Minimalism and Control[1]

Norbert Hornstein (University of Maryland)
Jairo Nunes (Universidade de São Paulo)

This paper discusses control within minimalism, focusing on general properties that any minimalist theory of control should have. Contrasting the Movement Theory of Control and PRO-based approaches to control, we argue that the MTC fares much better than its competitors in that it not only covers more empirical ground, but does so by relying on key architectural features of the Minimalist Program.

1. Introduction

This chapter discusses what properties a minimalist theory of control should have and how close extant proposals are, in meeting these desiderata. In particular, we concentrate on movement and PRO-based approaches to control, taking Hornstein 1999, 2001 and Boeckx, Hornstein and Nunes 2010 to be representative of the former, and Landau 2000, 2004, the latter.[2] Though the review is intended to be dispassionate, the reader should be familiar with the biases of the authors. We are of the vociferous opinion that a minimalistically respectable account of control will necessarily have some version of the Movement Theory of Control (MTC) at its core. Thus, in what follows the star is the MTC, the PRO-based approaches exploited as a useful foil (think Holmes and Watson).

Before getting down to some detail, we would like to outline the form of the argument in what follows. It has three steps. First, we show that many of the salient properties of obligatory control follow if we assume that it involves A-movement of the controller from the position of "PRO".[3] Second, we show how the MTC heavily relies on central minimalist assumptions. Third, we argue that standard PRO-based accounts of control violate one or another minimalist stricture. The conclusion is that if minimalism is on the right track, then some version of the MTC must be correct.

2. What any theory of control should account for

Any adequate theory of control should meet at least four desiderata. First, it must specify the kinds of control structures that are made available by UG and explain how and why they differ. Assuming, for instance, that obligatory control (OC) and non-obligatory control (NOC) are different, their differences should be reduced to more basic properties of the system.

[1] We would like to thank the editors of this volume and especially Yosuke Sato, for comments and suggestions on an earlier version of this paper. We would also like to acknowledge the support received from CNPq (grant 309036/2011-9; second author).

[2] Due to space limitations, we will only examine syntax-centered approaches to control. For detailed criticism to semantic-based accounts, see in particular Boeckx, Hornstein and Nunes 2010:chap. 7.

[3] We use scare quotes, i.e. "PRO," as the MTC denies that OC PRO exists. Thus, the usage here is purely descriptive.

Second, the theory must specify the nature of the controllee: what is its place among the inventory of null expressions provided by UG? Is it a formative special to control constructions or is it something that is independently attested?

Third, the theory must correctly describe the configurational properties of control, accounting for the positions that the controller and the controllee can occupy. In addition, it should provide an account as to why the controller and the controllee are so configured. Assuming, for instance, that the controllee can only appear in a subset of possible positions (e.g. ungoverned subjects), why are controllees so restricted?

Finally, the theory must account for the interpretation of the controllee, explaining how the antecedent of the controllee is determined and specifying what kind of anaphoric relation obtains between the controllee and its antecedent (in both OC and NOC constructions) and why these relations obtain and not others. For instance, assuming that controllers must locally bind controllees in OC constructions, why is the control relation so restricted in these cases?

As mentioned above, these desiderata hold of any approach – be it minimalist or not – that aims to *explain* the central features of control, rather than simply listing or stipulating them. Of course, additional strictures also come into place once these goals are explored against a minimalist setting. In the sections that follow we will discuss how the MTC and PRO-based accounts fare with respect to the four tenets listed above, once the mechanisms they rely on are examined using minimalist guidelines.

3. Control and the Duck Principle

The starting point of our discussion will be the useful methodological maxim expressed in (1):

(1) *The Duck Principle*: If something walks, talks and defecates like a duck, the default position is that it is a duck: i.e. If constructions α and β have the same properties, the grammar should generate them in the same way.

3.1. Warming up

Bearing the Duck Principle in mind, let us consider the data in (2)-(7) for starters:

(2) a. *[It$_i$ was expected [t$_k$ to shave himself$_k$]]
b. *[It$_i$ was hoped [PRO$_k$ to shave himself$_k$]]

(3) a. *[[John$_1$'s sister] was hired t$_1$]
b. *[John$_1$'s campaign hopes [PRO$_1$ to shave himself]]

(4) a. *[John$_1$ seems [that it was likely [t$_1$ to shave himself]]]
b. [John$_k$ convinced Mary$_i$ [PRO$_{i/*k}$ to leave]]

(5) a. *[John$_1$ seems [(that) t$_1$ will travel tomorrow]]
b. *[John$_1$ said [(that) PRO$_1$ will travel tomorrow]]

(6) a. John seems to be cooperative and Bill does too
b. [John$_1$ wants [PRO$_1$ to win]] and [Bill does too]

('… and Bill wants himself to win'/*'… and Bill wants John to win')

(7) a. *[John$_1$ strikes Bill$_2$ [t$_{1+2}$ as jealous of each other]]
b. *[John$_1$ asked Bill$_2$ [PRO$_{1+2}$ to shave themselves/each other]]

Given that the expletive is not a suitable antecedent for the anaphor in (2a), its ungrammaticality shows that A-traces cannot simply pick up their antecedent in discourse, but rather require a syntactic antecedent. In turn, (3a) shows that such syntactic antecedent must be in c-commanding position. (4a) further shows that minimality also matters: there can be no proper interveners between an A-trace and its (c-commanding) antecedent. (5a) shows that an A-trace (in English) cannot occupy the subject position of a finite clause. Finally, (6a) and (7a) respectively show that an A-trace receives sloppy interpretation under ellipsis and cannot take split antecendents.

The *a*-sentences in (2)-(7) are textbook illustrations of configurational and interpretive properties ascribed to A-traces. What is crucial for our discussion is that the same properties describe OC PRO. Thus, (2b), (3b), and (4b) jointly show that OC PRO also requires a local c-commanding antecedent; (5b), that OC PRO cannot be the subject of a finite clause (in English); and (6b) and (7b), that OC PRO also trigger sloppy interpretation under ellipsis and cannot be licensed by split antecedents.

The *b*-sentences in (2)-(7) illustrate some of the general properties of OC,[4] but in no way depict all of the empirical diversity associated with OC. For instance, the property illustrated in (5b) is not universal. In Brazilian Portuguese, for example, the embedded null subject of a finite (indicative) clause has the same interpretive properties as OC PRO.[5] The sentence in (8) below, for instance, shows that the empty category in the embedded subject position cannot freely pick up an antedecent in the discourse (indicated by the *w*-index), but must be interpreted as co-indexed with a local c-commanding DP. Hence, it must be interpreted as the *m*-indexed phrase *[o irmão d[o João$_k$]]$_m$*; not as *Pedro$_i$* because it is not local and not as *João$_k$* because it is not in a c-commanding position.

(8) *Brazilian Portuguese:*
[[o Pedro]$_i$ disse [que [o irmão d[o João$_k$]]]$_m$ estava achando
the Pedro said that the brother of-the João was thinking
[que *ec*$_{m/*i/*k/*w}$ deveria ganhar uma medalha]]]
that should receive a medal
'Pedro said that [João's brother]$_m$ was thinking that he$_m$ should get a medal.'

The interesting thing to point out is that languages that admit sentences like (5b) also allow sentences analogous to (5a). In other words, once "PRO" is permitted in the subject position of a finite clause in Brazilian Portuguese, so is an A-trace. Thus, hyper-raising sentences such as (9) are also possible in Brazilian Portuguese:[6]

[4] This is a subset of the relevant properties of OC. A fuller description is found in Boeckx, Hornstein, and Nunes 2010:chapter 3. We ignore other properties here for reasons of space.
[5] For relevant data and discussion, see Ferreira 2000, 2009, Rodrigues 2002, 2004, Nunes 2008, 2010, and Petersen 2011.
[6] See Ferreira 2000, 2009, Martins and Nunes 2005, 2010, and Nunes 2008, 2010.

(9) *Brazilian Portuguese:*
[os estudantes]$_1$ **parecem/acabaram** que t_1 **viajaram** mais cedo
*the students seem.**3PL**/finished.**3PL*** *that traveled.**3PL*** *more early*
'The students seem to have traveled earlier.'/'The students ended up traveling earlier.'

If we examine the data in (2)-(9) in light of the Duck Principle, the conclusion is inescapable: the grammatical mechanisms involved in generating A-movement are also involved in generating OC.[7,8] It should be noted that this conclusion is by no means new or intrinsically related to minimalism. Already in Chomsky (1977:82), for instance, we find the remark that "trace and PRO are the same element; they differ only in the way the index is assigned – as a residue of a movement rule in one case, and by a rule of control in the other". So, the task before us now is to investigate which mechanisms available in our minimalist arsenal can be resorted to in order to capture the phenomena that fall under these two rubrics in a unifying way. For the MTC, the answer is straightforward: OC is simply A-movement. Exploring (a version of) Chomsky's (2000, 2001) Agree operation, Landau (2000, 2004) in turn takes OC to be the output of an agreement relation triggered by PRO's feature under-specification.

