

Linear Compression as a Trigger for Movement¹

Andrea Carlo Moro

1. ‘Beyond Explanatory Adequacy’: How the World Shapes Grammar

A new challenge has been addressed in generative grammar in a recent paper by Chomsky: “in principle ... we can seek a level of explanation deeper than explanatory adequacy, asking not only *what* the properties of language are, but *why* they are that way” (Chomsky 2001: 2). More specifically, the idea within a biolinguistic approach is that the initial state of language acquisition is not fully genetically-determined² but it is also a function of the general property of organic systems and more generally it reflects the properties of the physical world.

There are different ways to approach global issues in science. One can discuss the fundamental aspects of a theory dissecting all aspects in detail; on the other hand, a totally different strategy, and surely a fruitful one in science, is to concentrate on a single simple problem and see what consequences comes from an attempt to solve it. In this paper I will follow this second path focussing on two straightforward questions and proposing a unified answer to them. The aim of this paper is in fact twofold: on the one hand, I will reinforce a non-standard theory of movement; on the other, I will suggest that this theory is a plausible candidate to study the possible connection between constraints on grammar and some properties of the organic systems. I have organized the argument in five sections, excluding the present one: a section illustrating the two questions (section II) immediately followed by a section illustrating the answer offered by the standard theory (section III); a further section illustrating the alternative theory and the unified answer to the two questions (section IV); another section indicating some general empirical consequences of the theory (section V) and a final section speculating on a recent interpretation of form and function in grammar within the Minimalist Program (section VI). The guiding thread of this discussion is the search for a further level of adequacy getting beyond “explanatory adequacy”.

2. Two Questions on Movement

All grammars must include some notion of displacement. To put it in Chomsky's words, the fact that some lexical items appear displaced from those positions where they receive interpretation is "an irreducible fact ... expressed somehow in every contemporary theory of language" (Chomsky 1995: 222).³ If one adopts a multi-level grammar, more specifically a transformational grammar, displacement can be captured by assuming that a phrase moves from one position at a given level to a different one at another level.

The theory of movement which lies at the very core of grammar is a complex theory aiming at explaining different things: what moves, where can elements move, how far can elements move, what triggers movement, why there is movement, etc. Each of these conceptually distinct issues raise many questions. Let us focus on the following two simple ones:

- (1) i. What triggers movement?
- ii. Why does movement involve deletion of phonological features?

There is no a priori reason to assume that the answer to these questions is unique. In fact, in the standard Minimalist framework stemming from Chomsky 1995 up to Chomsky 2001, for example, the phenomena referred to in (1) are treated separately. Let us then start by considering them separately.

3. The Standard Theory of Movement: Morphology as a Trigger for Movement

What triggers movement? Synthetically, the core idea of the Minimalist theory of movement is that it is triggered by the necessity to present the interface with the conceptual-intentional module (LF) with interpretable features only (Full Interpretation). Technically, this treatment is grounded on three conceptually independent assumptions: first, following Chomsky's own words "It is clear that there are uninterpretable features" (Chomsky 2001: 11); second, pairing an uninterpretable feature with a feature (of the same type) in an appropriate local relation deletes the uninterpretable feature by hypothesis (cf. for example Chomsky 1995 and Chomsky 2001); third, movement is just one way to implement pairing, in fact a last resort to achieve pairing.⁴ In conclusion, within the standard theory of movement, "... uninterpretable

features are the mechanism for displacement” (Chomsky 2001: 14). Let me synthetically summarize the three hypotheses upon which the standard theory of movement is grounded and briefly comment on them (taking Full Interpretation for granted as an property of human language):

- (2) The standard theory of movement:
 - i. There are uninterpretable features.
 - ii. Pairing features deletes the uninterpretable feature.
 - iii. Movement is a last resort operation to achieve pairing.

Clearly, the three statements in (2) are axioms. There is no obvious way to derive any of them from principled grounds. Their force can only be tested by evaluating the empirical predictions that the theory makes. A priori, any other set of axioms could equally well be adopted.

Consider (2i) first. It seems to me that there is so far no independent way to prove that a feature is not interpretable. This is quite a delicate point. When the notion of “interpretable feature” was first proposed, it was indeed assumed that one could observe the effect that such a feature has on the system, i.e. movement, but if one wants to show that movement is the effect of the existence of uninterpretable features, this kind of evidence cannot be exploited, otherwise the argument would be circular. The core question thus remains: is there an intrinsic independent way to say which features are uninterpretable? Chomsky’s statement that “It is clear that there are uninterpretable features” (Chomsky 2001: 11) seems to me to be too strong.

Note in passing that we cannot rely on any “intuition” about the interpretation of features (for example their contribution to meaning): the very fact that a feature is not interpretable *by a linguist* does not imply that the same feature is not interpretable *by the system*. The inquiry on the functional role of elements within a code is reminiscent of a different realm, i.e. genomics. For several years, entire segments of DNA were considered as uninterpretable by biologists; nowadays, advanced research in the field has shown that this cannot be simply the case: “junk DNA” turned out to be useful to protect genes from mutations due to external intervening factors and in certain cases it revealed an important function in delimiting the border of genes in the sequence. The system appeared to “interpret” things differently from the scientist (see Mayr 1988). Similarly, we cannot immediately assume that a feature is not interpretable because we do not understand its function. Rather we should find an independent way to establish whether it is indeed so. Thus, saying that Case features, for instance, are not interpretable qualifies as an axiom rather than a theorem.

Similar considerations arise with respect to (2ii), i.e. the idea that pairing interpretable features (such as number on nouns) with uninterpretable features⁵ (such as number on verbs or adjectives) deletes the uninterpretable features. This is not a fact, it seems to me to be rather a problematic hypothesis for at least three conceptually distinct reasons.⁶ First, for this theory of deletion to work, an extra assumption must be made, namely that uninterpretable morphological features are deleted on *all* copies not just on the one in the proper local configuration with a feature of the same type otherwise copying would be totally useless, in fact counterproductive. Notice that deletion “at a distance” should happen even if two (or more) copies belong to distinct phases (or cycles) yielding non-trivial problems in a derivational perspective. Second, if the features are deleted they must be deleted only after PF otherwise we should never pronounce them, contrary to the facts. Take for example Case features (which are considered to be always uninterpretable) in languages with fully overt Case morphology such as Latin; were they deleted before PF, we should not pronounce them, unless, of course, we assume ad hoc that we pronounce only the physical, i.e. phonological, support of the features (such as /em/ in *Caesarem* [Caesar, Accusative] or /i/ in *Caesari* [Caesar, Genitive] for example) while the uninterpretable Case features proper have been deleted. In other words, this would amount to assuming that *overt* features cannot be the trigger for *overt* movement yielding paradoxical consequences. For example, this generalization would make treatment of movement of noun phrases in languages with overt fully Case morphology hard to explain. Consider the limited domain of Romance clitics: it would be hard to explain object clitic movement, since the Accusative features are clearly manifest on the clitic itself after movement as in Italian *Gianni le ha viste* (Gianni them-fem.plur.-has seen-fem.plur.). Third, the notion of “feature deletion” requires a rather complex formulation. In fact, feature deletion appears to affect copies of an item in quite a different way depending on the type of feature involved. Deletion affects all copies of an item when it comes to morphological features but it affects *all copies but one* when phonological features are considered. Thus, besides the fact that one still needs a principled reasons as to why deletion of phonological features is involved in movement, it is the notion of deletion itself that becomes problematic.

As for the third statement (2iii), namely, that movement is a last resort to achieve pairing, the problem is empirical: if this is correct, then one should not expect to see an element staying in situ if it can move in a structure stemming from the same basic lexical array. More formally, suppose the lexical array L contains the item X endowed with an uninterpretable feature.

Suppose that X needs to move to delete the uninterpretable feature. Then, given the same lexical array L, there could not be an alternative structure where X does not move and (say) Y moves.⁷ We will come back to this issue in the section dedicated to the consequences of the alternative theory presented here. However, for the sake of clarity the core part of the argument concerning this issue can be anticipated as follows. If one can show that there is a set of structures in languages, call them “mirror structures”, where the same lexical array yields both the order ...X...Y... and ...Y...X... crucially via movement, then the idea that movement is a last resort operation driven by the necessity to delete the uninterpretable features of either X or Y would diminish in plausibility. I will try to show that such a set exists both across languages and across categories.

All in all, (2i) and (2ii) qualify as hypotheses not as facts and (2iii) is not empirically adequate if the argument proposed in the next sections is correct. Of course, this does not mean that the theory embodied in (2) is inconsistent, simply that one should be aware of this when choosing among competing theories, as we are going to do in the rest of the paper. For the sake of clarity, let us henceforth synthetically dub the standard minimalist theory of movement a “morphological theory of movement” since the major burden of labour in this framework falls on (uninterpretable) morphological features at the LF interface.⁸

Clearly, the morphological theory of movement offers no immediate answer to question (1ii), i.e. “why does movement imply deletion of phonological features?”. This point requires some comment. Deletion of phonological features wasn’t even posited in the theories that included (indexed) traces, i.e. instances of empty categories, as syntactic objects. Movement simply generated traces as links of a chain and no problem concerning phonological features was raised. The question has been revived in the Minimalist framework for movement has been reinterpreted as merging the same constituent in distinct positions. Leaving aside the various reasons which led to this revised form of generalized transformations (such as the solution to reconstruction problems involving unwanted downward movement, etc.) this theory – standardly referred to as “copy theory of movement” – now constitutes the standard approach to movement. For example, in a sentence like *What books did John read?* the theory states that the phrase *What books* is merged in two distinct positions: as the object of *read* and as the specifier of *did* (disregarding vP).⁹ Accordingly, question (1ii) can be properly paraphrased as follows: why does movement involve deletion of phonological features of all occurrences (or “copies”) of an element but one?”; more specifically, why is only the topmost occurrence of *what books* pronounced?

This question has not been directly approached in a systematic way in Chomsky's original papers stemming from Chomsky 1995. Nevertheless, it has indeed been independently addressed at least since the mid Nineties by several authors. A brief survey of the original sources (excluding those related to the proposal defended here; cf. endnote 1) would include among others: Brody 1995, Groat & O'Neil 1996, Pesetsky 1997, Nuñez 2001 and Bobaljik (to appear). It goes without saying that any answer based on a "naïve" notion of economy excluding pronunciation of more than one copy would not have any explanatory force, at least not until a principled theory of economy of pronunciation is proposed.¹⁰

In conclusion, a major theoretical issue should be pointed out: clearly, both in the standard morphological theory stemming from Chomsky's seminal proposals and in the works just cited here the two questions in (1) do receive conceptually separate answers. Movement is thus regarded as a heterogeneous phenomenon: on the one hand, it is triggered by interpretative reasons (Full Interpretation: specifically, feature readability at LF); on the other it affects the articulatory-perceptual module forcing deletion of phonological features of lexical items at PF. The specific proposal defended here, by contrast, is an attempt to answer both questions by referring to a single interface and thus conceptualize movement as a homogeneous phenomenon: both the trigger for movement and the deletion of phonological features are traced back to the interface conditions with the phonological-articulatory component. It is worth noting that if the proposal defended here proves correct, this theory would qualify as an attempt to derive some properties of grammar from the structure of an organic system (in that it refers to the articulatory system) according to the lines of thought suggested by Chomsky 2001. Whether or not the proposal is correct, of course, only future empirical research will tell. Nevertheless, it is worth noticing here that in the alternative proposal suggested in the next section, the role of the organic or physical world seems to be clearer than in the morphological theory, at least in so far as linearization is considered a physical requirement of human language.

