Syntax 2010 DOI: 10.1111/j.1467-9612.2010.00140.x

On Labeling: Principle C and Head Movement

Carlo Cecchetto and Caterina Donati

Abstract. In this paper, we critically reexamine the two algorithms that govern phrase structure building according to Chomsky (2008). We replace them with a unique algorithm, the Probing Algorithm, which states that the Probe of any kind of Merge always provides the label. In addition to capturing core cases of phrase structure building, this algorithm sheds light on Principle C effects and on the syntax of *wh*-constructions, which we analyze as cases of conflict between two Probes. In these two configurations a lexical item (which should become the label, being endowed with an Edge Feature that qualifies it by definition as a Probe) is merged with a syntactic object that, being the probe of the operation, should also become the label. In one case, this conflict produces two alternative outputs (a question or a free relative) that are both acceptable. In Principle C configurations, one of the resulting outputs (the one where the lexical item ''wins'') produces an object that is not interpretable. This way, Principle C effects are reduced to cases of mislabeling, with no need to postulate a specific condition to rule them out.

1. Introduction

One important assumption in the Minimalist Program, initially formulated by Chomsky (1995), is the Inclusiveness Condition, according to which narrow syntax merely operates on lexical items and cannot ''add'' interpretative material. This is usually interpreted as meaning that semantically active material such as indexes, bar levels, or labels cannot be inserted in the course of a derivation.

Still, there is an important theoretical notion that does not seem to be dispensable, namely, that Merge yields labeled syntactic objects: when Merge forms a syntactic object, the features associated with one and only one of the assembled items can trigger further computation.[1] If the inclusiveness condition is to be taken seriously, this cannot be captured through the insertion of a new object distinct from the items

Preliminary versions of this work were presented at the XXXII Incontro di Grammatica Generativa (University of Florence, March 2006), Interphase Conference (University of Cyprus, May 2006), NELS 38 (University of Ottawa, October 2007), as well as in seminars at the University of Siena and the University of Milan–Bicocca. We thank the audience of these meetings, Gennaro Chierchia, and Sandro Zucchi for useful comments and observations, and Carlo Geraci, Andrea Moro, and *Syntax* reviewers for detailed comments on previous versions of this manuscript.

[1] Collins (2002) sketches a theory of syntax in which labels can be completely dispensed with. However, his polemical objective is the notion of label as an extra object distinct from the two items that are merged, as was in Chomsky's (1995) version of bare phrase structure theory. In that early version of the theory, the output of merging of X and Y was not the minimally simple object {X,Y} but was either {X,{X,Y}} or {Y,{X,Y}}, depending on which category projects. We believe that once a label is defined as a subset of the features of one of the two merging objects, the quest for simplification argued for by Collins can be satisfied. Still, differences between Collins's approach and ours remain. They do not arise so much in the area of phrase structure theory, given that the notion of label is replaced in Collins's theory by the closely related notion of Locus, as Collins himself notices (p. 48), nor in the theory of subcategorization, for Collins assumes that lexical features like ± V, ± N do exist, although they do not project at the phrasal level. The area in which differences arise is the theory of locality, because a label-less theory *à la* Collins requires a reformulation of the Minimal Link Condition, with potentially different empirical predictions. We cannot make a complete comparison between the two approaches due to reasons of space. See also Seely 2006 for a different attempt to eliminate labels.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

that are merged, such as a label in standard X-bar theory. Rather, we define label as a subset of features, as in (1).

(1) Label: features of a syntactic object (SO) that can trigger further computation

Therefore, syntax should have a simple, automatic way to calculate the label of any syntactic object. Following Chomsky (2008), we call this the labeling algorithm. In this paper we discuss how this algorithm should be defined, keeping with the Inclusiveness Condition and taking seriously the unification of syntactic operations put forward in recent works, reducing movement to a special instance of Merge.

The paper is organized as follows: section 2 focuses on the issue of labeling from a theoretical point of view. We first discuss the two algorithms proposed in Chomsky 2008, providing a criticism and then proposing a new unified algorithm that can cover both External and Internal Merge (i.e., movement). Being defined on the notion of Probing, the system predicts that cases of conflict arise where more than one Probe is involved, which gives conflicting predictions on labeling. Two such case studies are discussed in the remainder of the paper. The first, discussed in section 3, concerns a conflict arising with External Merge: a case where the tension between two Probes derives what is standardly known as Principle C; the second, discussed in section 4, illustrates the same kind of conflict in connection to Internal Merge: interrogatives and free relatives are the case in point. Section 5 discusses an empirical prediction made by the analyses proposed in sections 3 and 4, when they are combined. Section 6 concludes the paper.

2. The Labeling Algorithm(s)

Chomsky (2008:145) proposes that the two algorithms in (2) and (3) are necessary and sufficient to yield labeled syntactic objects in most derivations:

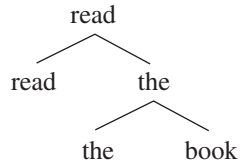
(2) In {H, α}, H a lexical item (LI), H is the label.

(3) If α is internally merged to β forming {α,β}, then the label of β is the label of {α,β}.

The status of these two principles is very different, as is their likelihood as syntactic primitives. Let us discuss them briefly in turn.

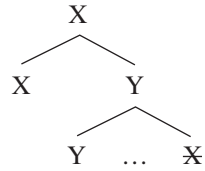
The concept of lexical item (LI) that is implicit in the algorithm in (2) is minimally simple—an LI is an item listed in the lexicon as such: a word. Rephrased in standard X-bar terms, (2) claims that it is always a head that projects. Under minimal assumptions on the relation of syntax and lexicon, (2) is a very likely candidate of a syntactic primitive, defining the centrality of words in syntactic derivations. To illustrate how (2) works, consider a case of External Merge of an LI to a syntactic object (SO), where SO is defined as the output of a Merge operation: as illustrated in (4), by virtue of (2), the SO generated by merging the LI with the SO gets the label of the LI (i.e., recall, a subset of its features: see (1)).

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(4) 

read the book

However, labeling is an issue concerning any kind of merge. If by merge we mean not only External Merge but also Internal Merge (i.e., movement), then we expect the algorithm in (1) to work indistinguishably in cases like (4) and in cases where movement is involved. Consider, for example, the abstract derivation in (5), where a simple lexical item is internally merged to a syntactic object.

(5) 

By (2), X provides the label. As such, (5) illustrates an interesting consequence of the algorithm (2) when applied to Internal Merge: (2) predicts that what is traditionally called head movement has the property of modifying the label of its target. The algorithm in (3) is exactly meant to avoid such a consequence and ensure that "in all movement operations it is always the target that projects." However, (3) explicitly sets apart External Merge, basically stipulating a residual of a "movement theory." This stipulation goes against the unification of syntactic operations that is explicit in the definition of movement as Internal Merge. As such, (3) is a severe departure from minimalist assumptions and ideally should be discarded. However, the algorithm in (2) alone is not enough to provide the computational system with an automatic device for labeling the core cases of syntactic objects created by Merge. Whereas we might expect labeling to be not always univocal, leaving some work to the interfaces, (2) alone provides too much indeterminacy and many suspicious and even wrong predictions. Let us see some of them in detail.

First of all, a system working with one and only one algorithm as (2) would have nothing to say about the very first step of any derivation, when two lexical items get merged, as in (6).

(6) {saw, John}

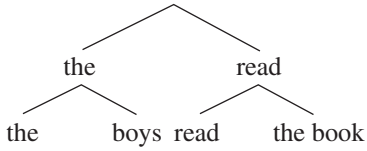
This would give us a weird grammar, in which any computation automatically runs at least two parallel derivations given any pair of lexical items, depending on which provides the label. In fact, this problem also arises if one assumes the pair of algorithms (2) and (3) proposed by Chomsky (2008). Chomsky discusses and acknowledges this problem but claims that a multiple spell-out system like the theory of phases ensures that the "wrong" derivation will crash early enough (Chomsky

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

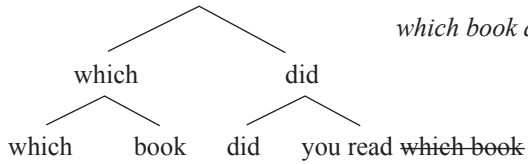
2008:145). Still, the system would introduce the computational burden of maintaining two parallel derivations up to the next higher phase even in trivial cases like (6) that are not temporarily ambiguous in any reasonable sense.

More problematic cases arising in a system containing only (2) are illustrated in (7) and (8), for External Merge and Internal Merge, respectively ((7) is a simplification because we do not represent the vP layer, but the same problem arises if a more detailed structure including vP is considered).

(7) *the boys read the book*



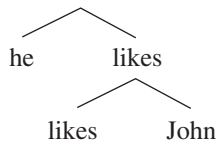
(8) *which book did you read*



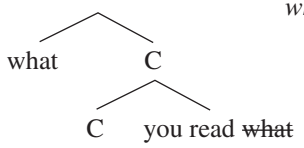
Both in (7) and (8), two objects are merged, neither of which is a lexical item: (2) might be taken to mean that they do not have any label; a clearly unwanted result. Alternatively, a system that has (2) as its only labeling algorithm might be taken to mean that labeling cannot be decided in such cases, and this is equally unsatisfactory.

Finally, the system yields wrong or at least very suspicious results in a number of contexts where a lexical item gets merged with a syntactic object, as illustrated in (9) and (10) for External Merge and Internal Merge, respectively.

(9) *he likes John*



(10) *what you read*



In both (9) and (10) the algorithm in (2) predicts that the label should be provided by the lexical item: a clearly wrong result in the case of (9), which is interpreted as a clause, not as a DP; a very suspicious result in (10), which can be interpreted as clausal in nature, not (necessarily) as a DP. Notice that the case in (10) is the reason why Chomsky (2008) stipulates the algorithm (3): to ensure that movement never

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

changes the label of its target.[2] In addition to being an unjustified stipulation, as already discussed, the algorithm (3) does not solve the problem of (9), which does not involve movement.

This quick review of some representative cases of Merge clearly shows that a system that contains only (2) as a labeling algorithm is unsatisfactory. A closer look at the problematic cases offers a simple solution, though. Consider the cases of "first merge" in (6). It is clear that the two lexical items selected from the numeration are not playing the same role in the computation: simply put, a transitive verb like *saw* selects a direct object like *John*, but *John* does not select *saw*. A classical way to describe this asymmetry is to say that *John* saturates *saw*, not vice versa. Given the strong unification thesis—that the operation responsible for movement and for structure expansion is one and the same (i.e., Merge)—it becomes appealing to frame this asymmetric relation between the two members of a merging pair in terms of a Probe–Goal relation: in this spirit we might say that *saw* has an unvalued feature (a selectional feature)—the Probe—that gets valued by some feature of *John*, the Goal. Capitalizing on this asymmetry, we might propose the following algorithm, which should replace (3) and complement the algorithm (2).

(11) The label of a syntactic object $\{\alpha, \beta\}$ is the feature(s) that act(s) as a Probe of the merging operation creating $\{\alpha, \beta\}$.[3]

What (11) basically says is that Merge is always asymmetrically triggered and is governed by the features of the items involved. To illustrate, in (6) *saw* provides the label because both the algorithm (2) and the algorithm in (11) converge: (the categorial feature of) *saw* is the probe referred to in (11) and *saw* is a lexical item, in compliance with (2).[4] Let us go back now to the other problematic cases. Suppose we (externally) merge a syntactic object with another syntactic object (the case in (7)): the algorithm in (2) has nothing to say because no LI is involved. But there will always be one (and by hypothesis only one) of the two syntactic objects that has triggered the operation needing the valuation of its selectional feature. So, in (7) the selectional/categorial feature of *read* will label the output. No difference arises when the same configuration is given by Internal Merge (i.e., movement): in (8) the operation is triggered by some feature of C, and the output ends up having label C.