Given that Agree is taken to be a subcomponent of Move in many minimalist approaches to movement (see e.g. Chomsky 2001), it is not surprising that by and large, the two competing approaches cover the same empirical terrain and, in particular, account for (2)-(9) with a comparable degree of success. Thus, both the MTC and Landau's PRO-based approach to control rely on c-command, minimality, some version of Chomsky's (2000) Activation Condition, and the Parallelism Requirement, for instance. Take the contrast between English and Brazilian Portuguese, for example. Suppose for the sake of the argument that finite T in Brazilian Portuguese may be ϕ-defective.[9] If so, finite clauses may define "porous" domains for *both* A-movement and Agree. In other words, contrasts such as (5) and (8)/(9) in themselves do not provide evidence for one approach over the other, for each approach can equally well incorporate comparable provisos to handle special cases such as (8)/(9).

It is very important to stress this point, for departures from standard cases are often taken to invidiously distinguish PRO-based accounts from the MTC, the exceptions taken as being problematic for the latter but not the former. The control differences between *convince* (cf. (4b)) and *promise* (cf. (10a) below), between *ask* (cf. (7b)) and *propose* (cf. (10b)), and the phenomenon of control shift illustrated in (11) are emblematic in this regard. However, *if* both movement and Agree are subject to minimality, then, as a point of logic, both accounts should in principle be empirically equal as regards controller selection in such cases. The theories don't diverge in their

[7] Please note how this statement is worded. It does *not* identify raising and control. It simply indicates that whatever operations underlie raising *qua* A-movement are also operative in OC configurations.

[8] The Duck Principle in fact invites us to go further and reanalyze anaphoric binding as species of movement, given that Principle A enforces virtually the same conditions on the relation between antecedent and anaphor that OC does. For reasons of space, we won't be able to explore this issue here. For specific proposals and relevant discussion, see e.g. Lidz and Idsardi 1997, Hornstein 2001, Zwart 2002, and Drummond, Hornstein and Kush 2011.

[9] See e.g. Ferreira 2000, 2009, Rodrigues 2004, and Nunes 2008.

conceptions of minimality or in their assumption that the control relation is syntactically mediated by an operation subject to minimality. Thus a problem for either is a problem for both and a remedy for one is likely to heal the other.[10]

(10) a. [John$_k$ promised Mary$_i$ [PRO$_{k/*i}$ to leave]]
b. [John$_1$ proposed to Bill$_2$ [PRO$_{1+2}$ to help each other]]

(11) a. [John$_k$ begged Mary$_i$ [PRO$_{i/*k}$ to leave the party early]]
b. [John$_k$ begged Mary$_i$ [PRO$_{k/*i}$ to be allowed to leave the party early]]

A phenomenon that is taken to favor PRO-based theories over the MTC in a less trivial way is partial control, illustrated in (12) below. The ungrammaticality of (12a) is due to the fact that *gather* requires a semantically plural subject. In turn, the grammaticality of the OC control structure in (12b) indicates that the plurality requirement of *gather* is somehow met in the embedded clause, for the antecedent of PRO is singular. Thus, the mismatch in number between controller and controllee in (12b) appears to show that the controllee cannot be the same as the controller, which would be problematic for the MTC, but may be accommodated in PRO-based theories.

(12) a. *The chair gathered at three.
b. The chair hoped [PRO to gather at three]]

In his in-depth study of partial control, Landau (2000, 2004) notes that only a subset of control structures supports partial control. The complement of implicative verbs such as *manage*, for instance, does not allow it, as exemplified in (13) below. This leads Landau to propose that tense is what is relevant in the licensing of partial control, only tensed infinitives such as the complement of desiderative verbs like *hope* (cf. (12b)) being able to do it.

(13) *The chair managed [PRO to gather at three]]

Putting aside technical problems with Landau's implementation of this licensing of a plural PRO by tense,[11] it is not at all clear that partial control is dependent on tensed infinitival T heads or, more broadly, that it is even a control phenomenon. As observed by Rodrigues (2007), one also finds "partial control" effects where no infinitival complements are involved, as illustrated in (14b) and (15b) below, which have predicates that require semantically plural subjects (cf. (14a) and (15a)). Rodrigues's conclusion is that what is relevant in the licensing of plurality in (12b), (14b), and (15b) is not tense, but modality.

(14) a. *The chair met at 6.
b. The chair can only meet tomorrow.

[10] For extensive discussion of the exceptional cases in (8), (10), and (11) and their analyses within the MTC, see Boeckx, Hornstein, and Nunes 2010:sections 4.4, 5.5, and 5.6.2 and references therein.
[11] See Boeckx, Hornstein, and Nunes 2010:sec. 2.5.2 for detailed discussion.

(15) a. *The chair applied together for the grant.
b. The chair cannot apply together for the grant.

What matters for our current discussion is that under the predicate internal subject hypothesis, the sentences in (14b) and (15b) are to be represented as in (16).

(16) a. [The chair]$_i$ can only [t_i meet tomorrow]
b. [The chair]$_i$ cannot [t_i apply together for the grant]

Thus, when we compare (12) with (14) and (15), we have a Duck Principle effect before us again, as OC PRO and A-traces are behaving alike. So, whatever accounts for the plurality interpretation in (16) should in principle be extended to (12b). Based on the fact that the plurality requirement at stake may also be satisfied via a commitative structure, as illustrated in (17) below, Boeckx, Hornstein, and Nunes (2010) in fact propose that "partial control" effects involve the licensing of a null commitative complement (perhaps by a modal element along the lines of Rodrigues' proposal), as sketched in (18). [12] Notice, in particular, that in (18a) there is no mismatch between PRO and its antecedent, which makes it perfectly possible to analyze PRO as an A-trace.

(17) a. The chair gathered **with Bill** at three
b. The chair met **with Bill** yesterday]
c. The chair applied together for the grant **with Bill**

(18) a. [The chair]$_i$ hoped [PRO$_i$ to gather ***pro***$_{commitative}$ at three]]
b. [The chair]$_i$ can only [t_i meet ***pro***$_{commitative}$ tomorrow]
c. [The chair]$_i$ cannot [t_i apply together for the grant ***pro***$_{commitative}$]

Regardless of whether Boeckx, Hornstein, and Nunes's proposal is on the right track, the important point to emphasize here is that if the MTC has to say something special about partial control in (12b), so do PRO-based accounts with respect to "partial control" effects in monoclausal structures (cf. (14b)/(15b)). Moreover, as PRO-based accounts and the MTC need comparable provisos in terms of tense/mood licensing in order to account for partial control, they are on equal footing in this regard.[13]

Let us then discuss some cases where the Duck Principle may indeed distinguish the MTC from PRO-based approaches.

---

[12] There are non-syntactic ways of implementing the commitative analysis. A recent interesting proposal by Pearson (2012) provides a purely semantic version of the analysis, reducing partial control to a temporal containment principle (rather than modality as in Rodrigues). The paper has three important virtues: (i) it restricts partial control to embedded clauses in a principled way by making it a by-product of a certain kind of anaphoric tense dependency, (ii) (in Pearson's words) "it is compatible with any mechanism whereby PRO inherits ϕ-features from the controller" (under the MTC this is so as PRO is a trace/copy of the controller), and (iii) it immediately accounts for why partial control PRO cannot license plural anaphors, as does any commitative analysis.

[13] That mood properties may interact with control is clearly seen in Japanese, which has three mood particles associated with obligatory control: the "intentive" marker *-(y)oo* with subject control, the imperative marker *–e/-ro* with object control, and the "exhortative" marker *-(y)oo* with split control. See Fujii 2006, 2010 for detailed discussion and analysis.

3.2. Case issues

We may start by examining Duck Principle effects in the domain of Case and morphophological computations.

It has long been observed that the application of some sandhi rules may be blocked by certain syntactic empty categories, the most well known example of such being *wanna*-contraction in English. As illustrated in (19) below, *want* and *to* may contract across an intervening PRO, but not across an intervening A'-trace. Curiously, A-traces also allow similar contraction, as shown in (20).[14]

(19) a. Who$_1$ do you **want PRO to** banish t_1 from the room →
Who do you **wanna** banish from the room?
b. Who$_1$ do you **want t_1 to** vanish from the room →
*Who do you **wanna** vanish from the room?

(20) a. John$_1$ **has t_1 to** kiss Mary → John hasta kiss Mary
b. John$_1$ **used t_1** to kiss Mary → John usta kiss Mary
c. John$_1$ is **going t_1 to** kiss Mary → John usta kiss Mary

Given that one of the standard differences between A-traces and A'-traces is that the former is Caseless while the latter is Case marked, it is very reasonable to assume that this difference is ultimately responsible for contrasts such as the one between (19b) and (20).[15] That being so, PRO in (19a) should be Caseless, which is in consonance with the MTC and GB accounts of control,[16] but not with the major PRO-based accounts within minimalism.[17] Under the approach proposed by Chomsky and Lasnik's (1993) and developed by Martin (2001), for instance, PRO is assigned null Case, whereas for Landau (2004) PRO receives regular Case like any other DP. Both approaches face problems of their own. The former has to explain why only PRO can bear null Case, while the latter fails to account for why PRO cannot be phonetically realized like other DPs marked with regular Case.[18] These problems already hint that the special properties ascribed to PRO in PRO-based accounts may *track* some properties of OC by coding the properties to be accounted for in terms of lexical features, but do not *explain* them. But even if we put these problems aside, what is relevant for our current discussion is that in (19a) we again see that OC PRO walks and talks like a (Caseless) A-trace.

