THE DILEMMA OF MINIMALIST PRO-GRAMMAR
A MINIMALISTA PRO-GRAMMATIKA DILEMMÁJA

Szerző: Friedrich Tímea, angol-pedagógia szak

Konzulens: dr. Czeglédi Csaba, főiskolai tanár

(Humán Tudományi Szekcióban III. helyezés)

0. Introduction

In this paper, I seek to explore two related questions: (i) Can we adequately account for all the distributional properties of PRO, in other words, can we provide a principled explanation for its distribution? (ii) Can PRO escape methodological considerations of simplicity and economy that have become measures of theory evaluation within the Minimalist Program?

Under the Government and Binding (GB) approach to linguistic theory, the distribution of the phonetically null subject of infinitivals, PRO, was accounted for in terms of government, an essential grammatical relation of the GB framework. Given the requirement of the PRO Theorem that PRO must be ungoverned and the conception of government as a prerequisite for Case assignment, PRO was assumed to lack Case. Certain conceptual and empirical problems of this approach led to the proposal of an alternative Case-theoretic view on the distribution of PRO, namely that it bears null Case, a special Case that is restricted to PRO and can only be checked by nonfinite Infl in a Spec-head relation.

With the development of the Minimalist Program (MP), many of the basic concepts of the previous GB framework are reexamined in an attempt to find more suitable solutions and answers concerning fundamental properties of the human language faculty. Crucially, within MP an emphasis is placed on notions such as economy, elegance and simplicity as key factors in evaluating a theory. In particular, economy principles in terms of least effort conditions and methodological concerns are held to be essential criteria of an adequate theory. These minimalist ideas induce a reanalysis of the status of PRO and the whole Control module, and lead to the proposal of the movement theory of control, which dispenses with PRO and other theoretical concepts responsible for explaining its properties. The basis for such a move comes from viewing PRO as a theory internal formative, bearing properties that are quite distinct from those of other nominal elements.

In chapter 1, I discuss how the distribution-of-PRO problem is viewed under the GB framework, followed by a description of the null Case theory as formulated within early versions of minimalism. In chapter 2, I introduce Hornstein's movement theory of control and investigate some problems that the theory raises. In chapter 3, I outline a current minimalist analysis of PRO, and in chapter 4, some suggestions for further research are provided.

Chapter 1: PRO and Case Theory

In this chapter, I discuss issues related to the distribution of PRO: in section 1. 1., I review the GB approach to PRO; in section 1. 2., the problems posed by the GB analysis and the null Case theory are described; in section 1. 3., I outline Martin's (2001) version of the null Case approach.

1. 1. The Distribution of PRO and the Binding Theory

A fundamental concept of the GB approach is government, a structural configuration defined in terms of a mutual c-command relation between a governor α (that is a head) and an element β that is governed. Within GB, both Case-assignment and θ-marking take place under government, the former at the level of S-structure (SS), the latter at D-structure (DS).

Under standard accounts of the framework (Chomsky 1981, 1982, 1986a), the distributional properties of PRO are crucially determined by the notion of government, specifically by the principle below:

(1) *PRO Theorem*
PRO must be ungoverned.

The PRO Theorem is deduced from Principle A and Principle B of the Binding Theory, a module of GB that is responsible for interpreting the referential properties of nominal elements:

(2) a. *Principle A*: An anaphor must be bound in its governing category.
b. *Principle B*: A pronoun must be free in its governing category.

The notion *governing category* (GC) is defined as follows: α is a governing category for β if and only if α is the minimal IP, NP containing β and a governor for β; the expressions 'free' and 'bound' are construed in terms of a relationship that involves c-command and co-indexing between two categories, that is, a relation between α and β where α binds β in case α c-commands β and they are co-indexed. In effect, β is bound in its GC when it must be governed within its GC, and free when it is not governed within its GC.

Given the binding principles in (2), the theorem in (1) is derived as follows: Chomsky (1981, 1982, 1986a) argues that PRO shares some features with anaphors in that it has no specific independent reference, but it is also like pronouns in that it need not have an antecedent – therefore the properties of PRO are specified as [+ anaphor, + pronominal]. As a pronominal anaphor, PRO has to satisfy both Principle A and B of the Binding Theory: it must be both bound and free in its GC, a conflicting requirement which PRO can only escape if it lacks a GC and hence a governor, thereby not violating the principles of the Binding Theory. From this reasoning follows the condition on PRO that it must be ungoverned. Additionally, given the PRO Theorem in (1), PRO is assumed to lack Case, as Case-marking takes place under government within GB. Chomsky (1995) justifies these assumptions with empirical data:

(3) a. *we found PRO
b. *it seems [PRO to be intelligent]
c. John tried [PRO to defend Mary].

In (3a), PRO occurs in a position that is both governed and Case-marked, whereas in (3b) Case cannot be assigned to the position where PRO appears, but the position is governed; hence the ungrammaticality of the examples. (3c) is an acceptable sentence, as the position which PRO occupies is ungoverned, indicating that nonfinite Infl is not a governor for PRO. It should be noted, however, that there is a contrast between (3b) and (3c): why is PRO governed by the matrix verb in (3b), but not in (3c), given that the structure of the two sentences seems to be the same? Chomsky (1986b) provides a solution in terms of barriers: the embedded clause following the matrix verb in (3c) is CP, hence a barrier to government of PRO by the verb *tried*; by contrast, the clausal complement following *seems* in (3b) is IP, which is not a barrier to government of PRO by the matrix verb.

These empirical facts seem to indicate that the PRO Theorem provides an adequate account for the distributional properties of PRO; nevertheless, by broadening the empirical coverage, we will see that the issue is more complicated than it first appeared.

1. 2. PRO and Null Case

One of the puzzles posed in connection with the proposal, pointed out by Chomsky (1995), is related to the principle which attempts to integrate the Case Filter, a criterion that every lexical argument NP be Case-marked, with Theta Theory (Chomsky 1981). This principle, dubbed as the Visibility Condition by Chomsky (1986a), maintains that an argument chain (which is formed by the operation Move α) must contain a Case-position in order to be visible for θ-marking. As PRO is a θ-marked argument but lacks Case on the assumption that it is ungoverned, it violates the Visibility Condition. This in turn requires a conceptually undesirable modification of the Visibility Condition (which, moreover, does not solve the problem): a chain must contain a Case-position or must be headed by PRO in order to be visible for θ-role assignment.

Another conceptual reason for reconsidering the analysis presented in the previous section is related to the structure of infinitivals that contain PRO (see Hornstein, Nunes and Grohmann 2005). PRO ocuppies exactly the same position in the tree where a lexical subject NP occurs in a finite clause, i.e. the specifier position of nonfinite Infl. A lexical NP in the Spec position of a finite Infl is assigned nominative Case and governed by the Infl head, assuming a definition of government in terms of m-command (where α m-commands β if and only if α is a head and m-commands β). This said, on the null hypothesis PRO should also be governed under m-command by nonfinite Infl; nevertheless, to maintain the assumption that PRO is ungoverned, a distinction is made between finite Infl and nonfinite Infl within GB: arguably, it is posited that nonfinite Infl cannot govern.

Additionally, an empirical problem is posed by the fact that PRO, just like any other Case-marked argument, has to move from a position where no Case is assigned, and is not permitted to move from a position that is Case-marked, as the examples given by Chomsky (1995) illustrate:

(4) a. We never expected [PRO_i to be found t_i].
b. *It is unfair [PRO_i to talk about t_i].

In (4a) PRO moves from a non-Case position. (4a) alone does not pose a serious challenge to the analysis presented in the previous section: the movement of PRO could be accounted for under the assumption that as PRO must be ungoverned, its movement from a governed to an ungoverned position in (4a) is tolerated by the grammar and yields an acceptable result. In (4b), however, where PRO moves from a position that is both governed and Case-marked, the result is ill-formed: this shows that PRO, just like any other overt argument NP that must bear Case, cannot move from a position where Case is assigned, even if this move would protect PRO from being governed. Note that (4b) is incorrectly predicted to be grammatical by the assumption inherent in (1) that such a move should in general be licit.

The above considerations lead Chomsky (1995) to propose that PRO is indeed always Case-marked, just like any other argument NP, but its Case is distinct from the standard Cases that ordinary NPs bear: PRO is marked for null Case, a special Case that can only be realized by PRO and is licensed by nonfinite Infl via Spec-head agreement. Given this assumption, the distribution of PRO is straightforwardly derived: PRO is restricted to the subject position of infinitivals, as it cannot check off its Case-feature in other positions, ruled out by Case mismatch. It is crucial to note here that in the framework of Chomsky 1995, Case is no longer assigned to NPs at SS under government; rather, Chomsky proposes a checking theory of

Case assignment, defined in terms of a feature checking relation between two elements. It is assumed that lexical items enter the derivation with their Case-feature and other formal (i.e. morphological) features already specified, and these features must be checked off against the corresponding features of a local head. When a formal feature is strong (such as the Case feature of nominals), it must be eliminated in the overt syntax before PF; when a feature is weak, it must be deleted at LF. If such morphological features are not eliminated either at PF or at LF, they cause the derivation to crash. That is exactly what happens when PRO appears in positions other than the subject position of a nonfinite clause: in that case, no proper matching relation occurs, therefore PRO cannot check its Case-feature, which results in a crash.

If PRO is assumed to bear null Case, all the problems mentioned above can be resolved: as a Case-marked element, PRO is no longer different from other arguments with respect to the Visibility Condition; moreover, since a checking theory of Case does not invoke the concept of government, it is no longer necessary to make a distinction between finite and nonfinite Infl in terms of their ability to govern. Additionally, a Case-theoretic view correctly predicts that PRO is allowed to move from non-Case positions, but not permitted to move from a Case position. By conciliating PRO with Case Theory, it seems that the result we get is a theory that is both conceptually and empirically more adequate than the previous GB approach to the distribution of PRO. The sentences in (5a-b) below, however, raise the following problem: as the nonfinite Tense (T) head in infinitival clauses is assumed to be able to check null Case, in principle PRO is allowed to occur in ECM and raising structures like (5a-b), an unwanted result. How can we exclude PRO from these constructions? In the next section, I look at this problem more closely.

(5) a. *Mary believes [PRO to be happy].
b. *John seems [PRO to be sick].

1. 3. Martin's Proposal

To account for (5a-b) above, Martin (2001) claims that PRO is excluded from infinitval clauses embedded under ECM and raising predicates simply because the nonfinite T head in these constructions cannot check null Case on PRO, unlike the T head in control infinitivals. Martin argues that this distinction is based on the temporal features of control and raising infinitivals: the event time of infinitivals selected by control predicates is interpreted as future in relation to the event time of the matrix clause, whereas the event time of infinitivals embedded under raising predicates is the same as that of the matrix clause. Therefore nonfinite T in control structures bears the feature [+tense] which enables T to check null Case (similarly to finite T, which is also specified as [+tense] and checks nominative Case), while T in raising and ECM structures is specified as [-tense], hence cannot check Case. Further, nonfinite tense in control structures is a modal element, and "it corresponds most closely to *would* or, in certain contexts, *should*" (cf. Martin 2001: 147).