4. A Unified Non-standard Theory of Movement: Linear Compression as a Trigger for Movement

Let us assume that the following three principles hold for the syntactic component of Universal Grammar:

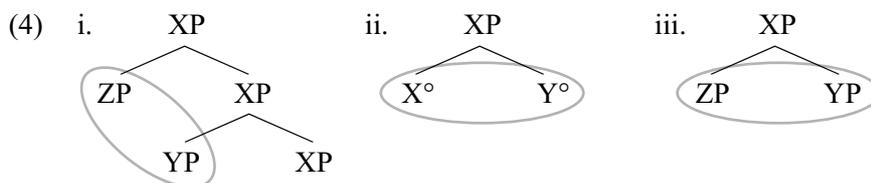
- (3) The alternative theory of movement:
 - i. Merge is unrestricted.
 - ii. The linear order of terminal nodes is established only when required, i.e. at spell-out.
 - iii. The linear order of terminal nodes is a function of hierarchy.

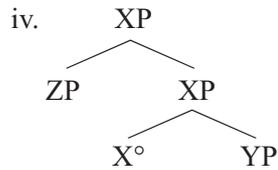
These principles have been adopted by different scholars in different frameworks (not uncontroversally) and are conceptually independent. Originally, (3i) and (3ii) were proposed in Chomsky (1995) and subsumed in many following works while (3iii) was proposed by Kayne (1994). Not only are these three principles independent, but crucially they are compatible with each other, i.e. one can imagine a consistent framework where these three principles hold simultaneously. In fact, such a framework was developed in Moro (1997b, 2000) suggesting an alternative way to think of movement as a consequence of the interaction of the three principles in (3). For the sake of clarity, I will summarize the idea as follows.

Let us preliminarily review these three principles in a synthetic way. The idea that Merge is unrestricted (i.e. 3i) expresses one of the essential properties of the operation Merge: the job carried out by this operation is simply to take two syntactic objects as an input yielding a third object as an output: that is, Merge takes the set of formal features of the syntactic objects X and Y yielding Z.¹¹ Indeed Merge does have restrictions, in particular it is assumed that Merge cannot add any further piece of information to those provided by X and Y (Inclusiveness Condition) and that it cannot create hybrid (and inconsistent) objects Z made by the intersection or the union of X and Y. Crucially, on the other hand, there is no upper limit as to how many times Merge can operate, in a given set of syntactic objects; this captures the essentially infinite, recursive and combinatorial character of human language.¹² The idea that linear order is established only when required (2ii), instead, was suggested by Chomsky (1995): if it weren't for the necessity to communicate sentences, hierarchical relations would be sufficient. Virtually all grammatical relations, ranging from agreement to Binding theory effects, are established hierarchically. Flattening structures into a linear sequence, thus, is just necessary for purely empirical, in fact physical, reasons

(pronunciation).¹³ In fact it is not hard to think of an organism which could be designed in a different way: had *homo sapiens* had two (or more) mouths, for example, we could easily imagine uttering parts of the same sentence simultaneously with no linear restriction. This is simply not the case. The third principle (2iii) establishes an explicit link between linear order of words in a sequence and hierarchy. Such a principle has been originally implemented by Kayne (1994). More specifically, in Kayne's (1994) theory of the Antisymmetry of syntax, precedence and antisymmetrical c-command are interconnected in a non-ambiguous way so that if a terminal node *x* precedes a terminal node *y* it must be the case that there is a non terminal node *X* dominating *x* asymmetrically c-commanding at least one non terminal node *Y* dominating *y* (Linear Correspondence Axiom). One of the major consequences of this approach is that it derives all major properties of the X-bar theory.¹⁴ In fact, as a corollary, taken together with the assumption that all pairs of terminal nodes features must be ordered, it follows that Merge cannot create "too symmetrical" structures,¹⁵ call them "points of symmetry", otherwise it would not be possible to put the terminal nodes contained in the symmetrical non-terminal nodes into a sequence.

The crucial step was to assume that these three principles in (3) are simultaneously active. This naturally led to an alternative theory of movement which will be synthetically presented here. Let us start from principle (3ii), i.e. let us assume that the LCA is active only when needed, namely at the interface with the articulatory-perceptual component (PF).¹⁶ Prior to that point, syntactic trees are free to contain points of symmetry as a consequence of principle (3i) which allows Merge to reiterate unboundedly. In fact, all types of symmetrical structures can be reduced to the basic types in (4i–iii).¹⁷ Given the definition of c-command proposed by Kayne (1994: 16)¹⁸, the constituents within the circle diagrams symmetrically c-command each other and the LCA makes it impossible to put the terminal nodes contained into such constituents into a linear order. Thus, the only derivable syntactic tree is the one in (4iv) (crucially eliminating intermediate X' projections; cf. Kayne 1994: 22–26):^{19,20}





Why should this weak version of the theory of Antisymmetry bear on movement? The specific proposal defended here is that movement can just be regarded as a way to rescue the structure at the interface with the phonological component in case a point of symmetry has been generated: movement deletes the phonological features of one element constituting a point of symmetry and copies it in a suitable position,²¹ thus solving the problem of linearization at PF. Technically, let us say that movement “neutralizes” a point of symmetry. To put it more generally, movement is regarded here as a consequence of the physical necessity to organize words into a linear order (call it “linear compression”) as opposed to the standard theory which considers it as triggered by uninterpretable features. This alternative theory of movement relying on a weak version of Kayne’s Antisymmetry theory has been labeled *Dynamic Antisymmetry*. It is important to notice that although this qualifies as an alternative theory with respect to the standard one, it is genuinely Minimalist in that it constitutes an attempt to trace back all syntactic principles to the interface with the two conceptually necessary levels of representation, namely PF and LF; the difference w.r.t. the standard theory is that the relevant interface here is not LF (where morphological features are interpreted) but PF (where words are put into a linear order).²² Moreover, as opposed to the standard theory, the crucial role of linearization (which again is only due to a property of the organic-physical world, rather than conceptual necessity) in the explanation suggests that a specific syntactic property may be related to the conditions imposed on the language faculty by the external world, as originally envisaged by Chomsky (2001).

In the next section I will focus on some general consequences of such theory leaving many questions and problems to the discussion in Moro (2000) and to future research. What is relevant here is that the approach based on the three principles in (3) leads to a unique answer to the two questions in (1) and it introduces into the theory an aspect of the physical world, i.e. the fact that structures must be flattened to produce utterances. In this theory, movement is triggered by the necessity to organize words into a linear sequence at spell-out;²³ deletion of phonological features turns out to be an obligatory property of movement, not a side phenomenon to be independently justified. Synthetically, adopting *Dynamic Antisymmetry* amounts

to pursuing the empirical and theoretical consequences of the following conjecture about Movement (i.e. copy and deletion):

(5) Movement intervenes if Merge generates a point of symmetry.

Of course, such a radically alternative theory raises many questions.²⁴ In fact, from a general point of view, a basic question arises as to why Dynamic Antisymmetry should be adopted, abandoning the standard morphological theory.

5. Some Consequences of the Alternative Theory

In this section, we will analytically review some consequences and expectation of Dynamic Antisymmetry both on theoretical and empirical grounds.²⁵ Let me begin with a simple list:

- (6) Some consequences and expectations of Dynamic Antisymmetry:
- i. Deletion of phonological features of lexical items is an obligatory property of movement, not a side phenomenon to be independently explained.
 - ii. Each structure generated by movement is in principle associated to a “mirror structure”.
 - iii. Lower copies of a moved XP can be spelled out only if they are reduced to X⁰s.
 - iv. If a point of symmetry is constituted by a base generated empty category then no movement is triggered.
 - v. *pro* cannot move.
 - vi. There is no covert movement (strong thesis) or movement is optional (weak thesis).
 - vii. Movement can vary across languages according to the X-bar status of lexical elements.

The first point presented here (6i) has in fact been discussed in the previous section and constitutes the core idea of the proposal defended here: deletion of phonological features is not to be regarded as a side phenomenon with respect to movement as an independent phenomenon. Rather, it constitutes an essential part of movement: movement *is* nothing but deletion of the

phonological features of either element constituting a point of symmetry and copy of the same element in a suitable LCA compatible position, pace locality conditions. This also gives us a principled reason as to why the top-most “copy” is pronounced: the “original” element in the low position cannot be pronounced because the LCA prevents the structure to be linearized. In fact, this is why movement takes place.²⁶

The consequences in (6ii) and (6iii) can be better understood by taking the structure of a point of symmetry given in (4) as a guideline. Consider (6ii) first. A point of symmetry has two defining properties: simplifying somewhat, it is constituted by two syntactic objects mutually c-commanding each other and the two objects must share the same categorial status, i.e. they are both either XPs or X^os. Moreover, for a point of symmetry to be an offending structure it must be the case that both elements are overtly realized, otherwise the LCA would not rule it out since null elements are not visible to linearization by hypothesis. We will come back to the case where an empty category is involved. Let us focus now on the case where they are both overt. We have two subcases to consider, related to (4i–iii) and (4ii), involving XPs and X^os respectively. Here we will consider points of symmetry made by XPs only.²⁷ A priori, since each point of symmetry is constituted by *two* elements X and Y one expects that each point of symmetry can be neutralized in *two* ways. Whenever X moves in a given structure where X and Y constitute a point of symmetry, an associated structure should exist where Y moves leaving X in situ. Of course, it seems at first glance that this expectation is not borne out: there must be factors that “obscure” such splitting processes which would otherwise be much more manifest in syntax. Nevertheless, a cursory cross-linguistic review suggests that there exists a core group of cases which naturally qualify for such an analysis. Consider for example the following paradigm:

- (7)
- a. [a picture of the wall] was [the cause of the riot]
 - b. [the cause of the riot] was [a picture of the wall]
 - c. John reads [books] of [these types]
 - d. John reads [these types] of [books]
 - e. [John] is [kind]
 - f. that is [kind] of [John]
 - g. Maria fa [riparare la macchina] a [Gianni]
(Maria makes repair the car to Gianni)
 - h. Maria [gli] fa [riparare la macchina]
(Maria him-makes repair the car)

- i [Gianni] [telefona] (Gianni telephones)
 l [telefona] [Gianni] (telephones Gianni)

Sentences of each pair share one characteristic: two bracketed constituents can appear in two possible orders.²⁸ I would like to suggest that this is a consequence of the “symmetry-breaking” nature of movement.

