In M. Kato and E. Negão (eds.), *Brazilian Portuguese and the null subject parameter*. Frankfurt am Main: Verbuert.
- Motomura, M. (2001). *Zibun as a Residue of Overt A-Movement*. Generals Paper, University of Maryland, College Park.
- Nunes, J. (1995). The copy theory of movement and linearization of chains in the minimalist program, PhD dissertation, University of Maryland, College Park.
- O'Neill, J. (1995). Out of control. In J. N. Beckman (ed.), *Proceedings of NELS 25* (pp. 361–71). GLSA, University of Massachusetts, Amherst.
- Perlmutter, D. (1970). On the article in English. In Manfred Bierwisch and K. E. Heidolph (eds.), *Progress in linguistics* (pp. 233–48). The Hague: Mouton.
- Pires, A. (2001). Clausal and TP-defective gerunds: control without tense. In *NELS 31*.
- Pires, A. (forthcoming). *The syntax of gerunds and infinitives: subjects, case and control*. PhD dissertation, University of Maryland, College Park.
- Polinsky, M. and Potsdam, E. (2000). Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory*, 19, 583–646.
- Polinsky, M. and Potsdam, E. (2001). *Backwards control*. MS, University of California San Diego and University of Florida, Gainesville.
- Rodrigues, C. (1990). *Deriving Brazilian referential null subjects from movement*. Generals Paper, University of Maryland, College Park.
- Rodrigues, C. (2000). Deriving Brazilian Portuguese null subjects from movement. Unpublished manuscript. University of Maryland, College Park.
- Rosenbaum, P. (1967). *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.
- Rosenbaum, P. (1970). A principle governing deletion in English sentential complementations. In R. Jacobs and P. Rosenbaum (eds.), *Readings in English Transformational Grammar* (pp. 20–9). Waltham, MA: Ginn.
- Roussou, A. and Manzini, R. (1997). Null arguments in early child grammars. *GALA Proceedings UK*. Durham, UK: University of Durham. (pp. 142–7).
- Saito, M. (1992). Long distance scrambling in Japanese. *Journal of East Asian Linguistics*, 1, 69–118.
- Salmon, N. (1986). Reflexivity. *Notre Dame Journal of Formal Logic*, 27, 401–29.
- Sells, P. (1987). Aspects of Logophoricity. *Linguistic Inquiry*, 18, 445–80.
- Sigursson, H. A. (1989). *Verbal syntax and case in Icelandic*. PhD dissertation, University of Reykjavik, Reykjavik.
- Sigursson, H. A. (1991). Icelandic case marked PRO and the licensing of lexical arguments. *Natural Language and Linguistic Theory*, 9, 327–63.
- Stowell, T. (1982). The tense of infinitives. *Linguistic Inquiry*, 13, 561–70.
- Thráinsson, H. (1986). On auxiliaries, AUX and VPs in Icelandic. In L. Hellan and K. K. Christensen (eds.), *Topics in Scandinavian Syntax*. Dordrecht: Reidel.
- Uriagereka, J. (1998). *Rhyme and reason*. Cambridge, MA: MIT Press.
- Williams, E. (1980). Predication. *Linguistic Inquiry*, 11, 203–38.
- Williams, E. (1982). The NP cycle. *Linguistic Inquiry*, 13, 277–93.
- Williams, E. (1994). *Thematic structure in syntax*. Cambridge, MA: MIT Press.
- Wurmbrand, S. (1998). *Infinitives*. PhD dissertation, Cambridge, MA: MIT.