So far, so good. The core cases of phrase structure construction seem to be captured by the interaction between (2) and (11). However, it is clear that a system based on just *one* labeling algorithm would be by far more minimal. Although we have shown

[2] Chomsky (2008:145) discusses the possibility that a conflict between the two algorithms might derive the ambiguity of (10), which can be either a free relative or an interrogative. This approach will be discussed in detail and confronted with ours in section 4.2.

[3] Something similar is proposed by Adger (2003:91), who reduces selection to a Probe–Goal relation and defines the head as the element that selects in any merging operation. Algorithm (11) is also reminiscent of Pesetsky & Torrego's (2006) Vehicle on Merge Requirement. Boeckx (2008:chap. 3) offers a detailed discussion on labeling, reaching similar conclusions. As will become clear, the system proposed here goes further, extending Probing to other relations not involving selection.

[4] The alternative derivation, in which *John* is the label of {*saw*, *John*}, obeys the algorithm in (2) but violates the algorithm in (11).

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

that (2) alone yields incomplete, contradictory, and even false predictions, we still have to explore whether a system including only (11) would fare any better. Although the intuition that lexical items are special, which motivates the algorithm (2), is sound, it might not require an ad hoc algorithm. Suppose we keep this intuition but reframe it in terms of features: capitalizing on a suggestion by Chomsky (2008:139, 144), we might say that every lexical item (with the exception of holophrastic expressions such as *yes* and *no* or interjections) is endowed with a feature, call it edge feature (EF), that forces it to merge with other material. If we assume that EF is what defines words as special entities permitting them to enter a computation, we can derive the effects of (2) without assuming it as a separate algorithm. More specifically, we propose that the EF of a word is to be identified with its categorial feature (after all, words come in different varieties because this allows them to combine according to rules of composition). If we assume this, any time an LI is merged, it qualifies as a Probe by virtue of its EF. This means that an LI, being a Probe by definition, always activates the algorithm in (11) and its categorial feature can provide the label. For example, each time a head (= LI) is merged with its complement, the categorial feature of the head is bound to project.

To illustrate how this system works, let us go back to the cases reviewed above. Let us start with "first merge": in (6), both *saw* and *John* are Probes, both being LI endowed with an EF. But the theory based on the algorithm in (11) still allows us to derive an asymmetry between them: the label of the syntactic object will be *saw* and not *John*, because *saw*, in addition to the EF, also carries a selection feature. This makes *saw* a "double Probe" with respect to *John*. So, assuming that a "double Probe" wins over a "single Probe," the label of {*saw*, *John*} will unambiguously be provided by the categorial feature of *saw*.[5]

Interestingly, however, it is not always the case that a selection feature is involved and interacts with an EF to define a lexical item as a double Probe. When adjunction is at play, Merge is by definition not triggered by selection (and no corresponding feature is involved) and EF alone drives Merge. Consider a simple derivation like (12).

(12) {{arrive, John}, early}

[5] We are assuming a derivational approach, where the notion of "double probe" is relativized to the step of the derivation in which the relevant Merge operation takes place. For example, if

- α is merged with β,
- α and β probe each other, and
- α has probed a third category γ in previous stages of the derivation, then

we do not expect α to be a double probe *with respect to β*. Such an abstract configuration can be illustrated in the case of T-to-C movement. Given the analysis of head movement that we give in section 4.1, T moves to the root when it is attracted by the syntactic object with label C (let us call it C′, using X-bar terminology for convenience). At this stage, in principle both T and C′ can project because they both qualify as a Probe (T is a probe by virtue of its EF and C′ is a probe by virtue of being the target). In fact, under normal circumstances, C′ will project. However, if T were a double probe by virtue of having probed the verb in previous stages of the derivation, we would incorrectly predict that T should be forced to provide the label. We thank a *Syntax* reviewer for pointing out the importance of giving a derivational definition of double probing.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

The only way for *early* to be merged is by virtue of its EF, which forces it to search for some material to attach to. As for the label of (12), the Probing Algorithm appears to (wrongly) predict that the adjunct itself should provide the label, being the probe of the operation. However, the issue arises only if *early* is merged cyclically, and it is widely assumed that adjunction can be a late, postcyclic operation. In fact, that adjuncts can be late merged is only natural if their insertion is not triggered by any feature of the material cyclically merged in the derivation. If *early* is late-merged, it attaches to a syntactic object that has its own label that cannot be changed (under some version of the No-Tampering Condition). Notice moreover that there are contexts where it can be shown that adjuncts *do* label the structure, as our theory predicts. This is the case of adjunct free relatives, which are crosslinguistically represented in a variety of languages (see Caponigro 2003 for an extensive crosslinguistic survey of adjunct free relatives). For example, (13) and (14) in Italian contain a free relative introduced by *come* and *dove* (the two *wh*-elements 'how' and 'where'). In (13) and (14) the adjunct ''projects'' by virtue of its EF, as our system predicts.

(13) Mangio come mi piace (mangiare).
eat as me pleases to-eat
'I eat how I like to eat.'

(14) Mangio la pizza dove tu mangi gli spaghetti.
eat the pizza where you eat the spaghetti
'I eat pizza where you eat spaghetti.'

This gives us the basis of a theory of adjunction, at least when adjuncts are lexical items, as *early* in (12), because only lexical items are endowed with an EF. But what about phrasal adjuncts? How can they be merged into a structure if they are not endowed with an EF (not being lexical items) and they are not selected by definition? A possible answer can be given in our system, and it exploits the same intuition that underlies a well-established explanation for why adjuncts are islands: Uriagereka's (1999) claim, in a multiple spell-out framework, that adjuncts are islands because the computational system treats them as lexical items. In the same spirit, we might argue that there is an option of ''lexicalizing'' syntactic objects and treating them as unanalyzable units, as such endowed with an EF. This option has a cost, though—the resulting object cannot be searched into and it becomes an island for extraction. This way islandhood of adjuncts follows from the way they are inserted into the structure.[6]

Let us close our discussion of the possible consequences of our proposal going back to the other cases we examined in relation to the original labeling algorithms proposed by Chomsky. Suppose we externally merge a syntactic object with another syntactic object: here no EF is present (none is a lexical item), and the label will be

[6] We thank an anonymous *Syntax* reviewer who led us to elaborate this approach to adjunction. What is proposed in the text is nothing but a sketch of a theory of adjunction, which we plan to develop in future work.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

provided by a subset of the features of the syntactic object that acts as a Probe of the operation. No difference arises when the same structure is generated by Internal Merge (i.e., movement).

An interesting consequence of this system is that, because the label is provided by the Probe, there can exist cases of labeling conflict if more than one Probe triggers the relevant merging operation. One such case, which we just saw, is (6), in which a double Probe wins over a single Probe. Other labeling conflicts, like (9) and (10), deserve closer attention. In both cases an LI is merged with an SO. The LI, as any LI, is provided with an EF, therefore is a Probe and should provide the label in compliance with the labeling algorithm (11). But the SO is the Probe of the operation, so it should become the label as well by the same algorithm. In these cases, a labeling conflict arises because there are two single Probes and they compete to become the label of the newly created syntactic object.

We devote the remainder of the paper to discussing these two cases in great detail in light of algorithm (11), henceforth referred to as the *Probing Algorithm*.

3. Principle C as a Case of Mislabeling

In this section we show that standard cases of Principle C can be reduced to symptoms of a mislabeling, dispensing with the canonical definition of Principle C, which is incompatible with the inclusiveness condition and is not minimally rooted as a syntactic primitive. For the purposes of this paper, we refer to the formulation in (15) as the canonical definition of Principle C.

(15) An R-expression cannot be c-commanded by a coindexed category.

The statement in (15) is a negative condition on the distribution of indexes. The tacit assumption is that DPs can be freely assigned identical indexes unless this is explicitly blocked. Principle C introduces one such blocking condition. Apart from the dubious status of indexes in the Minimalist Program, another possible concern with the canonical formulation of Principle C is that it is conceived as a primitive of the theory (whence the label *Principle* C), which is codified as such in UG. Although this is not unreasonable, given that Principle C is likely to be a language universal, if it were deducible from more primitive elements of UG, we would have an important simplification of the theory. In the same minimalist spirit, various attempts have been made to dispense with binding-theoretical principles. Chomsky (1993), Hornstein (2006), and Reuland (2001), among others, offer minimalist reformulation of Principles A and B. Kayne (2005) and Schlenker (2005) try to reduce Principle C from more primitive conditions. In Kayne's (2005) theory, every case in which a pronoun and its antecedent have the same semantic value is reduced to an instance of movement out of a clitic-doubling configuration. Principle C effects are then reduced to illicit cases of movement. In Schlenker's (2005) approach, Principle C (as well as the other binding-theoretic principles) follow from a nonstandard interpretive procedure, which can mimic the relation of c-command in the semantic component. The basic condition that replaces Principle C is an interpretative filter that prevents

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

any given object from appearing twice in any sequence of evaluation for a given sentence.

In this paper, we are concerned uniquely with Principle C and propose that its empirical coverage can be made to follow from the Probing Algorithm. In addition to being conceptually desirable, this is also preferable on empirical grounds, because a series of ''exceptions'' to Principle C (notably, identity sentences and clitic-doubling configurations) that require special stipulations in other accounts are naturally derived in ours.

3.1 Principle C Reduced to the Probing Algorithm

To deduce Principle C from the Probing Algorithm we introduce a special case of Probing, which we call *referential valuation*. The intuition that we would like to build on is that grammatical relations are asymmetric. For example, a DP values the agreement morpheme of the verb (and not vice versa). Similarly, a DP values the ϕ-features of an adjectival expression or it values the selection feature of a verb. We propose that something like that happens in a different domain—namely, referential properties of DPs. For example, if a referential expression like a proper name and a pronoun have the same semantic value (i.e., they pick out the same individual), this relation is asymmetric in the sense that it is the semantic value of the proper name that determines the semantic value of the pronoun (and not vice versa). Assuming a standard framework, one can say that a category *A* has an intrinsic semantic value—namely, it is a referential expression—if and only if its semantic value is independent from the function that assigns a value to free variables. It follows from this that, for example, a proper name has an intrinsic semantic value, but a pronoun does not. We define the notion of referential valuation as follows: *A* referentially values *B* if the semantic component receives an instruction from narrow syntax that has the effect that the semantic value of *B* must be the same as the semantic value of *A*.[7]

Given the similarities with other asymmetric relations, it should be clear that referential valuation is just another case of Probe–Goal matching, in which the Probe (a pronominal expression) searches for the Goal (a referential expression). As a result, we are widening the notion of Probe with respect to the way it is standardly conceived: Probe–Goal matching does not involve only valuation of ϕ-features, *wh*-features, and so forth, but also EFs and referential valuation: pretty much in the spirit of the strong unification we are trying to comply with in this paper.

Having introduced referential valuation, we are ready to discuss a standard case of Principle C violation like (16), in which *he* and *John* have the same semantic value (for the reader's convenience, here and in the rest of the paper we will continue to indicate that two categories have the same semantic value by coindexing them, but

[7] If a standard interpretative mechanism is assumed, a more precise definition of referential valuation goes as follows: *A* referentially values *B* if narrow syntax tells the semantic component to disregard all the assignment functions that do *not* assign to *B* the individual that is the intrinsic semantic value of *A*. However, any other semantic device that guarantees that if *A* referentially values *B*, then *B* gets the semantic value of *A* would work. We discuss how unbound pronouns are interpreted in sections 3.3 and 3.4.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

remember that this is just a notational device because we are assuming a system *without* indexes).