If PRO is a lexical element that receives structural Case (be it null or regular), one might expect it to function like A'-traces rather than A-traces with respect to *wanna*-contraction. As the data above indicate, this is incorrect. At the very least PRO-based accounts will have to explain why PRO, though Case-marked, functions like an A-trace and not as an A'-trace. As PRO-based accounts currently stipulate the distributional

[14] See e.g. Lightfoot 1976.
[15] See e.g. Jaeggli 1980.
[16] Recall that in GB must sit in an ungoverned position and Case assignment must take place under government (see Chomsky 1981); hence, PRO is bound to be Caseless.
[17] See Boeckx 2000 on this point.
[18] See Boeckx, Hornstein and Nunes 2010:sections 2.5.1 and 5.4.

properties of OC PRO, these data indicate that a rather articulated stipulation will be required.[19]

The conclusion is that here the Duck Principle does tease apart the MTC from PRO-based accounts within minimalism. It is not the case, like we saw in section 3.1, that the special provisos required by PRO-based accounts can be incorporated by the MTC. The MTC simply doesn't need them! Thus, the Duck Principle in tandem with Occam's Razor implicate the MTC.

3.3. Adjunct control

Let us now consider adjunct control. Of course, adjunct control involves adjuncts and adjuncts are perennial troublemakers. They always challenge attempts towards unification as there are many different types, which require different heights for merge, etc. Our aim here is not to explore adjuncts in depth, but to consider a subset of adjuncts, the ones which trigger OC. Take the data in (21), for example.

(21) a. John_i said [that [Mary_k's brother]$_m$ left [after $\text{PRO}_{m/*i/*k/*w}$ eating a bagel]]
b. *John_i watched TV [while PRO_i ate a bagel]
c. John_i left before PRO_i singing and Bill_k did too
'… and Bill_k left before he_k/*John_i sang'
d. *John_i called Mary_k after [PRO_{i+k} criticizing each other]

(21a) shows that PRO in this configuration requires a local c-commanding antecedent, (21b) that PRO cannot sit in the subject position of a finite adjunct (in English), (21c) that PRO inside the adjunct triggers a sloppy reading under ellipsis, and (21d) that PRO does not permit a split antecedent. All of these properties, the reader may recall, describe both OC PRO in complement control *and A-traces* (cf. (2)-(7)).[20] In other words, in adjunct control configurations OC PRO also quacks like an A-trace.

It should be noted that the similarities go beyond the orthodox cases. As mentioned in section 3.1, Brazilian Portuguese allows both an A-trace (cf. (9)) and an OC PRO (cf. (8)) in the subject position of a finite complement clause. Unsurprisingly, in adjunct control configurations in Brazilian Portuguese, the null subject of a finite adjunct clause behaves like OC PRO, as illustrated in (22).[21]

(22) *Brazilian Portuguese:*
[O pai do João$_i$]$_k$ cumprimentou o Pedro$_m$ [quando $ec_{k/*i/*m/*w}$ entrou na sala]
the father of-the João greeted the Pedro when entered in-the room
'[John_i's father]$_k$ greeted Pedro_m when $\text{he}_{k/*i/*m/*w}$ entered the room.'

[19] It is worth noting that this parallel between PRO and A-traces and the contrast between both and A'-traces is not tied to how sandhi effects are to be properly analyzed. What is relevant is that *whatever* the etiology, A-traces and PRO are treated similarly and that both are distinguished from A'-traces. For another approach to these "contraction" effects, see Anderson 2005:72ff.

[20] There are several other properties that both complement control and adjunct control display. For fuller discussion, see Boeckx, Hornstein, and Nunes 2010:section 4.5.1.

[21] See Rodrigues 2004.

Another telling pattern is found in instances of interclausal epicene agreement in Romance, as discussed by Rodrigues (2004, 2007).[22] The word for 'victim' in Italian, for instance, is invariably [+feminine] regardless of whether it refers to males or females. Accordingly, in raising constructions like (23), for instance, the adjectival predicate takes the feminine form even in the context where a man has been hurt.

(23) *Italian* (Rodrigues 2004):
La vittima sembra essere **ferita/*ferito**
*the victim seems be injured-**FEM**/injured-**MASC***
'The victim seems to be injured.'

Interestingly, the agreement seen in (23) is replicated in both complement and adjunct control, as shown in (24), but, crucially, not in NOC, as shown in (25).

(24) *Italian* (Rodrigues 2004):
a. La vittima ha cercato di essere **trasferita/??trasferito**
*the victim had tried of be transferred-**FEM**/transferred-**MASC***
alla stazione di polizia di College Park
to-thestation of police of College Park
'The victim tried to be transferred to the police station at College Park.'
b. La vittima mori' dopo essere **stata trasportata /??stato**
*the victim died after be been.**FEM** brought.**FEM** been.**MASC***
trasportato all' ospedale.
*brought.**MASC** to-the hospital*
'The victim died after being brought to the hospital.'

(25) *Italian* (Rodrigues 2004):
La vittima ha detto che essere ***portata/portato*** alla stazione
*the victim has said that be brought-**FEM**/brought-**MASC** to-the station*
di polizia non era una buona idea
of police not was a good idea
'The victim said that being brought to the police station was not a good idea.'

As Rodrigues reasons, if the null subject inside the infinitival in (24) is an A-trace, it must pattern with the A-trace in the embedded subject position of the raising constructions in (23) and the agreement morphology on the embedded predicate must match the gender feature of the antecedent of the embedded subject. Again, this should be so independently of the specific analysis one assumes for inter-clausal agreement in standard raising constructions. (25), on the other hand, cannot be analyzed as involving an A-trace in the subject of the infinitival clause, as the infinitival is a subject island. Once (25) cannot be analyzed in terms of an A-trace, inter-clausal agreement is blocked and the embedded predicate takes an (arguably default) masculine form.

The data above pose very serious problems for PRO-based analyses that rely on Agree such as Landau's (2000, 2004). Crucially, subjects are CED islands and whatever accounts for CED effects should prevent inter-clausal agreement to license OC PRO.

[22] See also Rodrigues and Hornstein 2013.

This is a good result in the case of the subject island in (25), for example, but not in cases of adjunct control such as (21a), (21c), (22), and (24b), for instance. To deny that these sentences involve OC because they do not instantiate an Agree configuration (see e.g. Landau 2000:section 5.1) raises the mystifying question of why the grammar should require additional mechanisms that yield the same effects as the ones related to PROs of complement control *and A-traces*. Ignoring the Duck Principle in face of similarities between OC PROs in adjunct control and A-traces such as the ones illustrated above comes, we believe, at a considerable cost.

At first sight, the same kind of problem faced by Agree-based analyses of OC PRO with respect to adjunct control should also haunt the MTC. After all, adjuncts are islands for movement and therefore movement out of the adjunct island in adjunct control constructions should also yield a CED violation. This is actually true if we are referring to GB-style grammars, but not to grammars with a *minimalist* architecture. Let's consider why.

Within GB, D-Structure provides the computational system with a unique root tree and all the syntactic computations after D-Structure must operate within this single rooted syntactic object. Thus, if we find an adjunct island between a trace and its antecedent, the movement that gives rise to this configuration must have incurred a CED violation. So, within the GB model incorporating the CED there is no way to generate adjunct control via movement.

Minimalist theories, on the other hand, dispense with D-structure (as it is not an interface level) and syntactic trees are constructed in a step-by-step fashion through (possibly) interleaved applications of Merge and Move.[23] Furthermore, Chomsky (1995) has argued that the computational complexity of syntactic derivations can be substantially reduced if we assume the Extension Condition, which requires that projecting operations work at the root node. Thus, whereas in GB the structure *[$_{VP}$ [$_{DP}$ the boy] [$_{V'}$ saw her]]*, for example, is generated in one fell swoop as part of the D-Structure corresponding to the whole sentence, in minimalism it is built from the (simplified) numeration N in (26a) through several applications of Select and Merge, as sketched in (26b-h).

(26) a. N = {the$_1$, boy$_1$, saw$_1$, her$_1$}
b. *Select:* N'= {the$_1$, boy$_1$, **saw$_0$**, her$_1$}
K = **saw**
c. *Select:* N''= {the$_1$, boy$_1$, saw$_0$, **her$_0$**}
K = saw
L = **her**
d. *Merge:* M = [saw her]
e. *Select:* N'''= {the$_1$, **boy$_0$**, saw$_0$, her$_0$}
M = [saw her]
O = **boy**
f. *Select:* N''''= {**the$_0$**, boy$_0$, saw$_0$, her$_0$}
M = [saw her]
O = boy
P = **the**
g. *Merge:* M = [saw her]

[23] Or E-merge/I-merge, both applications of the single merge operation.

Q = [the boy]
h. *Merge:* VP = [[the boy] [saw her]]

Although the final result is the same in both the one fell swoop and the step-by-step derivations, there is a crucial difference in how this result is obtained: in the derivation sketched in (26) the computational system must be able to handle more than one root syntactic object at a time. This is in fact trivially true for the first steps of any syntactic derivation. Take the derivational step in (26c), for instance. Before *saw* and *her* merge, they are independent root syntactic objects. Moreover, given the Extension Condition, the derivation of complex subjects and complex adjuncts invariably demands that the computational system deal with more than one root syntactic object at a time. For instance, the Extension Condition prevents an alternative continuation of (26e) where *boy* first merges with *[saw her]*, yielding *[boy [saw her]]*, and later *the* merges with *boy* in a noncyclic manner; hence, *the* must be selected and merged with *boy* so that the resulting structure merges with *[saw her]* (cf. (26e-h)). Interestingly, in the derivational step in (26f) there are three different root syntactic objects available to the computational system.

