Carlo Cecchetto
University of Milano-Bicocca
carlo.cecchetto@unimib.it

Caterina Donati
University of Roma-La Sapienza
donatica@yahoo.it

RETHINKING HEAD MOVEMENT

1. THE PROBING ALGORITHM

A minimal notion of syntactic label, as (1), seems to be unavoidable:

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

In a cyclic view, whereas categories embedded within a syntactic object can be probed, only the information represented at the root (=label) is the trigger of further steps of the computation.
If labels are necessary, there must be an automatic device that determines the label of any occurrence of Merge. Elaborating on Chomsky (2005), we will assume that the algorithm in (2) is necessary and sufficient to yield labeled syntactic objects (cf. Boeckx 2008 and Citko 2007 for discussion):

(2) **Probing Algorithm:** The label of a syntactic object {α, β} is the feature(s) which act(s) as a Probe of the merging operation creating {α, β}

(2) says that Merge is always asymmetrically triggered and governed by the features of the items involved. The simple algorithm in (2) can capture the core cases traditionally described by X-bar theory, if the notion of Probe includes selection. For example, the following are the fundamental steps of the derivation of a simple sentence like 'the boy ate the cake':

i. The label of {ate, {the, cake} } is the categorial feature of V
(the transitive verb selects for a direct object)

ii. The label of {*v* , {ate, {the, cake} }}is the categorial feature of *v*
(*v* selects for the VP)

iii. The label of {{the, boy}, {*v* , {ate, {the, cake} }}} is the categorial feature of *v*
(when the external argument is merged in Spec,*v* the feature which triggers the merging operation is the categorial feature *v*, which requires an external argument being merged)

iv. The label of {T, {{the, boy}, {*v* , {ate, {the, cake} }}}} is the categorial feature of T
(T selects for *v*P)

v. The label of {{the, boy}, {T, {{the, boy}, {*v* , {ate, {the, cake} }}}}} is the categorial feature of T
(when the subject is internally merged in Spec,T the feature which triggers this operation is the categorial feature of T -- we assume that the phi-features of T can be checked in situ via Agree, so they do not, at least directly, trigger merge of the external argument).

(2) can apply in a larger set of cases if, following Chomsky (2005: 6, 10), we assume that every lexical item (LI) is endowed with a feature, call it edge feature (EF), which forces the LI to merge with other material. It is very natural to identify the edge feature of a word with its categorial feature (after all, words come in different varieties because this allow 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 (2) and its categorial feature can

provide the label. For example, each time a head (=LI) is merged with its complement, the head is bound to project.
An interesting consequence of this system is that, since the label is provided by the Probe, there can exist cases of labeling conflict if more than one Probe triggers the relevant merging operation.

2. LABELING CONFLICTS: FIRST MERGE

One case of labeling conflict is (3)

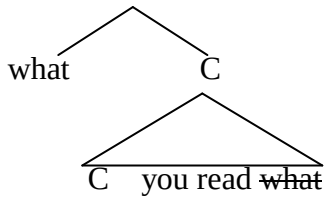
(3) {saw, John}

In (3), both 'saw' and 'John' are Probes, both being LIs endowed with an EF. But the theory based on the Probing Algorithm (2) can derive that the label of (3) will be the categorial feature of 'saw'. As a matter of fact, 'saw', in addition to its EF, also carries a selection feature, and this makes it a "double Probe". 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'. As we discuss in Cecchetto & Donati (2008), the core cases of X-bar theory can be captured by a system based on labeling algorithms like (2).

3. LABELING CONFLICTS: FREE RELATIVES

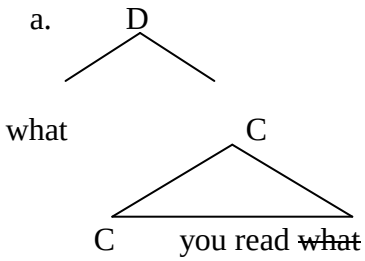
Another case of labeling conflict is (4):

(4) *what you read*

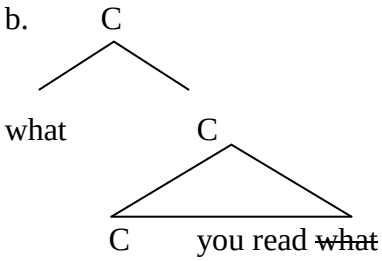


A labeling conflict arises in (4) since there are two Probes competing for labeling the newly created syntactic object. The LI 'what', as any LI, is provided with an EF, therefore it is a Probe and should provide the label. But the SO with which 'what' is merged, being the Probe of the operation, should also become the label. In fact, as predicted by our theory, two labeled outputs *are* available:

(5) a.



b.



