

# THE VIRTUES OF CONTROL AS MOVEMENT

*Cedric Boeckx and Norbert Hornstein*

*Abstract.* Control has played an important role in theoretical debates within the Minimalist Program. This is so because control implicates notions such as module,  $\theta$ -role, the Last Resort nature of syntactic operations, movement, binding, chains, Case, complementation, and more. Hornstein (1999) has controversially claimed that control is a subspecies of movement. That is, control is just like familiar instances of raising, except that it involves movement into an additional  $\theta$ -position. If correct, the movement analysis has important conceptual and empirical repercussions, some of which are examined here.

## 1. Preliminary Remarks

Typical minimal pairs like those in (1) and (2) reveal an asymmetry that Rosenbaum (1967) used to argue for a distinction between *raising* and *control* constructions.

- (1) a. John seems to have examined Bill.  
b. Bill seems to have been examined by John.  
*a and b are synonymous*
- (2) a. John tried to examine Bill.  
b. Bill tried to be examined by John.  
*a and b are not synonymous*

Since Rosenbaum's seminal study, asymmetries like the one just illustrated (see Davies & Dubinsky 2004 for a comprehensive survey of such asymmetries) have standardly<sup>1</sup> been treated as the result of distinct processes underlying raising and control. Whereas raising is a case of movement (3), control is the result of a rule of construal relating PRO to its antecedent (4).

- (3) John<sub>i</sub> seems to have [<sub>i</sub> examined Bill]
- (4) John<sub>i</sub> tried [PRO<sub>i</sub> to examine Bill]

Analyzing raising and control as disjoint processes seems to us to have missed an important fact: that the asymmetry observed in (1) and (2) is not an inherent property of processes (passivization, movement, construal, etc.) but rather results from the thematic difference between *seem* and *try*. Specifically, control

<sup>1</sup> For an early dissenting voice, see Bowers 1973.

verbs are thematically richer than raising verbs. In this brief essay, we claim (as we have done elsewhere, see Hornstein 1999, 2001, 2003; Boeckx 2000; Boeckx & Hornstein 2003, 2004) that once this thematic distinction has been factored out, raising and control can be treated in a grammatically uniform fashion, specifically as the results of movement. Put differently, we would like to do for raising and control (“A-movement”) what Chomsky 1977 did for “A'-movement.” Departing from a tradition that regarded the various A'-constructions as arising from fundamentally distinct processes, Chomsky argued that *wh*-movement, topicalization, *tough*-constructions, and comparatives all share a common core: all involve the application of an operator-movement rule. Note that Chomsky's unification strategy did not eliminate differences that exist between, say, *wh*-movement in interrogatives and *wh*-movement in relatives: for example, asymmetry with respect to extraction from factive complements, constraints on animacy of the *wh*-pronoun, different behaviors in *wh*-island contexts, different landing sites for *wh*-pronouns, availability of resumptive strategies, and so on. Rather, these differences were traced to other sources (e.g., the nature of the moved operators, the positions to which they moved). We adopt a similar strategy regarding raising and control. Both are formed via movement, albeit to positions with different thematic properties.

The existence of residual asymmetries is not specific to the “*wh*-movement” unification or our own. It is true of unification quite generally. In linguistics, for example, the locality unification achieved by Rizzi 1990 under the label of “relativized minimality” for A-movement, A'-movement, and head-movement does not eradicate the fact that head-movement is strictly local and clause-bounded, whereas A'-movement spans clauses, while A-movement falls somewhere in between. Similarly, in physics, the unification of electromagnetism, the weak and the strong interactions is only achieved at very high energies and/or extreme external conditions (high temperatures, densities, etc.). The clear differences in the phenomena observed at relatively low energies are in no way in contradiction with the fundamental underlying unity proposed by physicists.<sup>2</sup> If the residual differences can be attributed to some independent factor, unifying previously disjoint sets of phenomena increases explanatory depth. All things being equal, if we can explain the same domain of data with fewer principles, each principle has that much more independent empirical support. Thus, because reduction always enhances explanation where successful, it is *always* favored (if attainable) and, *ceteris paribus*, always enjoys the methodological high ground.