Another relevant difference between GB and minimalism is the copy theory of movement, which reinterprets Move as the output of the interaction between the more basic operations Copy and Merge.[24] Under the copy theory, the derivation of a sentence such as (27) below, for instance, proceeds along the lines of (28), where the computational system creates a copy of *John*, merges it with the previously assembled TP, and deletes the lower copy in the phonological component. Again, notice that in a system that has Copy as a basic operation, the computational system must be able to handle more than one root syntactic object, namely, the copy newly created and the root syntactic object containing the replicated material (cf. (28b)).

(27) John was arrested.

(28) a. K = [$_{TP}$ was arrested John]
b. *Copy:* K = [$_{TP}$ was arrested **John^{i}**]
L = **John^{i}**
c. *Merge:* M = [$_{TP}$ John^{i} was arrested John^{i}]
d. *Delete:* P = [$_{TP}$ John^{i} was arrested ~~**John^{i}**~~]

What is relevant for our discussion is that if the computational system can operate with more than one root syntactic object at a time and if movement is understood as the interaction between the basic operations of Copy and Merge, "sideward movement" becomes a logical possibility within the system. That is, given two root syntactic objects K and L in (29), the computational system may copy α from K and merge it with L.[25]

[24] The difference between copies and occurrences is immaterial for present purposes (see Larson and Hornstein 2012 for relevant discussion). For concreteness, we will frame the following discussion in terms of copies, which will be annotated by superscripted indices. We could recast the discussion in terms of E/I-merge but we leave this translation as an exercise for the fastidious.

[25] The theoretical option of sideward movement as a licit application of Move/I-Merge was first mooted in Bobaljik 1995, Nunes 1995, Bobaljik and Brown 1997, and Uriagereka 1998. For developments,

(29) a. K = [… α …]
L = […]
b. *Copy*: K = [… **α^{i}** …]
L = […]
M = **α^{i}**
c. *Merge*: K = [… α^{i} …]
P = [α^{i} [$_{L}$ …]]

Terminological metaphors aside, note that there is no intrinsic difference between the "upward" movement seen in (28), for instance, and the "sideward" movement sketched in (30) with respect to the computational tools employed. In both cases, we have trivial applications of movement, viewed as Copy plus Merge. Sideward movement is therefore *not* a novel operation or a new species of movement. [26] This point is worth emphasizing, as it has been consistently misunderstood. The fact that α in (29) does not merge with the structure that contains the "source" of the copy, as opposed to *John* in (28), may have independent explanations. First, (28) differs from (29) in an obvious way: the copy of *John* in (28) has only one syntactic object to merge with, whereas the copy of α in (29) has two. But more importantly, it may be the case that Last Resort licenses merger of the copy of α in (29) with L but not with K.[27]

Bearing these differences between GB and minimalism in mind, the derivation of an adjunct construction such as (30) under a sideward movement approach should proceed along the lines of (31).[28]

(30) John$_{i}$ saw Mary after [PRO$_{i}$ eating lunch]

(31) a. VP PP
6 6
saw Mary after John eating lunch

b. *Copy + Merge:*
VP PP
6 6
John^{i}saw Mary after **John^{i}** eating lunch

c. *Merge:*
VP

---

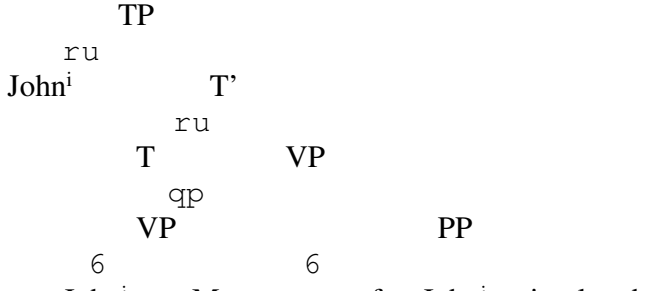
applications, and detailed discussion on how overgeneration is prevented, see e.g. Nunes 2001, 2004, 2012, Hornstein 2001, Hornstein and Nunes 2002, and Drummond 2009.

[26] As Chomsky is wont to say concerning E/I merge, *preventing* the option of sideward movement requires extra stipulations and hence, significant empirical motivation.

[27] Sideward movement is similarly compatible with an E/I-merge account, which dispenses with a Copy operation. Further, this view of things comes with a plausible cost accounting for why sideward movement is less preferred than upward movement and E-merge. Here, however, is not the place to elaborate on these, no doubt, cryptic comments.

[28] See Hornstein 1999, 2001 and Boeckx, Hornstein, and Nunes 2010.

d. *Copy + Merge:*



e. *Delete*: [TP John[i] [T' T [VP [VP ~~**John**~~[i] saw Mary] [PP after ~~**John**~~[i] eating lunch]]]]

Once VP and PP in (31a) are assembled, the computational system makes a copy of *John* from PP and merges it with VP (cf. (31b)), an instance of sideward movement that allows the external θ-role of the matrix clause to be discharged. After PP adjoins to VP (cf. (31c)) and the subject moves to [Spec, TP] (cf. (31d)), the lower copies of *John* are deleted in the phonological component (cf. (31e)) and the structure surfaces as (30).

Notice now that at the derivational step where *John* moves from PP to VP (cf. (31a-b)), PP is *not* an adjunct. Crucially, *adjunct* is not an absolute, but relational notion: a given expression is an *adjunct of* another. In (31a) PP is just a root syntactic object. Assuming that syntactic computations operate in a local fashion, the fact that later on PP will become an adjunct is irrelevant *at the derivational step where movement takes place*.[29] In other words, there is no island configuration in (31a) that would prevent copying. In fact, the copying and merger seen in (31a-b) is no different from the copying and merger found in licit instances of "upward movement" (cf. (28b-c)): in both circumstances, copying proceeds from a configuration that is not an island.

This approach correctly distinguishes licit cases of adjunct control like (30) from standard CED violations such as (32) below, for instance. Given that Extension Condition bars late adjunction, it must be the case that the PP in (32) merges with the matrix VP before the derivation builds the matrix TP. This being so, by the time the interrogative complementizer Q is merged, as sketched in (33), *which book* cannot move to check the strong feature of Q, as it is within an adjunct; hence the ungrammaticality of (32).

(32) *[[which book]$_i$ did [John [$_{VP}$ [$_{VP}$ call Mary] [$_{PP}$ after he read t_i]]]]

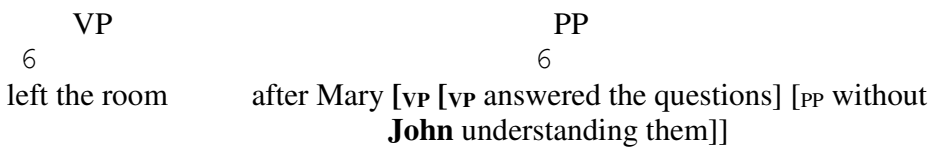
(33) [$_{CP}$ did+**Q** [John **[$_{VP}$ [$_{VP}$** call Mary] [$_{PP}$ after he read **[which book]**]]]]

[29] For relevant discussion, see e.g. Nunes and Uriagereka 2000, Hornstein 2001, Nunes 2001, 2004 and Hornstein and Nunes 2002.

Similar considerations apply to the illicit adjunct control construction in (34) below, with PRO taking the matrix subject as its antecedent. Under the relevant reading indicated by the brackets in (34), the PP headed by *without* is an adjunct of the VP headed by *answered.* The Extention Condition requires that these two constituents be merged before they become part of a larger structure. Thus, by the time the VP *left the room* is built, *John* is unable to undergo sideward movement to reach the matrix predicate, for it is inside an adjunct, as illustrated in (35).

(34) *[John$_i$ left the room [after Mary answered the questions without PRO$_i$ understanding them]]

(35)



Before exiting this subsection, let's briefly consider how to account for a very distinctive property of adjunct control, namely, that PRO must be controlled by the subject and not the object of the next higher clause, as illustrated in (36).

(36) John$_i$ saw Mary$_k$ after PRO$_{i/*k}$ eating lunch

Hornstein (1999, 2001) has argued that this subject-object asymmetry follows from economy computations. Consider the derivational step sketched in (37), for instance.

(37) N = {John$_0$, saw$_0$, Mary$_1$, after$_1$, eating$_0$, lunch$_0$}
K = [John eating lunch]
L = saw

In (37), *saw* must assign its internal θ-role and there are two potential candidates to receive it: *Mary*, which is still in the numeration, and *John* in the subject position of the gerundive clause. If *Mary* is selected and merged with *saw*, the derivation results in a subject control structure, after *John* undergoes sideward movement to [Spec,VP] (cf. (31a-b)). On the other hand, if *John* is copied and merged with *saw*, the derivation should give rise to an object control structure after *Mary* is plugged in as the external argument. Under the assumption that Merge is more economical than Move (see Chomsky 1995), the first option is enforced, yielding the subject-object asymmetry observed in (36). In sum, if economy independently restricts movement and sideward movement is just an instance of Move, then the restriction to subject control into adjuncts is what we expect (and find).[30]

[30] The result is actually a bit more robust than this. There are various ways of ensuring preference of merger over movement in these contexts. Under Nunes's (1995, 2001, 2004) system, for instance, the structure underlying (36) that could result in object control is independently excluded because it cannot be linearized, as the two copies of *Mary* do not form a chain and, accordingly, are not subject to deletion under Chain Reduction. For refinements and further discussion, see Nunes 2012:section 5.