Martin argues that VP-ellipsis provides syntactic evidence for the difference between raising/ECM and control infinitivals, as illustrated by (6a-b) (cf. Martin 2001):

(6) a. Kim isn't sure she can [$_{VP}$ solve the problem], but she will try [PRO [$_{T}$ to] [$_{VP}$ *e*]].
b. *Bill believes Sarah to [$_{VP}$ be honest], and he believes [Kim [$_{T}$ to] [$_{VP}$ *e*]] as well.

Martin claims that if we adopt an assumption under which only the complement of agreeing functional heads can be deleted under ellipsis, the grammaticality of the control structure in (6a) follows from the fact that the nonfinite T head agrees with PRO in its Spec position (i.e. checks null Case on PRO), hence the VP complement can be deleted. In the ECM structure in (6b), however, no agreement takes place between nonfinite T and the NP *Kim* in its Spec

position (i.e. T does not enter a Case-checking relation with the NP), yet the VP complement is deleted – accordingly, (6b) is ungrammatical. The empirical facts in (6a-b) are what we expect under Martin's assumption that T enters an agreement relation with its specifier only in control infinitivals (and consequently, T checks Case only in control structures).

In addition to the problem raised by (5a-b), a further issue in connection with the Case-checking properties of nonfinite T is posed by *want*-infinitivals: the predicate *want* can take an infinitival complement with either PRO or a lexical NP in subject position, and *want* can also take a *for*-infinitival complement. How can we account for these facts under the null Case theory? Martin argues that in all three cases, nonfinite T checks null Case on the embedded subject, as in (7a-c) below (cf. Martin 2001):

(7) a. John wants [PRO to win].
b. John wants [his team to win].
c. John wants [for his team to win].

For such an analysis to be viable, it is crucial for Martin to assume the following: (i) neither T in control infinitivals nor PRO bears φ-features (grammatical features like gender, number, etc.); (ii) φ-features on NPs are uninterpretable (a feature that is not legible at the interfaces and causes the derivation to crash at LF unless deleted prior to LF), therefore need to be checked; (iii) the complementizer *for* is associated with formal φ-features, but it cannot check Case on the NP following it. With these stipulations in mind, (7b-c) are accounted for as follows: in (7b), nonfinite T checks null Case on the lexical NP *his team*, but it cannot check the φ-features of the NP, therefore a null *for* precedes the NP to check the φ-features of the NP. In (7c), *his team* checks null Case in the embedded subject position overtly, then raises covertly at LF to the Spec position of *for* to check its φ-features. The ungrammaticality of (8a-b) below is also deduced from provisos (i), (ii) and (iii). In (8a), a lexical NP such as *John* cannot appear in the [Spec, TP] position of a control infinitival, as its φ-features remain unchecked in that position; similarly, in (8b) PRO is not allowed to co-occur with *for*, as it cannot check the complementizer's uninterpretable φ-features.

(8) a. *Mary tried John to open the door.
b. *Mary wants for PRO to win.

On closer inspection, however, Martin's analysis may prove to be problematic from both a conceptual and an empirical point of view. First, as argued by Baltin and Barrett (2002), it is theoretically suspect to assume a link between the presence of a PRO subject in an infinitival and the interpretive properties of that infinitival. Nevertheless, this is exactly what is suggested by Martin's proposal, which relates the modal interpretation of control infinitivals to the licensing of a PRO subject in these infinitivals. In other words, it is not reasonable to distinguish sentences associated with different semantic interpretations by a feature (in our case the [+tense] feature on T) that also has an effect on the structure of those sentences. Accordingly, although the time of the gerundial is interpreted as future in relation to the matrix predicate in (9a), while interpreted as identical with the event time of the matrix verb in (9b), it seems quite dubious to postulate the existence of a feature that would serve to distinguish the two interpretations syntactically as well (cf. Baltin and Barrett 2002):

(9) a. He avoided bringing the wine.
b. He continued bringing the wine.

Furthermore, as noted by Baltin and Barrett, there are control infinitivals with a PRO subject that do not display the semantic property attributed to them by Martin, as shown in (10a) where the infinitival is the complement of a psychological adjective, or in (10b) where it is the complement of a degree word. In both sentences, the event time of the embedded

nonfinite clause is interpreted as simultaneous with that of the matrix clause, a temporal relation between the two clauses that is supposed by Martin to be characteristic only of raising and ECM structures (cf. Baltin and Barrett 2002):

(10) a. Kerry was glad to see her best friend walk out the door.

b. John was too obtuse to be aware of his effect on others.

Likewise, in the case of VP-ellipsis, if we expand the empirical coverage to include raising constructions, it seems again that Martin's analysis makes the wrong empirical predictions. Interestingly, as Baltin and Barrett show, there are raising predicates that license ellipsis of their complement, as in (11a-b), contrary to what we expect under Martin's assumptions (cf. Baltin and Barrett 2002):

(11) a. Bob tries not to gain weight but he tends to_.

b. Although she hasn't spoken to the president yet, she is supposed to_.

Martin's analysis of *want*-infinitivals also raises certain nontrivial issues that need to be addressed. Let us begin by examining the status of provisos (i), (ii) and (iii) which are essential for Martin's account of these constructions to be tenable. Proviso (i) requires that both PRO and nonfinite T in control infinitivals lack φ-features entirely. Under the assumption that PRO does not bear φ-features, however, the agreement facts in (12) are left unexplained (cf. Chomsky 1995):

(12) a. I want [them to be officers].

b. *They want [me to be officers].

c. They want [PRO to be officers].

In (12a) and (12c), there is agreement in the embedded clause between the subject and the predicate, while the ungrammaticality of (12b) follows from the lack of φ-agreement between the embedded subject and *officers*. Chomsky (1995) illustrates with examples such as (12) that PRO must indeed bear φ-features for agreement, just like any other NP. Further, if the stipulation that PRO is Case-marked is to be retained, nonfinite T in control infinitivals also has to be assumed to carry a full set of uninterpretable φ-features. This follows from the fact that in the framework of Chomsky 2000, 2001a, 2001b, Case checking is a reflex of φ-feature agreement between a Case-assigning functional head and a nominal with an unvalued Case-feature that is valued and deleted through this agreement relation. Consequently, there must be a Case-assigning head with a complete φ-set in control infinitivals that values the uninterpretable Case-feature of PRO as a result of φ-agreement between PRO and this head, for which the only candidate is nonfinite T. The stipulations of proviso (i) therefore cannot be adopted.

Proviso (ii) requires that φ-features of NPs be checked in the course of the derivation, i.e. that φ-features on NPs be uninterpretable. In Chomsky's (2000) Agree-based system of feature checking, formal features of lexical items are either LF-interpretable, that is, they contribute to semantic interpretation at LF, or LF-uninterpretable, including features that have no semantic value at LF. Interpretable features are therefore legible at LF, in contrast to uninterpretable features that must be checked (i.e. valued and deleted under the operation Agree which is established by matching of uninterpretable features of *probe* α and interpretable features of *goal* β) prior to LF for the derivation to converge. Interpretability or uninterpretability of a feature is not stipulated, however: while uninterpretable features on a lexical item need explanation and justification, interpretability of a feature is attested by the fact that it is involved in the semantic interpretation of a linguistic expression. Accordingly,

Chomsky takes φ-features on nominals to be interpretable, and it is the φ-features of functional heads such as T and *v* that are uninterpretable (i.e. unvalued) and need to be valued in an agreement relation. Hence it is not reasonable to adopt proviso (ii).

Proviso (iii) defines the complementizer *for* as an element that bears uninterpretable φ-features and by virtue of this property, it enters an agreement relation with the φ-set of the subject NP in a *for-to* infinitival construction. Crucially, Martin assumes that the Case of the NP is not checked by *for* as a result of this φ-agreement relation; rather, it is nonfinite T that is supposed to check Case on the NP, assigning null Case to the nominal. Such an analysis, however, seems to be both conceptually and empirically inadequate. First, as we have seen above, Case checking is supposed to be assigned under φ-agreement between two elements, hence nonfinite T in principle cannot delete Case in a nominal if they do not φ-agree. Second, null Case cannot be allowed to be checked by lexical NPs: such an assumption fails empirically, as we will see below. Third, Martin's analysis cannot explain the strong dependency relationship apparent between *for* and the subject NP in *for-to* infinitivals, which seems to be related to Case-theoretic reasons. For illustration, consider (13a-f):

(13) a. For John to arrive on time would be preferable.

b. *John to arrive on time would be preferable.

c. It would be preferable for John to arrive on time.

d. *It would be preferable John to arrive on time.

e. *Mary prefers for usually John to arrive on time.

f. Mary prefers for him/*he to arrive on time.

Under the revision of (ii) above, the ungrammatical sentences can no longer be explained by alluding to the uninterpretability of the NP's φ-features. What appears to be crucial in these sentences is indeed the role of *for*, or rather its absence, as shown by (13b) and (13d), where the ungrammaticality must be due to the lack of *for*. (13e-f) also imply that the licensing of lexical subjects in *for-to* constructions relies crucially on the complementizer: in (13e), a sentence adverb is not allowed to intervene between *for* and the nominal, while in (13f) the close relation between *for* and the NP is attested by the obligatory overt Accusative Case marking of the pronominal subject. Given these empirical data, Hazout (2004) argues that in *for-to* infinitivals, the subject is licensed through a Case-checking relation by the complementizer *for* – in fact, there is no other candidate to check structural Case on the subject NP, as Case checking in a φ-agreement relationship is not possible between nonfinite T and a lexical NP, as shown by (13b) and (13d).

If, however, *for* is supposed to be responsible for the Case-licensing of the subject in *for-to* constructions, how do we account for the optionality of *for* in infinitivals selected by *want*-type predicates?

(14) a. I want (for) John to be successful.

b. We would prefer (for) him to stay at home.

One option is to assume that a null *for* is present to check Case on the NP in the constructions without an overt complementizer – in this case, however, it has to be explained why a null *for* is not possible in sentences such as (13d). Another option to account for (14a-b), as suggested by Radford (2000), is to suppose that the infinitival selected by a *want*-type predicate is a TP rather than a CP complement when the overt *for* is missing, and the Case of the subject NP is

checked by the matrix verb similarly to ECM structures.[130] An analysis along these lines seems to be motivated by empirical data in (15a-b): in (15a), a sentence adverb cannot intervene between *want* and *him*, whereas in (15b) the infinitival complement of *want* cannot occur in the focus position of a pseudo-cleft sentence. This seems to indicate that the ungrammaticality of both (15a-b) must follow from the inability of the matrix predicate *want* to check Case on the infinitival subject (cf. Radford 2000).

(15) a. *They wanted desperately him to win the race.

b. *What John wants is Mary to win the race.

