Book Review

Hornstein, Norbert (2001), *Move! A Minimalist Theory of Construal*
Blackwell Publishers, Malden (Mass, USA) and Oxford (UK)
Generative Syntax series, volume 5, 248 p.

Michael Brody (1995) argues that "assuming that chains and Move α express the same kind of relation, a theory that contains both concepts is redundant". Brody's proposal is to eliminate movement from the grammar and encode dependencies through a chain formation algorithm. In *Move!*, Hornstein goes the opposite way: he claims that the grammar contains no construal processes, all the relations they were taken to express being encoded now through movement. More striking that the proposal itself is the way in which Hornstein tries to achieve it: not by postulating extra characteristics of movement, but simply by allowing it to operate in a freer fashion than is usual in standard Minimalism. For instance, he drops the assumptions that the moved element must c-command its launching site, and that movement out of islands is impossible. He still retains two important conditions, namely, that movement can only be triggered by feature checking needs (the so-called *virus theory*), and that it is more costly than direct merger (the *merge-over-move preference*).

The book opens with a discussion of the Minimalist Program: what its goals are, how we got to it from the GB model, where it is heading… At the very beginning of ch. 1, Hornstein tells us that *Move!* is "based on *one* version of the Minimalist Program", namely, his own version. Therefore, it is not surprising that when, later in the chapter, he lists the ten features that his ideal model of grammar should have, some of them are different from Chomsky's (and other people's) own conceptions:

1) *Hornstein's Decalogue for a Syntactic Theory:*
   I) There are only two grammatical levels: LF and PF.
   II) Case and Theta theories apply in different domains (functional vs. lexical), but their licensing mechanisms are the same (i.e., feature checking).
   III) Features are checked in configurations licensed by phrasal structure.
   IV) Movement is greedy: an element may only move to check its own features.
   V) Movement is Copy & Merge (in turn, this implies that movement is more costly because it requires one more operation).

VI) Overt instances of merger and movement abide by the Extension Condition: they always apply at the root of the tree (Chomsky 1993).

VII) Theta roles are features (contra Hale & Keyser 1993a).

VIII) Binding theory applies at LF.

IX) The Theta Criterion is not maintained in its traditional sense.

X) Movement is not attraction.

These assumptions are first put to work in the second chapter ("Movement and Control", a slightly revised version of Hornstein 1999) with the aim to eliminate both PRO and the control module. He argues that PRO is an element that does not really fit in standard minimalist theories, since its existence is entirely based on the existence of D-Structure. The reasoning is as follows: DS is the level at which all lexical (i.e., thematic) information is introduced prior to movement. The ban against movement into theta positions follows from this, since all these position are already occupied beforehand. Thus, in a control sentence, PRO is necessary to get both the matrix and the embedded verbs' external theta roles assigned to the same participant without violating the ban against movement into theta positions. But, as Hornstein notes, once DS is eliminated from the theory, nothing forces us to assume such a ban. In other words, it should be possible to analyse control sentences in the same way as raising sentences, namely, by moving the embedded subject to the matrix SpecIP. The difference between both constructions would rely on the assumption that control, but not raising, verbs assign an external theta role to their subjects. Given this reasoning, it is still possible to maintain some weaker version of the Theta Criterion (e.g., requiring that all theta roles be assigned, which is probably independently required anyway by Full Interpretation), but not the clause banning movement to a theta position. Furthermore, Hornstein argues, the relation between PRO and its controller displays several characteristics that fall out if we assume that movement is involved:

2) *Some characteristics of control sentences:*
   a. PRO must have a controller (*It was expected* [*PRO to shave himself*]).
   b. The controller must be local (*John thinks* [*that it was expected* [*PRO to shave himself*]]).
   c. The controller must c-command PRO (*[*John's campaign*] *expects* [*PRO to shave himself*]).
   d. PRO can only have a sloppy reading under ellipsis (*John expects* [*PRO to win*] *and Bill does too* (=Bill win)).
   e. PRO cannot have a split controller (*John$_i$ told Mary$_j$* [*PRO$_{i+j}$ to leave together*]).