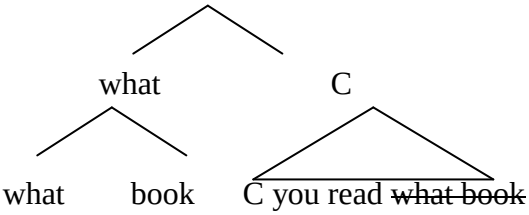
That two labels are available is shown by two possible distributions:

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

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

This kind of conflict never arises when a **phrase** is internally merged, as in (8).

(8) C *what book you read*

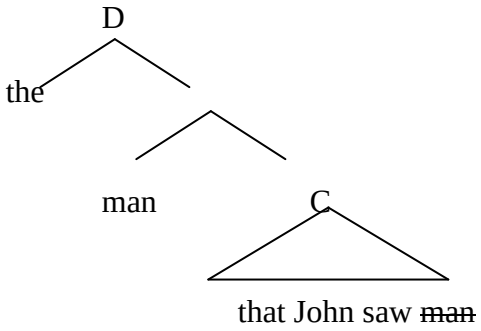


Here Merge holds between two SOs, and only one of them is a Probe: by (2), 'C', the probe of the merging, labels the entire construction.

4. LABELING CONFLICTS: RELATIVE CLAUSES

The raising analysis of relative clauses (Kayne 1994, Bianchi 1999), besides its many merits, has always been affected by a number of problems, which all relate to the relation between the external determiner and the nominal 'head' (Borsley 1996). What seems to happen is that the movement of the nominal "head" to the specifier of the clause turns the clause itself into a nominal object. This odd instance of 'projecting movement' can be interpreted and explained as a labeling conflict.

(9)



A labeling conflict arises in (9) when 'man' is internally merged with 'that John saw', since there are two Probes competing for labeling the newly created syntactic object. If the head of the relative is an LI, it is provided with an EF, therefore it is a Probe and should be the label. But the SO with which 'man' is merged, being the Probe of the operation, should also become the label. In fact, as predicted by our theory, two labeled outputs *should* be available: only one is however compatible with the selectional requirements of the external D head, namely the LI "man" must provide the label. Of course we must assume that the material that modifies the head noun in cases like (10) ("about Obama") must be late-merged, *after* the head noun has moved and "relabeled" the structure.

(10) The book about Obama that John read

This assumption makes so-called complements of nouns and adjuncts to the nouns more similar than it is usually thought. However, there is independent evidence that bridging the gap between these two categories may be necessary. For one thing, even so-called complements of nouns are never required for the structure to be acceptable, unlike the complements of transitive verbs. This is usually expressed by exempting the nouns from the theta criterion, but this is a tacit way to 'adjunctivize' the so-called complement of the noun. Secondly, the distribution of island effects supports this hypothesis. As a matter of fact, in the verbal domain there is an argument-adjunct asymmetry for extraction (only adjunct clauses are island). However, in the nominal domain no asymmetry arises, since both relative clauses and complement clauses of the noun are islands. The hypothesis that the material that modifies the head noun of the relative clause is late merged makes some predictions about reconstruction. Although data on reconstruction in relative clauses are notoriously shaky (cf. Bianchi, 1999, Cecchetto 2006, Munn 1994, Safir 1999), a sharp contrast arises between (11) and (12), which is exactly what the

late merge story predict:

(11) √ The professor of John$_i$'s that he$_i$ always praises
(12) * The professor$_i$ that he$_i$ always praises

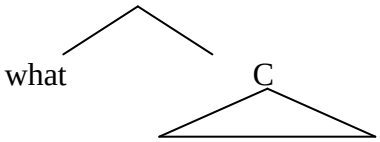
5. WHEN HEAD MOVEMENT SUCCEEDS AND WHEN IT FAILS

A rigorous minimalist perspective forces us to reconsider the dichotomy between phrasal and head movement. Whereas in a traditional perspective *wh*-movement in (13) would be classified as a case of phrase movement ('what' being embedded in an invisible DP), in a more minimalist perspective *what* must be considered an LI, unless there is strong empirical evidence that forces a departure from this simple assumption.

(13) What did you read?

This perspective opens an important question. We need to explain why head movement is not the *only* option in any environment, and in Wh-constructions in particular. This question is important since head movement is more minimal than phrasal movement in an intuitive sense, so economy considerations, if anything, should favor head movement over phrasal movement when a choice is given. For example, we need to explain why (14) is an impossible derivation in English and in many other languages.