In the present context, we think that the obvious empirical costs of unifying control and raising under movement are currently quite small and that the advantages—empirical and theoretical—are substantial. Before enumerating some of these, we wish to point out what we believe to be a methodological

<sup>2</sup> We would like to thank the University of South Carolina physicist Chaden Djalali for suggesting the phrasing we use in the text.

virtue of the movement theory of control (MTC). It is a simple theory based on a simple, well-understood idea. As such it is easy to understand and easy to use. Hence, the MTC is a very *useful* theory in the following sense. A theory is useful when its leading ideas and technical assumptions, once adopted, help narrow down analytic options. The MTC does exactly that: it rests on a well-understood construct—A-movement—and on a supposition—that movement into  $\theta$ -positions is licit. This may be right or wrong (and it is easy to see what could make it wrong), but it is relatively clear and easy to apply.

By contrast, the standard view of control brings with it 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 and whose structure, to our knowledge, is almost entirely opaque); 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, e.g., 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.<sup>3</sup> In short, in our opinion, the main problem with the standard view of control is its lack of structure, hence the limited possibility of deduction, and therefore explanatory depth. This lack of theoretical structure at once gives traditional accounts great descriptive latitude and robs them of any significant explanatory insight. As we also believe that the MTC is not overburdened with empirical inadequacies (in fact, we believe the opposite), we see no reason to limit our explanatory ambitions and believe that a lot can be learned by adopting the MTC as a working hypothesis.

In what follows we highlight a few virtues of the MTC. For a full-fledged discussion, see Boeckx, Hornstein, and Nunes, in preparation.

## 2. Virtues of Control-as-Movement

### 2.1 Demodularization

From an architectural point of view, as already stated, the MTC gets rid of an odd-looking element, PRO, by reducing it to an NP-trace (odd because it seems entirely bereft of interpretable features being both phonetically null and anaphorically dependent). The MTC does not need special government conditions (unlike the PRO theorem) or special features like null Case to license the distribution of PRO. Indeed, it is expected that PRO will appear in positions from which A-movement is licensed (generally speaking, non-Case-marked positions, though see the discussion below). The MTC gets rid of an

<sup>3</sup> This coincidence between movement and control properties was already noted in Bowers 1973 and Koster 1984.

entire GB module (the control module) by reducing the anaphoric dependencies typical of PRO to those witnessed in A-chains and this opens the door to the possible elimination of an entire class of rules (construal rules; see Hornstein 2001).

## 2.2 Phonetic Nullness

From an explanatory point of view, the MTC genuinely *explains* why PRO is always null at PF: copies left by movement are always null at PF (see below for what the MP has to say about why this is so). It also explains why it functions like an NP-trace with respect to Sandhi effects like *wanna* contraction (*I wanna go*) (see Hornstein 1999; Boeckx 2000).

- (5) a. John wants to → *wanna go* to New York.
- b. John has to → *hasta go* to New York.

## 2.3 Locality

### 2.3.1 Minimal distance

The MTC also *explains* the locality of control—specifically, the fact that PRO occurs only in the highest subject position, and also the fact that as Rosenbaum (1967) originally observed, the controller/PRO relation generally obeys the Principle of Minimal Distance (PMD). That is, PRO must be bound by the closest antecedent. For example, in (6) PRO must be controlled by the object, not the subject.

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

The PMD follows on a movement theory of control if one assumes that movement is governed by relativized minimality—a standard assumption.

To see this, consider what the derivation of (7) would have to be like were *John* the antecedent of PRO.

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

The copies of *John* mark the history of derivation, in accordance with the now standard copy theory of movement (see Chomsky 1993). 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 (8).

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

So, if PRO is the residue of A-movement, the PMD automatically follows.<sup>4</sup>

### 2.3.2 (Non-)finiteness

Still within the domain of locality, the MTC *explains* why “PRO” occurs in “porous” (i.e., nonfinite, tense- or  $\phi$ -defective) contexts: these are contexts that facilitate extraction (see Boeckx 2003b, 2005).

As is indeed well known, subject (A-)movement typically takes place from nonfinite, uninflected clauses.