(16) *He_i likes $John_i$.

As already mentioned, when the subject *he* is internally merged with the rest of the structure, there is a conflict between two Probes that are both potential labels of the newly created syntactic object: *he*, being an LI, is endowed with an EF, which by definition qualifies it as a Probe. So *he* should provide the label. The Label T, on the other hand, being the Probe of the Merging operation, should provide the label as well.

Let us consider the two possible derivations, starting with the derivation in which *he* wins and transmit its label. The definition of label in (1) determines that only the label can trigger further computation. So, *he*, being the label, can probe *John* for its referential valuation and the reading in which *he* and *John* have the same semantic value does arise. However, this derivation is obviously problematic. There are at least two (related) problems with it, both stemming from the fact that (16) would receive a nominal label, *but it is a sentence, not a DP*. The first problem arises if, as is commonly assumed, syntactic categories are mapped to a restricted set of semantic types. Although there is no rigid one-to-one mapping (i.e., CPs have different semantic types when they are independent sentences and when they are relative clauses), CPs, and possibly TPs, *but not DPs*, have the semantic type *t* of sentences.[8] So, if labels play a role at the syntax–semantics interface, (16) is unable to receive a sentential interpretation due to its nominal label.

On a purely syntactic plane, the nominal label in (16) is equally problematic. After T and the subject have merged, the structure is not completed yet, given that it lacks the Comp area. But a complementizer does not select for a nominal label, so (16) will never be selected by the ''right'' category. The trouble with the derivation in which *he* wins is a problem of mislabeling (a similar mislabeling problem arises at the vP label if *he* probes *John* for referential evaluation at this early stage of the derivation).

Let us now consider the alternative derivation in which the Label T wins. T unproblematically provides the label but, given the definition in (1), *he*, not being a label, cannot probe *John* (as indicated by the lack of coindexing in 17).

(17) He_i likes $John_j$.

Notice that the acceptability of (17) indicates that a pronoun is not forced to probe its sister node to get referentially valued by a matching Goal. Arguably, this introduces a difference with other cases of Probe–Goal relations. For example, T *must* (as opposed to *can*) search its sister for a matching DP category that values its ϕ-features. However, the basis for this difference between referential valuation and other cases of Probe–Goal matching is quite intuitive. If a pronoun is not referentially valued by Probe–Goal matching, nothing goes wrong in the semantic component, given that an independent procedure can assign a value to it. In standard treatments, the pronoun in

[8] Ignoring special cases of propositional DPs (e.g., *He knows the time* for 'He knows what time it is').

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(16) is interpreted through the mechanism of assignment functions to free variables. Other devices through which ''unbound'' pronouns can be interpreted have been proposed (see sections 3.3 and 3.4). What is essential at this stage of our discussion is that there must be *some* device that guarantees that a pronoun is interpreted even if it cannot be valued through Probe–Goal matching, and this is not controversial.

The treatment of (16)–(17) straightforwardly extends to cases like (18), in which *he* cannot have the semantic value of either *John* or *Bill*, because *he* would end up projecting.

(18) *$\text{He}_{i/j}$ said that John_i likes Bill_i.

The next step is to show that our account of Principle C effects does not extend inappropriately. Take sentences in (19) as representatives.

(19) a. He_i likes his_i friends.
b. John_i likes his_i friends.

If *he* and *his* were in a Probe–Goal relation in (19a), the reading in which *he* and *his* have the same semantic value should be ruled out by the same reasoning that rules (16) out (namely, (19a) should be another case of mislabeling). A similar problem would arise with (19b), if *John* and *his* were in a Probe–Goal relation. However, this problem does not arise, because the Probe–Goal relation is asymmetric. In every case of Probe–Goal matching, intrinsic features of the Goal value those of the Probe. A referential expression like a proper name has intrinsic referential features, whereas a pronoun is not intrinsically referential. Given that *his* cannot be a Goal in (19a), *he* cannot be the projecting Probe. Therefore, the mislabeling problem does not arise (a similar reasoning applies in (19b)). Of course, a legitimate question is how the relevant reading arises in these sentences. The importance of this question may not be completely apparent, because in a framework that assumes that indexes are freely distributed there is nothing special to say about them. What happens is that *he* and *his* (or *John* and *his*) receive the same indexes and this is the end of the story. But we are trying to avoid using indexes, in compliance with the Inclusiveness Condition. So, we must explain how the relevant reading arises in the sentences in (19) *in absence of indexes*. We do that in section 3.4.

It is worth stressing that our approach takes pronouns at face value—namely, we assume that pronouns are LIs (in fact, intransitive Ds, as proposed in Abney 1987), not complex syntactic objects. This assumption is not universally accepted; there are theories that take pronouns (and names as well) to be definite descriptions (cf. Elbourne 2008). Clearly, if a pronoun were a complex syntactic object, it could not project in (16) and the explanation based on the Probing Algorithm would not be viable. So, our theory is incompatible with theories that take a pronoun to *literally* be a definite description. However, because the Probing Algorithm looks at the lexical features of the objects that undergo Merge, not at the way syntactic objects are interpreted, our approach is in principle compatible with various theories concerning pronoun interpretation (see sections 3.3 and 3.3 for more discussion on this).

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

Finally, let us consider a case where the definition of Principle C in terms of mislabeling seems to fare worse than the canonical formulation that assumes Principle C as a primitive. This is the case of DP-internal Principle C effects, like (20).

(20) *his$_i$ picture of John$_i$

If *his* is a D, our approach to Principle C makes the wrong prediction: *his* should be able to project and hence probe *John* without yielding any mislabeling. The canonical approach, on the other hand, correctly derives the Principle C effect. However, if *his* is not a D, a Principle C effect is expected under both approaches.

Some languages show the categorial nature of elements like *his* more directly than English does. In Italian, the counterpart of *his* is not a D, because it occurs with a determiner (cf. the grammaticality of (21), if the possessive and the proper name do not have the same semantic value). In (21) the Principle C effect can be reduced to a case of mislabeling because a determiner like *la* cannot combine with an object that has *sua* as a label (*sua* being an adjective-like element).

(21) la sua$_{*i/j}$ foto di Gianni$_i$
the his picture of Gianni

If the underlying structure of the English DP were the same as the one overtly displayed by Italian but for the fact the D is null, our approach would have no problem in deriving the Principle C effect in (20).

Offering a crosslinguistic analysis of DP structure is clearly outside the scope of this paper, so we leave this issue to future research and acknowledge that this is an area in which our approach should be further tested. Note, however, that proposals reducing English (and French) to the structure overtly displayed by Italian do exist. This has been proposed by Valois (1991:64), for example, who claims that the fact that only Italian allows the overt co-occurrence of the determiner and the possessive pronoun relates to the availability in Italian of a ϕ-feature transmission mechanism by virtue of which the determiner can transmit its ϕ-features to the possessive (this is the same as the transmission mechanism between *pro* and its associate in the clausal domain—an area in which Italian, and French and English, differ as well).

3.2 When the Canonical Definition of Principle C and the Definition in Terms of Mislabeling Diverge

Until now, we have been arguing that the approach to Principle C in terms of mislabeling is to be preferred on conceptual grounds, because Principle C would not be a primitive and because any use of referential indexes would be avoided. Here, we try to make a case for the formulation of Principle C in terms of mislabeling to be empirically superior. We identify three areas in which the two alternative approaches to Principle C make clearly divergent predictions, and the approach in terms of mislabeling fares better.

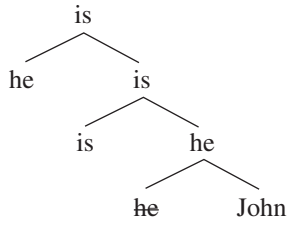
© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

The first domain is sentences of the form [DP is DP], which we call *identity sentence*s. The canonical formulation of Principle C makes an embarrassingly wrong prediction with identity sentences, given that (22) and similar sentences should be a patent violation of Principle C. In fact, they are perfectly OK.

(22) He_i is $John_i$.

Sentences like (22) are conveniently ignored in many discussions about Principle C.[9] The natural question is whether the formulation of Principle C in terms of mislabeling fares any better than more canonical approaches. We will show that it does. To see this, we have to focus on the initial step of the derivation of (22), when *he* and *John* are first merged. We assume that, at least in the case of identity sentences, the copula selects a headless small clause.[10] Given this structure, both *he* and *John* can provide the label. Let us focus on the derivation in which *he* does that. If *he* "projects," it can search its sister node for a Goal that can value its unvalued feature. This Goal is *John*. Given this derivation, the syntactic object created by merging *he* and *John* is a DP. Assuming that *he* later raises to T, (22) has the following structure:

(22′)



[9] Not always, though. Heim & Kratzer (1998:269–274) claim that identity sentences are in the same boat with "accidental coreference" cases like (i).

(i) Everyone likes John. Bill likes John, Mary likes John, Robert likes John. He_i likes $John_i$, too.

However, it is very dubious that (22) and (i) instantiate the same phenomenon. Rather special discourse contexts must be set up to bring out the judgments that coreference is possible in (i) and similar cases, but no special discourse context is required to make the same reading clear in (22). Heim (1998) also elaborates on the well-known distinction (due to Frege 1892) between the proposition expressed by an identity statement and its cognitive value. For example, (22) has two readings. The first is the tautological reading that states that John is identical to himself ($a = a$). The second reading ($a = b$) is more informative. Assuming that John can be associated to different guises (the guise 'Bill's best friend', the guise 'the person who is standing in front of me', etc.), the informative reading of (22) identifies two different guises as being associated to the same person. Specifically, (22) says that the person of whom the interlocutor has a current visual impression is the same person (called *John*) of whom the interlocutor carries in his/her memory an entry with various pieces of information. Schlenker (2005) elaborates on Heim's proposal to explain why the informative reading of (22) is not ruled out by Principle C. However, even if Heim's approach could be extended to the informative reading, it says nothing about the tautological reading of identity sentences, which *is possible*, contrary to what the standard formulation of Principle C predicts. Furthermore, as acknowledged by Schlenker, this approach runs into the risk of opening a Pandora's Box. If we introduce guises to explain the absence of binding violations in identity sentences, one can ask why we cannot *always* introduce different implicit descriptions to refer to a given individual, thus circumventing any kind of binding-theoretic violation.

[10] We are aware that, although solidly grounded and rather standard, this is not the only analysis for small clauses. See, among others, den Dikken 2006 and Adger & Ramchand 2003 for arguments against headless small clauses.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

In compliance with the Probing Algorithm, the label at each step of the derivation is determined by the following Probes:

(i) Label of {*he*, *John*} = categorial feature of *he* = D (*he* is a Probe, due to its EF)
(ii) Label of {*is*, {*he*, *John*}} = T (T is a Probe due to its EF and also because it selects the small clause)
(iii) Label of {*he*, {*is*, {*he*, *John*}}} = T (T is a Probe because it requires Internal Merge of *he*)

The critical step is (iii). The crucial observation is that the unvalued referential feature of *he* has already been valued in its base position. So *he* does not need to probe *John* at stage (iii) of the derivation, and the Probing Algorithm correctly dictates that the root can get a T label.

Clearly, what sets apart (22) and the Principle C configuration in (16) is that only in the former configuration can *he* probe *John* before moving to Spec,T without triggering any mislabeling at this early stage of the derivation. So, it is essential for our explanation that small clauses be allowed to receive a nominal label. We should double-check that this assumption is not problematic. Let us start by reflecting on the interpretive procedure. Assuming that the copy of *he* in its base position is not interpreted, the structure in (22′) reflects the fact that the copula identifies two categories that both have a DP label. As such, the structure in (22′) is compatible with the semantic analysis of copular sentences proposed by Partee (1987), who claims that in identity sentences an entity of type *e* is mapped onto the singleton set of entities identical with that entity. Thus, *John* is mapped onto the set of individuals who are identical with him (this set, of course, has just one element, John).