(14) C/D ** what you read book*



C you read ~~what~~ book

In (14) the LI 'what' is extracted from the phrase it labels/heads, and internally merged to the root. In this configuration the algorithm (2) 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: plainly, it is ungrammatical.
An obvious explanation for the ungrammaticality of (14) is that the label D of the DP is closer than the head D (we use X-bar terminology for simplicity), and thus the label acts as an intervener, forcing phrasal movement instead of (subextracting) head movement.
However sub-extraction is not always forbidden, since it does take place in familiar cases of verb movement: V-to-T or T-to-C movements are extractions of a head out of its constituent, at least under standard accounts. From this point of view, (15) provides a very interesting contrast.

(15) a. *I wonder what you read [~~what~~ book]
b. You read [~~read~~ that book]

Movement in the verbal domain and movement in the (wh-) nominal domain obey different locality constraints:
a. V movement is more constrained in its path: Head Movement Constraint.
b. Wh-movement is more constrained in its source : Left Branch Constraint.
In order to derive these differences and make sense of the contrast in (15) we need to go back to Labels.

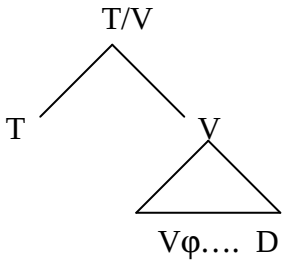
6. WHY HEAD MOVEMENT SUCCEEDS (WHEN IT SUCCEEDS)

Labels are features: see (1). In order for the algorithm (2) to be really effective we need a theory predicting which features exactly provide the label of a SO. For sake of simplicity, let us assume that features come into two categories:
- categorical features and the like
- phi features and the like.

We have illustrated the derivation of a simple sentence like 'the boy ate the cake' in section 1. The step of that derivation which is relevant for the present purposes is discussed below (we omit representing the *v*P level for simplicity):

V to T derivation:

(16) V to T movement

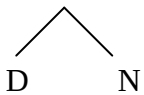


Here T probes for the phi-features of V. But the label of the verbal constituent is only the categorial feature V, since this feature is the Probe of the merging operation between V and the direct object. Therefore the label of the VP does *not* intervene between T and V and sub-extraction of V *is* allowed. What might intervene is any object along the path which is endowed with phi-features: this can derive the HMC.

7. LABELING INSIDE THE NOMINAL DOMAIN

In order to understand what goes wrong in (14), we need to discuss how labeling is determined in the nominal domain (an aspect we have neglected up to now).
Consider first what happens when a D and a N get merged and recall that we are assuming that nominal modifiers are late merged (see the discussion of relative clauses above).

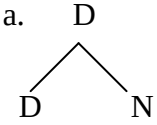
(17)



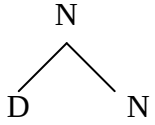
It seems that (17) is a labeling conflict configuration and that either LI can provide the label. However, D agrees with N, as is clearly shown in languages that have an overt morphology in the nominal domain. Suppose that an agreeing D is a Probe searching a Goal (the noun) to valuate its agreement features. This implies that an agreeing D necessarily provides the label to its constituent, since it is a double Probe. As a result, the label of a nominal constituent involving an agreeing D will be the categorial feature D, namely the resulting configuration will be (18a), not (18b).

(18)

a.

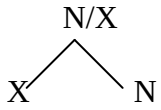


b.



In this story, what forces D to project is the existence of (possibly null) agreement. But consider a case in which N is merged with an LI (say, a modifier X) and N does not agree with X. In this context either X or N can transmit its label, since they are "single probe".

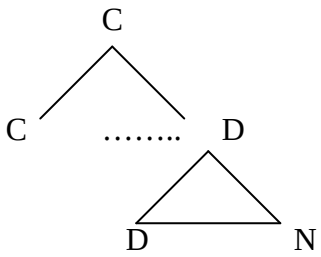
(19)



8. WHY HEAD MOVEMENT FAILS (WHEN IT FAILS)

We can finally consider the Wh-movement configuration underlying (14). If D provides the label, the following configuration obtains:

(20)



Here C probes for the categorial D/wh-feature of D. C cannot probe the LI D by-passing the label of the nominal constituent. Therefore the label *does* intervene and sub-extraction is *not* allowed. This explains the ungrammaticality of (14).
However, we have been assuming that D is forced to provide the label to the nominal constituent because it agrees with N. This means that, if N is merged with a determiner-like element X with which N does not agree, the label of the nominal constituent does not need to be the determiner-like element. Therefore, we expect that sub-extraction of the determiner-like element X be possible in this case.

9. WHY HEAD MOVEMENT SUCCEEDS (WHEN YOU MIGHT EXPECT IT TO FAIL)

That absence of agreement is the decisive factor that allows sub-extraction of the *wh* "determiner" is suggested by Italian data like in (21), which come close to being a perfect minimal pair. In Italian *quanto* ("how-many") agrees with a noun but does not agree with an adjective (this is morphologically visible). Crucially, *quanto* cannot be extracted and strand its associate when it agrees with it (when the associate is a noun, as in 21a). However, *quanto* can be extracted when it does not agree with its associate (when the associate is an adjective, as in 21b).

