A minimalist theory of A-movement and control^{*}

M. RITA MANZINI & ANNA ROUSSOU

Abstract

In this article, we point out some problems in the theory of A-movement and control within Principles and Parameters models, and specifically within the minimalist approach of Chomsky (1995). In order to overcome these problems, we motivate a departure from the standard transformational theory of A-movement. In particular, we argue that DPs are merged in the position where they surface, and from there they attract a predicate. On this basis, control can simply be construed as the special case in which the same DP attracts more than one predicate. Arbitrary control reduces to the attraction of a predicate by an operator in C. We show that the basic properties of control follow from an appropriate *Scopal* version of Chomsky's (1995) *Last Resort* and MLC and from Kayne's (1984) *Connectedness*, phrased as conditions on the attraction operation, or technically ATTRACT. Our approach has considerable advantages in standard cases of A-movement as well, deriving the distribution of reconstruction effects at LF and of blocking effects on phonosyntactic rules at PF.

1 Classical theories of A-movement and control and their problems

According to Chomsky (1981, 1982) the combination of ±anaphoric, ±pronominal features yields four different types of empty categories: A'-trace, A-trace, *pro* and PRO. PRO is identified with the +anaphoric, +pronominal empty category, which is subject to

^{*} Earlier versions of this paper were presented at the Conference on *Syntactic Categories* (Bangor, June 1996), the Girona Summer School in Linguistics (July 1996), the Research Seminar in York (January 1997), UCLA (February 1997), GLOW 20 (Rabat, March 1997), Scuola Normale Superiore (Pisa, March 1997), Lezioni di Dottorato, Universitá di Firenze e Siena (April 1997), MIT and the University of Maryland (October 1997). We are particularly grateful to the following people for comments and suggestions: Maya Arad, Elabbas Benmamoun, Valentina Bianchi, Bob Borsley, Misi Brody, Guglielmo Cinque, Annabel Cormack, Caterina Donati, Ken Hale, Ursel Luhde, Anoop Mahajan, Bernadette Plunkett, Luigi Rizzi, Ian Roberts, Heloisa Salles, Leonardo Savoia, Dominique Sportiche, Tim Stowell, Peter Svenonius, Anna Szabolcsi, and George Tsoulas.

both Principles A and B of *Binding Theory*. If both of these apply to a given instance of PRO, a contradiction arises, since *Principle A* requires PRO to be bound locally, i.e. in its governing category, and *Principle B* requires it to be free in the same domain. Therefore PRO can only be found in positions for which the notion of governing category is not defined, i.e. in ungoverned positions. In this way, no contradiction arises under *Binding Theory*. This result is known as the *PRO Theorem*. Thus under Government and Binding (GB), the distribution of PRO is predictable on the basis of independently motivated assumptions about empty categories and their binding properties.

Within the framework of Chomsky (1995) the theory of empty categories just sketched is effectively abandoned. In particular, traces of A'-movement are construed as copies of the moved material, rather than as -anaphoric, -pronominal empty categories, i.e. Rexpressions, as in the previous GB framework. Evidence from reconstruction phenomena provided by Chomsky (1995) favors this view. Similarly for Chomsky (1995), A-movement leaves traces that are copies, though we shall return shortly to the lack of evidence for reconstruction in these cases. Thus the classification of empty categories on the basis of ±anaphoric, ±pronominal features collapses. Independent developments in the theory also undermine the notion of government, as well as the idea that there exists a *Binding Theory* module within the grammar. In particular, properties of lexical anaphors can be derived via movement, as argued by Pica (1987), Reinhart and Reuland (1993). Assuming that traces are copies nevertheless leaves pronominal empty categories, namely pro and PRO, to be accounted for. In the present paper we will leave pro and the parametrization issues it implies aside, and we will therefore concentrate on PRO; for a compatible theory of pro see Manzini and Savoia (forthcoming).

Chomsky and Lasnik (1995) seek to maintain the basic descriptive generalization according to which PRO is found only in the subject position of non-finite sentences. According to their analysis PRO is associated with a special type of Case, called null Case. Since null Case is checked by only non-finite I, it follows that PRO is found only in the Spec of this latter category. Furthermore, because null Case is like any other Case from the point of view of Move, we also derive that PRO behaves like a lexical argument under A-movement. Thus in (1) PRO is generated in the internal argument position of *kill* and raises to [Spec, *to*] to check its (null) Case:

(1) John tried [PRO to be [killed PRO]]

A standard set of predictions about the distribution of PRO that followed from the *PRO Theorem* of Chomsky (1981), now follow from the stipulation of null Case. In particular

PRO cannot be found in object position nor in the subject position of tensed sentences, as in (2)-(3) respectively:

- (2) *John persuaded PRO
- (3) *John believes that PRO will eat

It seems to us that this approach has a number of problems. To begin with, null Case and PRO are only ever seen in connection with one another. In other words, there is no independent way of establishing the existence of either. Thus null Case does not appear to provide a genuine explanation for the distribution of PRO, but rather a way of stating the descriptive generalization concerning its distribution. This represents a step back with respect to the GB framework, where the distribution of PRO did indeed follow from the interaction of independently motivated assumptions, even if their abandonment seems now more than justified.

What is more, the minimalist approach to control inherits a number of problems from the GB analysis, not surprisingly since it essentially adapts it. In particular, as pointed out by Chomsky (1981), the distribution of PRO represents only one of the empirical facets of the theory of control. Even assuming that this can be accounted for in terms of null Case, the distribution of the antecedents for PRO, i.e. control proper, remains to be explained. The basic descriptive generalization concerning control, which appears to be accepted by Chomsky (1981), is Rosenbaum's (1967) *Minimal Distance Principle* (MDP), according to which the rule of *Equi-NP-Deletion*, or the empty category that succeeds it, namely PRO, is controlled by the closest available antecedent, where 'closest' is defined in terms of c-command. Thus in (4), the matrix object, *John*, obligatorily controls PRO; in (5) the intermediate subject, *John*, rather than the matrix one, does:

- (4) Mary [persuaded John [PRO to eat]]
- (5) Mary thinks that [John expected [PRO to eat]]

A small class of English verbs, including *promise*, provide a counterexample to this generalization, as exemplified in (6), where the matrix subject rather than the matrix object controls PRO:

(6) I promised John [PRO to eat]

This potential problem may be solved, however, by an articulate approach to the internal argument structure of *persuade*-type and *promise*-type verbs, of the sort proposed for instance by Larson (1991).

Within the GB framework of Chomsky (1981), there is no obvious way to reduce the MDP to some independently needed principle. Thus Manzini (1983) proposes a general resystematization of control theory, whereby PRO is taken to be a pure anaphor and its distribution determined by the lack of Case, rather than by the PRO Theorem. As an anaphor PRO is then subjected to a (suitably modified) version of Principle A. This theory can predict locality effects of the type in (5), though it invokes pragmatics to explain the distribution of control within the argument structure of the verb, as in (4) vs. (6). This latter assumption can however be dispensed with, given precisely the VP-shell conception of Larson (1991) just referred to. As we have already mentioned, within the minimalist framework a number of authors in turn take anaphoric binding to reduce to Move. Therefore any attempt at reducing control to binding would appear to translate into a reduction of control directly to movement. This step appears to be attractive to the extent that the locality condition on control, the MDP, is based on exactly the same notion of closeness as the basic locality condition on movement, i.e. the Minimal Link Condition (MLC). But such a radically minimalist approach to control theory is not explored by Chomsky and Lasnik (1995).

In recent years we know of at least three different proposals that explicitly subsume control under movement. Martin (1996) maintains along with Chomsky and Lasnik (1995) the existence of a null Case-marked PRO which, however, is subject to abstract movement exactly like a reflexive in the models of Pica (1987), Reinhart and Reuland (1993) mentioned above. O'Neil (1995) and Hornstein (1996) subsume control under overt DP-movement; thus the controller is generated in thematic position within the control sentence and then moves to its surface position through a higher thematic position. All of these theories radically differ from proposals, such as Bresnan's (1982), under which there is no syntactically represented subject at all for non-finite sentences. Bresnan (1982) goes on to argue that control into non-finite sentences is the product of a lexical, rather than a syntactic, operation. This approach denies some of the basic tenets of Principles and Parameters theory, in which the lexicon reduces to a list of primitive terms, and all operations relating them are carried out by the syntax.

The approach that we intend to take in turn differs from all the above, involving first a departure from the standard transformational theory of A-movement. We therefore turn to A-movement next.

Proposing that a DP can move through two thematic positions before reaching its Case position, as Hornstein (1996) does, amounts to a violation of the well-known generalization of Chomsky (1981) whereby movement is never to a theta-position. As

Hornstein (1997) carefully explains, this generalization is based on the GB concept of Dstructure, whose elimination represents one of the qualifying features of minimalism. Thus the fact that D-structure is by definition a "pure representation of GF-theta's" (Chomsky 1986) forces the conclusion that each theta-role is satisfied by an argument at the point of Merge; this in turn forces the postulation of PRO for the cases where lexical arguments are not available. Once the notion of D-structure is abandoned, this fundamental piece of theoretical justification for PRO becomes lost.

To be more precise, Chomsky (1995) holds on to the idea that there is a one-to-one match between arguments and theta-positions, but does so through an extra assumption. While properties of lexical items are in general lexical features, theta-roles are to be construed differently, namely as a configurational relation between a head and its specifier/complement. This configurational theory of theta-roles has the effect of barring movement to a theta-position exactly as Chomsky's (1981) *Theta-criterion* at D-structure does. But as Hornstein (1997) also points out, if theta-roles are instead construed as features, the null assumption under minimalist theory, then nothing prevents them from acting as attractors for arguments already merged; this in turn opens the way for an overt movement theory of control. What Hornstein (1997) fails to notice is that these same conclusions have implications that go far beyond control. It is these implications that interest us here.

If arguments and theta-roles need not be matched at the point of merger, it becomes possible to assume that arguments are generated not in thematic position but directly in the position where they surface, provided suitable means can be found to connect them to their theta-roles. If theta-roles are indeed features, the simplest such means suggested by the minimalist theory of Chomsky (1995) is Move-F. To illustrate these points, let us consider the simple sentence in (7). In Chomsky's (1995) terms, this is associated with a derivational stage of the type in (8a), irrelevant details omitted; by movement of the VP-internal subject to [Spec, I] the final string in (8b) is obtained:

- (7) John called.
- (8) a. $[_{VP}$ John called]
 - b. [_{IP} John I [_{VP} John called]]

In the alternative terms that we are suggesting here, the derivation of (7) would take a form much more similar to (9). In (9) the subject is merged directly in [Spec, I]. A theta-feature, notated provisionally as q, is moved from V to the IP domain, establishing the relevant thematic interpretation for the subject:

(9) $[_{\mathbb{P}} \text{ John } q\text{-}I [_{VP} \text{ called}(q)]]$

A number of problems are raised by the derivation in (9), but before considering them, we wish to establish the main empirical reasons that favor the covert movement approach in (9) over the overt movement approach in (8).