To conclude, once we note that PROs in adjunct control configurations also behave like A-traces, the Duck Principle invites us to analyze them as members of the same species. However, this guiding principle leads to a dead end if we have in mind a GB-style grammar, for movement out of an adjunct structure that is part of the bigger D-Structure representation that feeds the computation necessarily results in a CED violation. Another dead end is met if the interpretation of OC PRO is to be licensed by an Agree operation, as in Landau's (2000, 2004) PRO-based account. Given that an intervening adjunct island should bar an application of Agree, the only way out is to say that adjunct control PRO walks and talks like complement control PRO, but this is due to something else. Of course, Occam's Razor disfavors this route. In contrast, by relying on the combination of some key aspects of the Minimalism – namely, the abandonment of D-Structure, cyclicity as determined by the Extension Condition, and the copy theory of movement – the MTC is able to analyze adjunct control PROs as A-traces, like their complement control cousins. Crucially, by exploring the possibility of sideward movement made available by these minimalist architectural properties, the MTC manages to capture the fact that adjunct control PRO is also a residue of movement without incorrectly ruling in standard adjunct island violations.

The result of comparing PRO-based accounts and the MTC with respect to the Duck Principle is much stronger here than the one regarding *wanna*-contraction. As mentioned in section 2.2, PRO-based accounts that assume that OC PRO is Case-marked can always invoke special provisos to account for why OC PRO behaves like Caseless traces. But here there is simply no such escape for PRO-based accounts other than denying that adjunct control involves control. By contrast, adjunct control brings no turbulence to the MTC.

3.4. Phonetic realization

Let us finally examine how the MTC and PRO-based accounts fare with respect to the PF side of the grammar. Consider the (simplified) structure each approach assigns to a control sentence such as (38), for instance:

(38) John tried to work hard.

(39) a. *MTC analysis:* [John^{i} tried [~~**John**~~^{**i**} to work hard]]
b. *PRO-based analysis:* [John$_{i}$ tried [**PRO$_{i}$** to work hard]]

Both approaches account for the fact that the null subject of the embedded clause is phonetically null. However, this by itself does not put them on equal footing. If you ask *why* the embedded subject is null, the nature of the answers offered are completely different. The MTC will respond that this issue reduces to the more general question of why copies are deleted in the phonological component: whatever mechanism is responsible for deleting copies in other instances of movement is also put to work in the case of (39a).[31] In other words, the phonetic nullness of the controllee in (39a) is not a construction-specific property pertaining to control, but the product of a grammatical process that is shared by standard instances of movement. In turn, the PRO-based account

[31] Say, for instance, Nunes's (1995, 2004) Chain Reduction operation, which is triggered by linearization considerations.

will have to say that it is an irreducible (i.e. non-explainable) lexical property of PRO that it does not have phonetic content. Of course, there is nothing incoherent in stipulating that PRO has no phonetic content, but all things being equal, we would be a step closer to theoretical nirvana if this followed from deeper features of the system.

Interestingly, the answers the MTC and PRO-based approaches provide to the issue of the phonetic content of OC PRO are not only different on conceptual grounds, but also differ in their empirical coverage. An increasing body of literature has been showing that deletion of lower copies (our traditional traces) is not the only possibility found in natural languages. One may find cases where lower copies are pronounced instead of the head of the chain and even cases where more than one copy is phonetically realized.[32] These findings are completely orthogonal if we are examining control under PRO-based approaches, but become quite relevant if control is to be analyzed as in the MTC. Our friend the Duck Principle is ready to point out that if these unusual cases of copy realization exist, we should expect comparable cases in the domain of control. Haddad and Potsdam (forthcoming) discuss this and argue that the full spectrum of options is indeed attested. In addition to the familiar cases of forward control, where the controller (the highest copy) is phonetically realized (cf. (38)/(39a)), there are cases of backward control, where the controllee (a lower copy) is pronounced (cf. (40)), cases of alternating control where either the controller or the controllee is pronounced (cf. (41)), and cases of copy control, where both controller and controllee are phonetically realized (cf. (42)):[33]

(40) *Tsez (Polinsky and Potsdam 2006):*
~~kid~~ [**kid**-bā čorpa bod-a] y-oqsi.
girl.ABS girl-ERG soup.ABS make-INF II-began
'The girl began to make soup.'

(41) *Greek (Alexiadou, Anagnostopoulou, and Marchis 2010):*
(O Janis) emathe **(o Janis)** na pezi **(o Janis)** kithara **(o Janis)**
John-NOM learn-3SG John-NOM SUBJ play-3SG John-NOM guitar John-NOM
'John learned to play the guitar.'

(42) *San Lucas Quiaviní Zapotec (Lee 2003):*
a. R-cààa'z **Gye'eihlly** g-auh **Gye'eihlly** bxaady.
HAB-want Mike IRR-eat Mike grasshopper
'Mike wants to eat grasshopper.'
b. B-quìi'lly bxuuhahz **Gye'eihlly** ch-iia **Gye'eihlly** scweel.
PERF-persuade priest Mike IRR-go Mike school
'The priest persuaded Mike to go to school'
c. B-ìi'lly-ga' **Gye'eihlly** zi'cygàa' nih cay-uhny **Gye'eihlly** zèèiny.
PERF-sing-also Mike while that PROG-do Mike work
'Mike sang while he worked.'

[32] See e.g. Nunes (1999, 2004, 2011), Bošković and Nunes 2007, the collection of papers in Corver and Nunes 2007, and references therein.
[33] See Haddad and Potsman forthcoming for additional data, references, and more detailed discussion.

Moreover, one may even find the same kinds of restrictions that play a role in allowing or precluding unusual outputs of chain realization operating in these less familiar control constructions. Take contrast in (43) below, for instance. Given that Romanian is a multiple *wh*-fronting language, the expected pattern should be (43a) and not (43b). Bošković (2002) argues that appearances here are misleading and the *wh*-object in (43b) does undergo *wh*-fronting; however, a language specific PF constraint banning adjacent homophonous *wh*-phrases prevents the higher copy of the moved object from being realized and forces the pronunciation of the lower copy instead, as sketched in (44).

(43) *Romanian* (Bošković 2002)
a. ***Ce** **ce** precede?
what *what precedes*
b. **Ce** precede **ce**?
what precedes *what*
'What precedes what?'

(44) [ce ~~**ce**~~^{i} precede **ce**^{i}]

The same type of reasoning is used by Fujii (2006) to account for the contrast in the object control constructions in (45) below, where the controller is realized with nominative and not with accusative Case. Fujii argues that if the highest copy of the chain headed by *John* were realized, there would arise a violation of the Double-*o* Constraint in Japanese, which bans two instances of accusative marked expressions in the same VP domain; hence, the unacceptability of (45a). In (45b), on the other, a lower copy of *John* is realized (as nominative) instead of the head of the chain and the Double-*o* Constraint is circumvented, as sketched in (46).

(45) *Japanese (Fujii 2006):*
a. ??Taro-wa **John-o** [siken-ni too-ru-no]-o tetudat-ta
Taro-TOP John-ACC exam-DAT pass-PRS-C_{NO}]-ACC assisted
b. Taro-wa [**John-ga** siken-ni too-ru-no]-o tetudat-ta
Taro-TOP John-NOM exam-DAT pass-PRS-C_{NO}]-ACC assisted
'Taro assisted John to pass the exam.'

(46) [Taro-wa ~~**John**^{i} **-o**~~ [**John**^{i}**-ga** siken-ni too-ru-no]-o tetudat-ta]

As for cases with pronunciation of multiple copies, a common restriction is that the more morphologically complex a given copy is, the less likely it is for it to be pronounced more than once.[34] This is illustrated in (47), for instance, which shows that *wh*-copying constructions in German may allow pronunciation of multiple copies of simplex *wh*-element such as *wen*, but not of full phrases such as *wessen Buch*.

(47) *German (McDaniel 1986):*
a. ***Wen*** glaubt Hans ***wen*** Jakob gesehen hat?

[34] See Nunes 1999, 2004 for an account of this restriction.

whom thinks Hans *wen* Jakob seen has
'Who does Hans think Jakob saw?'

b. ***Wessen Buch** glaubst du ***wessen Buch*** Hans liest?
whose book think you *whose book* Hans reads
'Whose book do you think Hans is reading?'

Analogously, languages that allow copy control exhibit similar restrictions on how morphologically encumbered copies can be. Thus, although San Lucas Quiaviní Zapotec allows copy control with only a name (cf. (42)), it rules out copy control constructions like the ones in (48a), which involves a quantifier phrase, or (48b), which involves an anaphoric possessor.[35]

(48) *San Lucas Quiaviní Zapotec (Lee 2003):*

a. ***Yra'ta' zhyàa'p** r-cààa'z g-ahcnèe' **yra'ta'zhyàa'p** Lia Paamm.
every girl HAB-want IRR-help every girl FEM Pam
'Every girl wants to help Pam.'

b. *R-e'ihpyGye'eihlly **behts-ni'** g-a'uh **behts-ni'** bx:àady.
HAB-tell Mike brother-REFL.POSS IRR-eat brother-REFL.POS grasshopper
'Mike told his brother to eat grasshoppers.'