- (9) a. John is likely [*t* to be home]  
 b. \*John is likely [*t* is home]

Likewise, control is typically confined to nonfinite clauses.

However, Landau (2004) has gathered an extensive survey of so-called finite control phenomena, constructions where control takes place out of finite clauses. (See also Rodrigues 2004, which we discuss below.)

- (10) Hem kivu še yelxu ha-bayta mukdam. (Hebrew)  
 they hoped that will-go.3PL home early  
 ‘They hoped to go home early.’
- (11) I Maria prospathise na divasi. (Modern Greek)  
 the Maria tried.3SG C-SUBJ read.3SG  
 ‘Mary tried to read.’

As one would expect, there are severe restrictions on when finite control can take place. Careful examination of crosslinguistic data suggests that the finite

<sup>4</sup> There are well-known counterexamples to Rosenbaum’s PMD that (some claim) demonstrate that the PMD is incorrect. As the PMD is a consequence of the MTC, it too must be false. The argument form is impeccable but too simplistic in our view. The putative problems for the PMD come in two varieties. The first involves control into the complements of *promise*, illustrated in (i). The second involves cases of “control shift,” illustrated in (ii).

- (i) John promised Mary PRO to wash himself/\*herself.  
 (ii) a. John asked Mary to shave herself/\*himself.  
 b. John asked Mary to be allowed to shave himself/\*herself.

In both of these examples *John* appears to control the embedded PRO subject. The problem is that there appears to be an intervening nominal, *Mary*, that sits between PRO and *John*. This violates minimality, and so this reading should be unavailable. However, the minimality problem (and the argument against the PMD/MTC) disappears at once if the relevant interveners are DPs embedded inside a PP structure headed by a null preposition, as suggested in Boeckx and Hornstein 2003. Boeckx, Hornstein, and Nunes (in preparation) discuss data like (i) and (ii) at length and show that a PP analysis is not only adequate empirically, it also explains the judgment fluctuations and acquisition delay witnessed in these constructions.

character of clauses into which control can take place cannot solely be characterized in terms of the content of Infl. Instead, as Landau has shown, *both* Infl and C must be taken into consideration. Landau comes up with the classification given in Table 1.

Specifically, Landau documents numerous cases where control is allowed inside subjunctive clauses or is limited to third-person agreeing subjects. Taking finiteness to be expressed by the C-Infl pair (see Rizzi 1997; Chomsky 2005a,b; Miyagawa 2004; Boeckx 2003b, 2005, among others), we would like to argue that control, in our terms, movement, can take place out of a finite clause only if the finite clause out of which extraction takes place is temporally deficient (subjunctive) or  $\phi$ -deficient (limited to third person, i.e., [Number] agreement). The overall deficiency of finite clauses into which control can take place is clearly visible in Table 2 (adapted from Landau's chart to highlight the role of third-person agreement):

From this we can conclude that A-extraction of subjects of finite clauses is licit only if the intermediate valuing sites (C-T complex) is defective, tense-wise or  $\phi$ -feature-wise (i.e., whenever one of the features on the C-I complex is [-]). This is exactly the same generalization Boeckx 2003b proposed for raising out of finite (subjunctive) clauses in languages like modern Greek, which lacks infinitival embedded clauses (see Alexiadou & Anagnostopoulou 1999; for raising out of subjunctive clauses, see Uchibori 2000 for Japanese and Uriagereka 2001 for Romance).

- (12) Ta pedhia dhen fenonte na doulevoun. (Modern Greek)  
 the children not seem.3PL SUBJ-C work.PRES  
 'The children do not seem to work.'