(21)

a. ***Quanti** hai letto **libri**?
How-PL have-2s read books
'How many books have you read'

b. **Quanto** hai detto che sono diventati **alti**?
How-SING (you) have said that they have become tall-PLUR

In fact, sub-extraction of a *wh* determiner out of a *wh* phrase *is* possible in several other (well known) cases. Crucially, when sub-extraction is possible (22b and 23c) the *wh* determiner does not agree with N, since the latter is found within a prepositional phrase.

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

(23) a. **Welche Bucher** hat Johann gelesen?
Which-PL books has Johann read
b. *__Welche__ hat Johann **Bucher** gelesen?
Which-PL has Johann books read
c. **Was** hat Johann **für Bucher** gelesen?
What-NEUT has Johann for books read
'Which books has Johann read?'

Cases like (22b) and (23c) have been analyzed in a number of different ways in the literature, all assuming phrasal wh-movement of a *wh* "determiner", given the standard X-bar theory restrictions. In our system, we can give the simplest possible analysis to these cases: they are instances of licit head movement of a wh-determiner out of a complex phrase. Ultimately, the reason why this is possible in (22b) and (23c), but impossible in cases like (14), is that D does not need to provide the label because it is not involved in an agreement relation.

10. PREDICTIONS FOR FREE RELATIVES

We predict 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 (24).

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

In (24) the structure involving sub-extraction of the bare (non agreeing) 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, since 'combien' is never allowed in free relatives, for independent reasons:

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

Things are more interesting in German, since 'was' can indeed head a free relative:

(26) 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 (thanks to Uli Sauerland for these data):

(27) a. Seine Mutter kauft, **was** auch immer Johann **für Bucher** gerade liest.

His mother buys what also-ever Johann for books currently reads

b. ?*Seine Mutter kauft, **was für Bucher** auch immer Johann gerade liest.
His mother buys what for books also-ever Johann currently reads

Our analysis correctly predicts that only when the non agreeing 'was' moves alone as a head stranding its nominal complement a relative clause can be derived.
Another direction where to test our predictions is among those languages, notably Slavic languages, where Wh-elements can quite freely strand their associate. Polish is such a language . As predicted by our analysis, a relative reading only holds when the wh-element moves alone as a head: (28) vs. (29). Polish examples are based on example (61) in Citko (2007).

(28) a. * Odwiedze **które miasta** ty tez odwiedzisz.
visit-1sg which towns you also visit-2sg
'I will visit which towns you will also visit.'
b. **Które miasta** ty tez odwiedzisz ?
which towns you also visit-2sg
'Which towns will you visit?'

(29) a. Odwiedze **które** ty tez odwiedzisz **miasta**
visit-1sg which you also visit-2sg towns
'I will visit the towns you will also visit'
b. **Które** ty tez odwiedzisz **miasta**?
which you also visit-2sg towns
'Which towns will you visit?'

11. CONCLUSIONS

In this talk we have proposed a minimal theory of labeling and have touched on some of its consequences in different domains, ranging from the derivation of core cases captured by X-bar theory, to free and ordinary relatives, and to the locality conditions determining when the head of a constituent can be extracted out of it.

REFERENCES

Bianchi, V. (1999). *Consequences of antisymmetry. Headed relative clauses*. Berlin: Mouton de Gruyter.
Boeckx, C. (2008), *Bare Syntax*. Oxford: Oxford University Press.
Borsley, R. (1996). Relative clauses and the theory of phrase structure. *Linguistic Inquiry* 28: 629-647.
Cecchetto, C. (2006), "Reconstruction in relative clauses and the copy theory of traces" in P. Pica and J. Rooryck (eds.) *Linguistic Variation Yearbook N. 5*: 73-103, John Benjamin, Co., Amsterdam
Cecchetto, C. and C. Donati (2008). On Labeling. *Studies in Linguistics* 1, *CISC Working Papers*, University of Siena, 16-38.
Chomsky, N. (2005). On Phases. Manuscript, MIT.
Citko, B. (2007). *Missing Labels*. Ms, University of Washington.
Kayne, R. (1994). *The antisymmetry of syntax*, Cambridge, MA: The MIT Press.
Munn, A. (1994). A Minimalist account of reconstruction asymmetries. In *Proceedings of NELS*, 24: 397–410
Safir, K. (1999). Vehicle change and reconstruction in A'-chains. *Linguistic Inquiry*30: 581-620.