The crucial step is of course to show that movement is indeed involved (in a non trivial way) in generating the structures in (7): in other words, one must exclude that each pair is generated by inserting the bracketed elements directly in the order they display in the surface. There are robust empirical reasons to claim that the former is in fact the case. For most of the examples cited in (7), a movement analysis has been independently suggested in previous works. Since for space reasons the empirical motivations for these analyses cannot be illustrated here, let us simply reproduce the structures assigned to the sentences in (7) and briefly review the original sources of the corresponding analyses:

- (8) a. [a picture of the wall] was [_{SC} t [the cause of the riot]]
 b. [the cause of the riot] was [_{SC} [a picture of the wall] t]
 c. John reads [books] of [_{SC} t [these types]]
 d. John reads [these types] of [_{SC} [books] t]
 e. [John] is [_{SC} t [kind]]
 f. that is [kind] of [_{SC} [John] t]
 g. Maria fa [[riparare la macchina] a [_{SC} [Gianni] t]
 (Maria makes repair the car to Gianni)
 h. Maria [gli] fa [t H° [_{SC} t [riparare la macchina]]]
 (Maria him-makes repair the car)
 i. [_{SC} [Gianni] [_{SC} t [telefona]]] (Gianni telephones)
 l. [_{SC} [telefona] [_{SC} [Gianni] t] (telephones Gianni)

The structure of the sentences in (8a–b) were originally proposed in Moro (1988; see Moro 1997a for a comprehensive version) and they correspond to the so called “canonical” and “inverse” copular sentences involving raising of either the subject or the predicative noun phrase within the IP domain respectively; as for (8c–d) involving raising of the subject or the predicative noun phrase within the domain of DPs rather than IPs see Kayne (1994), Zamparelli (1995), Bennis, den Dikken & Corver (1998) and Moro (2000) among others;²⁹ (8e–f) illustrating inversion with an AP predicate was pro-

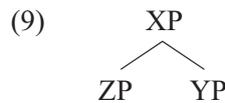
posed in a footnote in Moro (2000); the structures of causative constructions in (8g–h) have been suggested in Guasti – Moro 2001: they involve either movement of a predicative VP to the specifier of the prepositional complementizer *a*, or movement of the clitic subject out of the embedded clausal constituent through the specifier of the null prepositional complementizer H° (for a partially similar proposal see also Kayne 2001).³⁰ Let us skip (8i–l) for a moment and consider the structures from (8a) through (8h).

Assuming now that these analyses involving movement are indeed correct, i.e. there is no alternative way to produce the distinct orders by base generation, these examples stand out as clear cases of “mirror” structures, as expected by Dynamic Antisymmetry and will turn out to be a serious problem for the Minimalist theory of movement. In each and every pair, there are two distinct ways to neutralize the point of symmetry, i.e. moving either bracketed element. Since in all the analyses referred to here each element moved is linked by a predicative relation to the other, one could naturally describe the paradigm in (8) as generated by moving either the subject or the predicate to neutralize the point of symmetry. This approach of course raises several questions, above all: what type of point of symmetry is involved in predication? More specifically, is predication implemented in syntax as (3i), (3ii) or (3iii)? The answer to such a question is not trivial and involves rethinking many different issues. Since all pieces of evidence cannot be reproduced here I will simply illustrate the major lines of reasoning.

The structures illustrated here in (8a) through (8h) have all been treated as involving “small clauses”. The notion of “small clause” originally proposed by Williams (1975) (and critically revised in Williams 1994) is a problematic one. While, in actual fact, this label has been exploited for very different constituents (see Cardinaletti & Guasti 1995 and Graffi 1997 for a critical review) but there has been strong convergence over the years that it should be reserved to predicative linking, especially in non finite contexts. This gives us a clue to understand if and what type of point of symmetry is involved in the structures in (8a) through (8h) which all contain small clauses: in fact, the question can now be reduced to asking whether small clauses can be considered as points of symmetry.

Among the three options in (3), obviously (3iii) suggests itself as a natural candidate: in fact (3iii) essentially reproduces William’s original representation. Unfortunately, one cannot immediately consider all instances of small clauses recognized in the literature as instances of the structure in (3iii) otherwise, for example, one should expect raising from all small clause complements of *believe*-type verbs, contrary to the facts. In Moro (2000) I have suggested that the notion of “small clause” is insufficient to represent

those syntactic relations that have been standardly covered by such label, proposing that the notion of “small clause” should be split in two distinct sublabels: “rich small clauses”, which are projected by a head and are for example the complement of *believe*-type verbs, and “bare small clauses”, which are the complement of the copula, of D° and of C° .³¹ Now, since there are strong empirical arguments suggesting that bare small clauses are not projected by any head,³² the null hypothesis would be that only bare small clauses are indeed instances of (3iii) repeated here as (9) (see again endnote 10 for questions on the label of (9)):



This simultaneously gives us the reason why movement out of the small clause is necessary and the reason why there are two possibilities to neutralize each point of symmetry. The two phrases connected by a predicative relation (corresponding to ZP or YP) are merged yielding the bare small clause (YP) without the intervention of a head. In conclusion, movement of either phrase in each example in (8) neutralizes the point of symmetry instantiated by the predicative relations as implemented in the associated small clause.

As for the idea that the label of a small clause be distinct from that of either element constituting the small clause, this seems to be unavoidable. In fact, although it remains implicit in many frameworks, the fact that the distribution of the constituent resulting from merging a subject and a predicate can be neither the same as that of the subject nor as that of the predicate must be captured somehow by all grammars: in simple words, merging a subject DP with a(n inflected) predicate VP can neither yield a DP nor a VP; it yields a clause. Implementing such an intuition in a formal system has been a long-standing issue since the first works in generative grammar: for finite clauses the theory has stabilized in Chomsky (1986) proposing to consider clause structure as the X-bar projection of the C° – I° systems; for non finite clauses, instead, the issue has not reached similar convergence, especially in the case of non-verbal predication: since at least Stowell 1981 the problem of representing small clauses has been sharply debated; for a historical review see Graffi (1997).³³ Synthesizing, we will consider a bare small clause to be the distinct maximal projection generated by merging two maximal projections; furthermore, we will assume that a small clause is the implementation of all and only predicative linkings and it can be the complement of functional heads only.²²

Let us now turn to the residual pair proposed here, i.e. (8i–l): some further discussion is required. The structures assigned in (8i–l), involving a small clause structure linking the subject *Gianni* with the predicate *telefona*, rely on the analysis of clause structure proposed in Moro (2000): for independent reasons related to wh-movement in a Dynamic Antisymmetry framework, it was proposed that when the subject merges with IP the resulting label is a bare small clause rather than (a two segment) IP.³⁴ Whether or not this is correct depends on the analysis offered in the cited work, but if we assume so, some interesting consequences would follow. First of all, the alternation between preverbal and postverbal subject in Italian would be captured:³⁵ in fact, the alternation would correspond to the two distinct options for neutralizing the very same point of symmetry. This would make such an alternance a subcase of a more general paradigm illustrated in (8). But this would not be the only advantage. If one considers the canonical-inverse alternation affecting copular constructions as in (8a–b) one can easily conclude by a rapid recognition of the basic interpretative properties of these two sentences that the subject of inverse sentences is *obligatorily* focused. Notice that this cannot be captured by saying that the postverbal subject has raised to the specifier of the Foc° head of the Complementizer field in the sense of Rizzi (1997) (with successive remnant movement of the IP) because there are robust empirical reasons to assume that the subject of inverse copular constructions never moves (see Moro 1997a: 23–30).³⁶ Accordingly, the subject of an inverse copular sentence cannot even move to any VP-internal Focus position (for the existence of VP-internal focus positions see Belletti (1999), Longobardi (1999) and references cited there). Moreover, if this analysis were granted, it would qualify as no more an ad hoc solution: why should movement to such a position be obligatory for the postverbal DP of inverse copular sentences while crucially banned to the postverbal DP of canonical copular sentences?³⁷ All in all, leaving a principled explanation to future research, one can at least descriptively capture the situation by assuming that the position occupied by the subject of an inverse copular sentence in a bare small clause is unambiguously associated to Focus interpretation (we will come back to this in section IV).³⁸ The interesting fact to be noticed now is that this observation concerning Focus can be generalized to non copular sentences such as those in (8i–h): the postverbal subject in Italian is focused since it is in exactly the same configuration as the subject of inverse copular constructions, namely it is in situ within a bare small clause. This is so because the point of symmetry has been neutralized by moving the other element constituting it, namely the VP. When the subject is raised, instead, usual assumptions remain valid: Focus interpretation is optional and

arguably associated to further movement of the subject to the specifier of a Foc° head within the Complementizer field along the lines suggested in the pioneering work by Rizzi (1997).

Before moving to consider a further consequence of Dynamic Antisymmetry a major question we implicitly considered must be spelled out. I have illustrated here cases where movement exhibits *two* options related to the dyadic structure of point of symmetry yielding “mirror” structures across categories. As a matter of fact, as noticed before, it is easy to realize that it is not the case that each movement is always associated to two “mirror” structures. Why aren’t such “splitting” processes affecting points of symmetry pervasive? Obviously, this state of affairs would diminish the plausibility of Dynamic Antisymmetry unless one can find a way to explain why “mirror” structures are less pervasive than expected. In this section I will defend Dynamic Antisymmetry suggesting some reasons as to why splitting processes are often obscured in syntax.

Consider again the alternation between canonical vs. inverse copular constructions by focusing on the following pair:

- (10) a. [a picture of the wall] was [t [the cause of the riot]]
 b. [the cause of the riot] was [[a picture of the wall] t]

The point of symmetry constituted by two DPs c-commanding each other can be neutralized by moving either DP to preverbal position (usually labeled as specifier of IP). Now consider the following contrast:

- (11) a. [a picture of the wall] was [t [great]]
 b. *[great] was [[a picture of the wall] t]

The second sentence is ungrammatical, although we must assume for purely configurational reasons that a point of symmetry is constituted by the bare small clause selected by the copula. Why is it so? The reason is that for independent motives the immediate preverbal position is available to CPs/DPs only.³⁹ Of course this state of affairs cannot per se be a problem for Dynamic Antisymmetry: such a theory aims at explaining what *triggers* movement and at proving that it is not related to the interpretability of features but rather to the geometry of phrase structure. It goes without saying that it does not deny the relevance of features for movement: features can indeed select what type of solution is *compatible* with the structure although they do not trigger movement, as shown by examples like (11).

But feature compatibility is not the only reason that might selectively reduce the number of solutions for the neutralization of a point of symmetry: locality plays a crucial role. In the following paragraph, I would like to elaborate briefly on this point (which is treated in detail in Moro 2000).

In the cases examined in (8), we have seen that symmetry is generated by merging two maximal projections yielding a predicative nucleus (the bare small clause). What kind of symmetrical configuration would now trigger wh-movement? Is wh-movement indeed driven by the necessity to neutralize a point of symmetry? For the sake of simplicity, let us focus on a simple case involving a wh-object and reformulate the problem in a more concrete way: what is the point of symmetry which triggers movement of *which books* from the postverbal position in (12)?

(12) [what novels] has John written?