If we adopt the idea that in *for-to* infinitivals it is the C head *for* that checks Case on the subject NP, some further questions arise with respect to this Case-checking relation. One such question is related to the fact that the subject is assigned Case by C in [Spec, TP] position, not in a Spec-head agreement relationship with C (the canonical Case checking configuration on earlier assumptions). However, Case-valuing under such a long-distance agreement (LDA) relation is possible between two elements in Chomsky 2000, 2001a, 2001b, 2005a. In an LDA relation between the C head *for* and an NP, the operation Agree values and deletes the uninterpretable φ-features of C under matching with the interpretable φ-features of the nominal, and one effect of the operation is the structural Case marking of the subject NP in situ by C.

A further issue is posed by the proposal (Chomsky 2000, 2001a) that C selects T_{comp}, i.e. a T head with a full set of agreement features and the ability to assign Case - null Case if T is nonfinite, and Nominative Case if finite. As formulated in Chomsky 2001b, 2005a, 2005b, the true locus of Tense- and agreement features is indeed C, and T only inherits these features derivatively from C, hence only if it is selected by C. The problem that arises is the following: T_{comp} in nonfinite structures is incompatible with a lexical NP in its Spec position (which is presumably due to Case mismatch between nonfinite T and the overt NP), as shown by the impossibility of an overt subject in control infinitivals. As a *for-to* infinitival is a CP complement that is assumed to select T_{comp}, and we find a lexical subject in [Spec, TP] position, a question may arise regarding the properties of nonfinite T in these structures. There are two possible alternatives to overcome this conundrum. Alternative I is to suppose that as in *for-to* constructions it is the function of C to assign Case to the subject, T inherits only the φ-features of C, but is Tenseless, hence cannot assign Case and therefore no Case mismatch occurs. On this view, there are two ways to proceed:

(i) The φ-agreement relation between C and the NP is direct: the nominal first values the φ-features of nonfinite T and then also deletes the agreement features on C, as a result of which its Case feature is valued.

(ii) The φ-agreement relation between C and the NP is indirect: the NP values only the uninterpretable φ-set of T, and then T φ-agrees with C (this may be possible given Chomsky's (2001a) assumption that the uninterpretable φ-set of C is valued by T under the operation Agree). Under a principle of transitivity applied to φ-agreement relations (proposed in Frampton, Gutmann, Legate and Yang 2000), this has the effect of C-NP φ-agreement and structural Case checking on the NP by C.

Alternative II to get over the problem posed by the nature of nonfinite T in *for-to* infinitivals is to assume that the complementizer *for* selects for a TP with a defective head $T_{def,}$ associated with ECM and raising structures. T_{def} is φ-incomplete as it bears only the (unvalued) feature [person] and hence is unable to delete the Case feature of an NP (as desired). In this case, the

[130] However, see Lasnik and Saito (1991) for arguments against an analysis where the embedded subject in structures of the form *want someone to do something* raises to [Spec, VP] position (or to [Spec, AgrOP] on earlier assumptions) to check Accusative Case, as is supposed in the case of ECM structures.

uninterpretable agreement features of C are deleted under Agree by the interpretable φ-set of the subject NP (the effect of which is Case-marking of the NP by C), the only possible candidate to enter such an agreement relation with C.[131]

In view of the considerations above, let us now reexamine Martin's analysis of *want*-infinitival constructions. In (7a), null Case on PRO is checked as a result of φ-feature agreement between PRO and nonfinite T. In (7b), the Accusative Case on the NP is licensed by the matrix predicate *want*; additionally, there is no need to postulate the presence of a null *for* in the structure, as φ-features on NPs are interpretable, hence need not be checked. In (7c), the NP is assigned Accusative Case by *for*, and as the φ-set of the NP is interpretable, we do not have to stipulate covert raising of the subject at LF.[132] Further, the ungrammaticality of (8a) can no longer be explained under the assumption that the φ-set of the NP remains unchecked; indeed, if φ-features on NPs are interpretable (as assumed above), structures like (8a) are predicted to be acceptable under Martin's analysis, given that it allows lexical NPs to be assigned null Case by nonfinite T. In effect, contrary to what Martin supposes, it seems reasonable to attribute the unacceptability of (8a) to Case mismatch: null Case cannot be checked on a lexical NP. In a similar vein, as PRO is assumed to bear φ-features, (8b) can no longer be accounted for by alluding to the feature checking requirement of *for*, since the uninterpretable φ-set of the complementizer can in principle be deleted by the φ-set of PRO. In fact, if the postulation that PRO carries φ-features is adopted, (8b) is again predicted to be grammatical under Martin's analysis, since it allows null Case to be checked in *for-to* infinitivals. That possibility, however, needs to be excluded if we wish to account for (8b); in contrast to Martin's account, the ungrammaticality of the sentence is again due to Case-theoretic reasons: Accusative Case cannot be checked on PRO.

From the discussion so far, it seems that Martin's proposal is both theoretically and empirically inadequate. On the one hand, it leaves us without any principled explanation of some structures, given the counterexamples to the generalization related to the temporal properties of infinitivals. On the other hand, this temporal distinction does not appear to be motivated on strong conceptual grounds. On the face of it, this seems to be problematic not only for a null Case account of the distribution of PRO, but consequently also for a theory which maintains that a grammatical formative like PRO exists. In view of these considerations, we are faced with two options: either we attempt to save the null Case theory (along with PRO), in which case at least a different technical implementation of the theory is necessary for it to be tenable at all; or we may search for another alternative account of the syntax of control infinitivals that does not resort to the null Case theory and PRO. In effect, such a theory has been proposed recently within current minimalist approaches to grammar, and it is now in order to examine whether that theory proves to be preferable to previous accounts.

Chapter 2: A Movement Theory of Control (MTC)

In this chapter, I discuss Hornstein's movement analysis of control structures. In section 2. 1., I describe Hornstein's MTC, in section 2. 2., the problems raised by the theory are investigated and in section 2. 3., some concluding remarks are provided.

---

[131] For an alternative analysis of *for-to* infinitivals, see Pesetsky and Torrego (2001).

[132] That such an LF raising does not take place is in fact motivated on other grounds as well, as in the Agree-based system of Chomsky (2000, 2001a) there is no overt/covert distinction, and covert movement is reinterpreted in terms of overt agreement. Additionally, given LDA, a Spec-head configuration is not a prerequisite for feature checking and agreement, hence the NP would not have to raise to enter an Agree relation with *for*.

2. 1. Hornstein's MTC

The methodological concerns stressed within MP have motivated a minimalist effort to reexamine grammatical concepts and processes underlying standard GB accounts. In particular, issues related to optimal grammatical design have cast doubt on the status of theory internal constructs such as the non-interface levels DS and SS. Grammar internal formatives are considered problematic from a conceptual point of view within MP, and therefore such constructs should only be adopted if their postulation rests on strong empirical grounds.

Given this theoretical backdrop, Hornstein (henceforth H) (1998, 1999, 2000, 2001, 2003) and Boeckx and Hornstein (2003, 2004) argue for a reductionist view of the syntax of control infinitivals, which dispenses with the whole PRO module. Specifically, H argues that PRO is a grammar internal formative with highly idiosyncratic properties, and its postulation gives rise to various theoretical complications in the grammar: on the one hand, positing a PRO subject introduces a whole additional module into the grammar, the Control module, which serves to account for the interpretive properties of PRO in control structures. On the other hand, explaining the distribution of PRO requires undesirable theoretical adjustments in that we have to postulate a special null Case, which H claims to be both conceptually and empirically problematic: in effect, null Case is devised to fit only PRO. This follows from the assumption that the only nominal element that can be assigned null Case is PRO, and only the nonfinite T of control infinitivals can license this Case; further, PRO is the only nominal that is specified to bear but a single Case. Consequently, H maintains that the null Case theory is highly stipulative, and it nearly amounts to restating the observation that PRO occupies the [Spec, TP] position of control infinitives.

These considerations lead H to reexamine the distinction traditionally assumed to obtain between raising and control structures within the GB framework. Thus in the raising structure in (16a) *John* is supposed to have raised from the embedded to the matrix subject position, leaving a trace in its original site which is co-indexed with the moved expression. In (16b) the matrix subject position is assumed to be filled by lexical insertion, and the embedded subject position is occupied by a base-generated null argument PRO, which is controlled by *John* (cf. Hornstein 2003):

(16) a. $John_i$ seems [t_i to like Mary].

b. $John_i$ tried [PRO_i to like Mary].

In (16a), *John* is associated with a single θ-role, assigned to it by the embedded verb *like*, whereas in (16b) *John* is understood to be both the *trier* and the *liker*. This follows from the different thematic properties of the predicates: in contrast to the verb *try*, the verb *seem* does not assign a θ-role to the external argument position, as shown by its ability to host expletives in that site. The alleged structural difference between raising and control structures is traceable to the postulation of DS and the θ-Criterion within GB: DS requires that all thematic positions be filled by lexical insertion prior to movement operations, while the θ-Criterion stipulates that a chain contains at most one θ-position. Accordingly, adopting these two assumptions amounts to precluding movement between θ-positions and disallowing NPs to bear more than one θ-role. Given DS and the θ-Criterion, the structural distinction between (16a) and (16b) follows: in (16a) *John* is allowed to move from its DS θ-position to the matrix subject position (for Case-theoretic reasons), as *seem* does not assign a θ-role to *John*. In (16b), such movement would be illicit, as both the matrix and the embedded subject positions are θ-positions; consequently, a null PRO subject has to be assumed to occupy the embedded subject position.

Within MP, DS is dispensed with, since it is viewed as a non-interface theory internal level the postulation of which is not necessary from a conceptual point of view. H argues that adopting the θ-Criterion also raises methodological issues within the current minimalist approach to grammar. Once DS and the θ-Criterion are eliminated, the constraints that they imposed on grammar can also be abandoned. Accordingly, H claims that it would be reasonable to reconsider the traditional GB control/raising distinction and the status of PRO. Specifically, H proposes a movement-based theory of control, suggesting that in control structures like (16b) above, PRO is the residue of overt A-movement, hence it is identical to an NP-trace.[133] On such an account, the structure of (16b) is as follows:

(17) $John_i$ tried [t_i to like Mary].