Table 1. Classification of finite control (Landau 2004)

	Obligatory control			No control		
	<b>EC-infinitive</b>	<b>Balkan C-subjunctive</b>	<b>Hebrew third-person subjunctive</b>	<b>PC-infinitive</b>	<b>Balkan F-subjunctive</b>	<b>Indicative</b>
$I^0$	[-T, -Agr]	[-T, +Agr]	[+T, +Agr]	[+T, -Agr]	[+T, +Agr]	+T, +Agr]
$C^0$	[-T]	[-T]	[+T, +Agr]	[+T, (+Agr)]	[+T, +Agr]	$\emptyset$

Table 2. Classification of finite control (from a movement perspective)

	Obligatory movement			No movement
	<b>Infinitive</b>	<b>C-subjunctive</b>	<b>Third-person limitation</b>	<b>Declarative</b>
T	-T, - $\phi$	-T, + $\phi$	+T, non-max $\phi$	+T, + $\phi$
C	-T	-T	+T, (non-max $\phi$ )	+T, (+ $\phi$ )

Boeckx 2003b, 2005 argues that the reason movement can take place out of defective finite domains is because a chain can be extended up to its point of maximal checking, which in the case of A-chains is defined in terms of [full T(ense)–full  $\phi$  [Person + Number] (see Richards 2001, Rizzi 2004 for related ideas). One can think of this locality condition on chains as an interface condition that identifies maximal checking site with pronunciation site. Given that an element can only be pronounced once (Nunes 2004), once it reaches a maximal checking site (e.g., subject of declarative finite clause), the element is “frozen in place.”

Note, crucially, that a maximal checking site for an A-chain is not defined in terms of case checking (contra Landau 2004, this volume, who focuses on the issue of Case-checking to argue against a movement approach to control on the basis of finite control data). This is so because (a) there is nonnegligible evidence that multiple cases can be assigned to a single chain (see, e.g., the phenomenon of multiple Case assignment in languages like Korean; Yoon 1996 and much subsequent work; see also Boeckx and Hornstein, in press), and (b) Case is no longer central to current theories of checking, having been reinterpreted in terms of T- (Pesetsky & Torrego 2001) or  $\phi$ -checking (Chomsky 2001, 2004).

### 2.3.3 *Evidence from Brazilian Portuguese*

A version of the argument we proposed to handle finite control was developed in Rodrigues 2004 on the basis of Brazilian Portuguese (BP) data. We wish to briefly summarize her main argument here because she has provided some complex morphological evidence for the kind of derivation we have advocated on more abstract grounds.

Rodrigues’s starting point is the observation that BP is a partial *pro*-drop language. BP disallows null subjects except in two contexts: matrix first-person pronouns can be dropped, and embedded third-person pronouns can be, too. Rodrigues provides (to our minds) convincing evidence that matrix first-person pronoun drop is the result of a topic-drop process. We will say nothing more about it here. The cases of embedded third-person cases are more interesting, because Rodrigues (see also Ferreira 2004) provides a battery of tests to show that these are in fact residues of movement out of finite clauses, in a way similar to what we have been arguing for PRO in obligatory control contexts.

Rodrigues shows that a movement analysis straightforwardly explains all the salient properties of embedded third-person null pronouns. In particular, a movement analysis accounts for (i) why they are embedded (14), (ii) why they need an antecedent (15), (iii) why they are obligatorily bound by their antecedents (16), (iv) why the way they relate to their antecedents is subject to an intervention condition strongly reminiscent of relativized minimality (17), and (v) why they are excluded inside islands (18).

- (14) \**e* Está doente. (\*BP/√Romance *pro*-drop lang.)  
 is-3SG sick  
 'S/he is sick.'
- (15) \*Parece que *e* estava cansado.  
 seems-3SG that was-3SG tired-SG  
 'It seems that s/he was tired.'
- (16) O João<sub>1</sub> disse que *e*<sub>1/\*2</sub> estava cansado.  
 the João said-3SG that was-3SG tired-SG  
 'João said that he was tired.'
- (17) O Paulo<sub>1</sub> me contou que o João<sub>2</sub> disse que *e*<sub>\*1/2</sub>  
 the Paulo me told-3SG that the João said-3SG that  
 vai ser promovido.  
 going-3SG to-be promoted  
 'Paulo told me that João said that he (João) will be promoted.'
- (18) \*Quem<sub>1</sub> que o João encontrou a carteira que *t*<sub>1</sub> perdeu.  
 who that the João met-3SG the wallet that lost-3SG  
 'Who did João find the wallet that he lost?'