There are at least two different potential analyses under a Dynamic Antisymmetry approach. One is to explore the possibility that the wh-object itself is to be displaced to neutralize a point of symmetry which it constitutes *as a whole* with some other part of the VP structure. Alternatively, one could consider a different option, namely that the point of symmetry is *internal* to the wh-object and that the way the point of symmetry is neutralized forces further movement involving a form of pied-piping. A priori, there seems to be no logical reason to exclude either option although clearly if the first one were true we would face the problem of understanding why a non-wh object does not need to move. Indeed, a closer inspection of the typology of wh-movement both across and within languages suggests that the second analysis is more adequate on empirical grounds. It is a well-known fact that in certain languages wh-movement might involve “splitting” between the wh-element and the associated lexical item. A prototypical case would be the so called *was-fuer* construction in German, *wat-voor* constructions in Dutch, the *combien*-constructions in French and *cosa-di* constructions in low-register Italian where the wh-phrase can be raised alone leaving the lexical item in situ. Let us compare Dutch, German and Italian:

- (13) a. wat heeft hij voor romans geschreven
 (what has he for novels written)
 b. Was hat Johan für Bücher gelesen
 (what has Johan for books read)
 c. cosa ha scritto di romanzi?
 (what has he written for novels)

In general, these constructions have been analyzed as deviating from a more basic variant not involving *voor*, *für* and *di*, in Dutch, German and Italian respectively, such as in: *Welche boeken heeft hij geschreven?* (what books has he written) *Welche Bücher hat Johann gelesen?* (what books has Johann read) and *quali romanzi ha scritto* (what books has written)? However, one of the central concerns of the Dynamic Antisymmetry program of research is to show that splitting in wh-constructions is more pervasive across and within languages than usually thought. Accordingly, constructions involving splitting as the one in (13) have been reconsidered and it has been suggested that they are not exceptional; rather, they reveal the otherwise hidden symmetry-breaking nature of wh-movement. To understand the argument it is sufficient to recall one of the structures examined in (8), namely (8c–d) repeated here as (14a–b):

- (14) a. John reads [books] of [t [these types]]
 b. John reads [these types] of [[books] t]

In this case, movement is interpreted as a way to neutralize the point of symmetry in the small clause constituted by the predicative linking between the subject *books* and the predicate *this type*.⁴⁰ The idea then was to analyze (13a–c) as analogous to (14b) in that they all involve raising of predicates from a bare small clause:

- (15) a. wat heeft hij [t voor [romans t]] geschreven
 (what has he for novels written)
 b. Was hat Johan [t für [Bücher t]] gelesen
 (what has Johan for books read)
 c. cosa ha scritto [t di [romanzi t]]
 (what has he for novels written)

Two short comments on this analysis. First, this analysis is based on the idea that wh-phrases such as *what* are the interrogative counterpart of *this type* in (14), namely as predicates of the NP subject of the small clause rather than as realization of a D° ; whether or not this hypothesis is tenable is discussed in Moro 2000 and will not be reproduced here. Second, (14b) differs from (15a–b) in that the raised predicative element *this type* stops in the specifier position of the preposition/complementizer whereas *was/cosa* proceeds one step further up. Why is this? I have elsewhere proposed to consider this extra step as a consequence of the constitution of a further point of symmetry.

Recall that in Kayne's (1994) framework adjunct/specifiers c-command out of the maximal projection they are merged to. If we combine this with the reasonable assumption that *wat/was/cosa* are heads, we can derive the answer. Raising *wat/was/cosa* to the specifier position of the preposition/complementizer generates a new point of symmetry, since the non terminal node immediately dominating *wat/was/cosa* and the one immediately dominating the main verb would c-command each other: the only option is for *wat/was/cosa* to further raise to a suitable position in the left periphery. On the other hand, the reason why *this type* does not further raise to the same position as *wat/was/cosa* is related to the fact that *this type* is an XP not a head. As a result of this, c-command out of the phrase it is specifier/adjunct of is not problematic. There is asymmetric c-command between the verbal X° *reads* and the XP *this type*. That is sufficient for the LCA to allow linearization of the underlying words without further movement.

Let us now turn back to the main residual question of this section. What can obscure the splitting processes forced by Dynamic Antisymmetry? Now the role of locality conditions can be better understood. Let us concentrate on a simple contrast in Italian showing that no splitting process is possible:

- (16) a. **quali* ha scritto [t P° [racconti t]] ?
 b. [*quali* P° [racconti t]] ha scritto t ?

The wh-element *quali* cannot be extracted from the specifier position of the null preposition/complementizer P° . The specific proposal I have made is that this is due to locality condition on licensing of the empty category. In other words, splitting takes place in Italian too, but the process is obscured by the fact that P° cannot license the trace in its specifier position. The only option is pied piping the whole constituent as indicated in (16b).⁴¹ Notice that there is independent evidence that extraction from the specifier of a functional projection in the nominal domain might involve pied piping. Consider for example the following case:

- (17) a. **which* photographer's did John buy [t D° [t pictures of Rome]] ?
 b. [*which* photographer's D° [t pictures of Rome]] did John buy ?

The wh-phrase *which photographer's* cannot be extracted from the specifier position of D° ; the only option is to pied pipe the whole constituent paralleling (16b). That this is due to locality condition on empty category licensing is proved by the fact that if the trace is governed by a lexical head pied piping does not take place:

(18) which city did John buy [D° [pictures of t]] ?

It is not unreasonable, then, to assume that splitting is allowed only if there is a proper head which can license the empty category involved.⁴² This is realized as an overt P° in Dutch, German and Italian, such as *voor/für/di*; when P° is inert to government, on the other hand, as in the case of a null P°, pied piping is forced.⁴³

Interestingly, notice that this approach fits in well with similar considerations on proper government made by Rizzi (1990) who noticed the following symmetric contrast:

- (19) a. [how high] is that tower?
 b. *[how] is that tower [high]?
 c. [quanto] è [alta] quella torre?
 (how is high that tower)
 d. *[quanto alta] è quella torre?
 (how tall is that tower)

The contrast was traced back to the following independent opposition between the two languages:

- (20) a. that tower is [1000 meter F° [high]]
 b. quella torre è [F° [alta 1000 metri]]
 (that tower is tall 1000 meters)

The measure phrase would be extracted directly from the specifier position in English whereas it would be extracted from a lexically governed position in Italian. Thus the contrast in (19) follows as a consequence of the locality conditions on empty category licensing.⁴⁴

Summarizing so far, we have explored one exclusive expectation of Dynamic Antisymmetry among those synthesized in (6ii), namely that one should observe across and within languages “mirror” structures where two syntactic objects X and Y surface with both possible linear orders as a result of movement required by the neutralization of a point of symmetry constituted by X and Y. Interestingly, we have seen that a common characteristic of a broad class of examples involving the IP, AP and the DP domains is that X and Y are connected by a predicative linking providing indirect evidence in favor of a long standing proposal going back to Williams (1980) according to which the semantic notion of predication corresponds to a syntactic notion

of mutual command between maximal projections. The advantages of this proposal over the standard one based on the idea that movement is triggered by the necessity to delete uninterpretable features should be carefully evaluated. Here we have just illustrated some welcome consequences concerning clause structure. Notice however that the very existence of “mirror” structures such as the one in (8) constitute per se a challenging problem for the morphological theory of movement for the following reasons. In the standard Minimalist theory movement is considered not to be a costless operation: “Merge and Agree (or their combination) preempts Move, which is a “last resort”, chosen when nothing else is possible” (Chomsky 2000: 102). The alternation given by the “mirror” structures in (8) then constitute per se a challenging problem for such a view. If an element X can move to delete its uninterpretable features in a given structure, the prediction is that X cannot stay in situ allowing another element Y to move given the same basic structure. Obviously, this could not be consistent with the alleged “last resort” nature of the complex operation Move. Consider a concrete example: if the subject can raise in canonical copular constructions such as (8a), a last resort theory of movement predicts that there should not exist an inverse sentence associated to it such as (8b) where the same element remains in situ, contrary to the facts.⁴⁵ As for the empirical reasons why splitting processes can be obscured, we have isolated two distinct factors. The first type of reason is based on morphological considerations: Dynamic Antisymmetry suggests that movement is *triggered* by the necessity to linearize words at spell-out, nevertheless it does not deny that morphological features play a role in *selecting* what possibilities to neutralize a point of symmetry are viable. The second type of reason bears on locality conditions: we have seen that locality conditions on licensing empty categories play a dramatic role in allowing splitting processes, as in the case of wh-movement shown in the DP and AP domains.

Let us now turn to a further consequence of Dynamic Antisymmetry, i.e. to (6iii). Dynamic Antisymmetry does not imply that just one single occurrence of a moved syntactic object be pronounced. It rather states a weaker condition. “Copies” of an XP cannot be pronounced to the extent they violate the LCA at spell-out. This of course implies that a moved XP cannot be pronounced both in situ and in the displaced position; otherwise movement should not even take place according to Dynamic Antisymmetry. Indeed, the theory defended here leaves the logical possibility open that (intermediate) “copies” of the moved syntactic object be spelled out provided that no LCA problem is raised. More explicitly this could be possible to the extent syntax includes an operation of “phrasal reduction” preserving in a head only the

ϕ -features of the original phrase. This would remedy the LCA incompatibility because points of symmetry must be constituted by two categories of the same type. This abstract prediction naturally recalls a well-known empirical case which can be found for example in substandard Italian relative clauses and interrogatives. In these constructions, the *wh*-element can be “doubled” by a clitic, such as in *la stanza dove che ci metto i libri è fredda* (the room where that there-put the books is cold) where *dove* (where) is doubled by the clitic *ci* (there) or in *a quale donna pensi che Gianni le dà un fiore?* (to which woman think that Gianni her-gives a flower) where *quale donna* (which woman) is doubled by the clitic *le* (her) (for a discussion of such cases see Cinque 1988, 1990 and references cited there; for similar phenomena in Rumanian see Steriade 1980).⁴⁶

What matters here is that in both cases, the clitic pronoun is lower than the *wh*-phrase and shares (part of) its morphological features with the moved element, such as Case, number and gender features. Crucially, these constructions can not be easily interpreted within the morphological theory of movement regarding movement as a way to delete (uninterpretable) features. Disseminating features along the path of a *wh*-phrase is obviously inherently incompatible with a theory that makes feature deletion the engine for movement: why should an element move to a higher position if its features can be expressed in a lower position? Or why should *some* uninterpretable features, such as Case features, be left behind?⁴⁷ On the other hand, in the proposal defended here these cases are naturally captured: there is no intrinsic prohibition to expressing (an element sharing) the same features as the moved one in a lower position, for movement is not triggered by deletion of features.^{48,49} Let us now move to consider a further consequence expected under Dynamic Antisymmetry.

The consequences addressed in (6iv) and (6v) bear on the comparative issue in a rather non trivial way and they are strictly connected. For a point of symmetry to constitute a problem for linearization it must be the case that both phrases constituting the point of symmetry be overt, otherwise no problem for linearization at PF would raise by definition. Thus, if at least one phrase is base generated as a null element in a point of symmetry, prototypically *pro*, Dynamic Antisymmetry predicts that there should be no movement; correspondingly, this also implies that *pro* can never move. Before proceeding in illustrating an empirical case, it should be noticed that such an implication has non trivial consequences concerning parametrization of cross linguistic differences in movement. Since the inventory of null elements is parametrically determined (as a lexical property), Dynamic Antisymmetry suggests a way to parametrize movement itself: those languages which can

license *pro* should involve fewer movement operations because the points of symmetry involving *pro* do not need to be neutralized. In Moro (2000) I have discussed one case of this type, providing independent evidence in favor of such a hypothesis:⁵⁰ if movement were triggered by uninterpretable features, by contrast, one would expect movement not to be sensitive to the overt/null distinction, since there could well be either covert or overt uninterpretable features.