In (17) above, *John* has moved from the embedded to the matrix subject position, just like in a raising construction. Under this view, the operations underlying raising and control structures are the same (as both raising and control instantiate NP-movement, which is triggered by Case-theoretic needs), with the single distinction that in raising, movement of the NP is from a θ-position to a non-θ-position, whereas in control expressions move between θ-positions. The proposal entails that the subject position of control infinitivals is not a Case-marking position, hence A-movement from this position is licit. For such an analysis to be viable, H is forced to adopt the following assumptions:

(18) a. θ-roles are features on verbs/predicates, thus they are able to license movement, and an NP receives a θ-role by checking the thematic feature of the given predicate

b. there is no upper limit on the number of θ-features that an NP/A-chain can bear

c. movement is greedy, and Greed is interpreted as "enlightened self-interest" (ESI)

Positing (18a-b) is necessary for movement to be allowed between θ-positions; (18c) requires all movement to take place for reasons of feature checking (that is, be greedy), while ESI ensures that the checking needs of the target suffice to license movement.[134]

Given the assumptions in (18a-c), the derivation of a subject-control sentence under a movement approach is as follows (cf. Hornstein 1999):[135]

(19) a. John hopes to leave.

b. [$_{IP}$ John [$_{VP}$ John [hopes [$_{IP}$ John to [$_{VP}$ John leave]]]]]

In the derivation of (19a), *John* first merges with the embedded VP, thereby checking the external θ-role of *leave*. *John* then moves to the embedded [Spec, IP] position to check the EPP feature of *to*; the Case of *John* is not checked here, as no Case is assigned to the embedded subject position under H's theory. *John* then raises to the matrix [Spec, VP] position (which is again a θ-position), checking the θ-feature of the matrix predicate *hope,* thereby assuming a second θ-role – thus *John* bears both the *leaver* and the *hoper* θ-role. Finally, *John* moves to the matrix [Spec, IP] position, where it checks the EPP feature of I and its own Case feature, and the derivation converges with all relevant features checked.

---

[133] For a movement analysis of control structures, see also O'Neil (1997) and Manzini and Roussou (2000). For convincing arguments against such an analysis, see Culicover and Jackendoff (2001), Landau (2003) and Brody (1999, 2001).

[134] Under the early minimalist concept of Greed – in contrast to ESI - only the morphological requirements of the moved item are relevant. In Chomsky 2000, 2001a both the target (i.e. the probe) and the moved element (i.e. the goal) must be active for movement to take place.

[135] Given the copy theory of movement within MP, which takes traces to be copies of the moved element that are deleted in the phonological component, traces of A-movement are indicated by copies of *John* in (19b).

H argues that the assumptions in (18a-c) and the MTC allow for a significant reduction and simplification of the theory of grammar, since there is no need to postulate a special control module which accounts for the interpretive properties of PRO in control configurations; rather, the general characteristics of obligatory control (OC) and nonobligatory control (NOC) structures follow directly from the MTC. In other words, under the MTC we can explain the basic properties of OC and NOC structures. Consider the ungrammatical OC structures in (20) below (cf. Hornstein 1999):

(20) a. * It was expected PRO to shave himself.

b. * John thinks that it was expected PRO to shave himself.

c. * John's campaign expects PRO to shave himself.

d. * $John_i$ told $Mary_j$ PRO_{i+j} to wash each other.

(20a) shows that PRO must have an antecedent in an OC configuration, otherwise the sentence is not acceptable. (20b) indicates that the antecedent of PRO must be local in OC structures, whereas (20c) illustrates that this antecedent must c-command PRO: the ungrammaticality of the sentences follows from the fact that these requirements are not satisfied. (20d) shows that PRO is not allowed to have split antecedents in an OC structure. These properties of OC structures are derived under the MTC as follows: given that PRO is a trace of A-movement, it needs to have an antecedent, which has to c-command its trace and also be local to it, as is standardly supposed in the case of A-chains. The ungrammaticality of (20d) also follows under the MTC given the assumption that two NPs could not have moved from the same position, as would be required in (20d).

Still, there are grammatical NOC configurations where PRO need not have a local, c-commanding antecedent, and where split antecedents are allowed – surely, given the analysis of (20a-d) these cases cannot fall under the MTC. For this reason, H proposes that A-movement takes place only in OC structures, and what we have in the embedded subject position in NOC structures is the null pronominal pro, as in (21) below (cf. Hornstein 1999):

(21) It is believed that pro shaving is important.

H argues that such an assumption is confirmed by the interpretive properties of NOC constructions, since these structures can be paraphrased by replacing PRO (or pro, under H's view) with overt pronouns (as is expected if the null pronominal pro occupies the relevant position in NOC structures). PRO's distribution is also identical to that of a pronoun in NOC configurations in that it does not have to obey the locality conditions characteristic of A-traces and anaphors.

The question may arise as to why movement does not take place in these NOC structures. The reasoning is as follows: NOC PRO occurs within islands out of which movement is prohibited, hence movement from the embedded [Spec, IP] position is not an option in these structures. From this it follows that OC PRO and NOC PRO are in complementary distribution: when NOC PRO is permitted, OC PRO (i.e. movement) is barred. But does the reverse hold? In other words, how can we exclude NOC PRO from positions occupied by OC PRO? H argues that the problem may be solved in terms of an economy restriction: he assumes that movement (OC) is more economical than construal processes (NOC), hence preferred by UG. This means that if the derivation can converge through the application of movement operations (as in the case of OC), the more costly non-movement derivation where construal rules are employed (NOC) will be excluded by economy. This implies that the application of construal rules is a costly last resort operation, and therefore when movement is allowed to occur in a control structure, it must occur (the OC cases); only when movement is

barred do we expect to find an NOC configuration. Further, H treats the licensing of pro similarly to the phenomenon of *do*-support: he supposes that pro is a grammatical formative that is never present in the numeration (i.e. a set of lexical items selected from the lexicon that feed the derivation of a syntactic object), it is only added to the computation during the course of the derivation as a last resort option, thereby allowing a derivation to converge. In other words, the use of pro is costly, and is only permitted when convergence is not possible without pro. Additionally, pro is assumed to lack Case, but it is supposed to be able to check the relevant features of nonfinite Infl.

Let us finally review the empirical arguments put forth by H in support of the assumption that PRO is identical to an NP-trace and against a theory of null Case. The first argument is related to the phenomenon of *wanna*-contraction. Consider (22a-c) below (cf. Hornstein 1999):

(22) a. Who$_i$ do you want [t_i to vanish]? / * Who do you wanna vanish?

b. I want [PRO to leave]. / I wanna leave.

c. John$_i$'s going [t_i to leave]. / John's gonna leave.

In (22a), contraction of *to* onto *want* is blocked by the Accusative Case-marked *wh*-trace; in (22b), null Case-marked PRO fails to block contraction, just like the NP-trace in (22c). H argues that as PRO is supposed to bear null Case, it should behave similarly to a Case-marked *wh*-trace in blocking contraction, but it behaves like a non-Case marked NP-trace in contraction contexts, as shown by (22). If PRO is assimilated to an NP-trace, we can account for why both PRO and NP-trace allow contraction in (22), unlike the Case-marked *wh*-trace.

The second argument is directed particularly against a theory of null Case. H claims that a null Case theory is highly implausible given the sentence in (23), which H assumes to be a case of control, involving a PRO subject when the thematic object is not overtly realized. The problem is that in (23), PRO does not appear in [Spec, TP] position, hence cannot be assigned null Case by a nonfinite T head.

(23) Mary$_i$ washed PRO$_i$.

(23) is explained under the MTC as follows: as (23) is a case of OC, *Mary* is supposed to move from the object position of *wash* to the matrix subject position, thereby receiving both the internal and the external θ-role of the verb. This implies that reflexive verbs like *wash* can optionally suppress their Accusative Case in structures such as (23), given that two Cases (Accusative in the object postion and Nominative in the subject position of *wash*) cannot be licensed on *Mary*.

H claims that the MTC is superior to the standard account of the syntax of control infinitivals on both conceptual and empirical grounds. The theory, however, though not without merit, is also not without problems, as we will see in the next section.

2. 2. Some Problems for Hornstein's Theory

Let us first reexamine the two empirical arguments that seem to justify H's theory: *wanna*-contraction and PRO in positions other than [Spec, TP].

2. 2. 1. *Wanna*-contraction

Contraction is a phonological operation whereby finite auxiliaries or the infinitival particle *to* (i.e. a finite or infinitival T head) contracts onto a preceding matrix verb. H adopts the standard Case-marked trace analysis of contraction phenomena, which stipulates that non-phonetic Case-marked traces block contraction, while null arguments like PRO and non-

phonetic Caseless traces do not block contraction. Under this account, the contrast between sentences like (24a) and (24b) below is analysed as follows: the *wh*-trace in (24a) is Case-marked, hence blocks T-contraction, whereas the NP-trace in (24b) is Caseless, thus allowing the contraction of the infinitival T head onto the matrix V (cf. Bošković 1997).

(24) a. Who$_i$ do you want t_i to buy a car? / * Who do you wanna buy a car?

b. John$_i$ is supposed t_i to leave on Monday. / John is sposta leave on Monday.

c. I want PRO to buy a car. / I wanna PRO buy a car.

The above Case-marked trace analysis, as we have seen, raises the following problem with regard to (24c) if we assume a theory of null Case: PRO is marked for Case, unlike an NP-trace and like a *wh*-trace, but it does not block contraction, unlike Case-marked *wh*-traces and like Caseless NP-traces. H argues that this casts a doubt on a null Case approach.

Bošković (1997) and Radford (2000) argue, however, that the standard Case-marked trace analysis of contraction facts is problematic from both a conceptual and an empirical point of view. First, there is empirical data that may cast doubt on such an analysis (cf. Radford 2000, Bošković 1997):

(25) a. Who$_i$ do you think t_i is telling the truth? / Who do you think's telling the truth?

b. What$_i$ do you think t_i is happening there tomorrow? / What do you think's happening there tomorrow?

Although a Case-marked *wh*-trace apparently intervenes between the finite auxiliary and the matrix verb in both sentences, T-contraction is not blocked in either of the examples in (25). Under the standard analysis of contraction phenomena which H adopts, both sentences are predicted to be ungrammatical (contrary to fact), given that Case-marked traces are assumed to block contraction on that analysis.

The standard view of contraction also raises conceptual issues: in contrast to the Case-marked trace analysis, the null hypothesis would be to assume that given their uniform phonetic status, all empty categories (ECs) (*wh*-traces, NP-traces, PRO) behave similarly in contraction contexts, that is, we should expect all ECs to allow (or block) contraction. Theoretically, an optimal minimalist account of T-contraction seems to be to suppose that only overt categories block contraction.

These considerations cast doubt on the standard Case-marked trace analysis of contraction phenomena, which serves as the basis of H's empirical argument for reducing PRO to an NP-trace. The question then arises: how the contraction facts in (24) and (25) can be explained, if not under the standard Case-marked trace analysis? Radford (2000) claims that the phenomenon of T-contraction can be adequately accounted for under the phase-based theory of grammar developed in Chomsky 2000, 2001a, 2001b. Before proceeding with the details, however, it is necessary to outline the core set of assumptions that underlie the phase-based analysis of T-contraction.

In recent work, Chomsky argues that derivations are strictly cyclic, that is, the syntactic computation proceeds via derivational phases, which Chomsky takes to be CP and transitive *v*P. Phases are supposed to play a special role in the narrow syntactic computation by constraining all operations in terms of the Phase Impenetrability Condition (PIC), defined in (26) below (cf. Chomsky 2000, 2001a, 2001b):

(26) *Phase Impenetrability Condition (PIC)*

In a phase α with head H, the domain of H is not accessible to operations outside H, only H and its *edge* (its Specs) are accessible to such operations.