It is fairly obvious that no considerations pertaining to the operation of movement itself can distinguish the analyses in (8) and (9), since the same constraints operate both on phrasal and F-movement. The only difference between (8) and (9) concerns therefore the nature of the copy, or trace, left behind by the application of movement, which is fully phrasal in (8) and a mere feature in (9). On this basis, the predictions of the DP-movement and the F-movement models empirically differ with respect to reconstruction. Remember that since movement is Copy and Merge, DP-movement to [Spec, I] leaves behind a full copy of the DP in thematic position, exactly like *wh*-movement to [Spec, C] leaves behind a full copy of the *wh*-phrase. In the case of *wh*-movement this copying of phrasal material has been empirically justified by Chomsky (1995) on basis of the fact that it gives rise to reconstruction effects. Thus in a sentence of the type in (10) the anaphor *himself* can be interpreted as bound either by *John* or *Bill*:

(10) John wonders [which pictures of himself [Bill saw which pictures of himself]]

Under reasonable assumptions about anaphoric binding, the reading under which *himself* is anaphoric to *John* corresponds to the construal of the *wh*-phrase in its derived position; the reading under which *himself* is anaphoric to *Bill* corresponds to the construal of the *wh*-phrase in the trace position.

Chomsky (1995) himself argues, on the other hand, that DP-movement never gives rise to reconstruction. He accounts for it by assuming that reconstruction is a by-product of Operator-Variable interpretation at LF. Thus consider a sentence of the type in (11) under the standard DP-movement derivation:

(11) *Each other seem to them [each other to work]

The ungrammaticality of (11) requires that the conditions concerning anaphoric binding are computed with respect to the derived position of *each other*. If *each other* could reconstruct, then we would expect *them* to be able to bind it, giving rise to a well-formed reading. Crucially, our approach to raising and in general to A-movement predicts data of the type in (11) without need for any additional stipulation about reconstruction. Indeed the derivation of (11) takes the alternative form in (12):

(12) *Each other [q-I seem to them [to [work (q)]]]]

In (12) *each other* is merged directly into the [Spec, I] position in which it surfaces, and the only movement that takes place is that of q to its checking domain. Therefore we do not expect any effects connected with the reconstruction of *each other* to its q-position. We shall return to potential counterexamples in section 6.

The other fundamental argument in favour of the F-movement approach over the DPmovement one is represented by the contrast between the behaviour of A'-movement traces and A-movement traces with respect to syntactically conditioned phonological rules. While the reconstruction test in (10)-(11) shows that the presumed DP-copy plays no role at LF, and therefore is best abandoned at this interface level, the phonosyntactic rule test addresses the relevance of the presumed DP-copy at the other interface level, i.e. PF. As is well-known, a number of scholars have argued that A-movement traces, as opposed to A'-movement ones, do not block PF processes. Not surprisingly some of them, notably Postal and Pullum (1982), have concluded from this that DP-traces do not exist. However, the inexistence of DP-traces does not imply the absence of Amovement. We effectively agree that DP-traces do not exist; but A-movement does exist within the theory being proposed in the form of F-movement.

The blocking of phonosyntactic rules such as *wanna* contraction represents a classical argument in favor of the 'reality' of A'-movement traces. As the argument goes, the sentence in (13) has two possible interpretations, corresponding to the two structures in (14). In (14a) the trace of *wh*-movement is interpreted as occupying the object position of the embedded sentence; in (14b) it is interpreted as occupying the embedded subject position:

- (13) Who do you want to call?
- (14) a. [who do you want [PRO to call who]]
 - b. [who do you want [who-to call]]

If *want to* is contracted to *wanna*, as indicated in (15), the two interpretations reduce to just one, namely (14a); (14b) is blocked under contraction:

(15) Who do you wanna call?

A possible analysis of these data is that the phonological process responsible for the contraction of *want to* to *wanna* is sensitive to the presence of lexical material between the two in the shape of a *wh*-trace; hence it cannot apply in (14b). If so, a natural explanation for the lack of blocking effects in (14a) is that PRO, or whatever else PRO reduces to, is not lexical. Consider then A-traces, as seen for instance in (16). We take it

that in (16) *have to* does not define a control environment, i.e. one involving a theta-role both in the matrix and in the embedded clause, but rather a raising environment:

(16) a. I have to leaveb. I hafta leave

The lack of blocking effects in (16b) argues that what is traditionally construed as a DP-trace, as in (17a), is better construed as an F-trace, as proposed by the present theory and illustrated in (17b). Indeed (17b) suggests an immediate explanation as to why there are no blocking effects on contraction, namely that F-traces do not count as lexical, while (17a) is *prima facie* identical to (14b); we will return to these points shortly:

(17) a. [I have [I to leave]]b. [I q-have [to leave (q)]]

There also appear to be conceptual reasons that favor the F-movement construal over the phrasal construal of A-movement. To begin with, the mechanism illustrated for sentence-internal raising in (9) extends to long-distance raising across sentence boundaries of the type illustrated in (18):

(18) John q-I [seems [to work (q)]]

In (18), the q-role associated with the embedded verb *work* moves in one step to the checking domain of the DP *John* which is merged directly in the [Spec, I] position where it surfaces. From this construal of raising, in which the embedded [Spec, I] is not involved, we derive a first significant consequence for the grammar in general. The embedded subject position in raising contexts is the only one where according to Chomsky (1995), a D feature can be seen independently of Case. Indeed in Chomsky's (1995) analysis the *Extended Projection Principle* (EPP) requires the presence of a D-feature on the infinitival I, which forces a DP to pass through its Spec, as in the derivation outlined in (19). At the same time, the fact that the DP can and must move on to the matrix [Spec, I] is due to the impossibility of the infinitival I to check its Case:

(19) John I [seems [John to [John work]]

The analysis in (18) proposed here crucially differs from the analysis in (19), in that it does not imply movement through the Spec of the infinitival I, not even limited to the q-feature. This means that within the present theory we dispense with any D-feature

associated with a non-finite I. There is therefore no position within the present theory where the D-feature and the Case feature can be seen independently of one another. Thus the two features can be unified, eliminating the considerable redundancy between them in the grammar. We should also stress that according to Chomsky (1995), Case is the only feature which is non-interpretable on both the attractor and the attractee, casting doubts on its existence independently of the redundancy just noted (Roberts and Roussou 1997, Chomsky 1998). See Manzini and Savoia (forthcoming) and Roberts and Roussou (forthcoming) for alternative theories of the EPP compatible with the analysis proposed here.

To summarize, there is no evidence either at the LF or at the PF interface that Copy and Merge of lexical material, i.e. classical phrasal movement, is involved in Amovement contexts. This therefore seems to suggest that lexical arguments are generated directly in the position where they surface and an appropriate operation connects them to a theta-feature associated with the predicate. The appropriate operation would be Move-F in Chomsky's (1995) framework. In the following section we will, however, propose that there are conceptual and empirical reasons that favor an even more radical approach to A-movement.

2 The theory of A-movement

So far, we have simply suggested a reformulation of A-movement from Copy and Merge of a whole DP, i.e. DP-movement, to Copy and Merge of a single feature, i.e. F-movement. We have argued for this alternative on the basis of two sets of empirical arguments, namely the lack of reconstruction effects at LF and of blocking effects on phonosyntactic rules at PF. In the first case it is obvious that in the absence of a lexical copy we should not expect reconstruction effects. However, in the second case, one could argue that we should be able to detect some effects of the copy of the feature, since F-movement is Copy and Merge after all. In other words, the fact that there is an asymmetry between lexical copies and feature copies at PF has to be stipulated. This suggests to us that F-movement does not involve Copy and Merge of any sort and is therefore not movement in the sense defined by Chomsky (1995).

In the light of these considerations let us go back to the derivation in (9). Let us assume that *John* has some feature that needs to be checked by *called* and specifically by its theta-feature. If *John* and *called* were in the same checking domain, i.e. in practice in a head-Spec configuration, then checking could take place directly. The intervention of movement, if only of a feature as in (9), is due to the fact that *John* and *called* are not in the same checking domain. Suppose, on the contrary, we give up the notion of checking

domain; then whatever operation needs to take place between *John* and *called* can do so directly, as tentatively indicated in (20) by means of the italics:

(20) $[_{\mathbb{P}} John I [_{\mathbb{VP}} called(q)]]$

Chomsky (1998) explicitly recognizes the stipulative nature of the checking domain. Furthermore, he construes Copy and Merge, i.e. movement, which applies to phrases, as a separate operation from Attract, which applies to features. Thus the operation in (20) is a counterpart of the operation Attract in Chomsky's (1998) system; contrary to his analysis, on the other hand, we want to suggest that Copy and Merge is to be dispensed with. For present purposes we shall simply adopt the name ATTRACT for the operation in (20); this term is meant to capture both the fact that the operation plays the same role as Chomsky's Attract in the general economy of the grammar, and that it differs from it at least in one respect. Attract as defined by Chomsky (1998) still involves the merger of features of one lexical item (the attractee) into another (the attractor). The operation in (20) is conceived as preserving the integrity of both lexical items; the reasons for this lie in feature theory, a topic to which we shall return shortly.

Although we agree with Chomsky (1998) that the notion of checking domain has no theory-independent motivation, in Chomsky's (1995) system there is one more locality condition, namely the MLC, which is defined for the rule of movement itself. The MLC has very clear empirical consequences, namely the minimality effects discussed by Rizzi (1990) and Chomsky (1995). In order to capture these effects within our framework, we propose that the MLC applies to the operation ATTRACT. The original formulation of the MLC of Chomsky (1995) is based on the intervention of potential attractees. In other words, movement of attractee a to attractor g is blocked if there is an attractee b which is closer to g. This condition is reproduced in (21) referred now to ATTRACT:

(21) MLC

g ATTRACTS a only if there is no b, b closer to g than a, such that g ATTRACTS b.

(21) is trivially satisfied in (20), by *John* and q, since there is no other candidate thetarole and DP argument in the sentence. In the next section, we will motivate a modification of the MLC which will allow us to deal in a simple and elegant manner with more complex cases, including control. What matters to us at this point, however, is simply to establish the general feasibility of our approach.