Rodrigues also discusses an interesting gender agreement property active in BP that provides evidence for a movement analysis of the finite control construction under discussion. There are certain nouns in Romance that are invariably marked as [+feminine], though semantically they can refer to either male or female entities. One of these nouns is the Romance counterpart of *victim*. Thus, when *victim* is combined with the auxiliary verb *to be* followed by a participial form, the participial form records feminine gender agreement; witness the data in (19) from Italian and BP illustrate.

- (19) a. La vittima fu aggredita/\*aggredito dai  
 the victim-FEM was.3SG attacked-FEM/MASC by  
 fascisti. (Italian)  
 fascists
- b. A vítima<sub>1</sub> foi atacada<sub>1</sub>/?atacado<sub>1</sub> na rua. (BP)  
 the victim-FEM was-3SG attacked-FEM/MASC in.the street  
 'The victim was attacked by (the) fascists/on the street.'

(20) shows that feminine gender is also morphologically recorded by a universal quantifier *all* related to the [+fem] noun. Note, that it does not matter whether the quantifier floats or not, it is always [+fem].



- (20) a. Tutte/\*tutti le vittime arrivarono nello  
all-FEM/MASC the victims-FEM arrived-3PL in-the  
stesso momento. (Italian)  
same moment
- b. Todas/\*todos as vítimas chegaram no  
all-FEM/MASC the victims-FEM arrived-3PL at.the  
mesmo horário. (BP)  
same time  
'All the victims arrived at the same time.'
- (21) a. Le vittime arrivarono tutte/\*tutti nello  
the victims-FEM arrived-3PL all-FEM/MASC in-the  
stesso momento. (Italian)  
same moment  
As vítimas chegaram todas/\*todos. (BP)  
the victims-FEM arrived-3SG all-FEM/MASC  
'All the victims arrived (at the same time).'

Following much work on past participle agreement (Kayne 1989 and subsequent work), Rodrigues takes past participle agreement (and agreement on floated quantifiers) to be established locally (via Spec-head within an AgrP dominating VP, or under Agree, subject to intervention). Rodrigues shows that in obligatory control configurations, raising configurations, and configurations involving an embedded third-person plural null subject in BP, a quantifier or a past participle form within the embedded clause agrees in gender with the antecedent of embedded null subject, as represented in (22).

- (22) [S ...[DP<sub>Fem</sub>]<sub>1</sub>...[S<sub>inf</sub> PRO<sub>1</sub>/e<sub>1</sub>...Past Participle<sub>Fem</sub>/Quantifer<sub>Fem</sub>]]

Witness (23) (obligatory control), (24) (raising), and (25) (finite control).

- (23) a. La vittima<sub>1</sub> ha cercato di essere  
the victim-FEM had-3SG tried of be-INF  
trasferita<sub>1</sub>/??trasferito<sub>1</sub> alla stazione di polizia  
transferred-FEM/MASC to.the station of police  
di College Park. (Italian)  
of College Park
- b. A vítima<sub>1</sub> tentou ser transferida<sub>1</sub>/??transferido<sub>1</sub>  
the victim-FEM tried be-INF transferred-FEM /MASC  
para a delegacia de polícia de College Park. (BP)  
to the station of police of College Park  
'The victim tried to be transferred (to the police station at College Park).'

- (24) a. La vittima sembra essere ferita/\*ferito. (Italian)  
 the victim seems-3SG be-INF injured-FEM/MASC  
 b. A vítima pareceu estar ferida/\*?ferido. (BP)  
 the victim seemed-3SG be-INF injured-FEM/MASC  
 ‘The victim seemed to be injured.’

- (25) A vítima<sub>1</sub> disse que e<sub>1</sub> foi atacada<sub>1</sub>/??atacado<sub>1</sub>  
 the victim-FEM said-3SG that was-3SG attacked-FEM/MASC  
 na rua. (BP)  
 in.the street  
 ‘The victim said that he was attacked on the street.’

Rodrigues notes, correctly we think, that if the embedded null subject is a trace of its antecedent, then nothing is left to be explained about (22). The past participle and quantifier agree locally with the [+fem] DP, prior to the movement of this DP to the matrix clause.