The PIC ensures that once a phase α is completed (i.e. judged convergent), no further operations are allowed to access elements inside α below its head H; further, H and its edge are accessible to operations outside HP only until the next strong phase (CP or transitive *v*P) is completed. The conditions imposed by the PIC reduce computational complexity both for narrow syntax and the phonological component in that earlier phases of the derivation can be forgotten by the system.

The strong phase-level is also relevant with respect to Spell-Out, an operation that applies at some point in the computation to the structure already formed and sends it to the PF interface. This means that Spell-Out is cyclic, namely, it applies multiple times in the course of the computation: at every (strong) phase level, Spell-Out applies to the complement of the phase head and sends it to the interface levels where it is inspected for convergence.[136] The head of the phase and its specifiers are spelled out only after the next strong phase is constructed (i.e. H and its specifiers are spelled out together with the complement of the next higher phase head).

Additionally, in the current theoretical framework the operation Move is assumed to be the combination of the two operations Copy and Merge: when movement occurs, the system first makes a copy of the moved lexical item, then the lexical item is merged in the target position of movement.[137] What interests us here is again the role of Spell-Out: copies of the moved item are supposed to be deleted (except from the head of the chain) prior to the shipping of the relevant structure to the PF level, hence at the point when Spell-Out applies.[138]

Keeping in mind the above assumptions and the stipulation that contraction is blocked only by overt constituents, let us examine how the contraction phenomena in (24) and (25) can be explained under a phase-based approach to grammar. Consider sentence (25a), with the relevant structure illustrated in (26) below (cf. Radford 2000):

(26) [$_{v}$P who you [$_{v}$ ø] [$_{VP}$ [$_{V}$ think] [$_{CP}$ who[$_{C}$ ø] [$_{TP}$ who [$_{T}$ is] telling the truth]]]]

If we suppose that contraction of the finite auxiliary *is* onto the matrix V *think* takes place when the matrix VP containing *think* is spelled out on the *v*P phase, the grammaticality of the sentence follows straightforwardly. The trace of *who* in [Spec, TP] is deleted when the embedded CP phase is constructed and Spell-Out applies to its TP complement, whereas the copy in embedded [Spec, CP] position is deleted on the matrix *v*P phase. Accordingly, at the stage of the derivation when the matrix V head *think* is spelled out, there will be no overt constituent intervening between *think* and *is*, given that both copies of *who* have already been given a null spellout at this point of the derivation; hence contraction of *think* and *is* yields a grammatical sentence. The grammaticality of (25b) and (24b-c) can be accounted for in a similar fashion; let us rather turn to the ungrammatical case in (24a). On the assumption that (24a) is similar to an ECM structure, the complement of *want* is TP rather than CP.[139] The structure of the sentence is represented as in (27) below (cf. Radford 2000):

(27) [$_{CP}$ Who [$_{C}$ do] [$_{TP}$ you [$_{T}$ do] [$_{v}$P who you [$_{v}$ want+ø] [$_{VP}$ who [$_{V}$ want] [$_{TP}$ who [$_{T}$ to] win the race]]]]]

[136] The concept of cyclic Spell-Out ensures that the phonological cycle proceeds in parallel with the narrow-syntactic computation (i.e. there is in fact a single cycle in the grammar).
[137] This is the copy theory of movement already mentioned above.
[138] Alternatively, we may assume that copies are deleted prior to the application of Spell-Out by a copy deletion operation called Chain Reduction (CR), which, however, is not part of the narrow syntactic computation (it is rather an 'interface' operation - see Fujii 2005).
[139] See section 1. 3 for arguments that support an ECM analysis of structures of the form *want*+NP+*to*-infinitive.

As (24a) is an ECM structure, the matrix verb *want* has to raise from the head V position of VP to *v*, while *who* - as the object of the verb *want* - first moves to [Spec, VP], then raises further to [Spec, *v*P] position, which is an escape hatch for successive cyclic *wh*-movement. Finally, it raises to the matrix [Spec, CP] position. When the VP complement of the *v*P phase is spelled out, all traces are deleted within VP; however, the verb *want* is spelled out only on the matrix CP phase, at a point in the derivation when operations are no longer allowed to access items that belong to the matrix VP, given the PIC (which prohibits operations that cross phase boundaries). Accordingly, contraction of *to* onto *want* is blocked in (24a).

To sum up, it may be concluded that H's empirical argument which assimilates PRO to an NP-trace on the basis of contraction facts is not valid. This follows from the fact that the Case-marked trace account of contraction phenomena upon which H builds his argument is both conceptually and empirically problematic, as we have seen. Conversely, if we adopt a phase-based view of grammar and the assumption that only overt constituents block contraction, we can adequately account for contraction facts.

2. 2. 2. Control with Reflexive Verbs

H assumes that sentences involving reflexive predicates such as (23) above are formed by A-movement, whereby the moved NP receives both the internal θ-role (in the object position of the verb) and the external argument role (in subject position). Under this view, the structure of (23) is represented as follows (cf. Hornstein 1999):

(28) [$_{IP}$ Mary [past [$_{VP}$ Mary [wash Mary]]]]

According to H, (23) does not violate Case theory, as the NP *Mary* does not receive Accusative Case in the object position, given that the ability of *wash* to assign Case to its complement is optional; hence in (23) *wash* lacks a Case feature, and Case is checked on *Mary* only in the external argument position.

This analysis may pose a question as to how sentences such as (29a-b) below are formed.

(29) a. Mary washed herself.

b. John saw himself.

H argues that reflexives like *himself* in (29a-b) are the residues of movement, just like PRO. In other words, both PRO and reflexives are NP-traces, distinguished only by Case requirements: reflexives are formed when Case is obligatorily checked (just like in (29a-b) above). H has two reasons for such an analysis: (i) he wishes to eliminate Principle A of the Binding Theory, hence aims to reduce the distribution of anaphors to movement processes; (ii) once control is analysed as movement, local anaphors should also be reanalysed in these terms, given the empirical parallels between the two phenomena.

For illustration, consider how a reflexive like *himself* in (29b) is formed on H's assumptions: the computational system adds *self* to the derivation, which merges with *John* in the object position of the verb and is assigned Accusative Case. *John* then raises, and after its copy in the object position of *saw* is deleted, the pronoun *him* is added by the system to support the bound morpheme *self*.

Such a movement analysis of reflexives, however, raises both conceptual and empirical problems. First, H's analysis is incompatible with the Inclusiveness Condition (a principle which disallows the introduction of new features and objects in the course of the computation), as he assumes that neither the reflexive morpheme *self* nor the pronoun *him* are part of the lexical array that is selected to feed the derivation.

Second, H's account of reflexives requires that movement be allowed from a position that is Case-marked, violating the well-known generalization that arguments are not permitted to move from Case-marked positions. For illustration, consider (30) below (cf. Chomsky 1995):

(30) *We want John to strike t that the problems are insoluble.

As shown by (30), movement of *John* from its original Case-marked position is barred. Note that A-movement cannot take place from a Case position even when checking would otherwise fail in that position, as in (31):

(31) *She$_i$ seems to t$_i$ that Mary solved the problem.

Although Nominative Case cannot be checked on *she* in its original site, movement of the pronoun to a position where Nominative Case is assigned is not permitted (see Martin 2001).

Third, a movement analysis of local anaphors raises empirical problems as well. Consider (32) below:

(32) * John seems to himself that Mary solved the problem.

On H's theory, sentences such as (32) are predicted to be grammatical. In (32), no condition is violated under H's assumptions: *John* gets a θ-role as the indirect object of the matrix verb, then raises to the matrix subject position where it is assigned Nominative Case. The Accusative Case assigned to the (prepositional) indirect object position is checked by the reflexive morpheme *self*. So it seems that all relevant features are checked in the course of the derivation of (32), yet the sentence is not acceptable. Note that under H's movement hypothesis of local anaphors, the ungrammaticality of (32) cannot follow from θ-theoretic considerations: *John* does not lack a θ-role, as it is assigned a thematic role in the indirect object position. The derivation in (32) is also not blocked by the one in (33):

(33) It seems to John that Mary solved the problem.

The reason for this is that according to H, pronouns like *it* are not part of the lexical array, hence they are last resort grammatical formatives, inserted by the computational system only to save a derivation. H analyses movement – i.e. the formation of a reflexive – to be more economical than the insertion of a pronoun by the system (see section 2. 2. 4.). Consequently, (33) could only block (32) if the pronoun *it* was part of the array: in this case, merge of *it* would be preferred to the movement of *John* (for preference of Merge over Move, see Chomsky 2000). Let us now consider (34) below:

(34) John seems to himself to favour his son's drawing.

How can we account for the asymmetry between (32) and (34)? In other words, given the ungrammaticality of (32), why is (34) acceptable? The only possible answer seems to lie in θ-theoretic considerations: in (34), in contrast to (32), *John* can get a θ-role from the embedded predicate, raising from the infinitival subject position, whereas *John* in (32) cannot receive a θ-role. Therefore the ungrammaticality of (32) must follow from the fact that *John* lacks a θ-role, which in turn shows that *himself* in (32) cannot be formed by the movement of *John*. Consequently, a movement analysis of reflexives cannot be adopted, hence Principle A and the distribution of anaphors cannot be reduced to movement. If local anaphora cannot be reanalysed in movement terms, however, the analysis of control as movement also becomes dubious, given the empirical similarities between the two phenomena (which H himself notes, as we have seen above).

The problems raised by the analysis of (29a-b) also cast doubt on a movement-based approach to sentences with inherently reflexive predicates such as (23), which is proposed by H because he assumes (23) to be a case of control, with a PRO subject occupying the complement position of the reflexive verb *wash*. The assumption that (23) involves control in turn may pose a question with respect to the viability of a null Case account of the distribution of PRO, as in (23) PRO would occur in a position other than [Spec, TP]. But do we really

need to posit PRO in sentences like (23)? In fact, there is an alternative account of structures involving reflexive verbs, proposed by Reinhart and Siloni (R&S) (2004, 2005). R&S argue that despite the considerable crosslinguistic variation in the array of reflexive verbs, reflexivization is essentially the same phenomenon across languages, and is adequately explained by a single parameter – that is, reflexivization is either a lexical or a syntactic operation in languages. In English, reflexivization applies in the lexicon; it is a reduction operation that affects (namely, reduces) the internal θ-role of a transitive verb, thereby producing a reflexive predicate which can assign only an external argument role. The operation also involves Case reduction: it eliminates the (Accusative) Case checking ability of the verb. Although at the level of syntax only the external θ-role is realized in these structures, they require a reflexive reading at the level of semantic interpretation.[140]

What follows from this analysis is that if indeed no internal argument role is mapped onto the complement position of a reflexive verb, there is no motivation for postulating PRO in that position – this in turn would also have the conceptual advantage that we can maintain the close link observed between the Case- and θ-marking of a nominal.

On the whole, it seems that H's argument related to reflexive structures does not necessarily cast doubt on a null Case theory, whereas the proposed movement-based analysis of these constructions raises some problems that question the plausibility of such an account.