While the MLC takes care of what Cinque (1991) calls weak islands, another condition on movement is needed to account for strong islands. We follow Manzini

(1992, 1994) in concluding that the best theoretical account of the empirical data is provided by Kayne?s (1984) *Connectedness Condition*. In accordance with our formulation of the MLC in (21), we take *Connectedness* to hold directly of ATTRACT as well. In (22) we maintain Kayne's (1984) term of g-projection, though the relevant notion is defined with reference only to complementation, rather than government, as shown in (22b). Furthermore, although Kayne (1984) postulates a directional asymmetry, we do not, in line with later proposals by Bennis and Hoekstra (1984), Longobardi (1984). See also Manzini (in press) for a discussion of how the notion of g-projection could actually be subsumed under a reformulation of Merge:

(22) Connectedness Condition

- a. Let b ATTRACT a. Then b together with a and the g-projections of a must form a connected subtree.
- b. g is a g-projection of a if it is a projection of a, or a projection of some d such that a g-projection of a is a complement of d.

In the case of (20), the condition in (22) is satisfied by *John* and *q*. Indeed the g-projection set of V includes VP and IP. *John* in [Spec, I], V and the g-projections of V, namely VP and IP, form a connected subtree, in this case the whole sentential tree.

We can now consider questions, carefully skirted so far, concerning the nature of the feature theory that underlies the attraction of theta-roles by DP arguments. First, we need to specify exactly what kind of feature q is. Second, we need to specify which feature of the DP argument ATTRACTs q. Third, we need to specify what the properties of ATTRACT are with respect to notions such as strength and interpretability introduced by Chomsky's (1995) grammar. We shall then be able to trace our steps backward to a resolution first of the control problem and then a number of residual problems concerning A-movement itself.

Consider first the role played by such general notions as Chomsky's (1995) interpretability and strength in the present theory. Remember that according to Chomsky (1995) there are essentially two cases in which a feature attracts another feature. The first case arises when the attractor is strong; in this case attraction is overt and independent of the other properties of the features involved, in particular their interpretability. The other case in which a feature acts as an attractor, independently of its strength, is when it is non-interpretable; in this case another feature moves to check it, even if only abstractly.

Chomsky (1998) criticizes the notion of strength on grounds of its conceptual complexity, since it is construed as a feature of a feature. This, however, is not to say that strength does not have empirical relevance. Following Roberts and Roussou (1997)

and Manzini and Savoia (forthcoming), we assume that strength is in fact to be understood as an instruction to lexicalize a certain feature. Since lexicalization is tipically done by Merge within the present framework, the issue of strength is largely irrelevant for ATTRACT. For instance, in (20) we take it that I is associated with a D feature that needs to be lexicalized; in a non-null subject language like English the lexicalization is accomplished by merging a full DP, *John* in (20), in [Spec, I].

The issue of interpretability, on the other hand, requires more discussion. It seems to us that maintaining Chomsky's (1995) theory of interpretability within the present framework amounts to nothing more than a terminological gimmick. Thus we could associate every D with a non-interpretable q feature, which would ATTRACT an interpretable q feature associated with V; but it is certainly simpler to maintain that every D directly ATTRACTs q. The intermediate step of postulating a non-interpretable q feature is entirely *ad hoc*. In turn, if D directly ATTRACTs q, this amounts to saying in Chomsky's (1995) terminology that checking is well-defined between two interpretable features. This conclusion, and in fact the stronger conclusion that all features in the grammar are interpretable, has been independently motivated in the literature, notably by Brody (1995), Roberts and Roussou (1997) and we shall accept it here without further discussion. If so, checking not only can, but also must involve interpretable features.

Summarizing so far, in a simple sentence of the type in (20) the intepretable D feature of *John* ATTRACTs the interpretable q feature of V. As we already saw, the basic locality principles, i.e. the MLC and *Connectedness*, are directly defined for the operation ATTRACT and obviously satisfied in (20). Before considering more complex cases than (20) and eventually control, we need to go back to the question of what the nature of the q feature is. More specifically, we can ask whether the present theory of A-movement really chooses between the feature-based theory of theta structure that we have adhered to so far and the configurational view advocated by Chomsky (1995). As it turns out, ATTRACT, contrary to F-movement, does not really choose between these two options.

Hale and Keyser (1993) argue compellingly that the notion of theta-role is deeply flawed. In particular, it is easy to show that it is not restrictive enough, since there is no principled reason why the basic repertory of Agent, Theme, Source, Goal, etc. should not be further expanded; nor is there any reason why two or more of those specifications could not be combined; and so on. Face to these problems, Hale and Keyser (1993), followed by Chomsky (1995) among others, propose that theta-roles should be replaced by elementary predicate-argument configurations within a VP-shell. Thus instead of (20), where V is associated with a feature q, we have a structure like (23), where it is the presence of a single VP layer that indicates the nature of Theme of the one argument:

(23) $[_{\mathbb{P}} John I [_{VP} left]]$

The italics in (23) indicate that the operation ATTRACT ought to proceed as before. Indeed the theory of ATTRACT that we have now defined does not require the two features involved in this operation to have any special properties in relation to one another. Therefore there is no reason why the operation ATTRACT could not be defined directly between D and V. In this way, we reconstruct the basic intuition of the configurational theory of theta-roles, namely that they correspond to a relation directly defined between an argument and a predicate. At the same time, it becomes unlikely, if not impossible, that the attraction of V by D should involve merger of V under D, as under Chomsky's (1998) Attract. Hence we confirm our conception of ATTRACT as maintaining the lexical integrity of the items involved.

The analysis in (23) can also be extended to predicates with two or more arguments. In the pure VP-shell notation employed by Hale and Keyser (1993) a lower V represents the basic predicate of the sentence, and a higher V represents an abstract causative predicate in eventive verbs. In the rather more articulated conception of the VP-shell drawn by Chomsky (1995), the higher V corresponds to the light verb v, as in (24):

(24) [v[V]]

Suppose we adopt the structure in (24). Sentences which embed (24), such as (25), raise the problem of the derivation associated with the object, since so far only derivations associated with a subject have been considered:

(25) John killed Mary.

The peculiar problem posed by (25) under ATTRACT is that if *Mary* is generated within the lower VP, then it would mean that arguments can be generated in thematic positions after all. But if we want to maintain that the subject, *John* in (25), is not generated in a thematic position, then we need a general conception of predicate properties as 'weak', or in present terms not lexicalized. The alternative therefore would be for *Mary* to be generated in [Spec, v]; but this is even more problematic, since in this case it is not clear why *Mary* would ATTRACT the lower V rather than v itself, under the MLC.

We believe in fact that this set of problems is an artifact of Chomsky's (1995) conception of v, depending in turn on his device of multiple Spec's. Suppose that in line with a number of theorists, notably Kayne (1994), we enforce a single Spec constraint for each head. Then we are forced to attribute the two properties of v in Chomsky's

(1995) system to two different heads. On the one hand we can postulate a head, v proper, associated with a D feature; accusative case will be a morphological reflex of the lexicalization of this feature (see Manzini and Savoia 1998 for a precise realization of this general idea). On the other hand, a higher V head will be tipically associated with the CAUSE predicate and its relation to the DP in [Spec, I] will correctly establish the argumental interpretation of the latter. In other words, transitive sentence like (25) will have the structure in (26):

(26) [John I [V [Mary v [V]]

Needless to say, the correct word order for (26) in English requires the verb *killed* to be in the highest V. Arguments for this are also independently found in the literature (cf. Bobaljik's 1995 discussion of his Stacking Hypothesis).

Before concluding on this point, we would like to note that the conception of argument structure of Hale and Keyser (1993) and Chomsky (1995), while not necessarily having recourse to abstract predicates, owes an important debt to lexical decomposition frameworks. In essence the number of argument positions and their properties are restricted by the number of primitive predicate types and their possible combinations. An alternative theory based on the idea that argument structure is aspectually characterized is argued for by Tenny (1994), Borer (1994), Salles (1997), Arad (1998); Manzini and Savoia (1998; forthcoming) go as far as denying the existence of VP-shells. We believe that this approach avoids the problems associated not only with standard theta-roles but also with lexical decomposition and is therefore to be preferred. The relevant lines of empirical argument, however, cannot be pursued here; we shall therefore be satisfied with (26) for the purposes of this article.

In what follows we will provide empirical and conceptual arguments in addition to those already presented in favour of our view of 'A-movement'. The first set of arguments comes from the domain of control. We have independently argued that the classical analysis of control, by means of the empty category PRO, has conceptual and empirical drawbacks, which surface in different forms within the GB and minimalist frameworks. Within the present framework we will show that all of these difficulties can be overcome.

3 The theory of control

The major proposal that we will put forward in this section is that within the framework of assumptions laid out in section 2, control corresponds to a derivation in which one argument DP ATTRACTs two (or more) different predicates. We shall argue that this

model allows for an elegant account of control which overcomes many of the problems noted for previous frameworks in section 1.

Consider the basic control sentence in (27):

(27) $[_{\mathbb{P}} John I [_{\mathbb{VP}} tried [_{\mathbb{P}} to [_{\mathbb{VP}} leave]]]$

In (27) *John* is directly merged into [Spec, I] position to lexicalize the D-feature of I. In order to achieve the desired control interpretation, it must be the case that the DP ATTRACTs both the matrix V and the embedded V. Under the MLC in (21), the DP is clearly allowed to ATTRACT the higher V since this is the closest attractee. We may wonder, on the other hand, what allows DP to ATTRACT the lower V as well. Remember that under Chomsky's (1995) MLC, once movement of a given feature or phrase occurs, its trace ceases to be visible for the purposes of the MLC, allowing for movement of a lower feature or phrase of the same type to take place. Technically, all that is needed in the present framework is an adaptation of this convention under which a DP can ATTRACT a V across another V, as long as the latter is itself ATTRACTed by the same DP. It seems to us, however, that this technical solution comes short of a real answer to our problem.

One conclusion that can be drawn from the difficulty just noted is that contrary to what the MLC states, potential attractees do not really interact with one another. Rather it is potential attractors that interact, as independently suggested by Manzini (1997). Adapting slightly the formulation in Manzini (1997), we therefore define the MLC in scopal terms, as in (28):

(28) Scopal MLC

Feature F ATTRACTs feature F_A only down to the next F' that also ATTRACTs F_A .

Given the *Scopal MLC* the argument DP in (27) automatically ATTRACTs all of the V's that it has in its scope, where its scope extends as far down in the tree as the next DP. Since there is no other DP than *John* present in (27), the scope of *John* includes both the matrix and the embedded predicate, as desired.

An analogous problem is posed by *Last Resort*, in the sense of Chomsky (1995). Clearly, in (27) *Last Resort* allows the DP to ATTRACT the higher V. However, in Chomsky's (1995) formulation, it does not automatically allow for the lower V to be ATTRACTed by the DP as well. Note that there is at least one case in which Chomsky (1995) explicitly allows for a single attractor feature to be checked by more than one attractee. Indeed in connection with his discussion of multiple Specs, he admits the

possibility that a non-interpretable feature can survive checking by the first attractee and therefore act as an attractor for a further attractee. Technically speaking, all that is needed to allow for the analysis of control suggested in (27) is an assumption parallel to Chomsky's (1995) whereby an attractor, such as DP, can ATTRACT just once or n-times. Once again it seems to us that this solution, whatever its justification may be within Chomsky's (1995) system, does not represent a conceptually adequate answer to our problem.