In sum, if OC is a residue of movement, as advocated by the MTC, and if movement is to be understood in terms of the copy theory, as minimalism does, the Duck Principle leads us to expect that the full range of options available for copy pronunciation in standard movement operations should also be available in the case of control. Haddad and Potsdam provide substantial evidence that this expectation is realized.[36]

This line of reasoning has one important implication. There is really no way of combining backward control or copy control together with PRO-based accounts. To account for backward control, PRO-based theories would require base generating OC PRO in a position c-commanding its antecedent, as sketched in (49) below. However, this should lead to a violation of Principle C and thus should be impossible.

(49) $[PRO_1\ V\ [DP_1\ VP]]$

As for copy control, PRO-based approaches would have to incorporate rules that copy phonological matrices from antecedents to PRO.[37] Such rules are conceivable, but if not treated gingerly, would appear to collapse into complex versions of MTC under the copy

[35] Hornstein, Boeckx, and Nunes 2008 for details and further discussion.

[36] One more expectation: just as there are cases of backward control and copy control, we should expect to find cases of backward raising and copy raising. As Polinsky and Potsdam (2006, 2012) discuss in detail, this expectation is also met. See their papers for data, arguments, and references.

[37] Another option would be to allow the numeration to contain two copies of the controlled expression with some marked dependency between them. The problem then would be to explain why these do not induce a Principle C effect analogous to the ones found in (i):

(i) a.*$John_i$ managed for $John_i$ to win.
b. *$John_i$ wants $John_i$ to win.

theory. At any rate, as we hope to have made clear, the existence of cases of backward and copy control provides an interesting *novel* kind of evidence for the MTC and against PRO-based accounts.[38]

3.5. Wrapping up

Let us take stock in view of the discussion entertained above. Any adequate theory of control must:

(50) (i) specify the kinds of control structures that are made available by UG and explain how and why they differ;
(ii) correctly describe the configurational properties of control, accounting for the positions that the controller and the controllee can occupy;
(iii) account for the interpretation of the controllee, explaining how the antecedent of the controllee is determined and specifying what kind of anaphoric relation obtains between the controllee and its antecedent; and
(iv) specify what is the place of the controllee among the inventory of grammatical formatives provided by UG.

With respect to (50i), the MTC divides control into those cases parasitic on A-movement (OC) and those that are not (NOC). If "PRO" is a link in a well-formed A-chain, we have OC and "PRO" must have a local c-commanding antecedent, for example (cf. (2)-(4)). Otherwise, we have NOC. In other words, NOC acts like a pronominal relation not subject to the strict restrictions characteristic of A-chains.[39] In (51) below, for instance, we see that NOC PRO may have no antecedent (cf. (51a)), a nonlocal one (cf. (51b)), or a non c-commanding one (cf. (51c)). This is what we expect as the PROs in (51) sit within subject islands and so movement is impossible and no chain can relate the PRO within the subject gerund to any position outside it higher up.

(51) a. It is believed that [PRO_{arb} washing oneself once a week] is hygienic.
b. $John_1$ thinks that Mary said that [PRO_1 shaving himself] is vital.
c. $John_1$'s friends believe that [PRO_1 keeping himself under control] is vital.

The MTC also offers a straightforward answer to (50ii), in particular with respect to the distribution of "PRO." If "PRO" is actually a residue of movement, then we expect it to appear where A-traces (i.e. deleted copies in A-chains) are licit and to exhibit the properties that A-traces generally manifest. In languages like English, this coincides with caseless positions. Moreover, if minimality regulates movement, then we expect that no c-commanding DP can intervene between links of the OC chain. This implies that in the more familiar cases of forward control, OC "PRO" is phonetically null (being an A-trace), that it must be the highest DP of its clause, and that its antecedent must be the "closest" available DP. In addition, language specific rules may also trigger the

[38] Bacwkard control and copy control are also problematic for Manzini and Roussou's (2000) movement approach, according to which the "controller" is merged where it appears and attracts features of the controlled predicate.
[39] We have analyzed NOC PRO as essentially a phonetically null pronoun. Space limitations bar further elaboration. See Boeckx, Hornstein, and Nunes 2010:chapter 6 for discussion.

pronunciation of a lower copy instead of the head of the chain or pronunciation of more than one copy, yielding backward control and copy control, respectively.

As for (50iii), the MTC provides a precise answer to the issue of identifying which DP can serve as the antecedent of OC PRO:[40] the antecedent is the head of the A-chain of which OC PRO is a link. Given standard requirements on A-chain, this implies that controller selection under the MTC will comply with Rosenbaum's (1970) Principle of Minimal Distance and pick the closest c-commanding DP as the antecedent for PRO. In a subject control construction such as (52), for example, *Mary* must be the antecedent for PRO as it is the closest DP and movement of *John* from the position of PRO across *Mary* violates minimality.

(52) [John$_1$ expects [Mary$_2$ to try [PRO$_{2/*1}$ to wash]]]

Finally, with respect to (50iv), the MTC takes OC PRO to be not a lexical item with idiosyncratic properties, but a garden-variety trace of movement. So, whatever properties one may ascribe to PRO, they should be reduced to properties associated with movement operations.

It is worth observing that the MTC is the *only* current approach to control that derives the answers to the issues in (50) from more general grammatical principles. All PRO-based theories end up stipulating the properties to be captured in the guise of lexical features. Take the null Case and the Agree-based accounts, for concreteness. In Chomsky and Lasnik (1993), the distribution of PRO is tied to assignment of null Case. However, null Case is carried exclusively by the T^0 found in control clauses, and it is a Case that only PRO can realize. In turn, in Landau's (2004) approach, the distribution and interpretation of PRO are ultimately related to his assignment of [+R] and [-R] features to functional categories, where [+R] and [-R], when associated with a DP, are meant to indicate whether or not it may support independent reference. To the extent that they succeed, this type of account can *track* the distribution and interpretation of OC PRO, but does not *explain* why OC PRO has this specific distribution and interpretation and not others.[41]

4. Further Architectural Issues

4.1. The elimination of D-Structure

The MTC rests on one key assumption, namely, that movement into θ-positions is grammatically viable. In other words, the MTC is at odds with D-Structure. D-Structure, recall, is the syntactic level where all and only θ-relations are coded. It is also the input to all transformation processes (e.g. movement). Together, these two properties (i) prohibit movement into θ-positions and (ii) require that all argument DPs begin their derivational lives in θ-positions. The MTC is clearly incompatible with (i) and thus its theoretical viability requires the elimination of D-Structure as a grammatical level. As disposing of D-Structure (a methodologically unwelcome grammar-internal level) is a central architectural feature of the Minimalist Program, there exists a very tight conceptual connection between the Minimalist Program and the MTC. Not only does MTC imply the

[40] Recall, there is no theory of antecedent selection for NOC PRO as it does not require an antecedent.

[41] Landau (2004:842) in fact describes his R-assignment rule as an "honest stipulation" that plays the role of Case in previous models.

absence of D-Structure, but the absence of D-Structure is sufficient for the MTC given standard ancillary assumptions. Specifically, once D-Structure is eliminated as a grammatical level, nothing *prohibits* movement into θ-positions. Thus, eliminating D-Structure is both a necessary and sufficient condition for the MTC. Thus, to the extent that the elimination of D-Structure is a central feature of Minimalist Program, the MTC is quintessentially minimalist. If this is correct, the reader may be asking, why has this not been observed previously?

The main reason is that eliminating D-Structure does not necessarily imply removing all of D-Structure conditions from the grammar. Here's some Whig history: Chomsky's (1993) argument against D-Structure was actually quite narrowly focused. It only dealt with one of its properties, namely, that it is the input to the transformational component, thus preceding all movement operations. Chomsky (1993) describes this property of D-Structure in terms of *Satisfy*, an "all-at-once" operation that selects an array of items from the lexicon, arranges them in the X'-format, and presents the result to the computational system. Chomsky argues that Satisfy must be dispensed with and grammars must adopt generalized transformations that allow derivations to interleave operations akin to lexical insertion with operations akin to movement. This idea has been incorporated into the minimalist doctrine and was in fact the guiding intuition behind sideward movement, as seen in section 3.3. Recall that once generalized operations are resorted to, the system must be able to deal with more than one root syntactic object at a time; furthermore, once lexical insertion and movement are allowed to intersperse, a given expression may move from one root syntactic object to another before further lexical insertion proceeds.

The other defining property of D-Structure, namely, that it is the level where "pure GF-θ" is represented, was actually retained, but took another form. It was converted into the ban on movement into θ-role positions (Chomsky 1995:section 4.6) or the principle stating that "pure Merge in θ-positions is required of (and restricted to) arguments", where "*[p]ure* Merge is Merge that is not part of Move" (Chomsky 2000:103). However, neither translation of the "base-properties" of D-Structure fits snugly with other theoretical assumptions internal to the Minimalist Program. The most flagrant oddity in this revamping of D-Structure regards Merge. An unavoidable assumption within the system once Satisfy was dropped is that not only merger but also movement is a structuring building operation. In other words, Move must involve Merge as one of its components (cf. Chomsky's 2000 definition of pure Merge cited above) or is just another instantiation of Merge (cf. Chomsky's 2004 internal and external Merge). Now, if "pure"/"external" Merge is independently able to license θ-relations, why does it lose its powers when it is part of/related to movement? Whichever tack one takes, the prior differentiation between Move and Merge is conceptually difficult to retain and, correspondingly we believe, the prohibition against movement into θ-positions becomes theoretically awkward to enforce. There seems to be no reason for *why* this difference should exist if D-Structure does not. Thus, on both methodological and theory-internal grounds, we believe that there is every reason to retain the methodologically superior option (the complete elimination of D-Structure and its properties) that underwrites the MTC.