There are various things to note with in this approach. First of all, the requirement that the controller must c-command PRO (2c) falls out under the usual requirement that a moved element must c-command its launching site. It does not fall out in Hornstein's system, though, since, as we shall see,

he allows movement to non c-commanding positions as a crucial part of his analysis of parasitic gaps. Second, he is forced to treat theta roles as features. Recall that he assumes that movement can only be triggered by feature checking. Therefore, movement into theta positions can only take place if theta roles are features. This seems a somewhat strange conclusion to me since, while other formal features (Case, Tense, Agreement, Focus, Wh-…) are morphologically –or prosodically- expressed in at least some languages, theta roles are not. Moreover, given Greed (point IV in the Decalogue) the only option is that theta features reside in the DP itself, and not in the verb "assigning" them. This implies that any given DP must enter the numeration already "marked" to occupy a certain thematic position, whereas, on the other hand, this information does not seem to be available in the Lexicon.

Third, (2) listed the characteristics of obligatory control (OC) sentences. There also exist, though, non-obligatory control (NOC) sentences, which do not show any of the characteristics in (2). For these constructions, Hornstein argues that the embedded subject position is filled by a null pronominal similar to Romance *pro*. *Pro* is not an element present in the numeration, and it is inserted as a last resort operation to salvage sentences that would otherwise crash because movement from the embedded to the matrix IP is banned. This solution raises several problems, though. To begin with, Hornstein does not explain why *pro* cannot be inserted in the sentences in (2) to render them grammatical. Also, the dependency between *pro* and its antecedent is not mediated by movement, so is not easily expressible (if at all) in Hornstein's system. Finally, it is simply not true that OC sentences show the characteristics listed in (2) whereas NOC ones don't. San Martín (2002), following Landau (1999), argues that the fundamental difference between OC and NOC is that the former never allows an ambiguous reading of PRO –it must always be co-indexed with its antecedent. They give examples of irrefutable OC where a controller in a phrasal matrix object fails to c-command PRO (3a), or where split antecedents are allowed (3b, c). Hornstein obviates these exceptions.

3) a. It spoiled [Mary's$_i$ mood] [PRO$_{i/*arb}$ to listen to the news]
   b. John$_i$ promised his son$_i$ [PRO$_{i+j}$ to go the football game together]
   c. John$_i$ persuaded Mary$_i$ [PRO$_{i+j}$ to kiss in the library]

Chapter 3 ("Adjunct Control and Parasitic Gaps") tries to extend the movement approach to adjunct control and parasitic gaps. This chapter is mainly based on ideas that stem from Nunes (1995), as Hornstein himself acknowledges. The kind of sentences he deals with here are exemplified below:

4) a. John$_i$ saw Mary [before PRO$_i$ leaving the party]   *(adjunct control)*
   b. Which book did you read *t* [before Fred reviewed *t*]?   *(parasitic gap)*

The derivation of both sentences is similar. It involves constructing the adjuncts in parallel to the main clauses. After the adjuncts are completed, *John* and *which book* move out of the adjunct and attach to the main verb, still unconnected to the adjunct, as in (5). This is what Hornstein refers to as *sideward movement*, after Nunes' terminology (Bobaljik & Brown 1997 call it *interarboreal*, which I believe is a much less misleading term. Anyway, I will stick to *sideward* in this review). Later, the adjunct merges with the main clause and the derivation proceeds as usual.

5)   a. [VP John [saw Mary]]   [before John leaving the party]

   b. [VP read [which book]]   [before Fred reviewed which book]

Note that there is one difference between the two constructions: in the case of adjunct control (5a), *John* enters the main clause as the specifier of VP; in the parasitic gap example (5b), *which book* moves to the complement position of the verb. In the former case, this derivation could be enforced simply by appealing to the merge-over-move preference: *saw* requires a nominal complement. The two possible options are merging *Mary* directly from the numeration, or moving *John* out of the adjunct. This last option is blocked, though, since the former is more economical. As for (5b), Hornstein acknowledges it is a flagrant violation of the economy condition that yields (5a). He claims, though, that this derivation is allowed because, otherwise, the "real" gap would be in a position where it would c-command the parasitic gap. Given that traces of A-bar movement are subject to Condition C, this configuration would result in a Binding Theory violation. Nonetheless, I cannot see how this result can be achieved without resorting to look-ahead, though this problem can be minimized by limiting look-ahead to apply only within phases (Jairo Nunes, p.c.).