As for core syntax, no obvious reason prevents a category in the postcopular position to carry a nominal label. In fact, DPs normally sit in postcopular positions even when the canonical subject position is filled by an expletive-like element (e.g., *There were many boys*), so the null assumption seems to be that the copula can select for categories of D type.[11]

One might wonder why the small clause formed by merging *he* and *John* does not normally appear in positions in which DPs are allowed to appear, for example (23).[12]

(23) *I kicked [$_{DP}$ him John].

Arguably Case theory and θ-theory concur in explaining why (23) is out, because the two DPs *him* and *John* each need Case and a θ-role and neither sits in a position in which it can get one. In this respect, cases like (23) are different from cases like *I believe John to be wise*, in which, although two DPs are present, one DP (*John*) can

[11] As is well known, the hypothesis that *there* is an expletive has been challenged, most notably by Moro (1997). However, the gist of our proposal is that Principle C effects are avoided anytime the relevant sentence has the structure 'DP *is* DP' and this can probably be expressed, no matter if *there* is analyzed as an expletive or not.

[12] We thank an anonymous *Syntax* reviewer for this important question.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

check Case via exceptional Case marking and somehow share it with the other DP (details about the Case-sharing mechanism are irrelevant for our purposes in this paper).

Let us move to other cases of copular sentences to double-check if the approach that we are pursuing can account for them as well. First, let us focus on (24).

(24) He is [the friend of John].

Let us first consider the illicit reading in which *he* and *John* have the same semantic value. For this reading to arise, *he* has to probe *John*. This can happen either when *he* is first merged or raises to Spec,T. Regardless of when *he* probes *John*, a problem arises, though. In fact, if probing is constrained by a locality requirement, the reading in which *he* and *John* have the same semantic value is blocked by an intervention effect, because the closest DP that *he* can probe is the DP *the friend of John*, instead of the DP *John*.[13] This also explains why *he* and *the friend of John* can (in fact, *must*, given the semantics of copular sentences) have the same semantic value.

We think that the explanation in terms of intervention for the pattern in (24) is very intuitive. However, the concept of intervention is syntactic in nature. Therefore, this simple explanation can only be maintained if the referential valuation of the pronoun is the result of a syntactic operation, like Probing is. In this sense, the pattern in (24) is evidence for the approach that claims that referential valuation takes place as a result of a syntactic operation.

Let us now focus on a predicative copular sentence like (25). In such constructions, the obviation of Principle C effects seen with identity copular sentences is not observed.

(25) *He$_i$ is [envious of John$_i$].

This can be explained in our approach as follows. For the relevant reading to arise, *he* has to probe *John*. If this happens when *he* raises to the Spec,T, the familiar mislabeling problem arises because the sentence incorrectly gets a D label. If *he* probes when it is first merged, a different problem arises, because the small clause formed when *he* is merged with *envious of John* will get a D label but it must be interpreted as a predicate. If truly DPs, unlike NPs, cannot be mapped to predicates (cf. *I consider him *some/*every/*no boy*, **John is every boy (in this room)*), a mismatch takes place at the syntax–semantics interface.

We turn now to the second area in which the approach to Principle C in terms of mislabeling is superior to the traditional one. This can be illustrated by cases in which an expletive subject pronoun (*pro* as in the Italian sentence (26) or *il* as in the French sentence (27)) illicitly c-commands a postverbal subject. Sentences (26) and (27)

[13] The fact that *he* and *John* cannot have the same semantic value can be reduced to a Relativized Minimality effect, if intervention is defined in terms of containment (in addition to the classical definition in terms of c-command, due to Rizzi 1990). Descriptively, the intervention effect exemplified by (24) is a classical violation of the *i*-within-*i* filter.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

should be Principle C violations under a traditional approach but they are OK. Furthermore, the problem is not limited to subject pronouns, as shown by (28), a sentence which is grammatical in the varieties of Spanish (like River Plate Spanish) that allow clitic doubling, even if it should be out under standard assumptions.

(26) *pro*$_i$ È arrivato John$_i$.
is arrived John

(27) Il$_i$ est arrivé [un garçon]$_i$.
he is arrived a boy

(28) Lo$_i$ vimos a Juan$_i$.
him we-see to Juan

In a canonical framework, this unwelcome prediction may be blocked by some ad hoc assumption, such as by stipulating that expletive pronouns and/or clitic doubling configurations are somehow exempted from Principle C. However, as we will show, if Principle C is reduced to a mislabeling case, no special stipulation is needed.

According to a popular analysis of doubling cases (Torrego 1995, Uriagereka 1995, Cecchetto 2000, Belletti 1999, Boeckx 2003, among others) the clitic and the double are originated in the same phrase (a "big DP") and the former moves to its final landing site stranding the latter in the base position. If in the original configuration the clitic c-commands the double, the double can referentially value the clitic in the base position.[14] The clitic correctly transmits its label to the big DP. When it moves to its final landing site, the clitic does not need to probe the double for referential valuation, given that referential valuation has already taken place, so no mislabeling arises. This makes the cases in (26)–(28) very much like the case of identity sentences discussed earlier. In both configurations, referential valuation takes place at an early stage, so no mislabeling arises when the pronoun moves to its final landing site.

The third area in which the approach to Principle C in terms of mislabeling and the traditional one make divergent predictions is exemplified by sentences like (29) and (30).

(29) My father voted for my father.

(30) *He$_i$ voted for [my father]$_i$.

The canonical definition of Principle C rules out both (29) and (30). On the other hand, our approach excludes (30) as a case of mislabeling but does not preclude (29). In fact, (29) does not contain any pronominal expression, so the reasoning based on referential valuation simply cannot apply here. We would like to argue that this consequence of our approach is welcome. It is certain that (29) is odd, probably

[14] The clitic c-commands the double if the latter is inserted from the lexicon in the complement position of the big DP and the clitic is the head of the big DP (see Cecchetto 2000 for evidence showing this).

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

because a grammaticalized way to express the relevant information exists—namely, the sentence *My father voted for himself*. Still, the status of (29) cannot be equated to the status of (30). This becomes particularly clear in contexts that remove the oddity of (29) but cannot rescue the ungrammaticality of (30):

(31) In this election, each person voted for himself. This means for example that…
  a. √my father voted for my father.
  b. *he voted for my father.

In this section, we have shown that there are at least three areas in which the canonical definition of Principle C is problematic, whereas our approach fares better.[15]

3.3 Semantic Binding without Indexes

We have introduced the notion of referential valuation, which we propose to be the result of the syntactic configuration of Probe–Goal matching. At the semantic interface, referential valuation is read as an instruction to assign the very same individual to the category that gets valuated (the Probe) and to the one that valuates (the Goal).

We will now discuss how our approach can fit in a general theory of anaphora. A popular theory stemming from Tanya Reinhart's work includes two fundamental notions: semantic binding and (accidental) coreference. In this section we discuss the former.

A semantically binds *B* if *A* reduces the assignment dependency of *B*. Binding can be defined as the procedure of closing a property, which can be implemented as binding a free variable to a λ-operator, namely:

(32) *A* binds *B* iff *A* is the sister of a λ-predicate whose operator binds *B*.

One can ask if our approach requires semantic binding, in addition to referential valuation. The answer is positive. This is shown, for example, by the fact that we must explain how *John* and *his* can have the same semantic value in (19b), repeated as (33). In a framework like ours that does not allow indexes to be freely assigned (in fact, our framework, following a minimalist insight, bans indexes at all) the only way for *his* to become semantically dependent on *John* is through semantic binding:[16]

[15] A reviewer asks about cases like (i).

(i) *I believe [him$_i$ to hate John$_i$].

Because *believe* can take a DP complement, there should be no problem projecting *him*. So, our approach seems to predict no Principle C effect in (i). Our answer is that what goes wrong in (i) is that, if *him* projects, the embedded clause becomes a DP that must check Case and no Case is left for *him*. In other terms, the ungrammaticality of (i) is due to the impossibility of exceptional Case marking if the reading traditionally excluded by Principle C arises.

[16] By treating (33) as a case of semantic binding, we assume that proper names can undergo QR. See Heim & Kratzer 1998:chap. 8 for motivation.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(33) John likes his friends.
John (λx (x likes x's friends))

Of course, the same binding operation takes place in sentences like (34), in which the binder is a quantificational expression:

(34) [Every boy]$_i$ thinks that he$_i$ can win the competition.

Having said this, we must address a general question. How can the pronoun be bound by the λ-operator in absence of indexes? As a matter of fact, in popular treatments, like Heim & Kratzer 1998, indexes play the role of binding a free variable to a certain λ-operator. If indexes are eliminated, how can variables be bound by "their" λ-operator? Although the project of building an index-free syntax–semantics interface is outside the scope of this paper, we can show that in some core cases λ-operators can bind variables with no mediation by indexes (however, we continue using indexes as convenient notational devices in informal representations). For example, in (33), one can say that the λ-operator binds any variable that happens to occur in its c-command domain. Given that in (33) *his* is the only variable, it will turn out to be bound by (the λ-operator associated to) *John*. Indexes may seem to play a more substantial role when there are two potential binders, as in (35).

(35) [Every man]$_i$ said that [every boy]$_j$ likes his$_{i/j}$ picture.

However, at closer inspection it turns out that (35) is not a compelling argument for the existence of indexes, either. After all, indexes are necessary if one wants a tight correspondence between a specific bindee and a specific binder. However, (35) is ambiguous, which is exactly what one expects in a system that allows a λ-operator to bind whatever variable happens to be in its domain. More specifically, in absence of indexes, both (the λ-operator associated to) *every man* and (the λ-operator associated to) *every boy* can in principle bind *his*. The two available readings reflect these two options. Interestingly, cases like (35) are sharply different from cases of traces left by movement:

(36) A newspaper which *t* publishes every article that I like *t*.

The gaps in (36) are not free to choose their binder. For example, the gap in the object position of *like* cannot be associated to the relative pronoun *which*. Do we need indexes to capture this? Arguably not; the unwanted configuration is independently excluded (a chain between the object position of *like* and *which* would violate several well-established syntactic constraints). So, in (36) and similar cases, the syntax/semantics does not need to include indexed structures, as long as it contains the relevant information about chain formation (or as long as "traces" are copies).

Another configuration in which indexes are normally used is illustrated by sentence (37), in the reading in which *he* is unbound and *his* is bound by (the λ-operator associated to) *every boy*.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(37) [Every boy]$_i$ said that he$_j$ likes his$_i$ father.

Getting this reading in a system with indexes is straightforward (*he* can be assigned an index different from that of *every boy* and *his*). But if a λ-operator binds any variable in its c-command domain, how can we exempt *he* from being bound? Although examples like (37) may be seen as a challenge for a project of an index-free syntax–semantics interface, note that the problem originates only if unbound (or ''referential'') pronouns are treated as free variables. An alternative that has been recently revamped in formal semantics is treating unbound pronouns as indexicals whose extension is directly fixed as a function of the context, rather than through variable assignment. For example, Kratzer (2009) defends an approach along this line and explicitly rejects the idea that unbound third-person pronouns are free variables in the sense of formal logic. If Kratzer's proposal is on the right track, (37) stops being a problem for an index-free Logical Form because *he* is not treated as a variable to begin with. We will return to the question of how unbound pronouns are interpreted in section 3.4.

For now, we have to go quickly back to Principle C effects. If we assume (as we must) semantic binding, another problem seems to arise. Couldn't the illicit reading that we have excluded as a case of mislabeling *when Probe–Goal matching takes place* result from semantic binding, in absence of Probe–Goal matching? The standard assumption is that the binder must c-command the bindee and in Principle C configuration this does not happen, at least at Spell-Out. However, the proper name in (16) or the quantification expression in (38) in principle should be able to undergo QR and reach a position from which they c-command (and bind) the pronoun.