We can now concentrate on a non trivial empirical case involving *pro*. Let us start from the well-known fact that in Italian *pro* cannot be focused whereas its overt counterpart can. So for example, the following sharply contrasting judgements is found:

- (21) a. * pro_{+Foc} telefona
 (pro telephones)
 b. lui_{+Foc} telefona
 (he telephones)

Why cannot *pro* be focussed?⁵¹ We must first exclude some potential explanations. The impossibility to focus *pro* cannot be referred to intrinsic lack of referential capacities; after all *pro* can be the source of reference of anaphors as in *pro amano se stessi* (pro love themselves); this cannot even depend on the left periphery assuming Rizzi's (1997) split Comp theory, i.e. there would be no immediate reason to block raising of *pro* to the same position as *lui* with focus interpretation (spec-Foc^o);⁵² it cannot even be due to the impossibility to stress *pro*, since clearly *lui* is obligatorily focused in *telefona lui*, notably with or without stress on *lui* as opposed to preverbal *lui* which is focussed only if stressed (and raised to the spec-position of Foc^o in the split-Comp field, as proposed by Rizzi 1997). This is a crucial point. Recall that in *telefona lui* the subject *lui* is in the same position as the subject of an inverse copular constructions, i.e. it is in situ immediately dominated by a small clause (cf. 81). In other words, the following simplified representation holds for postverbal subject in Italian:

- (22) a. [_{SC} [telefona] [_{SC} lui t]]
 b. [_{SC} [la causa della rivolta] è [_{SC} lui t]]

If, on the other hand, one adopts Dynamic Antisymmetry, a natural answer would immediately be available. In fact, *pro* cannot be focused because the point of symmetry it constitutes does not need to be neutralized (since this element is already invisible to the PF component) and hence there would be

no problem of linearization at spell-out. In other words, only (23a) is a legitimate representation of the sentence *telefona*, not (23b):

- (23) a. [_{SC} pro [telefona]]
 b. [_{SC} [telefona] [_{SC} pro t]]

The conclusion is that *pro* can never be left in situ in a bare small clause constituent and it cannot even raise to the spec of Foc° in the split-Comp field (cf. Rizzi 1997): this explains why *pro* cannot be focused. Notice that the impossibility for *pro* to occur postverbally, in the object position, can be extended to unaccusative constructions. In the classic analysis going back to Burzio (1986), *pro* could be licensed only in spec-IP, thus the object position cannot be occupied by *pro*. Moreover, if one adopts the idea that the complement of a verb in unaccusative constructions is not a noun phrase but a small clause, suggested in Moro (1997) (see also Hale & Keyser 2002: 189, for a recent discussion) the impossibility of postverbal *pro* in unaccusative constructions would follow for the same reasons as those blocking *pro* as a subject of inverse copular sentences (see endnote 45 here).⁵³ For the correlation between movement and interpretation we will come back to it in section IV.

As for the sixth consequence (6vi), which is presented here in a twofold complex format (the “strong” and “weak” version) it cannot be discussed in full here.⁵⁴ Nevertheless, I would like to point out that there is an interesting convergence with independent work carried out in the original antisymmetrical framework by Kayne:

I have argued that in a number of cases where covert movement had been postulated it is possible and advantageous to dispense with covert movement (including feature raising ...) and replace it with a combination of overt movements [footnote omitted]. The strongest interpretation of this conclusion is that the cases explicitly considered ... are typical, and that it is not accidental that those cases lend themselves to analysis in terms of overt movement. It is rather that UG leaves no choice: Scope must be expressed hierarchically [footnote omitted], there are no covert phrasal movements permitted by UG, and neither can the effect of covert phrasal movement be achieved by feature raising. Scope reflects the interaction of merger and overt movement.

(Kayne 1994: 183)

Whether such a convergence with the “strong thesis” in (6iv) is accidental, or my specific proposal simply wrong, cannot be discussed here: I will simply leave the topic aside referring to the tentative discussion in Moro (2000): chapter 4.

As for the last consequence of Dynamic Antisymmetry approach (6vii), i.e. that movement can vary across languages according to the X-bar status of lexical elements, I will only briefly comment on it, suggesting the lines of thought that this theory would lead us to take by referring to a simple example. As we have just seen, a crucial aspect of a Point of Symmetry is that two overt elements of the same X-bar status c-command each other. Simplifying somewhat, by considering the three basic types, we could say that a Point of symmetry is made of pairs of heads or pairs of non-heads (which are offending when they are both overt). So for example an object of a verb V° does not move when it is realized as a full maximal projection but it needs to move when it is realized as a head/clitic as in *Dante fotografa Beatrice* (Dante photographs Beatrice) vs. *Dante la fotografa* (Dante her-photographs; see Moro (2000) for an explanation of this difference). Let us concentrate on this last example. If we abstract away from the X-status of the object of a verb, we should conclude that Italian is both a VO and OV language. Of course, no one would seriously adopt a parameter in this case. If the theory proposed here is correct, the rearrangement of phrases is due to the necessity to present PF with linearizable structure and the possibility to have different orders is a function of this principle. This fact suggests that difference in movement across (and within) languages can be managed with in terms of Dynamic Antisymmetry provided that languages differ with respect to how they realize lexical categories in terms of X-bar status. For example, if in a language specifiers could be heads, then if a phrase with an overt specifier were governed by a lexical head this would constitute a Point of Symmetry, since the two heads would c-command each other.⁵⁵ If so, movement would intervene to rescue the structure and yield an LCA-compatible tree as opposed to the language where specifiers cannot be heads. If we combine this with the observation made concerning (6iv) and (6v), namely that movement does not take place if it is constituted by base generated empty category such as *pro*, it would not be unreasonable to speculate that Dynamic Antisymmetry might explain some differences across and within languages without assuming specific order parameters.⁵⁶

Clearly, this important issue cannot be developed here, nevertheless, I thought it was important to observe that potentially Dynamic Antisymmetry could treat order-parameters by tracing them back to the X-bar structure. Clearly, the morphological theory of movement cannot immediately manage with these fact, nor the original version of Antisymmetry proposed by Kayne (1994) since the LCA applies at all levels.

Summarizing so far, we have explored some aspects of one of the exclusive expectations of a Dynamic Antisymmetry theory of movement, namely

that one should observe across and within languages “mirror” structures where two syntactic objects X and Y surface with both possible linear orders as a result of movement required by the neutralization of a point of symmetry. Interestingly, we have seen that a common characteristic of a broad class of examples involving the IP, AP and the DP domains is that X and Y are connected by a predicative linking providing indirect evidence in favor of a long standing proposal going back to Williams (1980) according to which the semantic notion of predication corresponds to a syntactic notion of mutual command between maximal projections.⁵⁷

The advantages of this proposal over the standard one based on the idea that movement is triggered by the necessity to delete uninterpretable features should be carefully evaluated. Here we have just illustrated some welcome consequences concerning clause structure. Notice, however, that the very existence of “mirror” structures such as the one in (8) constitutes per se a challenging problem for the standard theory, in which movement is considered not to be a costless operation: “Merge and Agree (or their combination) preempts Move, which is a “last resort”, chosen when nothing else is possible” (Chomsky 2000: 102). The “mirror” structures in (8), then, constitute a challenging problem for such a view. If an element X must move to delete its uninterpretable features in a given syntactic context, the fact that it can stay in situ in the same syntactic context while some other element Y moves is inconsistent with the standard theory.

6. A Neo-functional Theory of Movement?

In this last section I would like to make some speculative remarks concerning Chomsky’s (2001) theory of movement. From a technical point of view, as far as the mechanism of movement is concerned, there are no substantial differences in that paper with respect to the essentials of the theory of movement of the earlier versions presented here (corresponding to Chomsky 2000). The theory which stabilized in Chomsky (2000) is in fact essentially reproduced in Chomsky (2001) and is grounded on two major hypotheses: first, movement is in fact internal Merge; second, movement is triggered by the necessity to delete uninterpretable features (see Chomsky 2001: 7–11). Looking at things from the point of view of the typology of Merge, one can say that argument structure is associated with external Merge (base structure); everything else, scopal and discourse-related (informational) properties in particular, instead, with internal Merge (derived structure) (see Chomsky 2000: 10). This view leads to some new interesting thoughts concerning the

existence of movement in natural language: “It is hard to think of a simpler approach than allowing internal Merge (a “grammatical transformation”), an operation that is freely available. Accordingly, displacement is not an “imperfection of language; its absence would be an imperfection ... This ‘copy theory of movement’ is sometimes regarded as a controversial innovation. It is not: it is the null hypothesis [footnote omitted]” (Chomsky 2001: 8). These kinds of considerations, including the view of movement as internal Merge, however, do not affect the proposal defended here and I will not further discuss them. In fact, the very idea that movement is an instance of Merge is incorporated into Dynamic Antisymmetry itself (which only aims at offering an alternative explanation as to what *triggers* movement). Indeed, even if one adopts the idea that Movement is just internal Merge, still a theory of trigger for movement would obviously be required. In this section, I would rather like to focus on a different aspect of Chomsky’s (2001) discussion on movement.

Despite such similarities with previous versions, in the same paper Chomsky deepens his analysis of the nature of movement by emphasizing its “functional” role:⁵⁸ plainly, movement takes place only if “not otherwise expressible” interpretations are to be expressed. For the sake of clarity, let me reproduce Chomsky’s words:

- (24) Movement provides new interpretations which would not otherwise be expressible: it affects non-theta theoretical aspects of meaning only (scopal and discourse-related properties, e.g. new/old information, specificity, etc.). (Chomsky 2001: 10–14).

This view is clearly highlighted in the paper when the mechanism of trigger of movement is illustrated from a formal perspective; it can be synthesized as follows. Movement of α to a certain position, say the specifier of H° , must be triggered by endowing H° in the lexicon with an uninterpretable feature (prototypically, an “EPP-feature”) which makes such a non theta-position available. This would force movement of a lexical item containing a feature of the same type to move to the specifier of H° to allow deletion of the uninterpretable features along the lines illustrated in section I of the present paper. Generalizing, Chomsky labels such uninterpretable features triggering movement “OCC” meaning: “I must be the occurrence of some β ” (Chomsky 2001: 10). The crucial point linking movement to interpretation can now be cited directly: “Optimally, OCC should be available only when necessary: that is when it contributes to an outcome at SEM that is not otherwise expressible...” (Chomsky 2001: 10). It seems to me that such an

explicit link between interpretation in the broad sense and movement qualifies Chomsky's (2001) theory of movement as a "functional" theory of movement. I will not discuss here the idea that movement is triggered by the occurrence of an OCC feature nor the idea that an OCC feature is to be inserted in the lexicon to allow "not otherwise available" interpretations. Rather, I would like to show that the very idea that movement and interpretation are linked in a non trivial way can be approached within a Dynamic Antisymmetry framework in a quite different way. In the remaining of this section I will sketch out the lines of reasoning implied by the view defended here.

Consider again the alternation between canonical and inverse copular sentences as in (8a–b), repeated here as (25a–b):

- (25) a. [a picture of the wall] was [t [the cause of the riot]]
 b. [the cause of the riot] was [[a picture of the wall] t]

We have so far pursued the idea that movement is a way for grammar to rescue those structures that are too symmetrical to be linearized. A given structure involves movement only if the structure could not be otherwise linearized at spell-out.⁵⁹ As noticed in section III, although these two sentences are construed from the very same lexical array, they are not at all equivalent from an interpretative point of view: in fact, the subject in (25b) is obligatorily focused unlike the one in (25a). Clearly, distinct movements are associated with distinct informational interpretations here, much in the sense that passive and active sentences are. From this point of view, then, the idea that movement is linked to non-theta related (i.e. informational) meaning (as suggested by Chomsky 2001) is preserved. The real difference is that there is no need to assume that meaning is obtained by endowing an item with an "uninterpretable OCC" feature in the lexicon; meaning here is rather associated to structural configurations in a non ambiguous way.

It is worth emphasizing here that if this analysis proves tenable, (informational) meaning turns out to be associated to structural configurations much in the sense that theta-roles are associated to structural configurations in the seminal work by Hale & Keyser's (1993) theory (see also Hale & Keyser 2002 for a more comprehensive theory). In this theory theta roles are not primitives; rather, they are the configurations where arguments end up being in a given structure. Agent, Patient, Goal, etc. are just labels for configurations (reminding us the way they are interpreted at the interface with the semantic component).⁶⁰ Interestingly, moreover, notice that the "type" of meaning affected here by movement is typically discourse-related (informa-

tional) and non theta-related, as suggested in Chomsky (2001). Whether or not this can be extended to predication in general (as a prototypical non theta-related notion)⁶¹ surely remains an interesting topic for future research.