One more loophole in Hornstein's argumentation is that, at several points in the book, he argues that considerations of economy of movement are not at stake in the case of sideward movement. Adopting a Relativised Minimality-like view of economy of movement –roughly, don't move x past a place where x could have landed-, sideward movement is perfectly legal. The reason is that, since we are moving to a different tree, the material in the tree containing the launching site does not count for RM purposes. Following this reasoning, it would also be possible to make a copy of a non-subject DP in the adjunct in (5a) and move it to the matrix subject position. In other words, Hornstein's system predicts that (6a) should be grammatical:

6)   a. *John$_i$ saw Mary [before Fred gave the book to t_i]
   b. [VP John [saw Mary]]   [before Fred gave the book to John]

In the derivation of sentences like (5b), Hornstein assumes that *which book* must raise to the SpecCP of the adjunct clause before undergoing

sideward movement (though I cannot see why this step should be necessary, once sideward movement is allowed), Crucially, he does not assume anything similar for adjunct control sentences. One could try to do away with this split and assume that *John* in (5a) also has to go through the adjunct SpecCP. In this way, (6) could be excluded on the grounds that there is a suitable element, namely *Fred*, closer to SpecCP. Therefore, movement of *John* is blocked by economy considerations. I do not think this hypothesis could be workable, though, because –assuming one single SpecCP slot, which seems to be necessary anyway in order to account for island effects-, it would block the derivation of the one construction that Hornstein does not discuss in the chapter: the "classical" sentences involving both adjunct control and a parasitic gap (*Which book did you file t* [*before PRO reading t*]). This problem does not appear in Nunes (2001) analysis of parasitic gaps, since he adopts an "old-fashioned" approach to control, so only *which book* would be competing for the SpecCP slot. For the same reason, it does not appear either in the rest of the constructions dealt with in this chapter (noun-complement sentences, relative clauses, *tough-* constructions and purpose clauses), since they involve movement of a single element out of the adjunct.

Chapter 4 ("Attract and Sideward Movement") consists of a short discussion of the mechanisms of movement, specifically, an argument against an Attract operation. Assuming the existence of Greed, it can never be the case that the target of movement also contains the trigger, that is, no element can attract another. Also, given the existence of sideward movement, an attraction-based view of movement would also be problematic, since attract seems to be irreducibly based on c-command, i.e., an element can only attract constituents within its c-command domain. But once we acknowledge the possibility of moving to non c-commanding positions, an operation like attraction does not make much sense.

These are theory-internal reasons against Attract. To give them some empirical support, Hornstein focuses on a domain in which, allegedly, it makes sense to dissociate movement from attraction, namely, Superiority (non)-violations in multiple wh- constructions. In an attraction-based system, sentences like *What did who see?* Would be ruled out because C has overlooked the closest available wh- word and has attracted a more distant one. In order to explain Superiority effects without recurring to attraction, Hornstein argues that they actually are instances of weak cross-over, as already claimed in Hornstein (1995). The idea is that wh- in situ is interpreted functionally, rather than via movement. More specifically, "a pronoun cannot be linked to a wh- trace or variable to its right", where *linked* means, roughly, that all wh- words of these constructions bear the same index –though this does not mean they are co-referential. Functionally interpreted wh- words are equivalent to pronouns bound by quantifiers. Given this hypothesis, the two paradigms below can receive a unified account:

7)   a. Everyone$_i$ loves his$_i$ mother.
　　 b. *Who$_i$ does his$_i$ mother love t_i?

8) a. Who$_i$ bought what$_i$?
   b. *What$_i$ did who$_i$ buy t_i?

Moreover, Hornstein argues, this hypothesis allows us to account for sentences where the presence of a third wh- word salvages what otherwise would have been a Superiority violation, as in (9):

9) a. *John wonder what who saw t then.
   b. John wonders what who saw t when.