2. 2. 3. A Featural View of θ-roles

Let us now consider the stipulation that is crucial for H's theory to be tenable, namely, the featural view of θ-roles. H takes θ-roles to be features on predicates that are checked by a nominal expression, whereby the θ-feature of the verb is transferred to the nominal. A θ-feature on a predicate is able to license movement, hence (i) movement between θ-positions becomes possible, and (ii) (which follows from (i)) NPs are no longer restricted to bearing a single θ-role.

A-movement in OC structures is therefore motivated, on the one hand, by the θ-feature checking requirement of predicates, and on the other hand, by the Case-theoretic needs of the A-moved element. It is clear that such a movement analysis of control constructions is possible only if movement out of the embedded clause in control infinitivals is allowed, as is assumed by H in the case of OC structures. However, if we follow Chomsky (2000, 2001a) in positing that control structures are CP complements (in contrast to H's analysis, which takes them to be TPs – see Hornstein 1999), and also adopt a phase-based view of grammar, movement out of an embedded control infinitival does not seem to be an option.[141] For illustration, consider the derivation of (35a) (shown in (35b)) until the point when the embedded CP merges with its TP complement:

(35) a. John tried to leave.

b. $[_{CP}\ [_{C}\ ø]\ [_{TP}$ John $[_{T}$ to$]\ [_{VP}$ John leave$]]]$

John first merges with the embedded VP *leave*, thereby obtaining the *leaver* θ-role, then it raises to [Spec, TP] position to check the EPP feature of *to*. When the embedded CP is built, however, *John* has to raise from [Spec, TP] position to [Spec, CP], otherwise it is spelled out on the CP phase, and cannot raise further to the matrix [Spec, VP] position. Additionally,

[140] The analysis of reflexive derivations is indeed slightly modified in R&S 2005, where reduction involves the unification of the two θ-roles (an operation called 'bundling' by R&S) associated with the nominal in the external argument position at the level of interpretation, forming the complex θ-role [Agent-Theme]. The reflexive verb assigns this one unified complex θ-role to the nominal in subject position.

[141] See Rizzi (1982) or Radford (2004) for evidence of the CP status of control infinitivals.

given the PIC, *John* is only accessible to a probe outside CP if it moves to [Spec, CP] position, the only position from which movement out of the CP phase is possible. This movement, however, requires *John* to move from an A-position (Spec-TP) to an A'-position (Spec-CP), which is not possible according to Chomsky (2005a). Prohibiting movement from Spec-TP to Spec-CP is crucial for Chomsky to derive the A/A'-distinction (see Chomsky 2005a).

These considerations not only cast doubt on the plausibility of a movement-based approach to control, but may also prompt us to reconsider its conceptual foundations, in this case the assumption that θ-roles are features. The standard view of θ-roles, which is in effect the opposite of the featural view, is the configurational theory maintained by Chomsky (1995, 2000, 2001a, 2001b). Under the configurational approach, θ-roles are radically distinguished from morphological features like Case and agreement; that is, θ-roles are not taken to be formal features, hence they cannot be checked and cannot license movement (in contrast to morphological features). This conception of θ-theory ensures that there is no raising to a θ-position; rather, the θ-role of an argument "is determined by the position of first Merge" (Chomsky 2001a: 33), which means that θ-role assignment takes place in configurations of initial Merge of an argument NP. Specifically, Chomsky proposes the following principle (cf. Chomsky 2000: 103):

(36) Pure Merge in θ-position is required of (and restricted to) arguments.

(36) requires that arguments are first merged into a thematic position when they enter the derivation, and prohibits pure Merge (i.e. Merge that is not part of the operation Move) of arguments in non-θ positions – those positions can be filled by argument NPs only if they move there. For illustration, consider (37) below (cf. Chomsky 2000):

(37) a. * I expected [t to be a proof discovered].

b. I expected [there to be a proof discovered].

c. I expected [a proof to be discovered].

The ungrammaticality of (37a) follows from the principle in (36): the NP *I* cannot be selected from the lexical array to check the EPP-feature of *to*, as in this case the initial Merge-position of the argument (Spec-TP) would be a non-θ position.[142] The embedded [Spec, TP] position can only be filled by Merge of *there* (if the numeration contains the expletive), as in (37b), or by moving *a proof*, as in (37c).

To sum up: given the principle in (36), θ-roles cannot be assumed to be features. In other words, the configurational conception of θ-theory follows from (36): if an argument must receive its θ-role by pure Merge into a thematic position, then θ-roles are never assigned under Move, hence there is no movement to θ-positions, which in turn is inconsistent with a featural view of θ-assignment.

2. 2. 4. NOC Constructions

H argues that movement from the embedded infinitival subject position in NOC constructions is not allowed, hence what we find in these structures in that position is not an NP-trace left by A-movement. He assumes that the subject position in these infinitivals is filled by pro, the empty pronominal supposed to be licensed in the subject (and in some cases, the object) position of finite clauses in so-called null subject languages (NSL) like Italian. As we will see below, however, this analysis raises several problems.

[142] But see Hornstein, Nunes and Grohmann 2005 for an account of the ungrammaticality of sentences such as (37a) that does not resort to the θ-theoretic principle (36).

First, as noted by Brody (1999: 216), "if pro, the empty pronominal exists, then why not PRO, an empty anaphor" – in other words, assuming the existence of a phonetically null pronominal, pro, needs as much empirical evidence and support as positing the existence of a null anaphor, PRO.

Second, H's analysis crucially relies on the assumption that the subject position in control infinitivals is not a Case-checking position, as in this case movement from that position would be illicit in OC structures. That is why he holds that pro does not bear Case, but is able to check the relevant features of nonfinite I. This analysis, however, raises the conceptual problem already discussed in connection with PRO in section 1. 2.: pro is a nominal argument, receiving the θ-role of the embedded infinitival predicate, but in contrast with other nominal elements, it lacks Case under H's theory. H explains this peculiarity of pro by postulating that only phonetically visible elements require Case. If we examine the properties of pro in NSLs, however, we will find that it seems reasonable to assume otherwise. Consider the Italian example below (cf. Chomsky 1982):

(38) pro parla

'he is speaking'

We have to suppose that in (38), pro is inserted with a full set of interpretable φ-features, otherwise there would be no element that could value the uninterpretable φ-set of finite I (see Holmberg 2005). Under current assumptions, Nominative Case is assigned to pro by I as a result of this Agree relation between pro and finite I. The Case-marking of pro in (38) is also supported by the fact that it occupies a Case position (i.e. the Spec position of finite I, associated with Nominative Case assignment). These considerations not only suggest that H's explanation for why pro lacks Case in terms of phonetic visibility is inadequate, but also imply that a theory under which pro is Caseless cannot be correct.

Third, H's analysis is in conflict with the Inclusiveness Condition. H assumes that the licensing of pro in NOC structures can be modelled on a par with the phenomenon of do-support in English, that is, pro is not part of the lexical array selected to feed the derivation, rather it is a grammatical formative inserted by the computational system to save a derivation that would otherwise fail to converge. However, the assumption under which pro is not part of the array drawn from the lexicon violates the Inclusiveness Condition, which requires that "no new objects are added in the course of computation apart from rearrangements of lexical properties" (Chomsky 1995: 228). In other words, the Inclusiveness Condition bars the introduction of new elements during the derivation. H's analysis of NOC constructions is clearly inconsistent with such a principle.

Another point that could be made is the following: for the MTC to be tenable, it is crucial to exclude the licensing of pro in OC constructions, which in principle is allowed, as pro can check whatever features nonfinite I bears in control infinitivals. For this reason, H introduces the assumption that movement (taking place in OC structures) is more economical than construal processes that apply in NOC structures, hence when movement is allowed to occur, it must occur (resulting in OC), and only when movement is barred do we get a NOC structure. Accordingly, H suggests that the application of construal rules is a costly last resort operation, blocked by economy conditions if movement alone can yield a convergent result. But why would movement be more economical than construal processes? Such an assumption in itself – without independent empirical motivation and support – is only an ad hoc stipulation, which, however, is essential for H to make his hypothesis work: without this stipulation, there would be nothing to prevent the licensing of pro in OC structures. However, it is not reasonable to adopt such theses without independent empirical justification.

It is worth noting another point in connection with such economy considerations. The main thrust of the MTC is methodological: to reduce the inventory of empty categories, the types of rules, the class of assumptions, etc. that account for empirical facts in order to formulate simpler theories. This emphasis on methodological concerns motivates H to reconsider the phenomenon of displacement within grammar (i.e. that expressions are pronounced in one position and interpreted in another). Within GB, there are two rule types that are assumed to relate different positions within a sentence: movement processes and construal rules. H argues that economy considerations would favour only one rule type to account for displacement phenomena. In other words, we should explain all instances of displacement in a unified manner: either any case of displacement is movement, or any case of displacement is construal. As a result, we should dispense with either construal processes or movement, thereby allowing a reduction within the inventory of grammatical operations. Given that movement operations are supposed to be ineliminable, H claims that we should dispense with construal rules, which should rather be analysed as movement. But such a reduction is possible only if it is empirically viable, in other words, only if we can account for all the properties of the relevant constructions without resorting to construal processes in our explanation of the empirical facts. That does not seem to be so in the case of the MTC, however: to account for what kind of operation – if not movement – relates PRO (pro under H's analysis) to its antecedents *John* and *Mary* in the NOC structure in (39), H has to postulate the existence of some sort of construal operation that establishes a co-reference relation between these elements (cf. Hornstein 2003).

(39) John_1 told Mary_2 that PRO_{1+2} leaving each other was important to Bill.

If that is indeed the case, it is plausible to suppose that not only movement processes are ineliminable, as H assumes, but also construal rules. What else could explain the fact that the relevant expressions in (39) (i.e. PRO, *John* and *Mary*) are interpretively related to each other? In other words, there must be some kind of interpretive construal rule by which the relevant positions are related within a sentence if movement is barred. It seems, then, that the inventory of rule types is not reduced under the MTC: we cannot completely eliminate construal rules from the grammar, as we still need them to properly account for the properties of certain constructions. Put differently, if we need both rule types (construal and movement) to adequately explain empirical phenomena, we cannot reduce the class of rules only on the ground that such a move would be methodologically preferable. Consequently, grammar does not become more simplified in this respect on the MTC.

A further argument against the analysis of NOC structures can be formulated in connection with the last resort nature of pro under the analysis. H assumes that pro is not part of the lexical array and is only inserted by the computational system in order to save the derivation if all other grammatical alternatives fail. H, however, extends this analysis to overt pronouns, presumably for the following reasons:

(i) The null hypothesis is that all pronouns, either overt or phonetically null (like pro), have the same properties and behave similarly. This entails that if the empty pronominal pro is supposed to be introduced by the computational system only as a last resort option to yield convergence, in principle we have to assume that all pronouns can be analysed on a par with pro, hence that no pronouns are part of the numeration.