(38) *He$_i$ likes [every boy]$_i$.

So, the question remains open of why semantic binding is impossible in (16) and (38). Of course, we are dealing with a well-known problem, namely Strong Crossover. In the government-and-binding framework (Chomsky 1981), Strong Crossover cases were reduced to Principle C violations, because the trace left by the binder was equated to a referential expression. However, we are *not* assuming Principle C as a primitive, so we cannot take the easy way out by reducing Strong Crossover to Principle C. A natural explanation in terms of intervention is available, though. In fact, one can argue that what is wrong with (38) is that, after QR, the pronoun in the subject position creates an intervention effect for the chain that relates the quantificational expression and its copy:[17]

[17] The approach based on the intervention effects recasts the intuition underlying the Bijection Principle (Koopman & Sportiche 1982), which states that what is wrong with Strong (and Weak) Crossover configurations is that there is just one binder for two categories that need to be bound. However, a literal version of the Bijection Principle is untenable, for there are cases, like (i), in which an operator unproblematically binds two variables.

(i) [Every boy]$_i$ said that he$_i$ hates his$_i$ brother.

The approach based on intervention effects can distinguish between (i) and (38), because only in the latter case a chain created by movement (Internal Merge) gets disrupted by an intervention effect.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(39) *[Every boy]$_i$ he$_i$ likes ~~every boy~~
binder intervener bindee

In this section, we considered some core cases that suggest that indexes are not necessary to mediate semantic binding, at least if the popular treatment of unbound pronouns as free variables is abandoned in favor of the hypothesis that unbound pronouns are some sort of indexicals, as recently proposed. We also showed that a natural account of Strong Crossover in terms of intervention is easily available.

3.4 Dispensing with (Accidental) Coreference

The second notion that is normally assumed in the well-established theory of anaphora inspired by Reinhart's work is (accidental) coreference. It is assumed that two categories *A* and *B* corefer when they denote the same individual as a consequence of the value that the assignment function assigns to free variables. Coreference is thought to be necessary in all the cases in which a pronoun and the category it is semantically dependent on are not (and cannot be) in a formal configuration that can explain their semantic relation. Maybe the clearest case is intersentential anaphora, like (40).

(40) John$_i$ worked long hours. He$_i$ was very tired.

In (40) *John* cannot semantically bind *he*; they are not even in the same sentence. Still, the anaphoric reading is possible. Another case in which accidental coreference is usually called for is a sentence like (41).

(41) His$_i$ mother loves John$_i$.

In (41) the anaphoric reading may not be fully natural, but it is surely better than in (42).

(42) ??His$_i$ mother loves [every boy]$_i$.

Example (42) shows that semantic binding (after QR of *every boy*) results in a degraded output. Given that no such degradation is present in (41), the anaphoric reading cannot be originated by semantic binding and must be attributed to some other mechanism: typically, (accidental) coreference.[18]

The theory of coreference that we are summarizing assumes that indexes are freely assigned to NPs, so it can happen that any two NPs can corefer because they are fortuitously given the same index. It should be clear that this very idea goes against

[18] A reviewer asks why Weak Crossover cases like (42) are less deviant than Strong Crossover cases like (38), if both configurations are treated as cases of intervention. We cannot elaborate on Weak Crossover for reasons of space, but we can note that our approach easily fits theories, like Pica & Snyder's (1995), that do *not* regard Weak Crossover effects as intervention effects and offer an alternative explanation.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

the minimalist attempt to eliminate indexes. However, one does not need to be minimalist to recognize that (accidental) coreference is an extremely powerful mechanism, which, if left unrestricted, would make any constraint on referential dependencies totally vacuous. In the literature, the standard way to restrict (accidental) coreference is Grodzinsky & Reinhart's (1993) Rule-I, which in turn elaborates on the approach initially proposed by Reinhart (1983):

(43) Rule-I
NP *A* cannot corefer with NP *B* if replacing *A*, at LF, with a variable A-bound by *B*, yields an indistinguishable interpretation.

As discussed in the literature, Rule-I can get the desired result in many cases (e.g., it can block coreference in Principle C configurations like (16)). However, we are exploring the plausibility of a computational system without indexes, so we need to do our best to avoid coreference. Doing this is a long-term project and all we can do in this paper is sketch a possible approach and apply it to some core cases.

This is our line of attack to the problem. Imagine a system in which for any two categories to be in an anaphoric relation there must always be a formal operation that licenses this relation. This is exactly the opposite of the idea that indexes are freely distributed and that coreference is always available unless something (namely, Rule-I) blocks it. That a system like the one we are favoring makes sense is suggested by the fact that the supposedly free coreference mechanism can be shown to be sensitive to structural configurations. For example, as initially observed by Calabrese (1992), cross-sentential anaphora is affected by the position of the subject in Italian sentences like (44) and (45).

(44) John$_i$ ha lavorato. Poi *pro*$_i$ è andato al cinema.
John has worked later has gone to-the movie

(45) ??Ha lavorato John$_i$. Poi *pro*$_i$ è andato al cinema.
has worked John later has gone to-the movie

In (44) in which the subject *John* is preverbal in the first sentence, the null subject in the second sentence can be anaphoric to it. In (45), in which *John* is postverbal, the anaphoric reading is much harder. A similar pattern is found in (46) and (47), another case in which semantic binding is impossible since the potential binder *John* is in an 'if' clause and should escape this island to bind the null subject of the matrix clause. So, (46) is supposed to be a coreference case. However, if the subject is postverbal, as in (47), the anaphoric reading becomes much more difficult.

(46) Se John$_i$ viene licenziato, *pro*$_i$ si deprime.
if John is fired gets depressed

(47) ??Se viene licenziato John$_i$, *pro*$_i$ si deprime.
if is fired John gets depressed

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

Why should the position of the subject matter if indexes are distributed freely? We need a mechanism that makes the anaphoric reading dependent on the structural analysis of the sentence but, clearly, the unconstrained mechanism of coreference is nothing like that (note that Rule-I has nothing to say on the contrast between (44)/(45) and (46)/(47) because the pronoun does not sit in a position in which a variable can be bound). Let us thoroughly explore this issue by sticking to the framework that we adopt in this paper.

In our system we have just two ways to connect two DPs such that one of them (the pronoun) is semantically dependent on the other. The first way, ordinary semantic binding, does not apply in the sentences we are considering, as we just saw. The other device is referential valuation, introduced in section 3.1. Does referential valuation apply in (40) and (41) and in the corresponding Italian cases? Apparently not; the pronoun (the potential Probe) never c-commands the Goal that might referentially value it (*John*). This is easily shown: in (40) *he* does not c-command *John* because they are not even clausemates. In (41), *his* and *John* are clausemates, but the former does not c-command the latter. The same lack of c-command is observed in the Italian cases.

However, the Italian cases give us a clue. It is well known (see Cardinaletti 1997 and Belletti 2001, among others) that in Italian the preverbal subject position is associated with given information whereas the postverbal subject position is associated with new information. So, although the preverbal subject is interpreted as a topic, the postverbal subject is interpreted as Information Focus, adopting Kiss's (1998) terminology. In much recent work stemming from Rizzi's (1997) analysis of the left periphery, it is assumed that the left periphery hosts a TOPIC head, which can attract a topic phrase to its specifier position. By building on this type of analysis, we propose that, even if no topic phrase overtly moves to the TOPIC head, the TOPIC head can act as a Probe that looks for a Goal that can referentially value it. For example, this is what happens in sentences like (44) and (46), in which the TOPIC head is referentially valued by *John*. Because *John* is new information in (45) and (47), the TOPIC head cannot be referentially valued by it.

We can now link this discussion to the claim made in section 3.3 that unbound pronouns are a special kind of indexicals whose extension is fixed when a discourse context is supplied. The crucial question is what we mean by discourse context. Clearly, the notion of "context relevance" for fixing the reference of unbound pronouns is not limited to the familiar notion of context consisting of the set of speaker, hearer, time, and place of utterance that is sufficient for interpreting a typical deictic pronoun. The relevant notion of context will have to include individuals referred to by an act of pointing but also (crucially, for our purposes) individuals that are familiar from the previous discourse. For concreteness, let us call *discourse store* the set made up by the latter individuals. A way to make the intuitive notion of discourse store more precise is by assuming that only categories that referentially value a TOPIC head can enter the discourse store. So, in the alleged cases of accidental coreference in (40), (41), (44), and (46), because the DP *John* has referentially valued the TOPIC head, it has successfully entered the discourse store,

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

and by doing so, it can be referred to by the pronoun in the given context. In (45) and (47), *John* is new information, so it has not referentially valued the TOPIC head. Therefore, it has not entered the discourse store and cannot be referred to by the pronoun.[19]

Our account is still lacking in one respect. We have to explain why the mechanism based on access to the discourse store through probing by the TOPIC head does not apply in canonical cases of Principle C violations like (16). To answer this question, we would like to build on the same intuition that inspired Rule I—namely, if a grammaticalized way to establish an anaphoric relation exists, this blocks a discourse-based procedure. In (16) *he* and *John* can get the same semantic value as a result of Probe–Goal matching (in fact, in (16) the derivation crashes after referential valuation takes place, but in other cases—namely, identity sentences, clitic-doubling sentences, and in the examples in section 5—referential valuation is harmless). So, *he* cannot access the discourse store to get the interpretation that it gets as a result of referential valuation. Why is the discourse-based procedure blocked by the syntactic operation of probing? The intuition is that syntax consists of costless, automatic procedures whereas accessing the discourse store is a more global operation that is done only as required. If you can do something the easy way, you do not even try the difficult one (see Reinhart 1983 for an early defense of this claim).

This perspective can also explain why (41), repeated here as (48), is acceptable.

(48) His_i mother loves John_i.

In (48) *his* cannot probe *John* for referential valuation, due to the lack of c-command. So, the strategy based on accessing the discourse store is not blocked by the presence of a grammaticalized way to establish the anaphoric relation. *His* can be covalued with *John* as long as *John* has been probed by a TOPIC head in the previous discourse.

In summary, we have argued for a radical shift of perspective. In a framework that assumes coreference, an anaphoric relation is established for free, unless something blocks it. On the contrary, we propose that an anaphoric link is never free but only emerges if it is licensed by a formal operation. We first assumed two such formal operations: ordinary semantic binding and referential valuation. We then proposed that only a topic, formally defined as a Goal that referentially values a TOPIC head in a standard Probe–Goal configuration, can enter the discourse store. Finally, unbound

[19] Kiss's (1998) traditional distinction between Identificational Focus (which expresses exhaustive identification) and Information Focus (which conveys new information) plays a role in this respect. In fact, whereas Information Focus is inherently incapable of valuating a TOPIC head, Identificational Focus, like a category of the type "only X," does not need to be new information. For example, a sentence like *Only John came to the party* might be felicitously uttered in a discourse context in which John has already been mentioned, as shown in (i).

(i) John and Mary were both invited. However, only John came to the party. He enjoyed it.

In (i) *John* may enter the discourse store when the first sentence in the discourse (*John and Mary were both invited*) is uttered. Therefore, the pronoun *he* in the last sentence of the discourse can pick out *John*, even if *only John* is an Identificational Focus in the second sentence.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

pronouns can be given only the values of categories that are contained in the discourse store. It is worth stressing that we only sketched here the basis of a system that avoids indexing. Much work is needed to see if such a system could work outside the core cases considered here. However, the project of an index-free syntax–semantics interface is not especially tied to the specifics of our proposal regarding Principle C effects and should be pursued by anyone who accepts the minimalist insight underlying the Inclusiveness Condition. In fact, the hypothesis that we outlined is not the only index-free research program one can conceive. For example, Sauerland (2007) has proposed a system in which binders are unindexed λ-operators and bound elements are definite descriptions. As far as we can see, our approach is compatible with Sauerland's. We leave to future research a critical evaluation of these two alternatives. Our point here was to argue that such a project is much needed, and feasible as well.