Before we leave this discussion, it should be observed that the residue of D-Structure clothed as a ban on movement into θ-positions or the requirement that

arguments can only receive a θ-role in their first merge has also been put into empirical service in the account of contrasts such as the one in (53).

(53) a. *John expected [*t* to be [someone in the room]]
b. John expected [someone to be [*t* in the room]]

The EPP-feature of the embedded T is checked after insertion of *John* in (53a) and movement of *someone* in (53b). Given a Merge-over-Move approach, (53a) should trump (53b) *if* they were both convergent. Chomsky (1995) proposes that (53a) does not converge because *John* cannot receive the external θ-role of *expected* by moving to its Spec. Once (53a) crashes, it does not compete for economy purposes with the convergent derivation of (53b), where *John* gets a θ-role when it is first merged.

Note however that the contrast in (53) can also be derived if *someone* cannot have its Case checked by the matrix verb in (53a) due to the intervention of the trace of *John* (see Nunes 1995, 2004)[42] or if nonfinite clauses do not have TP specifiers (see Castillo, Drury, and Grohmann 1999 and Epstein and Seeley 2006). In other words, it is not obvious that we are forced to resuscitate D-structure restrictions in order to account for data like (53).

In sum, in Chomsky 1993, the elimination of D-Structure is only partial. The MTC requires that it be complete: not only must Satisfy be rejected, but the segregation of functions between lexical insertion and movement (the first being designated to satisfy θ-relations, the latter to satisfy all the other grammatical dependencies) should be given up as well.

4.2. The nature of PRO

Generative grammar has generally analyzed control properties as grammatical by-products for good reasons. Only in this way are its properties amenable to explanation. For example, in the Standard Theory, PRO is a phonetic gap that results from deletion under Equi. Why on this view is "PRO" phonetically null? Because it is the product of a deletion operation. Why is OC PRO anaphoric? Because deletion here is deletion under identity. Taking "PRO" to be the product of a grammatical deletion operation thus allows for an explanation of its semantic and phonetic properties.

The same holds for the EST conception of PRO as [$_{DP}$ *e*]. This is a permissible grammatical option in a model that distinguishes between phrase structure rules and lexical insertion operations: a "PRO" is what the grammar generates when the DP phrase structure rule applies but is *not* followed by a lexical insertion operation. This analysis also provides an account for PRO's phonetic and semantic properties. It is phonetically null because it has no lexical content and that requires an antecedent because having no content it has no interpretation of its own. Once again, this analysis of "PRO" reflects the

[42] If *be* in (53a) assigns Case to *someone* (see e.g. Belletti 1988 and Lasnik 1995), the comparison with (53b) becomes irrelevant, for in (53b) *someone* could not have undergone A-movement if it had its Case deactivated by *be*. Moreover, if *someone* is Case-licensed by *be* in (53a), *John* can have its Case valued as accusative by the matrix verb while it is in the embedded subject position. But if that happens, *John* cannot undergo any further A-movement, regardless of whether or not the target is a θ-position.

view that control facts (should) directly follow from basic operations and organizing principles of grammars.[43]

So, how is PRO to be described in a minimalist setting? It can be a grammar-internal formative or a primitive lexical item. There is no third alternative. In particular, the Inclusiveness Condition forbids PRO from being a non-lexical expression inserted during the course of the derivation and bare phrase structure eliminates the option of identifying PRO as [$_{XP}$ *e*]. Let us then briefly examine each of the options available.

In consonance with the Duck Principle, the option explored by the MTC takes the similarities between PRO and traces to their logical conclusion: PROs are actually traces! In particular, PRO is what we call the A-trace of an element that has wandered into a θ-position. As copies replace traces in the Minimalist Program, PROs are accordingly reanalyzed as copies, with significant empirical gain, we saw in section 3.4. What is critical to note here is that within minimalism copies are perfectly well defined in consonance with bare phrase structure: a copy is either a lexical item or a phrase built from lexical items. Moreover, the properties of control structures are expected to derive from general principles of grammar, as control relations – like A-trace dependencies – are grammatical products formed by movement. So, following a venerable tradition, the MTC embodies the assumption that the properties of control configurations derive from (and so directly reflect) the underlying operations and principles of UG.

Under the option of treating PRO as a lexical item, PRO is in turn expected to behave like *the*, *dog*, *bring*, *this*, etc. That is, it lives in the lexicon and it can merge and move, just like any other lexical item or phrase. Notice that there are no problems with bare phrase structure on this conception because PRO functions like any other (nominal) expression drawn from the lexicon. However, it is worth considering for a moment how radical a departure this is from the classical conceptions of control.

Since the early 1980s, generative grammarians have assumed that constructions do not exist as grammatical primitives. The idea is that the fundamental principles of grammar operate independently of the lexical items that they manipulate. For example, relative clauses are not islands because they involve particular lexical heads or contain particular lexical items but because they instantiate particular structural dependencies. Likewise, topicalization, focalization, or relativization do not obey islands because they involve topic, focus, or relative heads, but because they all involve (A')-movement and movement is subject to island effects. In other words, grammatical operations and restrictions have the properties they do not because of the functional features of the "constructions" in which they apply, but because of the formal properties that these constructions instantiate. It is in this sense that constructions do not exist; they are not the fundamental units of syntactic analysis. The problem with treating "PRO" as a lexical item is that it amounts to analyzing control configurations as constructions: control properties follow from the unique properties (often sets of stipulated features) of the lexical item *PRO*, which defines the construction. In effect, the "control construction" directly reflects the idiosyncratic properties of a distinctive lexical item, rather than the basic operations and organization of the grammar. Landau's (2004) featural specification of PRO is a good example. What drives the requisite operations is PRO's feature make-up. And PRO has the features it does because of the control facts attested. Were the control facts different, all that would be required is a different feature make-up for PRO.

[43] See Boeckx, Honrstein, and Nunes 20120:sections 2.3 and 2.4 for further discussion.

So if one asks: why does PRO have these features and not others? The answer is: just because. It is a brute fact about the properties of PRO, not the reflections of the operations of the grammar.

Indeed, many (if not all) the properties of the "lexical" item PRO cannot even be identified independently of the grammar. PRO needs a local, c-commanding, syntactic antecedent and can only be licensed within (tense- or ϕ-) defective domains. How are these requirements to be stated in purely "lexical" terms? How can they be expressed except by adverting to grammars, their structures and their basic operations and principles? They cannot be. PRO's requirements are *grammatical* licensing requirements. Postulating PRO makes no sense except in a grammatical context. Its requirements are entirely grammar-internal. Even describing what they are requires reference to principles and operations of the grammar. Consequently, the analysis of PRO as a lexical element is subject to the minimalist antipathy towards constructionism inherited from GB and so renders PRO a suspect element, given minimalist standards. In the end, postulating lexical elements like PRO to account for the attested properties of control cannot yield explanations of these properties (descriptions yes, explanations no), for a lexical item like *PRO* codes as part of its content the very properties that are supposed to be explained. This is the (very high) cost of treating PRO as a lexical item.

5. Conclusion

The MTC is unique in unifying PRO's distribution and antecedent selection under a single mechanism. Precisely the same theory that accounts for where OC PRO can appear determines which of the potential DP antecedents controls it. OC PRO is a link in a well-formed A-chain. The head of the chain is PRO's antecedent. That's the theory and it fits the facts, to a very good first approximation. Thus, among the alternatives on offer at present, only the MTC has the capacity to move beyond description to explanation. The reason is that only the MTC evades constructionism and tries to derive the properties of control structures from general principles of grammar rather than from the special licensing conditions of a peculiar lexical item. These theoretical ambitions are thwarted if one assumes that PRO is a primitive lexical item. On this constructionist view, its special licensing requirements are simply lexical quirks.

As discussed above, there is also a very close conceptual connection between the Minimalist Program and the MTC. The elimination of D-Structure, which is one of the central tenets of the Minimalist Program, is also a necessary and sufficient condition for the MTC to be viable:

(54) MTC ↔ no D-Structure

This picture sharply contrasts with what is found with PRO-based approaches to control within minimalism. They do not rely on any distinctive minimalist assumptions and thus, though they might be compatible with the Minimalist Program, their theoretical apparatus (though not the technology used to express control dependencies) is largely independent of it. Moreover, their constructionist bias is quite at odds with the explanatory ideals of the Minimalist Program.

That said, one should not conclude that because the MTC fits well with the Minimalist Program that the MTC is correct. However, it does suggest that those with

minimalist aspirations should smile on the MTC and that the burden of proof must be with those that reject it. Furthermore, if the fit between the Minimalist Program and the MTC is as tight as we have suggested, then the evidentiary bar relevant to rejecting the MTC should be quite high. To put things differently, *if minimalism is on the right track,* then some version of the MTC must be correct, i.e. from a minimalist perspective, the MTC is everything it's quacked up to be!

Further reading list

Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes. 2010. *Control as Movement*. Cambridge: Cambridge University Press.