(ii) H wishes to eliminate the Binding Theory from UG, and once Principle A is dispensed with (see section 2. 2. 2.), then, methodologically, the status of Principle B also becomes dubious; in fact, once Principle A is done away with, Principle B should also be eliminated. This said, H aims to reduce not only the distribution of local reflexives

to the theory of movement (in that they are analysed as the residue of A-movement), but also that of pronouns: bound pronouns – just like pro – are supposed to occupy positions from which movement is not allowed.

(iii) Once local anaphors are analysed as the residue of A-movement (i.e. as spelled-out NP-traces), and pronouns are analysed on a par with pro, the fact that pronouns and anaphors are in complementary distribution follows straightforwardly. In other words, the relation of anaphors to pronouns can be modelled on the relation assumed to obtain between OC and NOC structures. This means that when movement is possible and Case is checked, what we get is a local reflexive; when movement is barred and Case is checked, the computational system inserts an overt pronominal. This analysis entails that the use of reflexives is more economical and hence preferred over the use of pronouns, as reflexives are formed via movement.

For the above reasons, H is forced to extend the last resort analysis of pro to overt pronouns, under which pronouns are "not true lexical expressions but are grammar internal 'elsewhere' formatives which are costly to use but licensed if needed" (cf. Hornstein 2001: 152). This analysis, however, as we will see below, fails both conceptually and empirically, and consequently, it also casts doubt on an analysis of pro in these terms.

Let us now consider H's analysis of (40a-c) below (cf. Hornstein 2001):

(40) a. Everyone likes himself.

b. *Everyone$_i$ likes him$_i$.

c. Everyone$_i$ likes him$_j$.

(41) {everyone, likes, assorted functional material}

The lexical array selected for the derivation of all the sentences is shown in (41) – H assumes that neither *self* nor the pronoun *him* is part of the lexicon, and their use is not permitted unless needed for the derivation to converge. H argues that the unacceptability of (40b) follows from the fact that the derivation of (40a) is more economical than that of (40b), hence is preferred by UG. The reason for this is that (40a) is formed by movement, which, on H's assumptions, is more economical than the insertion of a pronoun by the computational system in (40b). In other words, the convergence of (40a) blocks the derivation of (40b) by means of an economy restriction: when movement is allowed to occur (as in (40a)), it must occur. On this account, the derivation of (40c) should be blocked by the more economical derivation that yields (40a). That this is not so is explained by H with the suggestion that the use of deictic pronouns (as in (40c)) is allowed as they support a stress feature.

Under the account of (40a-b), the complementary distribution of local reflexives and bound pronouns is straightforwardly derived, just as in the case of the OC-NOC distinction. Nevertheless, it seems that bound pronouns and anaphors do not always occur in complementary distribution, as shown by the acceptability of both (42a-b) below:

(42) a. John$_i$ wondered which picture of himself$_i$ Bill liked.

b. John$_i$ wondered which picture of him$_i$ Bill liked.

Both the reflexive *himself* in (42a) and the bound pronoun *him* in (42b) are coreferential with *John* in the matrix subject position, and they occupy the same position within the two sentences. How can we account for the empirical facts in (42) on H's assumptions? In fact, given sentence pairs such as (42a-b), H's analysis of reflexives/bound pronouns fails empirically. The reason for this is the following: as we have seen above, H takes movement (i.e. the licensing of reflexives) to be more economical than the use of pronouns, which is

derivationally costly, hence when a derivation can converge without the use of a pronoun, its use is forbidden. From this it follows that the convergence of (42a) should block the derivation of (42b), as (42a) is a more economical derivation, and consequently, it is preferred by UG, whereas (42b) should be barred by economy conditions. Thus, H's theory predicts that (42b) is unacceptable, since it is blocked by the convergence of the more economical derivation in (42a). Additionally, an assumption under which the reflexive *himself* is formed by movement in (42a) also seems to be problematic, as it occurs inside a *wh*-island out of which in principle no constituent can be moved.

Indeed, the empirical data in (42) shows that we cannot assume that the licensing of reflexives is cheaper than the use of pronouns, which is therefore barred whenever movement (i.e. the formation of a reflexive) is possible. Consequently, pronouns cannot be supposed to be last resort expressions which are added to a derivation only at a cost, hence an assumption under which pronouns are not part of the lexical array is untenable. Methodologically, this casts doubt on a last resort analysis of pro – why would the properties of phonetically null pronominals be so diverse from that of overt pronouns? On the null hypothesis, we have to assume that the empty pronoun pro is part of the array if that is the analysis that seems to be correct in the case of overt pronouns. If this is so, however, then what prevents the licensing of pro in OC constructions? If pro is part of the lexical array that feeds the derivation, and is able to check whatever features nonfinite I bears in OC structures, then nothing precludes pro from being licensed in OC structures. This in turn entails that we have no evidence for the existence of movement within OC, hence that an analysis of control in terms of movement cannot be correct.

Besides this empirical failure, a conceptual problem also emerges in connection with the analysis of (40a-b): why would the use of reflexives be more economical than the use of pronouns? As we have seen above, the reason for this according to H is that reflexives are formed via movement, which is more economical than the insertion of a pronoun by the computational system, which is derivationally costly, as pronouns are last resort expressions that are not part of the lexicon, permitted only if required for convergence. But let us now consider the steps involved in the formation of the reflexive in (42a) (shown in (43a)) and the pronoun in (42b) (shown in (43b)):

(43) a. *himself*: - insertion of *self* by the computational system (*self* merges with *John*)

- movement of *John* (after which the copy is deleted)

- insertion of the pronoun *him* by the computational system (to support *self*)

b. *him*: -insertion of the pronoun *him* by the computational system

As (43a) shows, the licensing of the reflexive *himself* involves at least three steps. Movement is just one component of the process of reflexive-formation: for a reflexive to be licensed, the computational system has to introduce a non-lexical expression in the course of the derivation two times – once the reflexive morpheme *self* and later on in the derivation the pronoun *him*. In the formation of the pronoun *him* (shown in (43b)) only one step is involved: the system adds the non-lexical formative *him* to the derivation. As we can see in (43), the insertion of the pronoun *him* is only part of the process of reflexive-formation, i.e. only a suboperation applied in the licensing of the local reflexive, whereas in the case of the formation of the pronoun, that same single step suffices. Why, then, should we assume that the licensing of reflexives is more economical than the use of pronouns? Such an assumption appears to be quite problematic, as the licensing of reflexives clearly involves not only more steps than pronoun formation, but it also contains pronoun insertion only as a subpart. In fact, given these considerations, we may conclude that it is pronoun insertion which is more economical,

and not the formation of reflexives. Evidently, therefore, a theory under which the use of reflexives is more economical than the use of pronouns is untenable on conceptual grounds as well. What does that entail for H's analysis? Consider once again (40a-b), repeated here as (44a-b):

(44) a. Everyone likes himself.

b. *Everyone$_i$ likes him$_i$.

Given the above concerns, we surely cannot account for the contrast in (44a-b) in the manner proposed by H: the reason for the complementarity observed in the distribution of local anaphors and bound pronouns in sentence pairs such as (44) can no longer be explained in terms of an economy condition. How to account, then, for the complementary distribution of the reflexive and the pronoun in (44)? It seems that we can only resort to the binding principles. This implies that we cannot completely eliminate the Binding Theory, and cannot reduce it to the theory of movement: the binding principles must be retained in some form to account for the facts in (44a-b). But if the Binding Theory cannot be reanalysed in terms of movement processes, a reanalysis of control as movement also becomes dubious. This follows from the significant empirical parallels observed between local anaphors and OC which require that the two phenomena are treated parallelly, i.e. accounted for by the same principles and mechanisms of grammar.

To summarise, a close examination of H's analysis of NOC constructions has shown that it fails both conceptually and empirically. This in turn calls the MTC into question, as we have seen.

2. 3. Conclusion

Given the considerations above, we may conclude that the MTC has proven to be problematic from both theoretical and empirical perspectives. H's theory would require the adoption of assumptions that are unattractive from a conceptual perspective and are not motivated on strong empirical grounds (such as the last resort nature of pro or the conception of θ-roles as features). Additionally, the empirical arguments H proposes in favour of his movement theory (wanna-contraction phenomena and control with reflexive verbs) are not valid, as we have seen. The discussion above therefore indicates that H's PRO-less account in terms of movement has not been found to be preferable to a PRO-ful account of the syntax of control infinitivals. In fact, the problems that the MTC raises make it untenable.

Chapter 3: A Recent Minimalist Analysis of PRO

Let us now return to the alternative treatment of control structures: a PRO-based theory of control, possibly along with the null Case theory. This section discusses a recent minimalist analysis of the distribution-of-PRO problem that seems feasible.

The minimalist view of PRO that emerges from Chomsky 2000, 2001a, 2001b, 2005a, 2005b is as follows: the hypothesis is that derivations proceed by phases, which are taken to be CP and transitive *v*P. Control infinitivals are analysed as CP complements, hence as syntactic objects constituting a phase. Within the phase-based computation, the two phase heads C and *v* are the operative elements, driving all operations. It is the phase heads C and *v* that are the true locus of tense and agreement. This means that the phase head C is inserted from the lexicon with a full set of uninterpretable φ-features, interpretable Tense and uninterpretable EPP-features (or edge feature in Chomsky 2005a), whereas T lacks these features in the lexicon. These features can be detected on T only if it is selected by C. Consequently, T inherits these features from C, and it serves "as a probe only derivatively by

virtue of its relation to C" (cf. Chomsky 2005a: 20). In this sense, C and T operate as a unit in inducing agreement. The φ-features that T inherits from the phase head C are uninterpretable/unvalued (i.e. not assigned a value in the lexicon), hence they must be valued and deleted for the derivation to converge. Uninterpretable features are eliminated from the derivation by the operation Agree, which establishes a matching relation between the uninterpretable φ-set of a probe P and the interpretable φ-features of a goal G, the result of which is deletion of the uninterpretable φ-features of P and structural Case assignment of G by P. Turning to control T, its uninterpretable φ-set serves as the probe that seeks a matching element with interpretable φ-features that could value the uninterpretable φ-set of the T head. The φ-features of nominals are complete and interpretable, hence capable of valuing the uninterpretable φ-set of a functional head. By assumption, then, the probe in T locates the nominal element PRO in [Spec, VP] position as goal (which must be assumed to be part of the lexical array, in compliance with the Inclusiveness Condition). This entails that PRO bears a full set of φ-features so as to be able to value the uninterpretable φ-set of control T, which is also complete. For Agree to apply, both probe and goal must be active, i.e. both probe and goal must have an uninterpretable feature. In our case, control T is rendered active by its uninterpretable φ-set, whereas PRO has to be assumed to bear an uninterpretable Case feature that activates it (so that it can establish an agreement relation with a functional head). Under the operation Agree induced by φ-feature matching between PRO and control T, the uninterpretable φ-features of T are valued and erased. By hypothesis, structural Case on nominals is deleted under a φ-agreement relation between the nominal and a functional head, and the Case assigned to the nominal depends on the probe it agrees with: finite T is associated with Nominative Case, *v* with Accusative, and control T with null Case. Accordingly, as a result of the φ-agreement relation between PRO and control T, the uninterpretable Case feature of PRO is valued and deleted, i.e. PRO is assigned null Case by the T head. The Agree relation between PRO and control T presumably involves the raising of PRO to [Spec, TP] position, so that the uninterpretable EPP feature of T is deleted.