Example (9a) can be ruled out in the same way as (7b) and (8b). For (9b), Hornstein proposes a representation in which the trace of *what* is linked to the right to *when*, and *who* is linked not to the trace of *what*, but to *when*. Thus, we can satisfy the condition that no pronoun/wh- in situ can be linked to a wh- trace/variable to its right. Hornstein argues that this effect cannot be achieved in an attraction-based theory of movement, since such a system would predict that Superiority would be violated in both sentences in (9). The reason is that, if the embedded C must always attract the closest wh- phrase it c-commands, then in no case can *what* move over *who*, independently of whether there is a third wh- word or not.

There are two points to comment on with regard to this hypothesis. The first one is that the notion of linking implies establishing a relation between two elements without resorting to movement. In fact, Hornstein himself makes it clear that any analysis based on movement would predict (9b) should be ungrammatical. It is not clear at all, then, how this fact can be reconciled with the main thesis of the book. Second, it must be stipulated that the linking operation is necessarily intransitive –though this conclusion is already present in Higginbotham's (1983) theory of linking. That is, if A and B are linked, and B and C are linked as well, it must not be the case that A and C are linked. Otherwise, in (b) we would get a situation in which *who* would be linked to the trace of *what*, rendering the sentence ungrammatical.

On a different level, the rejection of Attract as the proper way to think of movement has the interesting implication that the formal inadequacy driving movement must always be present in the moved element, rather than in the target. For instance, in the case of wh- movement, the wh- word would contain an uninterpretable wh- feature that would be checked off via movement to SpecCP. This contrast with other approaches (f.i., Boskovic 1999, or Pesetsky & Torrego 2001) that propose the opposite view of this issue: it is C that contains an uninterpretable wh- feature, which is eliminated by movement of a suitable wh- word.

Chapter 5 ("Is the Binding Theory Necessary?") tries to do away with the *Barriers*-style approach to binding. Hornstein argues that this approach is suspicious from the minimalist viewpoint because it makes use of the notion of "domain". First, domains for binding are usually defined in terms of government, a notion that has been eliminated in the Minimalist Program. Second, the Minimalist Program already has domains, though they apply to movement operations. Hornstein argues that a second kind of domain

applying to binding would be redundant and, therefore, unnecessary. This last point is further supported by the observation by Chomsky (1986, 1993) and Lebeaux (1983) that the locality conditions on anaphor licensing are identical to those on movement.

Hornstein takes up this idea by Chomsky and Lebeaux –that anaphors move up to their antecedents at LF- and tries to stretch it. He observes that the relation between an anaphor and its antecedent is the same one as between PRO and its controller, as sketched in (2). From this parallelism, he concludes that anaphors are the residues of movement of their antecedents:

10) *The parallelism between PRO and anaphors.*
   a. Anaphors need an antecedent (*John arrested herself*)
   b. The antecedent must be local (*John thinks that Mary arrested himself*).
   c. The antecedent must c-command the anaphor (*John's mother arrested himself*).
   d. Anaphors can only have a sloppy reading under ellipsis (*John likes himself, and Bill does too* (=like Bill)).
   e. Anaphors cannot have split antecedents (*John told Mary about themselves*).

The question at this point is, if PRO and anaphors are essentially the same thing (i.e., NP traces), why is it that anaphors are pronounced but PRO is not? Hornstein attributes this asymmetry to a difference in the case properties of the trace. Anaphors occupy a case-marked position, whereas PRO does not. This approach raises two problems: first, in a sentence like *John likes himself*, if the anaphor is actually a copy of the antecedent, why is it spelled out as *himself* rather that *John*? Second, if the upper copy of *John* has nominative case features, how can it check the accusative case of *likes*? Let us start with the latter question. Hornstein argues that it is *self* that actually bears the accusative case features. Thus, the derivation of the sentence would be like this:

11) [$_{IP}$ John$_i$ [$_{vP}$ self$_j$ [$_{VP}$ likes [[John$_i$] self$_j$]]]]

John raises all the way up to SpecIP to check nominative case. *Self* raises covertly to Spec$v$P to check accusative case (notice that this implies having a second derivation for LF). The only copy of *John* that survives linearisation is the upper one, and, since it has moved covertly, the copy of *self* that is spelled out is the lower one. Without further elaboration, the resulting sentence would be ungrammatical (*John likes self*). To solve this problem, Hornstein notes that *self* is a bound morpheme that needs a host to lean on. However, this host cannot be the lower copy of *John*, since this would induce a PF violation (i.e., pronouncing two links of the same chain). The solution is to insert *him* as a last resort operation, so as to provide *self* with a host. Notice also that *him* shows morphologically accusative case, even though it is part of a nominative chain.