4. Labels and Movement

In the previous section we discussed a case of conflict between two Probes competing for providing the Label, where only one of the two possible outputs is legitimate. However, another type of situation might in principle hold as well: one where, given a conflict, both possible outputs are acceptable and interpreted as (different) syntactic/semantic objects. This is what we shall discuss in this section.

4.1 Does "Head Movement" Exist in Syntax?

Recall we have been assuming the strong unification thesis—namely, that the operation responsible for movement and for structure expansion is the same: Merge, simply defined as an operation putting together lexical items or syntactic objects. One consequence that has been implicit throughout the paper is that any of the following sets should be available, for both Internal and External Merge.

(49) {LI, LI}[20] (order irrelevant assuming that the linear dimension
{SO, LI} falls within the phonological component)
{SO, SO}

On the "movement" side, this implies that what is traditionally called head movement and what is traditionally called phrasal movement should both be available to computation, contra recent attempts to ban head movement from syntax (e.g., Chomsky 2001).[21] These approaches crucially rest on the assumption that head movement is not to be considered as syntactic in nature because it lacks the essential

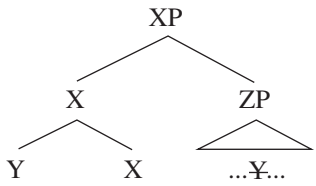
[20] Actually, this is not available to Internal Merge by definition: a lexical item cannot contain another lexical item to be internally merged to it.

[21] See Matushansky 2006 and Donati 2006 for a detailed criticism of this position, showing that it holds on wrong assumptions, both theoretical and empirical, that it is incompatible with other aspects of Chomsky's theory (e.g., phases), and that it brings undesirable consequences.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

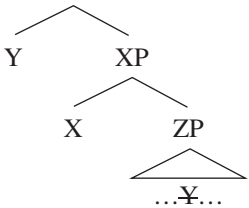
cyclic character of syntactic operations. This is certainly true of the standard head-adjunction configuration given in (50).

(50)



The traditional motivation for this configuration is that in core cases of head movement considered in the literature (V-to-T movement or V-to-C movement) the two heads conflate and behave as a single unit, hence the assumption that they form a sort of a "derived lexical item" represented in the head adjunction configuration above. But this is not the only configuration that head movement can, in principle, produce. Suppose we have an SO X endowed with a feature that needs to be valued. Nothing prevents the Internal Merge of (a copy of) a head (= LI) Y endowed with a matching feature to the root of the structure, as in (51).

(51)



The configuration in (51) is obtained by merging a new item to the root of the tree, hence complying with the extension condition. Given (51), the head conflation effect correlated with many cases of head movement can be the result of an independent process, perhaps phonological, call it affixation, which has nothing to do with head movement: affixation is something that can happen to two adjacent heads, independently from how and why they ended up being adjacent. The structure in (51) illustrates the kind of head movement predicted by the definition of movement in terms of Internal Merge—the kind of head movement we are considering here. So, from now on, by *head movement* we mean the movement of an LI, as shown in (51), and not the countercyclic configuration in (50).

4.2 Labeling Properties of Head and Phrasal Movement

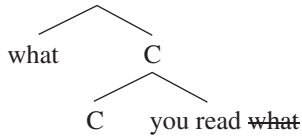
If we have two options available, head movement and phrasal movement, the question of what triggers the choice between the two becomes interesting. We propose that the Probing Algorithm can provide an answer. But let us proceed step by step.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

The most standard proposal for discriminating between the two movement options and accounting for their complementary distribution relies on locality differences. In a nutshell, head movement is claimed to be constrained by a specific locality condition (the Head Movement Constraint; Travis 1984), and thus available only in a very restricted set of cases—namely, when the Goal is a feature of the head of the complement of the Probe. However this account, which has been challenged on empirical grounds by many (Lema & Rivero 1990; Borsley, Rivero & Stephens 1996; Carnie 1995; Roberts 1994; Manzini 1994; among others) is really not an option given the minimalist approach to movement we are adopting. Recall that locality is a condition on the search procedure establishing Probe–Goal relations, which is only defined in terms of features: an unvalued feature acts as a Probe, and a matching feature gets searched as a Goal. Internal Merge is not part of this searching procedure but really a separate though parasitic operation triggered by some extra mechanism (generalized pied-piping; Chomsky 2007). As a consequence, the phrasal status of what is internally merged cannot be determined by the search procedure, nor by the locality constraints that affect it.

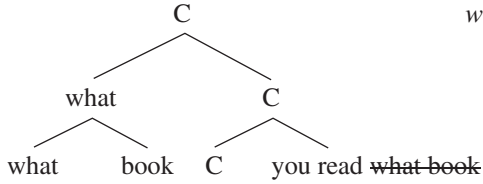
The solution to this problem lies in the Probing Algorithm: whereas head movement and phrasal movement cannot be distinguished on the basis of the Probe–Goal relation they establish, they have very different effects on labeling. Consider again the labeling conflict (10), repeated here as (52), from which our discussion began.

(52) *what you read*



Example (52) is derived by internally merging a single lexical item *what* (head movement) to the edge of a clause—an option, as we have seen, that we have no reason to exclude. The result is a conflict between two probes: *what*, being a lexical item, is by definition a probe (due to its EF) and should provide the label. On the other hand, C, being the Probe of the merging operation, should also provide the label. This kind of conflict never arises when a phrase is internally merged, as in (53).

(53) *what book you read*

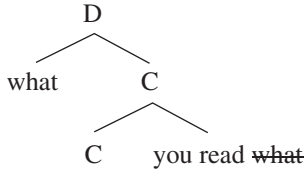


Here Merge holds between two SOs, and no conflict arises: by the Probing Algorithm (11), C, the Probe of the merging operation, labels the entire construction.

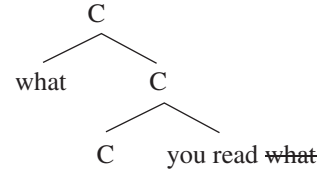
© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

The prediction is that the minimal difference between (52) and (53) should be reflected in the distribution and interpretation of the two structures: more precisely, (52) is predicted to have two possible labels, illustrated in (54), whereas the only labeling available to (53) is shown in the corresponding diagram.

(54) a.



b.



This prediction is reflected by the systematic ambiguity of a phrase like *what you read*: it can be interpreted as a free relative and be embedded under a verb selecting a DP.

(55) a. I read what you read/a book.
b. I read the thing that you read.

However, it can also be interpreted as an indirect interrogative clause and be embedded under verbs selecting for clausal complements:

(56) a. I wonder what you read/if the sun will shine tomorrow.
b. I wonder what book you read.

These two readings and distributions correspond to the two labeling possibilities: in (54a) the clause gets the D category of *what* and the structure of a (free) relative clause *à la* Kayne (1994) and Bianchi (1999); in (54b) the clause gets the C category and the structure of an interrogative clause.[22]

Crucially no ambiguity at all, neither in interpretation nor in distribution, arises when phrasal movement is involved: a clause corresponding to (53) can only occur in environments for clauses and can only be interpreted as a simple interrogative, as shown in (57).

[22] The idea that free relatives are derived through a projecting movement is not new, starting from Larson's (1998) first intuition: Donati (2006) developed a system similar to ours, but it included a residue of a phrase structure theory. In Donati 2006, the bare *wh*-element *what*, being both minimal and maximal, could move either as a phrase (hence not projecting and deriving an interrogative, as (56)) or as a head (hence projecting and yielding a free relative, as (55)). The system discussed here is more minimal in that *what* is always and only the thing it seems to be: an LI that moves as such. The ambiguity of the derived structure is explained by the labeling conflict, not by an ambiguity in the derivation. In Citko 2006, free relatives are interpreted as an instance of Internal Merge with a projecting Goal, but this option is not related to the phrase status of the moved element, the framework being that of a systematic exploration of all the logical possibilities available to a minimalist approach to labeling. Other proposals in the same direction not explicitly addressing labeling issues are Donati 1999; Iatridou, Anagnostopoulou & Izvorski 2001; and Bury 2003.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(57) a. I wonder what book you read.
b. *I read what book you read.

There is an (apparent) exception to this generalization: a class of free relatives that appear to allow phrasal *wh*-movement, illustrated in (58) and (59), respectively for Italian and English.

(58) Farò [qualsiasi cosa] vorrai ~~[qualsiasi cosa]~~.
I-will-do whatever thing you-will-want

(59) I shall visit [whatever town] you will visit *t*.

Battye (1989) offered a number of empirical arguments showing that these structures ought to be treated as "pseudo free relatives" (see also Kayne 1994:154n for a similar analysis proposed on different grounds). Although we cannot summarize all the arguments offered by Battye here, one revealing observation is worth mentioning—namely, that pseudo free relatives of the 'whatever'-type are compatible with the complementizer found in ordinary relatives in languages like Italian:

(60) Farò [qualsiasi cosa] **che** vorrai ~~[qualsiasi cosa]~~.
I-will-do whatever thing that you-will-want

This suggests an obvious analysis, which is that pseudo free relatives should be regarded as ordinary relatives in which the external determiner gets a special form. Assuming this, pseudo free relatives can be analyzed as ordinary relatives, extending to them the raising analysis proposed by Kayne (1994) and Bianchi (1999):

(61) [$_{DP}$ Qualsiasi [$_{CP}$ cosa] (che) [vorrai ~~cosa~~]].
whatever thing that you-will-want

The same analysis extends naturally to other "maximalizing relatives" (Grosu 2002) like (62), whose interpretation strongly suggests the presence of a silent *ever*-type determiner.

(62) I will read what books you will tell me.

We are now in the position of answering this question: what is the difference between the two movements available to syntax? We have been claiming that this difference has to do with their labeling consequences: whereas phrasal movement had no effect on the label of its target, head movement systematically gives rise to an extra labeling possibility (because heads are Probes by definition), in which the moved head "relabels" the target.[23]

[23] Or better "apparently relabels" the target: we have been claiming that head movement simply expands the structure, adding an extra label to the category it merges with.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

4.3 On the Distribution of Head and Phrasal Movement

In our system it is possible to distinguish between what are traditionally called head movement and phrasal movement. However, we have not yet derived the complementary distribution of the two movement options. In particular, if what head movement does is generate more labeling possibilities, why isn't it always selected in any derivation, especially given its more minimal status? Let us try to answer this important question.

Given what we saw in the preceding section, the alleged complementary distribution of the two types of movement is not that systematic: much of it is an illusion due to how phrase structure worked in preminimalist terms. In standard X-bar theory accounts, the head movement cases discussed above (i.e., free relatives) would be analyzed as phrasal movement cases, under the assumption that a simple lexical item cannot occupy a position where a phrase can sit. This way of looking at things is completely incompatible with the Inclusiveness Condition we have tried to adhere to in the paper. It is simply impossible to assume an ambiguity in the phrase structure status of a head: a head is a lexical item—that is, an element listed in the lexicon as such. The idea of a single lexical item being a phrase simply makes no sense, if phrases are defined as the syntactic objects obtained by merging *two* things.[24] This implies that each time we see a lexical item being displaced, we have to analyze it as head (= LI) movement.