Rodrigues’s analysis correctly predicts that gender agreement would fail if movement does not take place, as in nonobligatory control configurations, or in Romance languages that are not partial *pro*-drop (i.e., for those languages where *pro* is a distinct element, not the residue of movement). As (26) and (27) show, these predictions are borne out.

- (26) a. La vittima<sub>1</sub> ha detto che essere  
 the victim-FEM has-3SG said that be-INF  
 \*portata<sub>1</sub>/portato alla stazione di polizia non  
 brought-FEM/MASC to.the station of police not  
 e una buona idea. (Italian)  
 is-3SG a good idea  
 b. A vítima<sub>1</sub> disse que ser ?? levada<sub>1</sub>/levado<sub>1</sub>  
 the victim-FEM said-3SG that be-INF brought-FEM/MASC  
 para a delegacia de policia não é uma  
 to the station of police not is-3SG a  
 boa idéia. (BP)  
 good idea  
 ‘The victim said that being transferred to another city is not a good idea.’

- (27) La vittima<sub>1</sub> ha detto che *pro*<sub>1</sub> era \*stada  
 the victims-FEM has-3SG said that was-3SG been-FEM  
 aggredita<sub>1</sub>/stato aggredito<sub>1</sub> in strata. (Italian)  
 attacked-FEM/MASC in street  
 ‘The victim said that he was attacked on the street.’

Rodrigues points out that under a PRO-based analysis of control, the agreement contrast between obligatory control and nonobligatory control cases, (23) versus (26), is puzzling. In both cases, either PRO should be transparent to agreement by the antecedent or should block agreement by the antecedent, or should trigger agreement itself, but in any case, since PRO is assumed to be the subject of the relevant clause in both obligatory and non-obligatory control cases, no contrast is expected.

The concord facts discovered by Rodrigues are fascinating in their own right. But they bear emphasis in the context of the MTC, as similar (Case/agreement) concord facts from Icelandic and other languages have repeatedly been invoked (with no detailed analysis!) as providing lethal evidence against the MTC (see Landau 2003, 2004, this volume). We have argued elsewhere (Boeckx & Hornstein, in press) on the basis of data from Icelandic (see also Cecchetto & Oniga 2004 for a similar point on the basis of data from Latin) that on closer scrutiny the MTC offers the most straightforward analysis of such concord facts. In other words, what looked like a serious argument against the MTC in fact constitutes a good argument in its favor.

#### 2.4 Interpretation

The MTC also *explains* many aspects of the interpretation of obligatorily controlled “PRO.”<sup>5</sup> Specifically, as Hornstein (2001:39–41) has shown in detail, the interpretive restrictions found in (obligatory) control contexts (obligatoriness of a [local] antecedent, ban on split antecedents, *de se* interpretation) are precisely those we expect if PRO is in fact a copy left by A-movement. For example, the fact that (obligatory control) PRO requires a local c-commanding antecedent follows from the fact that PROs are traces in A-chains and share the properties that such traces enjoy. For example, they will necessarily have “antecedents” (viz. the head of the A-chain), and these antecedents will c-command them (as this is a well-formedness condition on chains). Similarly, the prohibition against split antecedents (\*John<sub>1</sub> asked Mary<sub>2</sub> PRO<sub>1+2</sub> to kiss each other<sub>1+2</sub>) follows from the fact that two elements cannot move from the same position (i.e. that traces cannot have split antecedents). These interpretive restrictions were largely stipulated in earlier approaches to control. Given the MTC, they are largely derivable.

In sum, MTC answers *why*-questions and provides a unified theory of PRO’s distribution and interpretation,<sup>6</sup> while at the same time simplifying the theory as a whole.

<sup>5</sup> On the nature of nonobligatory control, see Boeckx and Hornstein, to appear, where it is argued that the null category involved in *pro*.

<sup>6</sup> The MTC has other virtues such as those of offering a unified analysis of standard complement and adjunct control constructions, and accounting for PRO-gate effects. We will not go into these here. See Hornstein 2003, Kiguchi 2002, and Boeckx, Hornstein, and Nunes, in preparation.