Drummond, Alex, Norbert Hornstein, and Dave Kush. 2011. Minimalist Construal: Two Approaches to A and B. In *The Oxford Handbook of Linguistic Minimalism*, C. Boeckx (ed.), Oxford University Press, 396-426.

Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.

Hornstein, Norbert and Maria Polinsky (eds.). 2010. *Movement Theory of Control.* Amsterdam/Philadelphia: John Benjamins.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22:811-877.

References

Alexiadou, Artemis, Elena Anagnostopoulou, Gianina Iordachioiaia, and Mihaela Marchis. 2010. No objection to backward control. In Norbert Hornstein and Maria Polinsky (eds.) *Movement Theory of Control*, 89-117. Amsterdam/Philadelphia: John Benjamins.

Anderson, Stephen. 2005. Aspects of the theory of clitics. Oxford: Oxford University Press.

Belletti, Adriana. 1988. The case of unnacusatives. *Linguistic Inquiry* 19:1-34.

Bobaljik, Jonathan. 1995. Morphosyntax: The Syntax of Verbal Inflection. Doctoral dissertation, Massachusetts Institute of Technology.

Bobaljik, Jonathan and Samuel Brown. 1997. Inter-arboreal Operations: Head-Movement and the Extension Requirement. *Linguistic Inquiry* 28: 345-356.

Boeckx, Cedric. 2000. A note on contraction. *Linguistic Inquiry* 31, 357 366.

Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes. 2008. Copy-reflexive and Copy-Control Constructions: A Movement Analysis. *Linguistic Variation Yearbook* 8: 61-99.

Boeckx, Cedric, Norbert Hornstein, and Jairo Nunes. 2010. *Control as Movement*. Cambridge: Cambridge University Press.

Bošković, Željko. 2002. On Multiple *Wh*-Fronting. *Linguistic Inquiry* 33: 351-383.

Bošković, Željko and Jairo Nunes. 2007. The Copy Theory of Movement: A View from PF, in N. Corver and J. Nunes (eds.), *The Copy Theory of Movement*. Amsterdam/Philadelphia: John Benjamins, 13-74.

Castillo, Juan Carlos, John Drury, and Kleanthes Grohmann. 1999. Merge over move and the Extended Porjection Principle. *University of Maryland Working Papers in Linguistics* 8:63-103.

Chomsky, Noam. 1977. On *wh*-movement. In P. W. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax*, New York, NY: Academic Press, 71-132.
Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The View from building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. by Kenneth Hale and Samuel Jay Keyser, 1-52. Cambridge, Mass.: MIT Press.
Chomsky, Noam. 1995. *The Minimalist Program.* MIT Press, Cambridge, Mass.
Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In *Step by step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89-155. MIT Press, Cambridge, Mass.
Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A life in* Language, ed. Michael Kenstowicz, 1-52. MIT Press, Cambridge, Mass.
Chomsky, Noam 2004. Beyond Explanatory Adequacy, in A. Belletti (ed.), *Structures and Beyond*. Oxford: Oxford University Press, 104-131.
Chomsky, Noam and Howard Lasnik. 1993. The theory of Principles and Parameters. In *Syntax: an international handbook of contemporary research*, ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506-569. Berlin/New York: Walter de Gruyter.
Corver, Norbert and Jairo Nunes (eds.). 2007. *The Copy Theory of Movement*. Amsterdam: John Benjamins.
Drummond, Alex. 2009. How Constrained is Sideward Movement? Generals Paper, University of Maryland, College Park.
Drummond, Alex, Norbert Hornstein, and Dave Kush. 2011. Minimalist Construal: Two Approaches to A and B. In *The Oxford Handbook of Linguistic Minimalism*, C. Boeckx (ed.), Oxford University Press, 396-426.
Ferreira, Marcelo. 2000. Argumentos Nulos em Português Brasileiro. M.A. thesis, Universidade Estadual de Campinas.
Ferreira, Marcelo. 2009. Null Subjects and Finite Control in Brazilian Portuguese. In *Minimalist Essays on Brazilian Portuguese Syntax*, ed. by Jairo Nunes, 17-49. John Benjamins, Amsterdam/Philadephia.
Fujii, Tomohiro. 2006. Some theoretical issues in Japanese control. Docotral dissertation, University of Maryland, College Park.
Fujii, Tomohiro. 2010. Split control and the Principle of Minimal Distance. In Norbert Hornstein and Maria Polinsky (eds.) *Movement Theory of Control*, 211-244. John Benjamins, Amsterdam/Philadelphia.
Haddad, Youssef A. and Eric Potsdam. Forthcoming. Linearizing the Control Relation: A Typology. In Theresa Biberauer and Ian Roberts (eds.). *Principles of Linearization*. Berlin: Mouton de Gruyter.
Hornstein, Norbert. 1999. Movement and Control. *Linguistic Inquiry* 30: 69-96.
Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.
Hornstein, Norbert and Jairo Nunes. 2002. On Asymmetries between Parasitic Gap and Across-the-Board Constructions. *Syntax* 5: 26-54.
Jaeggli, Osvaldo. 1980. Remarks on *to*-contraction. *Linguistic Inquiry* 11:239-245.
Landau, Idan. 2000. *Elements of Control. Structure and Meaning in Infinitival Constructions*. Dordrecht: Kluwer.

Landau, Idan. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22:811-877.
Lasnik, Howard. 1995. Case and expletives revisited. *Linguistic Inquiry* 26:615-633.
Lee, Felicia 2003. Anaphoric R-expressions as Bound Variables. *Syntax* 6:84-114.
Lidz, Jeff and William Idsardi. 1997. Chains and Phono-logical Form. *UPenn Working Papers in Linguistics* 8:109-125.
Lightfoot, David. 1976. Trace theory and twice-moved NPs. *Linguistic Inquiry* 7:559-582.
Manzini, Maria Rita and Anna Roussou. 2000. A Minimalist Theory of A-movement and Control. *Lingua* 110, pp. 409-447.
Martin, Roger. 2001. Null case and the distribution of PRO. *Linguistic Inquiry* 32:141-166.
Martins, Ana Maria and Jairo Nunes. 2005. Raising Issues in Brazilian and European Portuguese. *Journal of Portuguese Linguistics* 4:53-77.
Martins, Ana Maria and Jairo Nunes. 2010. Apparent Hyper-raising in Brazilian Portuguese: Agreement with Topics across a Finite CP. In *The Complementiser Phase: Subjects and Operators*, 142-163, ed. by Phoevos E. Panagiotidis. Oxford University Press, Oxford.
McDaniel, Dana. 1986. *Conditions on wh-chains*. Doctoral dissertation, CUNY.
Nunes, Jairo. 1995. The Copy Theory of Movement and Linearization of Chains in the Minimalist Program. Doctoral dissertation, University of Maryland at College Park.
Nunes, Jairo 1999. Linearization of Chains and Phonetic Realization of Chain Links, in S. D. Epstein and N. Hornstein (eds.), *Working Minimalism*. Cambridge, MA: MIT Press, 217-249.
Nunes, Jairo. 2001. Sideward Movement. *Linguistic Inquiry* 31: 303-344
Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, MA: MIT Press.
Nunes, Jairo. 2008. Inherent Case as a licensing condition for A-movement: The case of hyper-raising constructions in Brazilian Portuguese. *Journal of Portuguese Linguistics* 7:83-108.
Nunes, Jairo. 2010. Relativizing Minimality for A-movement: ϕ- and θ-relations. *Probus* 22:1-25. 2010.
Nunes, Jairo. 2011. The Copy Theory, in Cedric Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*. Oxford: Oxford University Press, 143-172.
Nunes, Jairo. 2012. Sideward Movement: Triggers, Timing, and Outputs. In M. Uribe-Etxebarria e V. Valmala (eds.): *Ways of Structure Building*. Oxford University Press, Oxford.
Nunes, Jairo and Juan Uriagereka. 2000. Cyclicity and Extraction Domains. *Syntax* 3: 20-43.
Pearson, Hazel. 2012. A semantic theory of partial control. Presented at NELS 43, CUNY, New York.
Polinsky, Maria and Eric Potsdam. 2006. Expanding the scope of control and raising. *Syntax* 9:171-192.
Polinsky, Maria and Eric Potsdam. 2012. Backward Raising. *Syntax* 15: 75-108.
Rodrigues, Cilene 2004. Impoverished Morphology and A-movement out of Case Domains. Doctoral dissertation, University of Maryland at College Park.

Rodrigues, Cilene. 2007. Agreement and flotation in partial and inverse partial control configurations. In W. D. Davis and S. Dubinsky (eds.): *New Horizons in the Analysis of Control and Raising*, 213-229. Springer, Dordrecht.

Rodrigues, Cilene and Norbert Hornstein. 2013. Epicene agreement and Inflected Infinitives when the data is "under control": A reply to Modesto (2010). To appear in *Syntax.*

Rosenbaum, P. S. 1970. A principle governing deletion in English sentential complementation. In R. A. Jacobs and P. S. Rosenbaum (eds.), *Readings in English Transformational Grammar*, Waltham, MA: Ginn and Company, 20-29.

Uriagereka, Juan (1998). *Rhyme and Reason: An Introduction to Minimalist Syntax*. Cambridge, MA: MIT Press.

Zwart, C. J.-W. 2002. Issues Relating to a Derivational Theory of Binding. In S. D. Epstein and T.D. Seely, (eds.) *Derivation and Explanation in the Minimalist Program*. Oxford: Blackwell, pp. 269-304.