In fact, Chomsky's (1995) minimalist theorizing oscillates considerably between the conception of *Last Resort* that we accepted so far (implicitly or explicitly), and what he terms *Greed*. If *Last Resort* is based on the need of the attractor to have some feature checked, *Greed* is based on a parallel need of the attractee. According to Chomsky (1995, 1998), *Greed* is to be abandoned in favor of *Last Resort*, for conceptual reasons; for *Greed*, as opposed to *Last Resort*, requires look-ahead properties, that should be excluded by a minimalist grammar. It seems to us that this latter argument is compelling and therefore bars a return to *Greed* as a possibile solution to our predicament.

As it turns out, however, the *Scopal MLC* that we have just formulated in order to solve the locality problem suggests a solution to the *Last Resort* problem as well. Quite simply, our proposal is that *Last Resort* also functions as a *Scopal Last Resort* principle which requires (and therefore allows) a given attractor to ATTRACT all of the potential attractees down to the domain of the next attractor. To be more precise, remember that Chomsky (1995) unifies *Last Resort* and the MLC into a single principle. In (29) we provide our version of the *Scopal Last Resort* + *MLC*. Effectively, (29) represents a strengthening of (28) to a biconditional:

(29) Scopal Last Resort + MLC

F ATTRACTs all and only the F_A's that are in its scope

It is worth noticing that the attraction of V by DP within the present framework provides a natural translation of Chomsky's (1981) *Theta-Criterion*, whereby every argument must be assigned a theta-role and every theta-role must be assigned to an argument. Chomsky (1995) enforces the satisfaction of both of these clauses by his configurational definition of argument structure. The condition in (29) derives the same results as the *Theta-Criterion*, when applied to arguments and predicates, since it effectively requires that every DP be matched with all and only the predicates in its immediate scope. In recalling the *Theta-Criterion* we slightly weakened it by omitting the one-to-one correspondence between arguments and theta-roles present in the original formulation. This requirement does not hold by hypothesis under the present construal of control. To be more precise, every theta-role/predicate is associated with a single DP, but vice versa, a single DP can be associated with more than one theta-role/predicate. In fact, there is independent evidence from secondary predication for this weakening of the theory, as explicitly recognized by Chomsky (1986).

We are now in a position to consider whether the present theory derives the first empirical generalization on control reviewed in section 1, namely Chomsky's (1981) *PRO Theorem.* The relevant data are in (2)-(3). Within the framework of assumptions adopted here their relevant structure is as in (30)-(31) respectively:

- (30) [$_{IP}$ John I [$_{VP}$ persuaded [$_{vP}$ v [$_{VP}$ V]]]
- (31) $[_{\mathbb{P}} \text{ John I } [_{\mathbb{VP}} \text{ believes } [_{\mathbb{CP}} \text{ that } [_{\mathbb{P}} \text{ will } [_{\mathbb{VP}} \text{ eat}]]]]$

Consider (31) first. *John* cannot be merged directly into the matrix [Spec, I] position, because this implies a violation of the requirement imposed by the strong D-feature of the embedded I represented by *will*. Suppose then *John* is merged in the embedded [Spec, I] position. In the framework of Chomsky (1995) its movement into the matrix [Spec, I] position is straightforwardly blocked by the fact that its (non-interpretable) Case feature has already been checked and cannot therefore check the Case feature of matrix I. In section 2, however, we suggested that Case features do not have any syntactic import and should be subsumed under D features. This means that we must find some alternative means to rule out movement from the lower to the higher [Spec, I] in (31). The generalization that prohibits movement from one Case position to another can be reconstructed within the present framework as a prohibition against the same D(P) lexicalizing more than one D feature.

If we extend the constraint we just formulated so that any lexical element is allowed to lexicalize one and only one feature, we expect not to be able to copy it at all. This result is undesirable to the extent that we want to maintain a Copy and Merge derivation for the *wh*-movement cases. One possibility is that the reconstruction effects that argue in favour of this derivation according to Chomsky (1995) are themselves to be explained in some alternative way (cf. Kayne 1998). For present purposes, however, it is sufficient to adopt an intermediate generalization, whereby any given lexical item can lexicalize the same feature once and only once. This still allows for *wh*-movement, in that the *wh*-phrase lexicalizes a D feature at the position of merger and a different feature, namely wh/Q, in the position where it moves to.

(30) is blocked in much the same way as (31). Remember that we take v to be associated with a D-feature that needs to be lexicalized. If *John* is inserted in [Spec, v], it cannot be raised to [Spec, I] to lexicalize the D feature of I, because it has already lexicalized another D feature. If it is inserted directly in [Spec, I], then the D-feature associated with v is not lexicalized at all. In fact, with respect to the distribution of PRO

our results are not so much reminiscent of Chomsky's (1981) *PRO Theorem*, as of the alternative generalization of Manzini (1983), according to which the distribution of PRO is dictated by its lack of Case.

Another generalization that remains to be accounted for is that expressed by the MDP, as illustrated in (4)-(5) in section 1. The relevant structure of (5) is as in (32):

(32) $[_{\mathbb{P}} \text{ Mary I } [_{\mathbb{VP}} \text{ thinks that } [_{\mathbb{P}} \text{ John I } [_{\mathbb{VP}} \text{ expected } [_{\mathbb{IP}} \text{ to } [_{\mathbb{VP}} \text{ eat}]]]]]$

Notice that in (32) *Mary* lexicalizes the D feature of the matrix I and *John* lexicalizes the D feature of the intermediate I, as required; by hypothesis, non-finite I does not have a D feature. In order to capture the MDP, we need to establish the conclusion that the most embedded predicate *eat* is ATTRACTed by *John*, as indicated by the italics, and not by *Mary*. Under the *Scopal MLC*, *Mary* ATTRACTs only down to the scope of the next attractor, i.e. *John*. Therefore *eat* is not a possible attractee of *Mary*. The only possibility that the *Scopal MLC* allows for is the correct one, whereby *thinks* is ATTRACTed by *Mary* and the other predicates by *John*.

Classical MDP examples of the type in (4), with the structure indicated in (33), are accounted for much in the same way as (32):

(33) [$_{IP}$ Mary I [$_{VP}$ persuaded[$_{VP}$ John v [$_{VP}$ V[$_{IP}$ to [$_{VP}$ eat]]]]]

If we take v to represent a D-position, lexicalized by *John*, then the lower V?s cannot be associated with *Mary* across it, under the *Scopal MLC*, as desired. Remember that a separate ATTRACT operation connects *Mary* to the higher V position in the matrix VP-shell.

Within the present theory, control and raising differ only by the nature of the predicates involved. This holds for subject control and raising to subject, to which we will return, as well as for object control and raising to object, or *Exceptional Case Marking* (ECM). As argued by Johnson (1991), Chomsky (1995), the overt realization of *John* as the subject of the ECM infinitival in (34) below depends on the presence of what for them is a Case position and for us a D position in the VP-shell of *believe*. In present terms, this D position requires to be lexicalized by a DP, which in turn attracts the predicates in its immediate scope, as indicated in (35). Remember once again that *John* triggers a separate instance of ATTRACT, involving the higher VP-shell predicates. The effect is that of raising to object because of the nature of the predicates involved (as expressed for instance by. Chomsky's 1955 idea that *believe* forms a complex predicate with the embedded verb):

(34) John believes Mary to eat too much.

(35) John I [believes [Mary v [V[to [eat too much]]]]]

It is often assumed in the literature that contexts of obligatory control, such as the ones that we are considering, do not admit of split antecedents. In this respect, the *Scopal MLC* yields the conventionally accepted results, since the embedded predicate in (36) is in the attraction scope of *Mary* but not of *John*:

(36) John persuaded Mary to drink.

Transitive control predicates such as *persuade* also raise the question of the interaction between control and classical A-movement, such as passive. Consider (37):

(37) John was persuaded to eat.

Our theory predicts that the embedded and the matrix V's are both ATTRACTed by *John* only if there is no intermediate D position in the sentence skeleton. Thus *contra* Baker, Johnson and Roberts (1989), we are led to the conclusion that in short passives only one D is syntactically represented. This analysis has another consequence that turns out to be correct. It is often noticed in the literature (Williams 1980) that subject control verbs such as *promise* do not passivize, or to the extent that they passivize they require control by the derived subject, such as *John* in (38):

(38) John was promised to be hired.

Within our theory, the well-formedness of (38) with an 'object' control, rather than a 'subject' control reading, corresponds to a derivation whereby *hired* and *promise* are both ATTRACTed by *John*. As desired, this is the only derivation allowed by the theory.

Summing up so far, if what precedes is correct, the *Scopal MLC* yields the two basic properties of control, namely the *PRO Theorem* and the MDP, without need for extra stipulations in the grammar. In particular, within the present theory there is no need to state the MDP as a separate locality principle; nor to capture the distribution of control by using a specialized empty category PRO and/or the *ad hoc* notion of null Case.

Other properties of control remain nevertheless to be taken into account, notably arbitrary control. Before we go on to arbitrary control, it is worth pausing to note that the idea that there is no A-movement, and that the syntax provides operations for matching DPs inserted in D positions with thematic positions also characterizes approaches within

the categorial grammar framework. In particular, Cormack (1998) proposes that the [Spec, V] position is filled by an abstract element q, which passes on the relevant thematic properties to the subject DP, generated directly in Case position. Similarly, the distinction between raising and control in terms of a DP ATTRACTing one or two predicates is reminiscent of work within the categorial grammar tradition such as Jacobson's (1992). What distinguishes the present approach, is mainly the fact that through the *Scopal MLC* + *Last Resort* the advantages of the transformational approach are also retained.

4 Arbitrary Control

As we mentioned in section 1, the elimination of PRO in the contexts of obligatory control can also be achieved in a framework which adopts the standard construal of A-movement as DP-movement, by assuming that a DP can move from one q-position to another, along the lines suggested by O'Neil (1995), Hornstein (1996). Thus for instance in (39), control would be expressed by merging *John* in the embedded [Spec, V] position and moving it to the matrix [Spec, V], via the embedded [Spec, I], and then to the matrix [Spec, I]:

(39) John I [John tried [John to [John leave]]]]

Under this approach, Hornstein (1996) achieves a unification of the MDP with the basic locality principle in grammar, i.e. the MLC, making it equivalent to the present proposal in this respect.