### 2.5 Empirical Bonus

In addition to all these theoretical virtues, the MTC has offered the possibility of making sense of the phenomenon of backward control, which had been observed in the 1980s, but resisted an explanation. (One can almost say that the MTC helped linguists “(re)discover” backward control.)

Polinsky and Potsdam (2001) offer the most comprehensive analysis of backward control in Tsez, a language of the Caucasus. (Backward control has also been identified in Malagasy, Brazilian Portuguese, Bezhta, Tsaxur, Kabardian, Adyghe, Korean, Japanese, and possibly Jacalteco. See Polinsky and Potsdam, this volume, for further discussion). Backward control describes cases in which the controlled PRO (asymmetrically) c-commands its antecedent. (Observe that the PRO in (28) necessarily takes *kidba* as its antecedent.)

- (28) PRO [kidba      ziya      bisra]      yoqsi.  
         girl-ERG cow-ABS feed-INF began  
         ‘The girl began to feed the cow.’

Polinsky and Potsdam show in detail that the standard theories of control that involve PRO and binding cannot easily account for backward control constructions.

In fact, in PRO based approaches to control, BC should simply be impossible as the control configuration illustrated in (28) should lead to serious ungrammaticality once the control relation is established. For example, in the configuration in (28) PRO c-commands the antecedent that it is anaphorically dependent on. This violates both principles A and C on the assumption that (obligatory controlled) PRO is anaphoric and its antecedent in (28) is an R-expression. By contrast, backward control can be accounted for in a movement based theory in a rather straightforward manner, viz. as a case of pronunciation of a lower copy, which as the literature has revealed is not unique to control (see Nunes 2004, Bošković 2001, and references therein). In closing, let us note that backward control is one of these empirical bonuses that theoretical research hopes for—a phenomenon that does not fit in any framework except in the one that appears to be the most desirable one from a purely conceptual point of view. (To be precise, the phenomenon of backward control relies on the MTC and another property of grammar with many minimalist virtues, the copy theory of movement.)

### 3. Conclusion

The idea that the grammatical processes responsible for these two types of structures are less different than generally believed is *not* novel (see, e.g., Bowers 1973). What is different is the conceptual setting of the MP—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). If correct, the movement hypothesis offers a case of reductionism of the kind minimalism strives for.

Of course, we want our quest to be empirically responsive, but empirical responsiveness should not entail the sort of theoretical and technological profligacy that makes explanation impossible. In fact, as Chomsky has rightly observed, far too often we mistake technical redescriptions for explanations. In our view, the main problem with standard approaches to control is that they track the facts rather than explain them and this is because they avail themselves of an overly rich descriptive apparatus. In our view, Landau 2000, and, in an even more extreme fashion, Culicover and Jackendoff 2001, 2005 and Jackendoff and Culicover 2003 offer illustrations of what we have in mind.

At this stage we feel that the MTC offers a good combination of theoretical structure and empirical coverage. The problems that remain (such as the issue of control in nominals, see Culicover & Jackendoff 2001) do not seem lethal, at least at present (for reasons discussed in Boeckx & Hornstein 2003). Thus the MTC seems to us to be a very good working hypothesis, one that fits well with general minimalist ambitions in the following way. Chomsky has remarked that we should aim to find out *how* the faculty of language is optimally designed. This is to assume that it is. Why believe that? The best answer that we can think of is that, if it isn't, we will soon find out by trying to see how it might be. In short, a well-designed system should be our default position. Such types of systems enjoy a certain kind of epistemological privilege. They should form the bull's eye that theory aims to hit. The principle virtue is that having this sort of target will be illuminating even if it is missed. And that is a very good reason to aim for it.

*Cedric Boeckx*  
*Harvard University*  
*Department of Linguistics*  
*Boylston Hall, 3rd Floor*  
*Cambridge, MA 02138*  
*USA*

*cboeckx@fas.harvard.edu*

*Norbert Hornstein*  
*University of Maryland*  
*Department of Linguistics*  
*1401 Marie Mount Hall*  
*College Park, MD 20742*  
*USA*

*nhornste@umd.edu*