All in all, if this discussion is correct, Dynamic Antisymmetry appears not only able to capture some general aspects that the standard theory based on deletion of uninterpretable features does not capture (synthesized in (6i–iv)), it also seems to be able to share with the competing theory the capacity to deepen our knowledge of the link between some aspects of meaning in the broad sense and syntax.

In this paper I have defended an alternative theory of movement (Dynamic Antisymmetry) that considers this phenomenon as the result of the necessity to flatten hierarchical structures into a linear sequence at spell-out (linear compression). When a too symmetrical structure is generated by Merge, movement intervenes to rescue the structure by deleting the phonological features of an offending item and copying it into a suitable c-commanding position. Deletion of phonological features (of “copies”), thus, turns out to be an obligatory part of movement which would go unexplained in the standard minimalist theory of movement based on (deletion of) uninterpretable features. This alternative approach has been defended by highlighting some general consequences that the theory has, such as the existence within and across languages of “mirror structures”. Many questions remain unanswered, such as, above all, whether a Dynamic Antisymmetry approach can be extended to all types of movement (passive, raising, etc). Perhaps, this alternative approach raises even more questions than the competing one based on morphology. I will try to approach these issues in future research, confident that the new questions raised by a Dynamic Antisymmetry approach are worth exploring. Moreover, it has been observed that since this theory relies on a condition which is linked to the physical organization of the biological world, i.e. the fact that words must be put in a time sequence, it seems to partially fit the challenging suggestions made by Chomsky (2001), namely that the ultimate aim of a theory of language is to “seek a level of explanation deeper than explanatory adequacy, asking not only *what* the properties of language are, but *why* they are that way” (Chomsky 2001: 2).

Notes

1. This paper was presented at *Triggers* conference in Tilburg in 2002. The core idea of theory presented here was first proposed at GLOW 1996 and published at different stages of development as Moro (1997b, 2000 and 2003). The discussion proposed here focuses on the new Minimalist perspective suggested by Chomsky (2001). I am very grateful to the audience of *Triggers* conference for many stimulating and deep comments on this proposal. Thanks also to Giorgio Graffi, Giuseppe Longobardi, Luigi Rizzi, Orin Percus, Massimo Piattelli Palmarini, Alessandra Tomaselli and two anonymous reviewers for their criticism and the discussion of this proposal.
2. Chomsky uses the expression “genetically-determined”. I do not think that this is quite appropriate, as Medawar (1967) noticed, but a discussion of this specific aspect would take us too far (see Moro 2002 and references cited there).
3. The empirical issue of displacement was recognized in the XX century at least since the Fourties by Post-Bloomfieldian syntacticians and sometimes referred to as “discontinuous constituents” (see Pike 1943: 77, cited in Graffi 2001. For a critical and comprehensive survey of this matter see Graffi 2001: 300–305).
4. The choice between different ways to construe the proper local relation is based on economy considerations: “Merge and Agree (or their combination) preempt Move, which is a “last resort”, chosen when nothing else is possible” (Chomsky 2000: 102). Indeed, the process of “pairing” can be performed in different ways (crucially including the operation “Agree” and “Expletive insertion”, as proposed since Chomsky 1999): nevertheless, for what interests us here, movement is just one of the ways to construe the proper local relation to allow pairing and deletion.
5. When it comes to Case features, one should also say that pairing *uninterpretable* features with *uninterpretable* features deletes the uninterpretable features. I am referring here to number features.
6. To overcome the difficulty related to the notion of “interpretability”, it has been recently proposed to remedy this situation by shifting the terminology from “interpretable/uninterpretable” to “valued/unvalued” features (see for example Chomsky 2001). If we adopt this view, the existence of unvalued features, prototypically number features, is plausible and it just corresponds to the very traditional intuition that verbs agree in number with nouns but not viceversa. However, if different types of features are considered, such as Case features or wh-features for example, the idea of “valuation” appears to be problematic. For example, what is the equivalent of Case features which are considered to be always uninterpretable? Should they be considered to be always unvaluable? See Piattelli Palmarini & Uriagereka (2003) for a different point of view.
7. Strictly speaking, one should assume that not only L is identical but also the order external Merge applies to compose the structure is.

8. If we look at the development of the theory of movement in transformational grammar, the fact that movement evolved into a morphological theory can hardly be surprising. In fact, the idea that all instances of movement can be traced back to morphological requirements has been progressively pursued in different stages. A major step toward such a unified approach came from the unification of *wh*-movement and Case assignment. Originally, the two processes were captured by two conceptually distinct devices: the Case filter and the *wh*-criterion, respectively. From a theoretical point of view, the essential difference between a filter and a criterion is that a criterion is a filter imposing a one-to-one mapping between a (feature contained in a) head and a (feature contained in a) phrase. More specifically, a criterion requires a (biunique) spec-head relation between two elements containing the same type of feature. For example, criteria have been proposed for θ -relations (cf. Chomsky 1981: 36), *wh*-movement (cf. Rizzi 1996 and references cited there) and negation (see Haegeman & Zanuttini 1997). It is the reduction of Case assignment to spec-head agreement between a phrase and (a head containing) an Agr° that has led to a unified analysis of these phenomena, for it has become possible to consider Case assignment as a particular instance of the more general class of criteria on a par with *wh*-movement. Prior to the trend of research stemming from the so called split-Infl Hypothesis (see Pollock 1989 and independently Moro 1988), it was not in fact conceivable to reduce Case assignment to a criterion, since, for example, Accusative case was not assigned in a spec-head relation; in fact, Accusative was assigned under government by the lexical verb. After the split-Infl hypothesis stemming from Pollock's (1989) own work was extended to include an independent head for object agreement (i.e. Agr° -O: cf. Kayne 1989b, Belletti 1990, Chomsky 1995 and many related works), the reduction of Case filter to a criterion was made possible. Both Nominative and Accusative case turned out to be the same type of structure, namely a spec-head relation with an abstract Agr° . This first proposal has successively been fine-tuned, the elimination of Agr° by Chomsky 1995 being a major turn. Quantifier Raising has been treated in a similar way involving deletion of an uninterpretable feature in the Comp System (see Hornstein 1995 for a detailed and critical analysis of Quantifier Raising within the Minimalist Program).
9. This is in fact the approach of Chomsky (2001). For analogous proposals and the derivation of some restrictions on movement from Relativized Minimality conditions see Starke (2001).
10. Moreover, if economy is intended as a way to "spare" energy while speaking the distinction between competence and performance would be severely undermined.
11. I will leave the notion of "formal feature" unspecified here, i.e. it can include phonological, semantic, morphological as well as syntactic features. Strictly speaking, notice that syntactic features can include X-bar theoretical informa-

tion, more specifically “being a head” or “being a (maximal) projection”: for a critical discussion of this hypothesis see Moro (2000: 122).

12. As for the typology of Merge, i.e. internal and external Merge, see section vi.
13. Strictly speaking, flattening structures (both in the original antisymmetry theory proposed by Kayne, and obviously in a Dynamic Antisymmetry approach) would not be necessary when an individual produces sentences at the mental level only, i.e. without uttering them: I have no empirical arguments to approach such an issue. In general, as far as I know there is no comprehensive theory as to why we use the same restrictions as in spoken language when speaking at the mental level, including phonological instructions in the first place.
14. See the original work by Kayne (1994) and the critical illustration in Cinque (1996).
15. By “too symmetrical” I mean a structure where either two heads or two maximal projections c-command each other, as illustrated in (4) later in the text. For an illustration of how the LCA works see the original proposal in Kayne (1994), or Cinque (1996) and Moro (2000: ch. 2). Notice that the LCA is not compatible with the so called Bare Phrase Structure hypothesis (cf. Chomsky 1995). In the latter hypothesis, when say, two heads like *meet* and *John* are merged to form a VP, there is no intermediate projection to protect mutual c-command of the two heads. Thus, if we adopt the LCA, we must assume that there are intermediate empty nodes so that N” is merged with V°, not N° (cf. Kayne: 9 and Moro 2000: 85 ff. for further discussion)
16. Whether spell-out is a single operation or a multiple one does not affect the presentation of the core proposal here. For a critical discussion of the multiple spell-out hypothesis and references see Uriagereka (1999).
17. The structure in (4iii) is *prima facie* problematic: can a constituent have a label which is not the projection of either element? One possible solution (fully discussed in Moro 2000) comes from the formal interpretation of Merge given by Chomsky (1995). Let us focus on the crucial point. The label of a constituent K can be complex, provided that no extra information is added. More specifically, the label of adjunct structure is the ordered pair of the projecting element (i.e. $\langle \alpha, \alpha \rangle$). It seems reasonable then to assume that Merge allows a further combination, with the resulting label shown in (i):

$$(i) K = \{ \langle \alpha, \beta \rangle, \{ \alpha, \beta \} \}$$

This is also a formally acceptable option. Crucially, it does not conflict with the essential property of Merge, namely, not introducing extra information, specifically extra features of a constituent different from α and β . The format of this Merge output might seem to generate ambiguity, since from a purely formal point of view the mirror option, where the ordered pair constituting the label is inverse (i.e. $\langle \beta, \alpha \rangle$) is also possible. This problem can be solved by assuming

that the given label is intended to be the short form of the more articulated one given in (ii):

(ii) $\{\{\alpha, \{\alpha, \beta\}\}, \{\beta, \{\alpha, \beta\}\}\}$

In such case, the output is totally neutral with respect to α and β and, crucially, the essential requirement that Merge not introduce new information is preserved. (I am indebt to James Higginbotham for an extensive discussion on this topic).

18. X c-commands Y iff a. X and Y are categories (not segments of categories); b. no segment of X dominates Y; c. every category that dominates X dominates Y.
19. The representation in (4iv) can be misleading: since hierarchy is what counts to linearize terminal nodes, (4iv) repeated here as (i)a is totally equivalent to any of the following representations:
 - (i) a. [ZP [X° YP]]
 - b. [[X° YP] ZP]
 - c. [ZP [YP X°]]
 - d. [[YP X°] ZP]]

The linearization of any of these notational variant of the same hierarchical relations would in fact be the same under the LCA: the terminal nodes contained in ZP precede the terminal in X° and the terminal in X° precedes those in YP. The choice among (i)a–c is irrelevant and conditioned by the monodimensional, i.e. linear, nature of representation.

20. As for the label of (4iii) I will refer to Moro (2000) for a full argument. In a nutshell, the idea is that phrase markers are genuine and primitive syntactic entities (as suggested in Kayne 1994 contra Chomsky 1995). Thus a structure like (4iii) is nothing but the features associated to a maximal projection with no further categorial information (see also note 17).
21. The observation that traces are not visible to the PF component, independently of whether they are considered as copies or not, was explicitly made by Kayne (1994: 133, footnote 3) and Chomsky (1995: 337) but was not further pursued in those original papers.
22. Another difference between the two theories can be straightforwardly highlighted by adopting Van Riemsdijk's (1995) terminology. The standard theory is a "drag-chain" theory of movement whereas the alternative theory presented here is a "push-chain" theory of movement, in that the trigger for movement in the former is given by the "landing site" while the trigger in the latter is given by the "launching site" of movement.
23. An interesting domain to test the empirical hypothesis presented here (and more generally to test the theory of antisymmetry) is Sign Language (SL). More specifically, since SL appears to involve movement, it would be interesting to explore whether it could also be correlated to the linearization processes that occur in this domain. For a critical discussion of linearization processes in SL

(and the structure of SL in general) see Neidle, Kegl, Maclaughlin, Bahan & Lee (1999) and references cited there. I am indebted to Carlo Cecchetto and Sandro Zucchi for advice on this topic.