How is PRO excluded from raising and ECM infinitivals? Chomsky takes these structures to be TPs, selected by V. The T head of raising and ECM infinitivals is defective (T_{def}), i.e. not φ-complete, bearing only a partial φ-set, presumably only the uninterpretable feature [person] and an EPP feature. A defective probe is unable to assign a value to the uninterpretable Case feature of a nominal; only a φ-complete probe (like control T) is able to delete the Case feature of a matched goal. For this reason, successive-cyclic movement through the specifier position of T_{def} is possible: as the structural Case feature of the nominal reaching that position is not deleted, it remains active, hence it can undergo further movement and agreement. Movement of the infinitival subject in raising and ECM constructions to embedded [Spec, TP] position therefore leads only to partial agreement between the head T_{def} and the raised nominal, with T_{def} matching the nominal in only the feature [person]. The nominal in Spec-T_{def} position values and deletes the uninterpretable [person] feature of T_{def} and its EPP feature. The defective T head, however, cannot delete Case in the nominal, which raises further to matrix subject position in raising structures, or to [Spec, *v*P] position in ECM constructions where it is assigned Accusative Case. Given these considerations, PRO is not allowed to occur in raising and ECM infinitivals, as its Case feature could not be valued by T_{def}.

The above implementation of the null Case theory does not face the conceptual and empirical problems posed by Martin's conception of null Case assignment, and it seems to provide a principled explanation for all the structures under consideration.

Chapter 4: Conclusions and Some Residual Issues

In this paper, I sought to answer two closely related questions. Question (i) is concerned with the distribution-of-PRO problem: can we adequately account for all the distributional properties of PRO? For any theory which posits the existence of PRO, it is crucial to be able to do so. It seems that recent developments in minimalism provide us with the analytic tools and theoretical apparatus required to explain the distributional features of this empty element, yielding an account of PRO in terms of the null Case theory. This account does not face the empirical and conceptual problems of previous approaches to the distribution-of-PRO problem. Question (ii) is related to the fact that within MP, methodological concerns in terms of economy, simplicity, parsimony etc. have become standards in theory evaluation. These considerations are manifested in the MTC, which dispenses with PRO, and analyzes control as movement. An investigation of such an analysis has shown that it is both theoretically and empirically inadequate, hence must be rejected.

The two questions above raise the following issue: does PRO have a place in grammar? This paper argues for a positive answer. We may conclude that a PRO-ful null Case-theoretic analysis is superior to other alternative accounts of the syntax of control infinitival constructions.

As regards the minimalist theory of null Case, which serves to explain the distributional properties of PRO, let me suggest a possible avenue for further research. Recent work by Cecchetto and Oniga (C&O) (2004) points to empirical evidence from Latin and Italian which they claim proves that PRO carries standard structural Cases like Nominative or Accusative rather than a special null Case. Specifically, C&O argue that PRO always shares Case with its controller. They base their argument on Latin examples such as (45) and (46) below (cf. Cecchetto and Oniga 2004):

(45) Ego volo [PRO esse bonus].

I (NOM) want PRO to be good (NOM)

(46) Ego iubeo te [PRO esse bonum].

I order you (ACC) PRO to be good (ACC)

(47) Ego sum bonus.

I (NOM) am good (NOM)

In Latin, subject NPs overtly agree in Case, number and gender with the predicative adjective in copular structures, as shown by (47). C&O argue that as the adjective shows Nominative agreement in (45), PRO must be assumed to bear Nominative Case as well, agreeing in Case with the adjective. According to C&O, that is expected if PRO shares Case with its controller, which occupies the matrix subject position and bears Nominative Case. On similar grounds, PRO is supposed to be Accusative in (46), as its controller is Accusative as well – this is again justified by the overt Accusative agreement on the adjective. If that is the case, however, PRO cannot bear null Case, as two Cases cannot be checked on a single nominal.

An assumption under which PRO can carry the standard structural Cases raises several issues. Just to mention one: if PRO can bear Nominative or Accusative Case, what explains the unacceptability of sentences like (48a-b), where PRO occupies a Nominative and an Accusative Case position?

(48) a. *PRO saw John.

b. *John saw PRO.

(49) *John tried Mary to leave.

Note that the ungrammaticality of (48a-b) cannot be accounted for by suggesting that the φ-features of PRO are not full, hence cannot value the uninterpretable φ-set of finite T. This would pose a problem with regard to the valuation of the complete φ-set of the T head in control infinitivals. As regards (49), it is also predicted to be acceptable on the assumption that standard structural Case is checked in the subject position of control infinitivals.

We could also attempt to find an explanation for the empirical facts in (45) and (46) under the null Case approach. To overcome the problems posed for a null Case theory by these Latin examples, it may be suggested that the Phase Impenetrability Condition does not bar Case agreement between the adjective and the controller of PRO. This is possible if we suppose that the Case of the adjective is not valued through the usual Case checking procedure by a functional head, rather it only morphologically reflects the Case properties of the associated nominal element, hence its Case is valued through a concord relation with this nominal. We may suppose that PIC is not valid for Case agreement of this type. For this analysis to be tenable, we also have to assume that PRO does not block the agreement relation between its controller and the adjective by intervening between them. This seems reasonable, as PRO is not a PF object, hence, in principle, should not be a blocking category. There may be other options to deal with the empirical data above, but I leave these issues for further research.

References

Baltin, Mark and Leslie Barrett (2002). The Null Content of Null Case. Ms., New York University.

Boeckx, Cedric and Norbert Hornstein (2003). Reply to 'Control is not Movement'. *Linguistic Inquiry* 34: 269-280.

Boeckx, Cedric and Norbert Hornstein (2004). Movement under Control. *Linguistic Inquiry* 35: 431-452.

Bošković, Željko (1997). *The Syntax of Nonfinite Complementation: An Economy Approach.* Cambridge, Mass.: MIT Press.

Brody, Michael (1999). Relating Syntactic Elements. Remarks on Norbert Hornstein's 'Movement and Chains'. *Syntax* 2: 210-226.

Brody, Michael (2001). One More Time. *Syntax* 4: 126-138.

Cecchetto, Carlo and Renato Oniga (2004). A Challenge to Null Case Theory. *Linguistic Inquiry* 35: 141-149.

Chomsky, Noam (1981). *Lectures on Government and Binding: The Pisa Lectures.* Holland: Foris Publications. Reprint. 7th Edition. Berlin and New York: Mouton de Gruyter, 1993.

Chomsky, Noam (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Linguistic Inquiry Monograph Six. Cambridge, Mass.: MIT Press

Chomsky, Noam (1986a). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Chomsky, Noam (1986b). *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, Noam (1995). *The Minimalist Program.* Cambridge, Mass.: MIT Press

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. Cambridge, Mass.: MIT Press.
Chomsky, Noam (2001a). Derivation by phase. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press
Chomsky, Noam (2001b). Beyond Explanatory Adequacy. *MIT Occasional Papers in Linguistics* 20: 1-28. Cambridge, Mass.: MITWPL.
Chomsky, Noam (2005a). On Phases. Ms., MIT.

Chomsky, Noam (2005b). Three Factors in Language Design. *Linguistic Inquiry* 36: 1-22.
Culicover, Peter and Ray Jackendoff (2001). Control is not Movement. *Linguistic Inquiry* 32: 493-512.

Frampton, John, Sam

Gutmann, Julie Legate, and Charles

Yang (2000). Remarks on 'Derivation

by Phase': Feature Valuation,

Agreement, and Intervention. Ms.,

MIT.

Fujii, Tomohiro (2005). Cycle, Linearization of Chains, and Multiple Case Checking. In *Proceedings of Console XIII*, ed. Sylvia Blaho, Luis Vicente and Erik Schoorlemmer, 39-65. Leiden: Leiden University Centre for Linguistics.
Hazout, Ilan (2004). Long-Distance Agreement and the Syntax of for-to Infinitives. *Linguistic Inquiry* 35: 338-343.
Holmberg, Anders (2005). Is There a Little Pro? Evidence from Finnish. *Linguistic Inquiry* 36: 533-564.
Hornstein, Norbert (1998). Movement and Chains. *Syntax* 1: 99-127.
Hornstein, Norbert (1999). Movement and Control. *Linguistic Inquiry* 30: 69-96.
Hornstein, Norbert (2000). On A-chains: A Reply to Brody. *Syntax* 3: 129-143.
Hornstein, Norbert (2001). *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.
Hornstein, Norbert (2003). On Control. In *Minimalist Syntax*, ed. Randall Hendrick, 6-81. Malden, Mass.: Blackwell.
Hornstein, Norbert, Jairo Nunes, and Kleanthes K. Grohmann (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.
Landau, Idan (2003). Movement out of Control. *Linguistic Inquiry* 34: 471-498.
Lasnik, Howard, and Mamoru Saito (1991). On the Subject of Infinitives. In *Papers from the 27th Regional Meeting of the Chicago Linguistic Society, Part I: The General Session*, ed. Lise M. Dobrin, Lynn Nichols, and Rosa M. Rodriguez, 324-343. Chicago: Chicago Linguistic Society, University of Chicago. [Reprinted in Howard Lasnik, *Minimalist Analysis*, 7-24. Oxford: Blackwell (1999).]
Manzini, Maria Rita, and Anna Roussou (2000). A Minimalist Theory of A-Movement and Control. *Lingua* 110: 409-447.
Martin, Roger (2001). Null Case and the Distribution of PRO. *Linguistic Inquiry* 32: 141–166.

O'Neil, John (1997). Means of

Control: Deriving the Properties of

PRO in the Minimalist Program.

Doctoral dissertation, Harvard

University.

Pesetsky, David, and Esther Torrego (2001). T-to-C Movement: Causes and Consequences. In *Ken Hale: A Life in Language*, ed. Michael Kenstowicz, 355-426. Cambridge, Mass.: MIT Press

Radford, Andrew (2000). T-contraction in a Phase-based Theory of Grammar. *Essex Research Reports in Linguistics* 33: 21-32.

Radford, Andrew (2004). *Minimalist Syntax: Exploring the Structure of English*. Cambridge: Cambridge University Press.

Reinhart, Tanya and Tal Siloni (2004). Against the Unaccusative Analysis of Reflexives. In *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*, ed. Artemis Alexiadou, Elena Anagnostopoulou, and Martin Everaert, 159-180. Oxford: Oxford University Press.

Reinhart, Tanya and Tal Siloni (2005). The Lexicon-Syntax Parameter: Reflexivization and Other Arity Operations. *Linguistic Inquiry* 36: 389 – 436.

Rizzi, Luigi (1982). *Issues in Italian Syntax*. Dordrecht: Foris Publications.