However, the two theories can be told apart in at least one environment, namely arbitrary control. Arbitrary control is illustrated in (40) where the embedded infinitival verb is apparently not associated with any DP, and ends up being interpreted roughly as having a generic argument:

(40) It is hard [to [work]]

In the terms of Chomsky (1981), the infinitival [Spec, *to*] position hosts a PRO, which in the absence of any available controller, is assigned a free index, and hence arbitrary interpretation. For Hornstein (1996), in the absence both of PRO and of any overt DP argument, the only possible analysis consists in assuming that the embedded [Spec, *to*] position is occupied by the other pronominal empty category, *pro*.

There are both empirical and conceptual problems with this latter approach. Thus it is not explained why in English *pro* would be restricted to exactly the arbitrary control

environments. What is more, much current literature argues in favor of the elimination of *pro* even from the null subject configurations for which it was originally introduced by Chomsky (1982). In particular, Pollock (1996), Nash and Rouveret (1997), Alexiadou and Anagnostopoulou (1997) argue that V is sufficient to check the strong feature of I in null subject languages, while Platzack (1995), Manzini and Savoia (forthcoming) treat null subject languages as having a weak (non-lexicalized) D feature in I. For the reasons advanced by all of these authors, we conclude that *pro* is best abandoned; in this perspective the *pro* analysis of arbitrary control, and therefore the standard DP-movement approach to obligatory control that forces it, are undesirable.

What remains to be demonstrated is that our theory predicts the basic facts concerning arbitrary control without having recourse to the re-introduction of pronominal empty categories. In present terms, the problem that needs to be solved is that, although the embedded predicate in (40) is apparently left without an argument, this does not lead to a failure of interpretation; rather it leads to the interpretation already noted whereby *work* takes a variable of a generic operator of some sort as an argument. We can provide a syntactic formalization for this interpretation by assuming that an abstract adverb of quantification is available in sentences of this type, which ATTRACTs the predicate. More precisely, we can identify the position of the operator with that of finite C. The derivation for (40) therefore takes the form in (41) where finite C ATTRACTs the infinitival V:

(41) [*C* [it is hard [to *work*]]

One argument in favor of C acting as an attractor for the embedded predicate is provided by the fact, noticed in the literature (cf. Bresnan 1982), that predicates not associated with a lexical DP can be interpreted as having a specific, rather than a generic, argument given the appropriate context. Thus while a generic reading for the non-lexicalized argument is associated with the generic context in (41), a specific reading is associated with the non-lexicalized argument given the specific temporal context in (42):

(42) It was hard to work (on that beautiful sunny day).

In (41) and (42), then, the interpretation of the non-lexicalized argument varies in accordance with the temporal context. Suppose that the operator in C ATTRACTs Tense, exactly as it ATTRACTs the lexical predicate. If so, we predict that a generic operator in C determines a generic interpretation for both the non-lexicalized argument and for Tense in (41), while the presence of a specific operator in C in (42) determines a specific interpretation for them. Remember that it is independently argued in the

literature, starting with Enç (1987) that the interpretation of Tense depends on C; see Roussou (forthcoming) for more recent discussion. Indeed we can assume that the attraction properties of the operator in C are always satisfied by Tense, independently of whether they ATTRACT a lexical predicate or not.

It is important to stress that we predict a correlation between temporal and argument interpretation only when the argument is not lexical; only in this case do they both depend on C. If there is a lexical argument, as in all of the cases of ordinary control considered in previous sections, and it is the lexical argument itself that satisfies an argument place in the predicate, we predict no correlation between the generic or specific nature of the lexical argument and that of the temporal context. For instance, it has been noticed (cf. Brody and Manzini 1988), that if an overt argument is associated with the matrix predicate in (40), it takes on the role of obligatory controller, as in (43):

(43) It is hard for *us* [to [*work*]]

In present terms *us* in (43) ATTRACTs the embedded predicate, as predicted by the *Scopal MLC*. Indeed in (43) the predicate *work* is in the immediate attraction scope of the DP *us*. Since in (43) the interpretation of the argument position of *work* depends on *us*, while the interpretation of the matrix Tense depends on C, we expect no correlation between the two. In other words, in this case we correctly predict that it is possible to have a specific argument in a generic context. The interaction between temporal and argumental interpretation are strictly confined to arbitrary control, which seems to us exactly the correct empirical result.

A problem that we have not addressed so far concerns the role played by the expletive *it* and the matrix predicate in (40)-(42). Since *it* is a DP, we might expect it to count as the nearest attractor for the embedded, as well as the matrix, predicate. If so, however, we cannot predict the interpretation of the relevant sentences, which as we have seen depends on the matrix C acting as the attractor for at least the embedded predicate. One partial solution consists in saying that only an argument DP, and not an expletive, ATTRACTs a predicate. This solution, however, has both conceptual and empirical disadvantages. From a purely conceptual point of view we notice that the distinction between arguments and non-arguments is not a feature of DP's but rather a consequence of the derivation associated with them. Thus *it* is a non-argument if it has an argument associate, and an argument otherwise.These considerations are strengthened by the empirical evidence provided by examples such as (44). (44) contains an instance of *there* which is expletive with respect to another DP; in this latter case the associate DP is ultimately interpreted as the controller of the embedded predicate:

(44) There arrived several people [without telling us beforehand]

For reasons that have to do with the structure of adjuncts (cf. section 5 below), the interpretation of (44) cannot be obtained by having *several people* directly ATTRACT the embedded predicate. Rather *there* independently ATTRACTs both the embedded predicate and *several people* (see Chomsky 1995, Manzini and Savoia forthcoming for theories of the expletive-associate pair which involve attraction of some sort or another). This analysis confirms that it is really DP's that attract V's, independently of whether they are arguments, and not some +/-argument property of DP's.

Returning now to the examples in (40)-(42), we still wish to suggest that the reason why *it* allows C to ATTRACT the embedded predicate is connected to the fact that *it* is an expletive. In particular, note that the control sentence is the associate of *it*. In the framework of assumptions just adopted, this means that *it* ATTRACTs not only the matrix predicate but also the control sentence. Thus the control sentence comes to be interpreted as the argument of the matrix predicate. What is more, the predicate in the control sentence is not in the scope of *it* but only in the scope of the matrix C. Thus it is C that ATTRACTs it and determines the 'arbitrary' interpretation of its argument.

It is worth checking at this point whether the present set of assumptions still allows us to account for raising and superraising. In a simple raising sentence like (18), repeated here as (45), *John* ATTRACTs both the matrix and the embedded predicate, that are in its immediate scope, as indicated by the italics:

(45) [John I [seems [to [work]]]]

Consider then (46)-(47). (47) represents an example of superraising, classically accounted for by assuming that *it* intercepts A-movement of the argument across it; (46) is its well-formed counterpart:

- (46) It seems [that John was told [that Mary left]]
- (47) *John seems [that it was told [that Mary left]]

In (46), John ATTRACTs the predicate told; it ATTRACTs the predicate seem which is in its immediate scope and at the same time the associate, in this case the *that*-clause. The correct interpretation ensues, with the *that*-sentence taken as the argument of seem. Consider the superraising example in (47). The intended interpretation, under which John is an argument of told, is blocked by the Scopal MLC; indeed John cannot ATTRACT told because told is in the immediate scope of it. Other construals of the sentence are equally excluded. In particular, although seem is in the immediate scope of

John, and can be ATTRACTed by it on the basis of the *Scopal MLC*, the interpretation fails, since the interpretive requirements of *John* cannot be satisfied by a predicate like *seem* (cf. the ungrammaticality of **John seems*).

The analysis of control presented in this section and in the preceding one depends of course on the idea that the C position of the control complement does not have any operators that block attraction of the embedded predicate by a higher DP or finite C. If such attractors were present, we would not be able to derive obligatory control in (4)-(5), as well as in (43). Furthermore, with respect to arbitrary control we would not be able to predict that the interpretation of the non-lexicalized argument co-varies with that of the matrix Tense as in (40)-(42). The absence of appropriate operators in the infinitival C can in turn be connected to the absence of temporal properties in infinitivals. This conclusion directly contradicts the conclusions of Chomsky and Lasnik (1995), whose null Case for PRO is supported by temporal properties associated with to (cf. Martin 1996). This latter view is apparently confirmed by the observation of Bresnan (1972), Stowell (1982) that a future interpretation is associated with the controlled infinitival in examples of the type in (4) or (6). Though the observation is in itself correct, we believe that it does not argue for temporal properties associated with infinitival I. Rather, the future interpretation can be taken to support a modal characterization for infinitivals (see Roussou forthcoming). Recent work by Boškovic (1997) seeks to provide further arguments in favour of the temporal analysis of control clauses based on a comparison between English and French. Though a discussion of his data is outside the scope of the present paper, we believe that the modal-based approach, or eventually an aspectual one, accounts for them as well.

Let us summarize so far. On the basis of the *Scopal MLC* and of the appropriate assumptions concerning attractors, the theory is able to derive the following theorem: a predicate is necessarily ATTRACTed by the first argument DP or finite C that c-commands it. This in turn means that what in standard terms is an arbitrary PRO can only surface if there is no DP argument in the clause that immediately contains the control clause. Thus arbitrary control is impossible in (43), though forced in (40)-(42). The operator that we have postulated in C independently acts as an attractor for Tense and is associated with it in all finite environments.

Other empirical consequences can be made to follow from what precedes. Suppose we embed (40) under a higher sentence, as in (48):

(48) John thinks [that it is hard [to work]].

In (48), two interpretations are available for the predicate in the control sentence; remember that in (48) the control sentence itself is the associate of it and its scope is

therefore the same as the subject's. The first reading, the arbitrary one, depends on *work* being ATTRACTed by an operator associated with the intermediate C. There is, however, a second reading, with *John* interpreted as an argument of *work*. Our theory clearly predicts that this reading cannot depend on control by *John*, in the sense defined so far, that is on the attraction of the predicate by *John*. Indeed the temporal operator in the intermediate C acts as an attractor for *work*, as we have just seen, effectively preventing *John* from ATTRACTing it.

There is an alternative way to obtain the desired reading though. We have already assumed that the embedded finite C is associated with an operator that ATTRACTs *work*. We suggest that the interpretation whereby *John* is the argument of *work* also depends on this derivation. As we have seen, the argument of a predicate ATTRACTed by the operator in C can be interpreted as generic or specific; in this latter case we assume that it can be anaphoric to a DP. Hence the non-lexicalized argument of *work* can be interpreted as anaphoric to *John*. In this respect we follow a long tradition in control studies (Williams 1980, Bresnan 1982, and many subsequent works including Hornstein 1996), which suggests that the non-obligatory long-distance variety of control is to be treated in the same terms as arbitrary control. At the same time, contrary to the works just mentioned, the present theory introduces a parallelism between 'obligatory' and 'non-obligatory' control, to the extent that both phenomena reflect the (obligatory) attraction of a predicate by an appropriate attractor. The only difference is that DP is the closest attractor in the case of obligatory control.