Given the reasoning behind Condition A, the explanation of Condition B is quite straightforward. In the same way that NOC PRO was argued to appear only in contexts where OC PRO was not derivable by movement, pronouns appear in contexts where movement cannot form an anaphoric dependence. Example (12a) is wrong because, in this context, it is possible to form such a dependency via movement. Therefore, it is not necessary to resort to a pronoun. In (12b), since a DP can check case only once –we will return to this issue shortly-, the only way to get both nominative cases checked is by inserting a pronoun in the lower SpecIP. Similarly, Hornstein claims that bound and resumptive pronouns arise only when movement is not possible. For instance, (13a) would be a Left Branch Condition violation if *everyone* moved from the position of *his* to its surface slot. In (13b), *him* is embedded in an island, therefore *which person* cannot have formed a dependency with this position via movement.

12) a. *John$_i$ likes him$_i$
    b. John$_i$ said [that he$_i$ would leave soon]

13) a. Everyone$_i$ loves [his$_i$ mother]
    b. Which person$_i$ is it that John met a man [who likes *(him$_i$)]?

But, since Condition B can be equated to NOC PRO, the same problem we noted above rises up again here. That is, Condition B involves establishing a relation between two items without resorting to movement. In principle, such operations are forbidden in Hornstein's system, and I believe that not even the last resort status they have can hide the fact that they do not fit at all in this framework.

Finally, there seems to be an asymmetry between the binding examples in this chapter and the parasitic gap constructions of ch. 3, in that it is implicitly assumed that the former can check (nominative) case only once. Otherwise, nothing would prevent *John* in (12b) to move from the embedded to the matrix SpecIP, yielding the ungrammatical **John said that t would leave soon* (notice that this example cannot be independently excluded on the grounds that it involves illicit movement. Direct SpecIP-to-SpecIP movement wouldn't violate RM, and it doesn't seem that the trace/copy of *John* would cause any problems, given the present conception of Principle A. Alternatively, one could have an intermediate landing site in SpecCP, given that an A-A'-A chain has already been justified in the analysis of parasitic gaps). On the other hand, *which book* in examples like (4b) seems perfectly able to check able to check (accusative) case twice: once in the embedded, and once in the matrix clause.

The final chapter ("Case, C-command, and Modularity") consists of a short discussion of some methodological issues. To begin with, whether movement is "conceptually necessary" or not. Hornstein's answer is that it is, the justification being that movement is simply Copy & Merge. Merge is uncontroversially indispensable, and Hornstein argues that so is Copy. Therefore, an operation composed of two indispensable operations must in

turn be indispensable as well. I do not quite understand, though, Hornstein's argument for the necessity of Copy. It goes as follows: when an item is drawn out of the Lexicon, the latter does not shrink in size. This suggests that what actually happens is that we create a copy of the item in question and pull it out of the Lexicon. I believe, though, that at this point Hornstein is confusing Copy and Select. Making copies of the items in the Lexicon is Select, and this can be argued to be a necessary operation. On the other hand, Copy applies to the items present in the numeration, which does actually get smaller as the derivation proceeds. Actually, it could be argued that Hornstein's argument actually points to a no-movement view of grammar. One of the convergence requirements of sentences is that the numeration be reduced to an empty set, therefore an operation (Copy) that enlarges the number of items present looks suspicious and should be eliminated.