24. The first obvious one is whether the conjecture could be considered as an “if and only if” proposition, including all types of movement; of course this is the more interesting step to take and one that I am temptatively pursuing. Whether or not this is right, only further research will tell. I will consider it a success if Dynamic Antisymmetry will allow us to ask the right questions, more than ensure that we find the right answers.
25. The analyses illustrated here have been developed originally in Moro (2000) and Moro (2003) and are reproduced here to support the main thesis.
26. Strictly speaking, following Chomsky (2001) one should not use the terms “copy” and “original”: each occurrence of an element is the occurrence of the same element in distinct positions. I just refer to these terms as purely descriptive labels. As for intermediate traces, I will simply refer to Moro (2000).
27. For the analysis of points of symmetry constituted by heads see Moro (2000: 84–92).
28. In the causative sentence, the subject appears as a clitic in one case; I will refer to Guasti & Moro (2001) for the discussion of such specific characteristics.
29. English and Italian differ here, since in Italian it is not necessary that the two DPs agree in number as opposed to English (cf. **this type of books*). This form of agreement could be perhaps related to ECP reasons or to the mass/count distinction: I will not explore this issue here.
30. The hypothesis that the VP embedded in causatives has different positions related to the presence of *à* can be supported by a French dialect, as was pointed out to me by Luigi Rizzi. In that dialect, where the subject of the causative can be cliticized as in (i):
 - (i) Marie le fait [_{SC} t [laver la voiture]]
the object inside the causative can be cliticized onto the embedded verb only if *à* is absent:
 - (ii) a. Marie le fait t [la laver t]
b. *Marie fait [la laver t] à Jean
 One possible way of thinking to explain this fact is that there is movement of the embedded VP to a higher position and that *la* constitutes a further point of symmetry with the matrix verb: I will leave the elaboration of this explanation to future research.
31. If the copula is considered to be the spell-out of I°, the generalization suggests itself here that bare small clauses would be the complement of functional heads only, i.e. I°, D° and C°. Moreover, if small clauses really are the only implementation of predication in syntax, this would amount to saying that a predicative structure can only be the complement of a functional head. For a further refinement of the notion of “bare” and “rich” small clause see Pereltsvaig (2001a).

32. The empirical arguments I am referring to can be found in Moro (2000: 43–48): they involve the distribution of adverbs, cliticization and the occurrence of predicative markers such as *as* and its equivalents in Italian (for example, *come*). This refines a previous analysis that did not distinguish between bare and rich small clauses and wrongly assumed that bare small clauses contained a head: see Longobardi (1988), Moro (1988) and Cardinaletti & Guasti (1995) for a source of different analyses.
33. Notice that the idea that the relation between a predicate and a subject is not mediated by a head, or equivalently that the two c-command each other, essentially reproduces the original intuition by Williams (1980) according to which two such roles were just defined as two mutually c-commanding projections. Indeed Williams' original proposal appears to be still valid under different perspectives. For a detailed discussion concerning labelling and Merge as defined in the Minimalist framework see Moro (2000).
34. Remember that for the theory of Antisymmetry there are no intermediate projections, i.e. there is no distinction between specifiers and adjuncts. In particular here there is no I': when VP is merged with I^o, IP is yielded. This is not to say that the IP system is defective in that it cannot have specifiers/adjuncts; in fact it can, so for example adverbs can be specifiers of the IP system but not subjects. This view also has a non-trivial consequence that will not be discussed here, namely that IPs can be predicates.
35. The difference among languages will still be reduced to the possibility of properly governing the trace of the subject, i.e. to the possibility for pro to occur in subject position, as in the traditional theory.
36. I will not consider here some further important properties distinguishing left periphery Focus from postverbal one, such as contrastive properties etc.
37. Just to give one simple example: there is no focus on *the cause of the riot in a picture of the wall was the cause of the riot* but there must be focus on *a picture of the wall in the cause of the riot was a picture of the wall*.
38. The advantage with respect with the ad hoc solution we are refusing here is that we do not assume an intermediate Foc^o head dedicated to and thus just maintain the more parsimonious structure.
39. Of course, inverse copular sentences show up only when the predicate can occur in the same position as the subject, namely when the predicate is a noun phrase, for morphological reasons. This type of structure is to be kept carefully distinct from cases of "locative inversion". Hoekstra-Mulder (1990) for example suggested that unaccusatives can be analyzed as "locative inversion" constructions where spec-IP is occupied by a PP (i)a on a par with copular sentences like those in (i)b (see Hoekstra-Mulder 1990: 28 ff.):
 - (i) a. [_{IP} [_{PP} in the room]_i entered [_{SC} a man t_i]]
 - b. [_{IP} [_{PP} in the room]_i was [_{SC} a man t_i]]

This analysis, which explicitly subsumes the theory of *there*-sentences as inverse sentences proposed in Moro (1990) (cf. Hoekstra-Mulder 1990: 33 ff.), cannot be maintained for empirical reasons. Consider the following examples:

- (ii) a. [_{DP} the cause of the riot]_i is [_{SC} a man t_i]
 b. [_{DP} the cause of the riot]_i is/*are [_{SC} John and Mary t_i]
 c. [_{PP} in the room]_i is [_{SC} a man t_i]
 d. [_{PP} in the room]_i are/*is [_{SC} John and Mary t_i]

Assuming that agreement is invariantly established in spec-IP, it would be hard to assume that (i)a and (i)b instantiate the same type of structure. Rather, (i)b is an instance of the topic constructions that have been analyzed by Cinque (1990) and Rizzi (1997) involving some portion of the scattered CP layer (and movement of the verb to a higher functional head). This would explain why the equivalent of (i)b in Italian involves a locative clitic *ci* which is not allowed in inverse copular sentences:

- (iii) a. [_{IP} [_{PP} nella stanza]_i *(c') era [_{SC} un uomo t_i]]
 b. [_{IP} [_{DP} la causa della rivolta]_i *(c') era [_{SC} un uomo t_i]]

The status of *ci* with the copula is discussed in detail in Moro (1997a: ch. 2; and summarized in the Appendix of Moro 2000). As for the lack of *ci* in Italian unaccusatives (but not in many Northern Italian dialects as observed by Burzio 1986 and Poletto 1993 among others) see Moro (1997a).

40. Notice that *type* can occur as a predicate in a copular constructions such as in *a mammal is a type of animal*. Interestingly, however, *of* must show up in cases like *many books are *(of) this type*. For the role of prepositions as copulas (in noun phrases) see the seminal work by Den Dikken (1997).
41. Notice that if *di* is present, NP-stranding and agreement can cooccur:
- (i) *quali ha scritto di racconti?*
 (which-mas.plur. has written of novels-masc.plur.)
42. For a critical approach to government within noun phrases and its impact on the general system, see Giorgi-Longobardi (1986, ch. 2).
43. That the distinction between overt vs. null head is relevant for government has been accepted since at least Rizzi's (1990) theory of extraction of preverbal subject in embedded sentences (cf. *who do you think [_C e] t left* vs. * *who do you think [_C that] t left*) although the possibility to govern a trace was reserved to null C° (which is considered to be the agreeing complementizer in English as opposed to *that*), contrary to the case studied here.
44. I am not considering here the interesting case of exclamatives where sentences like *quanto alta è quella torre!* (how tall is that tower) appears to be partially acceptable. For a discussion on exclamatives see Portner & Zanuttini (to appear).
45. A potential counterexample to the theory presented here could come from a pair like the following:

- (i) a. *pro sono io*_{+Nom}
 (pro am I)
 “it’s me”
 b. **io sono*
 (I am)

Apparently the prediction made by the standard theory considering movement as a last resort operation is borne out: since *io* can stay in situ in (i)a (while the phonologically null element *pro* occurs in preverbal position) *io* cannot move in (i)b manifesting the alleged “last resort” nature of movement. No mirror structure is in fact generated here. This conclusion however is not correct. The reason why (i)b is ungrammatical is independent of movement. Let us consider the following sentences with the associated structure:

- (ii) a. **Maria considera [Gianni pro]*
 (Maria considers Gianni pro)
 b. **Maria considera [pro il colpevole]*
 (Maria considers pro the culprit)
 c. **io sono [t pro]*
 (I am)

In Italian *pro* cannot be licensed within a (bare or a rich) small clause (in fact the only proper environment is a relation with a rich verb inflection; see also Rizzi (1986) for non verbal *pro* licensing), thus (i)b is ruled out by the same reasons as (ii)a–b and has nothing to do with movement. The copula must always occur with two maximal projections linked by predication: since *pro* cannot be licensed in a small clause, (i)b just lacks one maximal projection for the copula to be interpreted. All in all, the contrast in (i) does not stand as a challenging counterexample to Dynamic Antisymmetry: indeed, as far as the data in (8) are concerned it seems that a theory assuming that movement is a last resort operation would not be empirically adequate. As for the contrast between *sono io* and *io lo sono*, i.e. for the contrast between pro-predicative *lo* and pro-predicative *pro* see Moro (1997) and references cited there.

46. I am indebted to Giorgio Graffi for pointing me out these cases.
 47. In Italian, clitics are assumed to be endowed with Case features, always uninterpretable.
 48. As for the possibility that only the lower XP be pronounced, such as in *wh*-in situ constructions, Dynamic Antisymmetry would force us to analyze them as involving “remnant movement”. Leaving aside the reasons which trigger this movement, take for example a simple case like the following echo question:

- (i) *hai visto cosa?*
 (have seen what)

The only analysis compatible with Dynamic Antisymmetry would be the following, crucially excluding adjacency between V° and *cosa*:

- (ii) a. *cosa* [hai visto t]
 b. [[hai visto t] H° [*cosa* t]]

In such a case, no LCA problem would raise since there would be no symmetrical c-command between the verbal head *visto* and the head *cosa* (for the analysis of *cosa* as a head see Moro (2000) and references cited there). Interestingly, notice that *cosa* would be focused, witness the normal intonational emphasis it receives when it is pronounced in situ: *cosa* is in fact in situ within a bare small clause as the subject of an inverse copular sentence (see discussion on in situ Focus in this section and in section IV).

49. Interestingly, it is never the case that a moved element is doubled by a stressed pronoun. If the analysis associating clitics to heads and stressed pronoun to full phrases XP is correct (see for example the seminal work by Kayne 1989; see also Cardinaletti & Starke 1994, Sportiche 1992 and Moro 2000 for convergent approaches to this idea) then Dynamic Antisymmetry would also explain why stressed pronouns do not occur in substandard Italian relative clause and interrogative “doubling” constructions: a stressed pronoun, being an XP, would reconstitute the point of symmetry whereas a clitic pronoun, being an X°, would not.
50. The case discussed in Moro (2000) involved *pro* in verbal agreement: see section 3.3.2. of that monograph.
51. Taking contrastive focus to be a test, one can easily construe the relevant minimal pair:
- (i) a. LUI telefona, non lei
 (he telephones not she)
 b. **pro* telefona, non lei/loro
 (telephones not she/them)

Notice however that one can have a contrast in sentences of the type in (ii):

- (ii) *pro* TELEFONA, non telefonano
 (telephones not telephone)
 (s/he telephones not they telephone)

The sentence in (ii) however is irrelevant because the contrast is in fact expressed by the verb (as mediated by inflection).