One question that remains to be addressed concerns the +human interpretation that appears to be necessarily attached to arbitrary PRO. This apparently provides a good argument in favor of the approach of Chomsky and Lasnik (1995), since their PRO can be endowed with a feature such as +human in the lexicon. In a language like Italian, furthermore, the feature specification of arbitrary PRO can include +masculine, +plural, since participles and adjectives are seen to agree with it in these features. However, if this is the case for arbitrary PRO, then there must be a separate lexical entry for non-arbitrary PRO which has as many combinations of lexical features as its antecedents. This result appears to be unparsimonious. Nor is it likely that a feature like +plural can be construed as a default, even if one were to overcome general theoretical objections to the notion of default itself (on these points see Manzini and Savoia forthcoming).

In this connection, Williams (1992) points out the similarity between arbitrary control and logophoricity. Thus elements that behave logophorically such as long-distance anaphors typically target as their antecedent, "the source of the report, the person with respect to whose consciousness or self the report is made, and the person from whose point of view the report is made" (Sells 1987). Naturally, all of these characterisations

point to a +human antecedent for logophors. We therefore conclude that an operator ATTRACTing a predicate leads to a logophoric reading of its argument; whence the surfacing of +human properties for arbitrary control. We shall return to the issue of logophoricity, and its syntactic or pragmatic nature, in section 6.

5 Other consequences for control

In sections 2-4, we established that the *Scopal MLC* is observed by ATTRACT involving predicates, in that the closest DP or C attractor, is obligatorily targeted. There is, however, a second major constraint on ATTRACT, whose effects on control we should in principle be able to observe, namely *Connectedness*, as formulated in (22). The relevant environments correspond of course to the classical strong island ones, namely adjuncts and subjects.

Consider first control into adjuncts. As is well known, adjuncts tend to display the same properties as complements with respect to control. Thus the only possible interpretation for the adjunct infinitival in (49)-(50) is one in which the embedded predicate is associated with the matrix subject:

- (49) John left before eating.
- (50) John left without asking.

One possible interpretation of this fact is that, as suggested by Kayne (1994), right adjunction is not allowed by the grammar; therefore the embedded sentences in (49)-(50) are attached as complements. This analysis, however, is in direct contradiction with a number of other properties of adverbials, as argued in detail by Manzini (1995). One property, which is of direct interest here, is that in general, a matrix object cannot serve as controller for an adverbial. Thus (51)-(52) have exactly the same control properties as their counterparts in (49)-(50) despite the insertion of an object in the matrix sentence:

- (51) John left us before eating.
- (52) John left us without asking.

If adverbials are generated in the most deeply embedded position in (51)-(52), as suggested for simple adverbs by Larson (1988), then control by the object should be possible and in fact obligatory by the MDP or the *Scopal MLC*. Therefore, as pointed out by Williams (1974), the subject orientation of most adverbials argues in favour of the conclusion that they are attached high enough not to be in the scope of the object.

Incidentally the same conclusion excludes that in (44) the associate could directly control the embedded predicate.

Thus adverbials are not in complement position, but rather adjuncts. We have independent evidence that in the case of wh-movement, extraction from adjuncts is blocked, as in (53) below. Therefore if control in (49)-(52) is to be construed as attraction of the embedded predicate by the matrix subject, the question arises why this is not blocked as an instance of adjunct island violation. However, adjunct islands can be circumvented by parasitic gaps. Thus for wh-movement, the adjunct island violation in (53) contrasts with the wellformed parasitic gap configuration in (54):

- (53) *Which article did you review my book [without reading t]
- (54) Which book did you review t [without reading e]

This suggests that subject control in (49)-(52) also reflects an underlying parasitic gaplike configuration. Consider indeed (49) in terms of the present theory. Assuming that the *before* sentence is generated as a sister to the highest VP-shell projection, the relevant structure is as in (55):

(55) John I [left [before I [eating]]]

In (54) we independently know that the DP *John* ATTRACTs the matrix predicate, which is on the main branch of the sentence. But if so, attraction of the embedded predicate, which is on the adjunct branch, by *John* creates a parasitic gap configuration, as desired. This can be most simply illustrated by comparison with (54), where a single antecedent, the *wh*-phrase, binds a trace in the main branch of the sentence and a trace inside the adjunct, as schematized in (56). Similarly in (55), a single DP, *John*, acts as an attractor for both the predicate in the main branch of the sentence and for the predicate inside the adjunct, as schematized in (57):



As we did for control into complements, we can now consider the interaction of control into adjuncts with passivisation. In the cases illustrated in (58)-(59) below, the argument DP in matrix [Spec, I] functions as the attractor for the embedded predicate. This is correctly predicted by our theory on the basis of the *Scopal MLC* as well as *Connectedness*. As can be seen by the comparison between the active forms in (b) and the passive forms in (a), we find a switch of the thematic properties of controllers, since in passives it is the patient that controls the embedded predicate, while in actives it is the agent that does so. Our theory thus predicts that notions such as agent and patient are irrelevant for control; instead the embedded predicate in the adjunct is attracted by the closest DP connected to it. As we have argued for the case of control into complements, in short passives, the agent is not syntactically projected and cannot therefore interfere with control.

- (58) a. John was fired after doing that.
 - b. We fired John after doing that.
- (59) a. John was hired without moving a finger.
 - b. We hired John without moving a finger.

One apparent counterexample to the above conclusions is represented by control into purpose clauses. Well-known examples of the type in (60) are generally taken to show that the agent is syntactically realized in passives, and acts as a controller for the adjunct (cf. Brody and Manzini 1988):

(60) The boat was sunk [(in order) to collect the insurance].

In order to account for (60), we are led to assume that the control sentence is adjoined to a projection higher than the VP to which we adjoined the adjunct sentences in (49)-(52). Indeed, it is necessary and sufficient for us to assume that the purpose clause is attached high enough in the matrix sentence to be outside the scope of the subject, in order to predict that it is not controlled by the subject under the *Scopal MLC*. Following fairly standard assumptions, we can associate the purpose clause with an IP-adjoined position, as in (61):



The fact that the embedded predicate in (61) becomes associated with the arbitrary interpretation, is in present terms due to it being ATTRACTed by the matrix C. This also follows form the *Scopal MLC*, since the matrix C is the first attractor that takes the adjunct predicate in its scope. Note that *Connectedness* is also satisfied, in that the matrix C equally ATTRACTs the matrix I for reasons of Tense interpretation. Therefore, the attraction path from matrix C to the adjunct predicate forms a connected subtree together with the matrix attraction path from C to I. This analysis does not take into account the intuition that the argument of the embedded predicate is interpreted as anaphoric to the agent of the passivized matrix predicate. Since we independently argued

that in short passives the agent is not syntactically represented, we are led to impute this intuition to pragmatic inference.

Note that if purpose clauses are adjoined to a position which is outside the scope of the matrix subject, we predict that their predicate is not controlled by it even in active environments, for instance in (62). Hence we are led to claim that the apparent control of the embedded predicate by the matrix subject in (62) reduces to arbitrary control and anaphora, essentially as in the case of apparent long-distance control in (48):

(62) We sank the boat [(in order) to collect the insurance]

In other words, the matrix C acts as an attractor for the adjunct predicate, licencing a specific interpretation, anaphoric to some argument in the sentence. Once again, we impute the fact that the second interpretation is in fact the only possible one to a pragmatic inference factor.

We are now in a position to consider control into the other major type of strong islands, namely subjects. The relevant data are of the type in (63):

(63) John believes [that [behaving badly at the party]I would bother Mary]

(63) is a typical context of non-obligatory control; indeed the non-lexicalized argument in the subject clause can be interpreted as arbitrary, or as anaphoric to either the matrix argument, *John*, or to the embedded one, *Mary*. Leaving aside the question of islandhood for a moment, this state of affairs is as expected, given that the nearest attractor of the appropriate sort for the embedded predicate is the operator in the intermediate C. The DP-argument *Mary* cannot act as an attractor because it does not take the subject sentence in its scope, while the DP-argument *John* is higher, and thus further away than the intermediate C in terms of the *Scopal MLC*. In turn, attraction of the control predicate by the intermediate C is compatibile either with an arbitrary interpretation or with an interpretation anaphoric to some individual(s) in the discourse, including *John/ Mary*.

As for the question whether the intermediate C is accessible to the gerund under *Connectedness*, remember that wh-extraction from a subject is barred; however, if a parasitic gap configuration is involved, *Connectedness* is satisfied, as seen in the contrast in (64):

- (64) a. *Who do [friends of t] admire you
 - b. Who do [friends of e] admire t

Therefore in (63), the gerund can be ATTRACTed by the intermediate C, in that this attraction path forms a connected subtree with the attraction path independently defined between C and the matrix I for reasons of Tense interpretation. This yields a derivation of the type in (65):



In short, both the fact that non-finite subject sentences are typical contexts of nonobligatory control and the fact that non-finite adjuncts typically display obligatory control can be explained on principled grounds. The alternation between obligatory and non-obligatory control depends on whether the closest attractor is a DP argument (as for adjuncts) or a C (as for subjects). As for the possibility to reach into strong islands, this arguably depends on the creation of parasitic gap-like configurations. Thus a DP subject can ATTRACT two predicates, one in the main branch of the sentence and one inside an adjunct. Similarly a C can ATTRACT an I in the main branch of the sentence and a predicate inside a subject sentence. The conclusion that control into adjuncts amounts to the creation of a parasitic-gap configuration is also put forward by Hornstein (1997), though it is embedded under a completely different analysis of parasitic gaps; furthermore, his analysis does not readily extend to control into subject sentences.

Before concluding the discussion of control, another notoriously problematic set of data is worth reviewing briefly. We have so far illustrated control with embedded non-finite sentences. However, it is well-known that in appropriate contexts, such as interrogatives and exclamatives, infinitivals are acceptable as matrix sentences, as for instance in (66)-(67):

(66) What to do?

(67) Ah, to go to the sea!

What is more, it appears that interrogative sentences embedded as complements to matrix verbs, do not display obligatory but rather optional control effects, as in (68):

(68) I asked how [to behave]

Consider first (66)-(67). Because there obviously is no tensed C which takes the infinitival in its scope, our general theory about the licencing of arbitrary control depends on assuming that some other operator can act as an appropriate attractor for the infinitival predicate. As it turns out, this operator can be identified with the one licensing exceptional matrix the apparently use of the infinitival. some sav interrogative/exclamative modal operator. If the same operator is present in the embedded context in (68), we fully predict that it will licence the arbitrary reading of the non-lexicalized argument of the embedded predicate, pre-empting control by the matrix subject.