The final issue I want to touch c-command. At this stage of the discussion, it is not clear what c-command could be really useful for. In most approaches, c-command is exploited mainly to establish dependencies (i.e., *x...y is a well-formed dependency iff y is (not) c-commanded by x*). However, most of these dependencies have been reduced to movement. The logical conclusion would be that c-command should be reduced to movement (an idea developed by Epstein 1999). But Hornstein thinks otherwise. After all, and unlike virtually any other approach, his system allows movement to non c-commanding positions. Therefore, c-command can hardly be a consequence of the structure building process. The fact that movement does at times take place to c-commanding positions is more an accident than a relevant property of the grammar. The only place left for c-command, then, would be linearisation along the lines of Kayne's (1994) LCA: if x c-commands y and x does not c-command x, then x precedes y. This too can be dispensed with, Hornstein believes. The original relevance of c-command for linearisation is that it was the only grammatical relation that could be made to be inherently asymmetric. But here it is argued that Merge is antisymmetric as well. The idea is that, if two constituents x and y merge, the projection of the label will establish an asymmetric relation that can be exploited for linearisation purposes (though, given Hornstein's reasoning, asymmetric merge cannot be used to determine the linear order with respect to a head of both specifiers *and* complements. One of them has to be left out to be handled by some version of the Directionality Parameter).

Move! aims to a new way to conceive the grammar, in which every single relationship between two elements must be mediated by movement. However, the success is only partial: the constructions in which a construal process is invoked (NOC PRO, linking, Condition B) are an embarrassment for Hornstein's goal. Similarly, I have pointed out several loopholes and contradictions in the argumentation. Despite its weak points, I find the leading idea in *Move!* very attractive. One must keep in mind that this book is the first step in a very ambitious research program, too wide to be covered –even partially- in a 200-odd page volume. Its success –or failure- will

ultimately depend on whether future research can solve the many problem raised in the book.

References.

Bobaljik, Jonathan, and Samuel Brown (1997), "Interarboreal Operations: Head Movement and the Extension Requirement", *Linguistic Inquiry 28*, pp. 345-356.

Boskovic, Zeljko (1999), "On Multiple Feature Checking: Multiple Head Movement and Multiple Wh- Fronting", in Epstein & Hornstein (eds.), pp 159-187.

Brody, Michael (1995), *Lexico-Logical Form, a Radically Minimalist Theory*, MIT Press, Cambridge, Massachussetts.

Chomsky, Noam (1986), *Knowledge of Language*, Praeger, New York.

Chomsky, Noam (1993), "A Minimalist Program for Linguistic Theory", re-edited in Chomsky (1995), *The Minimalist Program*, MIT Press, Cambridge, Massachusetts, pp. 129-166.

Epstein, Samuel (1999), "Unprincipled Syntax: the Derivation of Grammatical Relations", in Epstein & Hornstein (eds.), pp. 317-345.

Epstein, Samuel, and Norbert Hornstein (1999), *Working Minimalism*, MIT Press, Cambridge, Massachusetts.

Hale, Ken, and Samuel J. Keyser (1993a), "Argument Structure and the Lexical Expression of Syntactic Relations", in Hale & Keyser (eds.) *The View from Building 20*, Cambridge, Massachusetts, pp. 53-109.

Higginbotham, James (1983), "Logical Form, Binding, and Nominals", *Linguistic Inquiry 14*, 395-420

Hornstein, Norbert (1995), *Logical Form, from GB to Minimalism*, Blackwell, Oxford.

Hornstein, Norbert (1999), "Movement and Control", *Linguistic Inquiry 30*, pp. 69-96.

Kayne, Richard (1994), *The Antisymmetry of Syntax*, MIT Press, Cambridge, Massachusetts.

Landau, Idan (1999), *Elements of Control*, PhD dissertation, MIT.

Lebeaux, David (1983), "A Distributional Difference between Reciprocals and Reflexives", *Linguistic Inquiry 14*, pp. 723-730.

Nunes, Jairo (1995), *The Copy Theory of Movement and Linearisation of Chains in the Minimalist Program*, PhD dissertation, University of Maryland, College Park.

Nunes, Jairo (2001), "Sideward Movement", *Linguistic Inquiry 32*, pp. 303-344.

Pesetsky, David, and Esther Torrego (2000), "T-to-C Movement: Causes and Consequences", in Kenstowicz (ed.), *Ken Hale, A Life in Language*, MIT Press, Cambridge, Massachusetts, pp. 355-426.

San Martín, Itziar (2002), *On Subordination and the Distribution of PRO*, PhD dissertation, University of Maryland, College Park.