52. Notice that in our framework the impossibility for *pro* to move to spec-Foc° follows as a particular case of a more general restriction following from Dynamic Antisymmetry (as indicated in (6v)).
53. I am grateful to an anonymous reviewer for pointing me out this issue.
54. For a detailed discussion see Moro (2000: ch. 4). On a different approach to covert movement see also Bobaljik (to appear).
55. As for the possibility for specifiers to be heads, Cinque 1996 also noticed that the requirement “that a head cannot be a specifier is also derived, albeit via a further assumption (“that the highest element of a chain of heads must have a specifier” – Kayne 1994: 31) If a head, in order to be licensed, needs to project

(and discharge its θ -role(s)), it follows that the source of a head in specifier position must be a lower head position. But then the possibility arises of excluding its moving to a specifier position as a violation of Relativized Minimality (Rizzi 1990; or “Shortest Movement” – Chomsky 1995). A closer potential landing site (the head of the phrase it adjoins to) is skipped (this still does not prevent a head from becoming its own specifier)” Cinque (1996: 449, fn. 6). Thus, if Relativise minimality conditions can be refined to the extent that not all heads per se can function as intervenors for other heads, the possibility for a head to be a specifier cannot be excluded. Interestingly, if (spec-head) agreement between a head and another head as its specifier were allowed, intermediate activation of agreement in Romance languages could be interpreted without assuming that the intermediate copies/traces of clitics are full noun phrases.

56. Pushing this speculation to the limit, one could conclude that (at least when it comes to word order) the differences across languages should also be observable within languages, much in the sense that Italian is an OV and VO language, as suggested in the text.
57. See also Rothstein (1983) for some extension of the original proposal by Williams to secondary predication.
58. The term “functional” has not been used by Chomsky.
59. In a sense, this can be regarded as a “last resort” quality of movement, although it is so in a very different way w.r.t. the standard minimalist theory.
60. For an extension of the Hale & Keyser’s (1993) theory to unaccusatives see Moro (1997b: chapter 5) and references cited there. See also Hale & Keyser (2003, chapter 6).
61. That predication is to be kept conceptually distinct from theta assignment can be proved in several ways (see Moro 1991, 1997a and references cited there). A major argument in transformational grammar comes from the existence of active/passive alternation: for a pair like *John reads a book* and *a book is read by John* (where the subject is *John* and *a book*, respectively) one surely wants the same VP, i.e. the same underlying thematic structure, to generate the two sentences where the subject is the external and the internal argument, in the active and passive sentence respectively. Other cases can be observed in copular constructions where sentences like *this is Dante’s desire for Beatrice* shows that although the theta roles pertaining to the lexical head *desire* are exhausted within its maximal projection, still that maximal projection can be the predicate of another DP, i.e. *this*.

References

- Belletti, A.
1990 *Generalized Verb Movement*. Torino: Rosenberg & Sellier.
- Bennis, H.
1986 *Gaps and Dummies*. Dordrecht: Foris.
- Bennis, H., M. den Dikken and N. Corver
1998 Predication in Nominal Phrases. In *The Journal of Comparative Germanic Linguistics* 1: 85–117.
- Bobaljik, J.
to appear A-Chains at the PF-interface: copies and *covert* movement, *Natural Language and Linguistic Theory*.
- Brody,
1995 *Lexico-logical form: a radically minimalist theory*. Linguistic Inquiry Monographs. Cambridge Mass.: MIT Press.
- Burzio, L.
1986 *Italian Syntax*. Reidel, Dordrecht.
- Cardinaletti, A. and M.-T. Guasti (eds.)
1995 *Small Clauses*. Chicago: Academic Press.
- Cardinaletti, A. and M. Starke
1999 The Typology of Structural Deficiency. On Three Grammatical Classes. In *Clitics in the languages of Europe. Empirical approaches to language typology*, Riemsdijk, H. van (ed.), 20–5. Berlin/New York: Mouton de Gruyter.
- Chomsky, N.
2000 Minimalist inquiries: The framework. In *Step by step*, R. Martin, D. Michaels, and J. Uriagereka (eds.) Cambridge, Mass.: MIT Press.
2001 Beyond expository adequacy. *MITOPL* 20, Cambridge Mass.
- Cinque, G.
1988 La frase relativa. In *Grande grammatica italiana di consultazione*, Vol. I, L. Renzi (ed.). Il Bologna: Mulino.
1990 *Types of A-bar Dependencies*, Linguistic Inquiry Monographs 17. Cambridge, Mass.: MIT Press,
1996 The antisymmetric program: theoretical and typological implications. *Journal of Linguistics* 32: 447–464.
- Den Dikken, M.
1997 Predicate Inversion in DP. In *Possessors, predicates and movement in the determiner phrase*, Alexiadou-Wilder (eds.), 177–214. Amsterdam: Benjamins.
- Giorgi, A. and G. Longobardi
1991 *The Syntax of Noun Phrases*. Cambridge: CUP.
- Graffi, G.
1997 Frasi complete e frasi ridotte. *Lingua e Stile* 32, 2: 273–291.
2001 *200 Years of Syntax. A critical survey*. Amsterdam: John Benjamins.

- Groat, E. and J. O'Neill
 1997 Spell-out at the Interface: achieving a unified syntactic computational system in the minimalist framework. In *Minimalist Ideas*, W. Abraham, S.D. Epstein, H. Thráinsson and J.-W. Zwart (eds.).
- Guasti, M. T. and A. Moro
 2001 Causative sentences and Dynamic Antisymmetry. In *Current Studies in Italian Syntax. Essays offered to Lorenzo Renzi*, Cinque and Salvi (eds.), North-Holland 2001, Amsterdam.
- Haegeman, L. and R. Zanuttini
 1991 Negative heads and the Negative Criterion. *The Linguistic Review* 8: 233–251.
- Hale, K. and J. Keyser
 1993 On the argument structure and the Lexical expression of syntactic relations". In *The View from Building 20: Essays in Honor of Sylvain Bromberger*, K. Hale and J. Keyser (eds.). Cambridge, Mass.: MIT Press.
 2002 *Prolegomena to a theory of argument structure*. Linguistic Inquiry Monograph 39. Cambridge, Mass.: MIT Press.
- Hoekstra, T. and R. Mulder
 1990 Unergatives as Copular verbs. *The Linguistic Review* 7: 1–79.
- Hornstein, N.
 1995 *Logical Form*, Oxford, UK: Blackwell.
- Kayne, R.
 1989a Null Subjects and Clitic Climbing. In *La grammaire modulaire*, Ronat, M. and Couquaux (eds.), Édition de Minuit.
 1989b Facets of Romance Past Participle Agreement. In Benincà (1989).
 1994 *The Antisymmetry of Syntax*. Linguistic Inquiry Monograph 25. Cambridge, Mass.: MIT Press.
- Longobardi, G.
 1988 *Symmetry Principles in Syntax*. Padova: Clesp.
- Mayr, E.
 1988 *The growth of biological thought. Diversity, evolution and inheritance*. Cambridge, Mass.: Belknap Harvard.
- Medawar, P. B.
 1967 *The Art of Soluble*. Londra: Methen.
- Moro, A.
 1988 Per una teoria unificata delle frasi copulari. *Rivista di Grammatica Generativa* 13: 81–110.
 1996 *Dynamic Antisymmetry*. XIX GLOW Conference. Università di Atene.
 1997a *The Raising of Predicates*. Cambridge University Press, UK.
 1997b Dynamic Antisymmetry: movement as a symmetry-breaking phenomenon. *Studia Linguistica* 51: 50–76.
 1996 Dynamic Antisymmetry: movement as a symmetry-breaking phenomenon. *Working Papers* 13, Department of Cognitive Sciences, Istituto Scientifico San Raffaele.

- 2000 *Dynamic Antisymmetry*. Linguistic Inquiry Monograph 38. Cambridge, Mass.: MIT Press.
- 2002 Linguistica Mendeliana ovvero quali domande su genetica e grammatica? *Lingua e Linguaggio* 1: 39–58.
- 2003 The role of linear compression in human languages: two questions on movement. *Lingua e Linguaggio* 3: 31–68.
- Neidle, C., J. Kegl, D. Maclaughlin, B. Bahan and R. G. Lee
1999 *The Syntax of American Sign Language. Functional categories and hierarchical structure*. Cambridge, Mass.: MIT Press.
- Nuñez, J.
2001 Sideward Movement. *Linguistic Inquiry* 32: 303–344.
- Pesetsky, D.
1997 Optimality Theory and Syntax: Movement and Pronunciation. In *Optimality Theory: an overview*, D. Archangeli and T. D. Langendoen (eds.), 134–170. Malden, Mass.: Blackwell.
- Piattelli Palmarini, M. and J. Uriagereka
in press The Immune syntax: the evolution of the language virus. In *Variation and Universals in Biolinguistics*, Jenkins, L. (ed.). Oxford: Elsevier.
- Pike, K.
1943 Taxemes and immediate constituents. *Language* 19: 65–82.
- Pollock, J.-Y.
1989 Verb Movement, UG and the Structure of IP. *Linguistic Inquiry* 20: 365–424.
- Poletto, C.
1993 *La sintassi del soggetto nei dialetti italiani settentrionali*. Quaderni Patavini di Linguistica, Monograph 12. Padova: Unipress.
- Portner, P. and R. Zanuttini
to appear The semantics of nominal exclamatives. In *Ellipsis in non-sentential speech*, R. Elugaro and R. J. Stainton (eds.). Dordrecht: Kluwer.
- Sportiche, D.
1992 Clitic Constructions. In *Phrase Structure and The Lexicon*, J. Rooryck and L. Zaring (eds.). Dordrecht: Foris.
- Rizzi, L.
1986 Null Objects in Italian and The Theory of pro. *Linguistic Inquiry* 17: 501–558.
1990 *Relativized Minimality*. Linguistic Inquiry Monograph 16. Cambridge, Mass.: MIT Press.
1996 Residual Verb Second and the WH-criterion. In *Parameters and Functional Heads*, Belletti, A. and L. Rizzi (eds.), 63–90. Oxford/New York: Oxford University Press.
1997 The fine structure of the left periphery. In *Elements of Grammar: Handbook of Generative Syntax*, L. Haegeman (ed.), 281–337. Dordrecht: Kluwer.

- Rothstein, S.
1983 *The Syntactic Form of Predication*. Ph.D. diss., MIT: Cambridge, Mass.
- Starke, M.
2001 *Move is Merge*. Ph.D. Thesis, University of Geneva, Switzerland.
- Steriade, D.
1980 Clitic doubling in Romanian Wh-constructions and the Analysis of Topicalization. Papers from the 16th regional meeting of the Chicago Linguistic Society, University of Chicago.
- Stowell, T.
1981 *Origins of Phrase Structure*. Ph.D. dissertation, MIT: Cambridge, Mass.
- Uriagereka, J.
1999 Multiple Spell Out. In *Working minimalism*, D. S. Epstein and N. Hornstein (eds.). Cambridge, Mass.: MIT Press.
- Williams, E.
1980 Predication. *Linguistic Inquiry* 11: 203–238.
- van Riemsdijk, H.
1995 *Push chains and drag chains*. Ms., University of Tilburg.
- Zamparelli, R.
1995 *Layers in the Determiner Phrase*. Ph.D. Dissertation, University of Rochester.