6 Further consequences for A-movement

In our discussion of A-movement in section 2, we argued that the fundamental role of A-movement in the grammar is that of conveying theta properties to DP positions, where arguments are realised. In the alternative formalization that we provided, a similar role is played by the attraction of predicates by DP's. This construal of A-movement explains the lack of reconstruction effects at LF and the blocking of PF rules.

A number of other facts standardly associated with A-movement remain to be considered. One of them is the triggering of agreement. Let us for instance consider a standard unaccusative sentence in Italian including an embedded past participle. The subject overtly agrees with the past participle in number and gender, as in (69):

(69) Maria è partita. Mary is left-sg.fem."Mary has left."

Under standard assumptions about A-movement the derivation of (69) proceeds along the following lines: *Maria* moves through [Spec, v]; this configuration triggers agreement between *Maria* and the past participle, though *Maria* is ultimately realized in [Spec, I], as illustrated in (70):

```
(70) Maria è [Maria partita]
```

In the present framework, *Maria* is merged directly in [Spec, I], while the lexical verb is ATTRACTed by *Maria*, without Copy and Merge applying at any point in the

derivation. If that is the case, the question arises what triggers agreement between *Maria* and the past participle at all. We assume that the past participle, exactly like the DP, is associated with f-features in the lexicon, agreeing in this with Chomsky (1995). On the other hand, we have already established that for reasons entirely independent of agreement, DP ATTRACTs the lexical V. In this framework agreement of the f-features present on the DP and those present on the past participle can be seen simply as a reflex of the ATTRACT operation that involves these two elements. Manzini and Savoia (forthcoming) show that this approach also extends to configurations standardly taken to involve agreement with *pro*.

Another set of problems that we need to consider concerns one of the arguments that we put forward for our approach to A-movement, namely the lack of reconstruction effects. Though data of the type in (11) seem to provide a clear-cut argument against reconstruction of A-traces, other data have been held to support it. In particular, Belletti and Rizzi (1988) point out contrasts of the type in (71):

(71) a. Pictures of *himself* worry *John/him*.b. **Himself* worries *John/him*.

They argue that the grammaticality of (71a) is due to reconstruction of the derived subject, i.e. *pictures of himself*, into its thematic position, which they take to be lower in the VP-shell than the position of the Experiencer *John/him*. If so, (71b) can no longer be excluded by the impossibility of reconstruction. Therefore they suggest that Principles B and C of Chomsky's (1981) *Binding Theory* must be satisfied at S-structure, whereas Principle A can be satisfied in their terms either at D-structure or at S-structure, i.e. equivalently at S-structure or at LF in a theory that uses reconstruction. Under this set of assumptions, both (71a) and (71b) are well-formed with respect to Principle A. (71b) is independently excluded in that it represents a violation of Binding Principles B/C at S-structure, given that the object *John/him* is (locally) bound by the subject, contrary to (71a).

Crucially, Belletti and Rizzi (1988) make reference to two levels of representation, and this is not consistent with minimalist assumptions. What is more, this recourse to two different levels of representation cannot easily be mimicked within the more limited theoretical resources of minimalism. If we say that both the derived and the trace position count at LF, we can of course derive the ungrammaticality of (71b), but we have no way of deriving the grammaticality of (71a), since in its derived position *himself* is not bound by *John/him*. If the subject DP on the other hand is interpreted only in its reconstructed position, we have the original problem of not being able to account for the ungrammaticality of (71b).

Data of the type in (71a) can also be reproduced for standard raising configurations, as in (72):

(72) Pictures of *each other* seem to *them* [to be on sale]

We believe that the correct approach for this kind of examples is the one proposed by Reinhart and Reuland (1993). They argue in essence that *self* moves to the closest predicate, marking it as reflexive. Therefore the interpretation of *John likes himself* is determined by an abstract structure of the form *John self-likes him* where the reflexively marked predicate forces the coreference between *John* and *him*. Disjoint reference in *John likes him* corresponds to the absence of a reflexive marker on the predicate. Given this analysis, since the anaphoric construal of *himself* depends on *self* acting as a reflexivizer of the nearest predicate, apparent anaphors inside (*picture-*)DPsm bound across sentence boundaries as in (73a), must correspond to logophors. This conclusion is independently supported by a number of facts; for instance a first person anaphor can appear in *picture-*DPs freely, as in (73b):

- (73) a. Lucie thought that a picture of herself would be nice on that wall.
 - b. A picture of myself would be nice on that wall.

Note that though Reinhart and Reuland (1993) like Sells (1987) conceive the construal of logophors in pragmatic terms, this conclusion need not be subscribed to. Thus we could assume that *self* in its logophoric interpretation is ATTRACTed by a syntactic position, even though one different from the predicate. The point that at least the Icelandic logophor *sig* is subject to syntactic 'movement' is argued for by Manzini (forthcoming) on the grounds that elements like *sig* obey strong islands and show parasitic gaps effects. Though the distribution of logophoric *himself* is different from that of logophoric *sig*, we then assume that a syntactic account is equally possible. Remember that according to Williams (1992), arbitrary control itself represents an instance of logophoric interpretation. We suggest therefore that the correct syntax for logophoric *himself* might be roughly the same as for arbitrary control, i.e. attraction by the nearest finite C. On this understanding we expect logophoric *himself* to systematically appear inside subject islands, as in (73), exactly like arbitrary control.

In short, data of the type in (73) argue not for reconstruction, but for the existence of a logophoric *himself*. This in turn raises doubts concerning all arguments for movement as Copy and Merge based on the reconstruction of *himself*, including those presented by Chomsky (1995) for *wh*-movement and reproduced in (10). As we already remarked at the outset, a discussion of A'-movement is beyond the scope of the present paper. What

is relevant for present purposes is that the availability of a logophoric binding account for (73) and the like neutralizes the potential counterargument to our construal of Amovement.

Further evidence for reconstruction in DP movement, relating to bound pronominals, is presented by Lebeaux, as quoted by Brody (1996). Thus (75) is considerably worse than (74):

- (74) [His_i mother's]_j bread seems to [every man]_i to be known by her_j to be the best there is.
- (75) *[His_i mother's]_j bread seems to her_j to be known by [every man]_i to be the best there is.

The argument runs as follows. If in (74) *his mother's bread* is reconstructed in the intermediate subject position, *his* is in the scope of *every man* and *his mother* c-commands *her*; this yields a wellformed interpretation. In (75), on the other hand, if *his mother's bread* is reconstructed within the scope of *every*, it is also reconstructed in a position c-commanded by *her*, which means that the resulting representation is ruled out by Principle B. If *his mother's bread* is not reconstructed, the structure is ruled out by whatever principle requires bound pronouns to be in the scope of their operator.

We propose on the contrary that in (74)-(75) the binding of *his* is accomplished not by reconstruction of *his mother*, but rather by QR of *every*: this latter operation is successful when *every* raises from the matrix sentence, but not when it raises from an embedded one. We suggest that the reason for this is related to the fact that only *seem*, which as we shall motivated below has an event-less structure, is crossed in (74), while QR would cross the event domain of *know* in (75). Though our analysis is still tentative, it does make a clear prediction, namely that substituting any other pronoun for *her* in (75), as in (76), yields a sentence which is still worse than (74). Judgements are subtle, but native speakers at best assign to (76) an intermediate status between (74) and (75). We can therefore group (75) and (76) together, attributing any residual contrast to their different relative complexity:

(76) (*)*His* mother's bread seems to us to be known by *every man* to be the best there is.

In short, on the basis of the discussion that precedes, we uphold Chomsky's (1995) generalization that A-movement cannot be reconstructed. This in turn provides a straightforward and powerful argument in favour of our approach, as opposed to the standard DP- movement one. This conclusion does not hold of the reformulation of A-movement proposed by Sportiche (1996). Sportiche also recognizes the difficulty

represented for DP-movement by contrasts of the type in (71). Thus he proposes that the D head of the DP is merged in the position where it surfaces, essentially as proposed here; if the reflexive in (71b) corresponds to such a head, he predicts that it does not reconstruct. By contrast the NP predicate is always merged in the thematic position, and NP-movement takes systematically place from thematic to D-position. Thus an anaphor in the NP-predicate, as in (71a), can give rise to reconstruction effects. But as we have seen, all of the evidence in favor of DP-/NP-movement from reconstruction of anaphors can be circumvented, thus depriving Sportiche's theory of an important empirical basis. Furthermore, the lack of blocking effects on PF rules is not explained by Sportiche's approach either.

Another phenomenon that appears to involve the LF-interface has been explicitly argued to mirror the underlying structural difference between raising and control. Thus Chomsky (1981), Burzio (1986) point out a contrast between (77) and (78):

- (77) One interpreter each seems to have been assigned to the visiting diplomats.
- (78) *One interpreter each tried to be assigned to the visiting diplomats.

In order to explain the data in (77)-(78) a theory of *each*, hence of distributivity, is clearly needed. Though this is largely outside the scope of the present article, we shall nevertheless consider briefly what an account of (77)-(78) might involve. *Each* is construed not with *an interpreter*, i.e. the distributee, but with *the visiting diplomat*, i.e. the distributor. This is made morphologically obvious in Italian by the fact that *each* agrees in gender with the distributor, as in (79):

- (79) a. Un interprete ciascuno/ *ciascuna fu assegnato ai diplomatici An interpreter each-m/ each-f was assigned to the diplomats
 - b. Una guida ciascuno/ *ciascuna fu assegnata ai diplomatici A guide each-m/ each-f was assigned to the diplomats
 - c. Un interprete *ciascuno/ ciascuna fu assegnato alle mogli An interpreter each-m/ each-f was assigned to the wives
 - d. Una guida *ciascuno/ ciascuna fu assegnata alle mogli A guide each-m/ each-f was assigned to the wives

Following Beghelli and Stowell (1997) we assume that the structure of the sentence includes a projection that acts as a host for distributive quantifiers, namely DistP, which is the highest quantificationl position below the inflectional ones. For the sake of the present discussion we shall in fact identify the surface position of *each* with DistP. According to Beghelli and Stowell (1997) "an active Dist head selects a ShareP ... which

in turn requires that an existential QP appear in Spec of ShareP", corresponding to the distributee. In (77)-(78), however, the distributee, being in [Spec, I], is higher than ShareP and is therefore not available to fill its Spec. In such cases Beghelli and Stowell (1997) assume movement of a "covert existential quantifier over events" to this position.

Our idea is that in the raising case in (77) a successful interpretation is reached in that the matrix ShareP can be associated with a quantification over the embedded *assign* event across the event-less *seem*. In the control case in (78), on the other hand, a successful interpretation would require that both a quantification over the matrix and the embedded event be associated with the same matrix ShareP. This we take to be impossible, as desired.