We can illustrate this by briefly reflecting on the distribution of clitics in Romance languages. The literature on this topic is huge, reflecting the very intricate pattern of clitics in different varieties, and we cannot do justice to it here. However, we can focus on what is probably the most significant feature of clitics, which is that, although they are standardly analyzed as undergoing phrasal movement,[25] typically clitics are heads being displaced from their thematic position to an inflection-related position, as illustrated in (63). There is no reason (and no way) in the present approach not to analyze clitic movement as head movement.

(63) Maria lo conosce ~~lo~~.
Maria him knows him
'Maria knows him.'

Given this important proviso, we can just say that in canonical cases of *wh*-movement both head movement and phrasal movement can be displayed, in a distribution that is

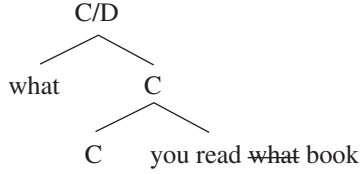
[24] Chomsky (1995) assumed that a head with the distribution of a phrase is an ambiguous element, being both minimal, because it is an LI, and maximal (he assumed a relational definition according to which a maximal projection is a category that does not project any further). Our system is simpler; it dispenses with the relational definition of maximal projection.

[25] Clitics in Romance have always been a problem. Given standard assumptions on phrase structure, clitics are clearly "maximal projections" in the position where they are generated, but they end up being affixed to an inflectional head. This tension is at the root of standard and influential analyses such as Kayne 1989, where the clitic starts up as a phrase but moves and adjoins as a head, changing its phrase structure status in the course of the derivation (violating a condition like that of the uniformity of chains; Chomsky 1995). A uniform head movement analysis is by far more minimal.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

not complementary at all. On the other hand, only head movement of a *wh*-element is able to yield free relatives as the one discussed in relation to (52).

However, we still need to explain why head movement is not the *only* option in any environment, and in *wh*-constructions in particular. This question is important because head movement is more minimal than phrasal movement in an intuitive sense (less stuff is involved), so economy considerations, if anything, should favor head movement over phrasal movement when a choice is given. So, we need to explain why (64) is an impossible derivation in English and in many other languages.

(64) 

**what you read book*

In (64) the lexical item *what* is extracted from the phrase it labels/heads, and internally merged to the root. In this configuration the Probing Algorithm (11) yields two possibilities, hence a conflict: C, the probe of the operation, should provide the label and the result should be an interrogative clause. On the other hand *what*, an LI, is also a Probe, due to its EF. So, it should be able to provide the label and the resulting structure should be a relative clause. The configuration should be ambiguous, but it is not. As shown in (65), it can neither be embedded under a context selecting for an interrogative clause nor under a context selecting for a relative clause: plainly, it is ungrammatical.

(65) a. *I wonder what you read book.
b. *I read what you read book.

There are at least two possible approaches to explain this restriction. One is to claim that (64) is an illicit movement in that it extracts a subconstituent out of a constituent. This would amount to excluding any instance of head movement that does not involve an intransitive head, in any context. This approach appears to be too strong in light of familiar cases of verb movement: V-to-T or T-to-C movements are exactly extractions of a head out of its constituent, at least under standard accounts. From this point of view, (66) provides an interesting contrast.

(66) a. *I wonder what you read ~~what~~ book.
b. [$_{TP}$ You read [$_{vP}$ [$_{VP}$ ~~read~~ that book]]].

Observing the contrast in (66) we might propose an alternative explanation that builds on a more precise characterization of the relation between the head and the label of a syntactic object. Let us go back briefly to the original definition given in (1), repeated here as (67).

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(67) Label: features of a syntactic object (SO) that can trigger further computation

By (67), only a subset of the features of a syntactic object survives the derivation and defines the label. Suppose a feature *A* is probed in a syntactic derivation: if *A* is shared by the head and by its label, the label by definition will always be closer than the head, and phrase movement will be triggered. If, on the other hand, the feature *A* does not belong to the subset of the features of the head that define the label (the features that percolate up to the label, to speak informally), we predict that the label does not intervene, and head movement is triggered. Let us see how this abstract system might derive the contrast in (66).

In (66a) the features probed by C are the categorial feature D and a *wh*-feature as well: these features plausibly project up to the label (phrases have a category and an interrogative status). As a result, the D-feature of the label is closer than the D-feature of the head and subextraction of the head alone is impossible; similarly, for the *wh*-feature. In (66b), on the other hand, the feature probed by T is a not a categorial feature, but rather a set of inflectional features on the verb itself, which do not project up to the VP label: as a result, the label does not act as an intervener and the subextraction of the head alone is possible.

Note that extraction of a *wh*-''determiner'' out of a *wh*-phrase, although difficult, is not completely impossible: the literature on *wh*-movement is full of cases like those illustrated by French sentence (68a) and by German sentence (69c).

(68) a. **Combien** as-tu lu **de livres**?
how-NEUT have-you read of books
'How many books have you read?'
b. **Combien de livres** as-tu lus?

(69) a. **Welche Bücher** hat Johann gelesen?
which-PL books has Johann read
'Which books has Johann read?'
b. ***Welche** hat Johann **Bücher** gelesen?
c. **Was** hat Johann **für Bücher** gelesen?
what-NEUT has Johann for books read

These data have been analyzed in various ways in the literature, all assuming phrasal *wh*-movement of a *wh*-''determiner,'' given the standard X-bar theory restrictions discussed above. In our system, we shall analyze all these cases as instances of licit head movement of a *wh*-determiner out of a complex phrase. Actually, our system predicts that determiners are disallowed to move out of a category only if they provide a label to this category, given that only in this case the label acts as an intervener. From this point of view the data above suggest that a crucial factor in determining whether a D provides the label of its constituent and is thus disallowed to move is to see whether it is involved in an agreement relation: in German in (69a,b) *welche* agrees with its NP complement and moving it as a bare head is impossible;

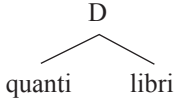
© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

in (69c) *was* does not agree with *Bücher* and such extraction is OK. The same holds in French (68): here *combien* can move as a head, but crucially it is not involved in any agreement relation. Something similar is visible in Italian in (70) versus (71): whereas *quanto* cannot strand its associate when it agrees with it (when it is an NP, as in (70)), it can be extracted when it does not agree with it (as with the AP in (71)).

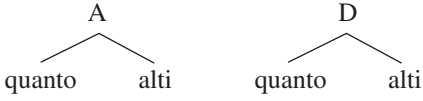
(70) a. **Quanti libri** hai letto?
how-PL books you-have read
'How many books have you read?'
b. ***Quanti** hai letto **libri**?

(71) a. **Quanto** hai detto che sono **alti**?
how-SG you-have said that they-are tall-PL
b. **Quanto alti** hai detto che sono?

Why should agreement play such a role? Suppose that an agreeing D is a Probe searching a Goal to valuate its agreement features. This presupposes that an agreeing D necessarily provides the label to its constituent (given the definition in (1), repeated in (67)). As a result, the label of a syntactic object involving an agreeing D always shares the relevant features with the D and is by definition closer to an external Probe than the D itself: this is illustrated in (72) for the Italian case. The D (or *wh*-) feature of the label is clearly closer to any external Probe and blocks subextraction of the head *quanti*.

(72)


Suppose, however, that a nonagreeing D is *not* a Probe. This implies that when it is merged with some category, nothing forces the D to provide the label, and two configurations are equally possible: either the merged constituent provides the label, or D itself does. In (73), for example, either the adjective *alti* provides the label, or *quanto* does.

(73)


In the latter case, the label intervenes, when a D (or a *wh*-) feature is externally probed (say from C): as a result, the entire constituent moves. In the former case, the label is not endowed with the relevant features, hence it does not intervene and head movement is triggered (see (71a) and similar cases of D extraction in French and German). So, it appears that our approach, far from being challenged by the optionality of D extraction out of a DP, naturally derives it as just another case

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

(in addition to the one illustrated by free relatives) of labeling ambiguity: two Probes compete to ''project,'' but the output is acceptable at the interface no matter which one wins and labels to the structure.

In summary, our system predicts that head movement is always possible as long as there is no label endowed with the relevant feature that is closer to the Probe than the head itself. This happens in at least two different cases: either when the probed features do not belong to the subset of features that define the label (e.g., V-to-T movement), or when the head whose features are probed does not provide the label to the object that contains it (e.g., the facts discussed earlier of subextraction of D).

In turn, this reinforces the conclusion that the distribution of ''head'' movement is much less limited than is usually assumed—a conclusion that allows a unified theory of phrasal and head movement operations. We predict, however, that a *wh*-construction can be interpreted as a (free) relative *in all and only those cases* where a D is allowed to move alone, as the ones discussed above. This prediction appears to be exactly fulfilled in Italian, as shown by the contrast in (74).

(74) a. Detesto **quanto** sono **arroganti**.
I-hate how they-are arrogant
b. ??Detesto **quanto arroganti** sono.

In (74) the structure involving extraction of the bare (nonagreeing) quantifier stranding its associate is compatible, as predicted, with a verb selecting for a nominal complement. Crucially, this is not so when the quantifier moves together with its phrase, again as predicted. The prediction cannot be checked in French, because *combien* is never allowed in free relatives, for independent reasons:

(75) *Je déteste combien ils dépensent.
I hate how they spend

Things are more interesting in German, given that *was* can indeed head a free relative:

(76) Seine Mutter kauft, was auch immer Johann gerade liest.
his mother buys what also ever Johann currently reads
'His mother buys what Johann currently reads.'

The prediction of our approach is that the following contrast should hold:

(77) a. Seine Mutter kauft, **was** auch immer Johann **für Bücher** gerade liest.
his mother buys what also ever Johann for books currently reads
b. ?*Seine Mutter kauft, **was für Bücher** auch immer Johann gerade liest.[26]

[26] Thanks to Uli Sauerland for these data.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

A nonagreeing *was* can either move alone as a head stranding its nominal complement, as in (77a), or pied-pipe the entire phrase, as in (77b): our analysis correctly predicts that only in the former case can a free relative be derived.

In this section we argued that the ambiguity of sentences involving bare *wh*-words, which are compatible with the distribution of both interrogatives and free relatives, can be reduced to cases of conflicts between two Probes. Because they arise only when head movement is involved, this viewpoint provides a new understanding of the very nature of this typology of movement as opposed to phrasal movement.

5. A Final Prediction

This paper investigated the effects of the Probing Algorithm in two distinct areas, *wh*-movement and Principle C configurations. In this section we show that the proposed analyses for these two configurations, when combined, make an empirical prediction that allows us to further test our approach.

As extensively argued in section 3, the problem with a Principle C configuration like the one exemplified in (78) is that the illicit reading arises only if 'he' transmits its label, but, when this happens, the sentence gets the wrong label.

(78) *[$_{DP}$ *pro*$_i$ Ha votato per John$_i$].
has voted for John
'He$_i$ voted for John$_i$.'

However, imagine an abstract configuration in which the pronoun could transmit its label with no harm. In that configuration, given the approach developed in section 3, the Principle C effect should be obviated, because the pronoun might successfully probe the proper name for referential evaluation: this, we claimed, is what happens in the derivation of identity sentences like *He is John* and in clitic-doubling configurations. Are there other configurations of this kind? A natural candidate is free relatives (we thank Marcel den Dikken for pointing out the relevance of this case), given that we analyzed them as cases in which a *wh*-determiner does transmit its label. So, cases like (79) allow us to test if, as our approach predicts, Principle C effects are really obviated when the potentially offending pronoun can ''project.''

(79) [$_{DP}$ Chi ha votato per John] è uscito dalla stanza.
who has voted for John has gone-out from-the room
'Who voted for John left the room.'

Maybe the best way to check this prediction is to imagine a context in which only one person voted for John and this person happens to be John himself. It seems that, for (79) to be true in that context, John must have left the room. This means that 'who' does not need to be referentially disjoint from 'John', unlike what happens to the pronoun 'he' in a classical Principle C configuration like (78). So, the Principle C

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

effect is obviated in (79) and this is consistent with (and indirectly supports) our approach to Principle C and free relatives.[27]

6. General Conclusion

One persistent goal of syntactic research in the past 15 years has been the attempt to simplify phrase-structure-building rules. The aim was to preserve the empirical coverage of X-bar theory by dispensing with its rich apparatus. A first step has been Kayne's (1994) approach, in which much of X-bar theory was reduced to a single axiom (the Linear Correspondence Axiom). A further step was bare phrase structure theory, which, starting from Chomsky 1995 has undergone various reformulations until the Chomsky 2008 version, in which only two algorithms govern phrase structure building. In this paper, we critically reexamined these two algorithms and claimed that the algorithm that dictates that a lexical item transmits its label when it is merged with another object conforms to minimalist assumptions, but it sounds like a stipulation. A second algorithm proposed in Chomsky 2008 does not obey minimalist requirements because it is specifically restricted to movement configuration and, by doing so, it does not allow reduction of movement to (Internal) Merge. Therefore, we proposed a system involving only one algorithm (the Probing Algorithm), which holds equally for Internal and External Merge: in a nutshell, the Probe of a Merge operation always provides the label. In addition to capturing core cases of phrase structure building, the Probing Algorithm enabled us to shed light on two distinct areas—Principle C effects and the syntax of *wh*-constructions, which we analyzed as cases of conflict between two Probes. What unifies these two configurations is the fact that a lexical item (which should provide the label being endowed by definition with an EF acting as a Probe) is merged with a syntactic object that, being the Probe of the operation, should also become the label in compliance with the Probing Algorithm. In one case, this conflict produces two alternative outputs (a question or a free relative) that are both legible at the syntax–semantics interface. In Principle C configurations, one of the resulting outputs (the one where the lexical item "wins" and projects, so the pronoun and the referential expression can have the same semantic value) produces an illicit object. This way, Principle C effects are reduced to cases of mislabeling, with no need to postulate a specific condition to rule them out.

We hope to have shown that the simplification of the apparatus, in addition to complying with minimalist assumptions, can reinforce the deductive power of the theory. In particular, we carried out a simplification of phrase structure theory rules that allowed us to adopt the very same explanation for two apparently unrelated phenomena, such as constraints on the interpretation of pronouns and the categorial status of *wh*-constructions.

[27] Note that, as extensively discussed by Caponigro (2003), ordinary free relatives, as the one in (79), semantically behave like definite DPs, not like quantified DPs (one test used by Caponigro involves adverbs of quantification). So, 'who' is interpreted as a definite determiner, not like a quantifier in (79). This is relevant, because the absence of Principle C in (79) cannot be reduced in any obvious way to a seemingly related fact observed in quantificational contexts—namely, that a sentence like *Everyone voted for John* entails that John voted for himself.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

References

Abney, S. 1987. The English Noun Phrase in its sentential aspect. Ph.D. dissertation, MIT, Cambridge, MA.
Adger, D. 2003. *Core syntax*. Oxford: Oxford University Press.
Adger, D. & G. Ramchand. 2003. Predication and equation. *Linguistic Inquiry* 34:325–359.
Battye, A. 1989. Free relatives, pseudo-free relatives and the syntax of CP in Italian. *Rivista di Linguistica* 1:219–250.
Belletti, A. 1999. Italian/Romance clitics: Structure and derivation. In *Clitics in the languages of Europe*, ed. H. van Riemsdijk, 543–579. Berlin: Mouton de Gruyter.
Belletti, A. 2001. Inversion as focalization. In *Inversion in Romance and the theory of Universal Grammar*, ed. A. Hulk & J.-Y. Pollock, 60–90. Oxford: Oxford University Press.
Bianchi, V. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: Mouton de Gruyter.
Boeckx, C. 2003. *Islands and chains: Resumption as stranding*. Amsterdam: John Benjamins.
Boeckx, C. 2008. *Bare syntax*. Oxford: Oxford University Press.
Borsley, R., M. L. Rivero & J. Stephens. 1996. Long head movement in Breton. *The syntax of the Celtic languages: A comparative perspective*, ed. R. Borsley & I. Roberts, 53–74. Cambridge: Cambridge University Press.
Bury, D. 2003. Phrase structure and derived heads. Ph.D. dissertation, University College London.
Calabrese, A. 1992. Some remarks on focus and logical structures in Italian. In *Harvard Working Papers in Linguistics 1*, ed. S. Kuno & H. Thráinsson, 91–127. Cambridge, MA: Department of Linguistics, Harvard University.
Caponigro, I. 2003. Free not to ask: On the semantics of free relatives and *wh*-words cross-linguistically. Ph.D. dissertation, UCLA.
Cardinaletti, A. 1997. Subjects and clause structure. In *The new comparative syntax*, ed. L. Haegeman, 33–63. London: Longman.
Carnie, A. 1995. Nonverbal predication and head movement. Ph.D. dissertation, MIT, Cambridge, MA.
Cecchetto, C. 2000. Doubling structures and reconstruction. *Probus* 12:93–126.
Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, N. 1993. A Minimalist Program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, ed. K. Hale & S. J. Keyser, 1–52. Cambridge, MA: MIT Press.
Chomsky, N. 1995. Bare phrase structure. In *Government and binding theory and the Minimalist Program*, ed. G. Webelhuth, 383–439. Oxford: Blackwell.
Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step: Essays in minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels & J. Uriagereka, 89–155. Cambridge, MA: MIT Press.
Chomsky, N. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. M. Kenstowicz, 1–52. Cambridge, MA: MIT Press.
Chomsky, N. 2007. Approaching UG from below. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. U. Sauerland & H.-M. Gärtner, 1–29. Berlin: Mouton de Gruyter.
Chomsky, N. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. R. Freidin, C. Otero & M.-L. Zubzarreta, 133–166. Cambridge, MA: MIT Press.
Citko, B. 2006. Missing labels. Ms., University of Washington, Seattle.
Collins, C. 2002. Eliminating labels. In *Derivation and explanation in the Minimalist Program*, ed. S. Epstein & T. D. Seely, 42–61. Oxford: Blackwell.
den Dikken, M. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*. Cambridge, MA: MIT Press.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

Donati, C. 1999. Movimento e proiezione [Movement and projection]. *Rivista di Grammatica Generativa* 24:27–44.
Donati, C. 2006. On *wh*-head movement. In *Wh-movement: Moving on*, ed. L. Cheng & N. Corver, 21–46. Cambridge, MA: MIT Press.
Elbourne, P. 2008. The interpretation of pronouns. *Language and Linguistics Compass* 2:119–150.
Frege, G. 1892. Über Sinn und Bedeutung [On sense and reference]. *Zeitschrift für Philosophie und philosophische Kritik* 100:25–50.
Grodzinsky, Y. & T. Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24:69–102.
Grosu, J. 2002. Strange relatives at the interface of two millennia. *Glot International* 6:145–167.
Heim, I. 1998. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In *The interpretive tract* (MIT Working Papers in Linguistics 25), ed. U. Sauerland & O. Percus, 205–246. Cambridge, MA: MITWPL.
Heim, I. & A. Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.
Hornstein, N. 2006. Pronouns in a minimalist setting. Ms., University of Maryland, College Park.
Iatridou, S., E. Anagnostopoulou & R. Izvorski. 2001. Observations about the form and meaning of the perfect. In *Ken Hale: A life in language*, ed. M. Kenstowicz, 189–238. Cambridge, MA: MIT Press.
Kayne, R. 1989. Facets of Romance past participle agreement. In *Dialect variation and the theory of grammar*, ed. P. Benincà, 85–105. Dordrecht: Foris.
Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Kayne, R. 2005. Pronouns and their antecedents. In *Derivation and explanation in the Minimalist Program*, ed. S. Epstein & T. D. Seely, 133–166. Oxford: Blackwell.
É. Kiss, K. 1998. Identificational focus versus information focus. *Language* 74:245–273.
Koopman, H. & D. Sportiche. 1982. Variables and the Bijection Principle. In *Projet sur les langues kru: Premier rapport*, ed. J. Kaye, H. Koopman & D. Sportiche, 176–202. Montreal: Université du Québec.
Kratzer, A. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40:187–237.
Larson, R. 1998. Free relative clauses and missing Ps: Reply to Grosu. Ms., Stony Brook University, New York.
Lema, J. & M. L. Rivero. 1990. Long head movement: ECP vs. HMC. In *Proceedings of NELS 28*, ed. J. Carter, R. M. Déchaine & K. Kusumoto, 219–234. Amherst, MA: GLSA Publications.
Manzini, R. 1994. Locality, minimalism and parasitic gaps. *Linguistic Inquiry* 25:481–508.
Matushansky, O. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37:69–109.
Moro, A. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
Partee, B. 1987. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. J. Groenendijk, D. de Jongh & M. Stokhof, 115–143. Dordrecht: Foris.
Pesetsky, D. & E. Torrego. 2006. Probes, Goals, and syntactic categories. In *Proceedings of the 7th Tokyo Conference on Psycholinguistics*, ed. Y. Otsu, 25–60. Tokyo: Hituzi Syobo.
Pica, P. & W. Snyder. 1995. Weak crossover, scope, and agreement in a minimalist framework. In *Proceedings of the West Coast Conference on Formal Linguistics 13*, ed. R. Aranovich, W. Byrne, S. Preuss & M, Senturia, 334–349. Stanford, CA: CSLI Publications.
Reinhart, T. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.
Reuland, E. 2001. Primitives of binding. *Linguistic Inquiry* 32:439–492.
Roberts, I. 1994. Two types of head movement in Romance. In *Verb movement*, ed. D. Lightfoot & N. Hornstein, 207–242. Cambridge: Cambridge University Press.

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd

Rizzi, L. 1990. *Relativized minimality.* Cambridge, MA: MIT Press.
Rizzi, L. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. L. Haegeman, 281–337. Dordrecht: Kluwer.
Sauerland, U. 2007. Flat binding: Binding without sequences. In *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, ed. U. Sauerland & H.-M. Gartner, 197–254. Berlin: Mouton de Gruyter.
Schlenker, P. 2005. Non-redundancy: Towards a semantic reinterpretation of binding theory. *Natural Language Semantics* 13:1–92.
Seely, T. D. 2006. Merge, derivational c-command, and subcategorization in a label-free syntax. In *Minimalist essays*, ed. C. Boeckx, 182–217. Amsterdam: Benjamins.
Torrego, E. 1995. On the nature of clitic doubling. In *Evolution and revolution in linguistic theory*, ed. H. Campos & P. Kempchinsky, 251–275. Washington, DC: Georgetown University Press.
Travis, L. 1984. Parameters and effects of word order variation. Ph.D. dissertation, MIT, Cambridge, MA.
Uriagereka, J. 1995. Aspects of the syntax of clitic placement in western Romance. *Linguistic Inquiry* 26:79–123.
Uriagereka, J. 1999. Multiple spell-out. In *Working minimalism*, ed. S. Epstein & N. Hornstein, 251–282. Cambridge, MA: MIT Press.
Valois, D. 1991. The internal syntax of DP. Ph.D. dissertation, UCLA.

Carlo Cecchetto
University of Milan Bicocca
Piazza dell'Ateneo Nuovo 1
20126 Milan
Italy

carlo.cecchetto@unimib.it

Caterina Donati
Sapienza University of Rome
Via Salaria 113
00198 Rome
Italy

caterina.donati@uniroma1.it

© 2010 The Authors
Journal compilation © 2010 Blackwell Publishing Ltd