Finally, it should be noticed that the present theory of A-movement is not compatible with the analysis of floating quantifiers argued for in Sportiche (1988), according to which quantifiers such as *tous* in French are stranded in their base-generated thematic position by leftward movement of the D. But though this analysis has proven very influential, there are several pieces of evidence that cast doubt on it, as argued by Bobaljik (1998) quite independently of the present theory.

7 Conclusion

To conclude, in section 1 we argued that there are empirical and conceptual problems for the classical construal of control, involving PRO, as well as for that of A-movement. In section 2 we argued that an alternative formalization for A-movement is possible within the minimalist framework, whereby DPs are merged directly in D-position, where they This alternative formalization for A-ATTRACT predicates from the VP-shell. movement, which effectively eliminates DP-traces, is supported by the lack of reconstruction effects at LF and of blocking effects on rules at PF. What is more, the theory proposed in section 2 allows for a straightforward account of control which dispenses with PRO. Quite simply, control corresponds to the case in which more than one predicate is ATTRACTed by the same DP argument. In section 4, we concentrated on one difference between our approach and O'Neil's (1995), Hornstein's (1996), for whom control reduces to DP-movement and arbitrary control must therefore involve an empty pronominal subject pro. In present terms, the empty category pro can be eliminated from the grammar as well. Arbitrary control reduces to the case in which a predicate is controlled by an operator in C, rather than by a DP-argument.

Throughout sections 3-4, we showed that the distribution of PRO, the distribution of antecedents for controlled PRO, and the distribution of controlled and arbitrary PRO are determined uniquely by the interplay of ATTRACT with the MLC, which we construe as *Scopal MLC*. Similarly, in section 5, we explained obligatory control into adjuncts as

a by-product of parasitic gap-like patterns of predicate attraction predicted by *Connectedness*. Arbitrary control into subjects amounts to the same, with the difference that an operator in C acts as the attractor, rather than a DP-argument. In section 6, we turned to A-movement again considering more complex data involving reconstruction and other LF effects.

References

- Alexiadou, A. & E. Anagnostopoulou (1997). Parametrizing Agr: Word-Order, V-movement and EPP-checking. Ms., FAS Berlin and University of Tilburg.
- Arad, M. (1998). VP-Structure and the Syntax-Lexicon Interface. Doctoral dissertation, University College London.
- Baker, M., K. Johnson & I. Roberts (1989). Passive Arguments Raised. Linguistic Inquiry 20: 219-251.
- Beghelli, F. & T. Stowell (1997). Distributivity and negation. In A. Szabolcsi (ed.). Ways of scopetaking. Dordrecht: Kluwer.
- Belletti, A. & L. Rizzi. (1988). Psych-Verbs and q-Theory. *Natural Language and Linguistic Theory* 6: 291-352.
- Bennis, H. & T. Hoekstra (1984). Gaps and parasitic Gaps. The Linguistic Review 4: 29-87.
- Borer, H. (1994). The Projection of Arguments. Umass Occasional Papers 17. GLSA, University of Massachusetts, Amherst.
- Bobaljik, J. (1995). *Morphosyntax: The Syntax of Verbal Inflection*. Doctoral dissertation, MIT, Cambridge, Mass.
- Bobaljik, J. (1998).GLOT.
- Boškovic, Z. (1997). *The Sntax of Nonfinite Complementation : An Economy Approach*. Cambridge, Mass.: MIT Press.
- Bresnan, J. (1972). *Theory of Complementation in English*. Doctoral dissertation, MIT, Cambridge, Mass.
- Bresnan, J. (1982). Control and Complementation. Linguistic Inquiry 13: 343-434.
- Brody, M. (1995). Towards perfect syntax. *Working Papers in the Theory of Grammar* 2. Research Institute for Linguistics, Hungarian Academy of Sciences.
- Brody, M. (1996). Some Restrictive Aspects of Perfect Syntax. Paper presentend at the University of Stuttgart.
- Brody, M. & M. R. Manzini (1988). On Implicit Arguments. In Ruth Kempson (ed.). *Mental Representations: the interface between language and reality*. Cambridge: Cambridge University Press.
- Burzio, L. (1986). Italian Syntax. Dordrecht: Kluwer.
- Chomsky, N. (1975 [1955]). *The Logical Structure of Linguistic Theory*. Cambridge, Mass: The MIT Press.
- Chomsky, N. (1981). Lectures on Government and Binding. Dordrecht: Foris.
- Chomsky, N. (1982). Some Concepts and Consequences of the Theory of Government and Binding. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1986). Knowledge of Language. New York: Praeger.
- Chomsky, N. (1995). The Minimalist Program. Cambridge, Mass.: MIT Press.

Chomsky, N. (1998). Minimalist Inquiries: the Framework. Ms., MIT.

- Chomsky, N. & H. Lasnik. (1995). The theory of Principles and Parameters. In N. Chomsky. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Cinque, G.(1990). Types of A'-dependencies. Cambridge, Mass.: The MIT Press.
- Cormack, A. (1998). *Definitions: Implications for syntax, semantics and the language of thought.* New York: Garland.
- Enc, M. (1987). Anchoring conditions for Tense. Linguistic Inquiry 18: 633-657.
- Hale, K. & S. J. Keyser (1993). On Argument Structure and the Lexical Expression of Syntactic Relations. In K. Hale & S. J. Keyser (eds.). A View from Building 20. 53-109. Cambridge, Mass.: MIT Press.
- Hornstein, N. (1996). On Control. Ms., University of Maryland.
- Hornstein, N. (1997). Control in GB and Minimalism. GLOT 8.2: 3-6.
- Jacobson, P. (1992). Raising without movement. In R. Larson, S. Iatridou, U. Lahiri & J. Higginbotham (eds.). *Control and Grammar*. 149-194. Dordrecht: Kluwer.
- Johnson, K. (1991). Object positions. Natural Language and Linguistic Theory 9: 577-636.
- Kayne, R. (1984). Connectedness and Binary Branching. Dordrecht: Foris.
- Kayne, R. (1994). The Antisymmetry of Syntax. Cambridge, Mass: The MIT Press.
- Kayne, R. (1998). Talk given at the University of Florence.
- Larson, R. (1988). On the double object construction. Linguistic Inquiry 19: 335-391.
- Larson, R. (1991). Promise and the Theory of Control. In Linguistic Inquiry 22: 103-139.
- Longobardi, G. (1984). Connectedness and island constraints. In J. Gueron et al. (eds.). *Grammatical representation*. 169-185. Dordrecht: Foris.
- Manzini, M. R. (1983). On Control and Control Theory. Linguistic Inquiry 14: 421-446.
- Manzini, M. R. (1992). Locality. Cambridge, Mass.: The MIT Press.
- Manzini, M. R. (1994). Locality, minimalism and parasitic gaps. Linguistic Inquiry 25: 481-508.
- Manzini, M. R. (1995). The position of adjuncts: A reply to Kayne. *Quaderni del Dipartimento di Linguistica dell'Università di Firenze* 6.
- Manzini, M. R. (1997). A minimalist theory of weak islands. In Peter Culicover and Louise McNally (eds.). *The Limits of Syntax (Syntax and Semantics)*. New York: Academic Press.
- Manzini, M. R. (in press). Dependencies, phrase structure and extractions. In D. Adger et al. (eds.). *Specifiers: Minimalist approaches*. 188-205. Oxford: Clarendon Press.
- Manzini, M. R. (forthcoming). Sentential complementation: The subjunctive. In P. Coopmans, M. Everaert & J. Grimshaw (eds.). *Lexical selection and lexical insertions*. HIL.
- Manzini, M. R. & L. M. Savoia (1998). Clitics and auxiliary choice in Italian dialects: Their relevance for the Person ergativity split. *Recherches Linguistiques a Vincennes*.
- Manzini, M. R. & L. M. Savoia (forthcoming). Parameters of subject inflection in Italian dialects. In P. Svenonius (ed.). *Subjects, Expletives and the EPP*. Oxford: Oxford University Press.
- Martin, R. (1996). A minimalist theory of PRO and control. Ph.D. Dissertation, University of Connecticut, Storrs.
- Nash, L. & A. Rouveret. (1997). Proxy categories in Phrase Structure theory. Ms., University of Paris VIII.
- O'Neil, J. (1995). Out of Control. *Proceedings of NELS* 25. GLSA, University of Massachusetts, Amherst.
- Pica, P. (1987). On the Nature of the Reflexivization Cycle. *Proceedings of NELS 17*. GLSA, University of Massachusetts, Amherst.

- Platzack, Ch. (1995). Null subjects, weak Agr and syntactic differences in Scandinavian. Working Papers in Scandinavian Syntax 53: 85-106.
- Pollock, J.-Y. (1996). Langage et cognition: Introduction au programme minimaliste de la grammaire generative. Paris: Presses Universitaires de France.
- Postal, P. & G. Pullum. (1982). The contraction debate. *Linguistic Inquiry* 13: 122-138.
- Reinhart, T. & E. Reuland. (1993). Reflexivity. In Linguistic Inquiry 24: 657-720.
- Rizzi, L. (1990). Relativized Minimality. Cambridge, Mass.: The MIT Press.
- Roberts, I. & A. Roussou. (1997). Interface Interpretation. Ms., University of Stuttgart and University of Wales, Bangor.
- Roberts, I. & A. Roussou. (forthcoming). The EPP as a Condition on the T-Dependency. In P. Svenonius (ed.). *Subjects, Expletives and the EPP*. Oxford: Oxford University Press.
- Rosenbaum, P. (1967). *The Grammar of English Predicate Complement Constructions*. Cambridge, Mass.: MIT Press.
- Roussou, A. (forthcoming). Control and Raising in and out of Subjunctive Complements. In B. Joseph, A. Ralli & M.-L. Rivero (eds.). Oxford: Oxford University Press.
- Salles, H. (1997). *Prepositions and the Syntax of Complementation*. Doctoral Dissertation, University of Wales, Bangor.
- Sells, P. (1987). Aspects of Logophoricity. In Linguistic Inquiry 18: 445-480.
- Sportiche, D. (1988). A theory of floating quantifiers and its corrolaries for constituent structure. *Linguistic Inquiry* 19: 425-449.
- Sportiche, D. (1996). A-Reconstruction. Paper presented at Rutgers University.
- Stowell, T. (1982). The Tense of Infinitives. Linguistic Inquiry 13: 561-570.
- Tenny, C. (1994). Aspectual Roles and the Syntax-Semantics Interface. Dordrecht: Kluwer.
- Williams, E. (1974). Rule Ordering in Syntax. Doctoral dissertation, MIT, Cambridge, Mass.
- Williams, E. (1980). Predication. Linguistic Inquiry 11: 208-238.
- Williams, E. (1992). Adjunct Control. In R. Larson, S. Iatridou, U. Lahiri & J. Higginbotham (eds.). *Control and Grammar*. 297-332. Dordrecht: Kluwer.