

Cartography, criteria, and labeling.

III. Labeling and criteria.

1. The labeling algorithm in Chomsky (2012).

Chomsky (2012) “Problems of Projection” (to appear in *Lingua*): how do categories created by Merge get a label? (see also Chomsky 2008, Donati & Cecchetto 2011)

- (1) Labeling algorithm: The category created by Merge inherits the label of the closest head.
- (2) Nodes must have a label to be properly interpreted: the interpretive systems must know what kind of object they are interpreting.

NB: Labeling may be seen as a particular case of minimal search.

NB: (2) is different from previous labeling assumptions, in which labeling was considered a prerequisite for further applications of Merge. With the new view, Merge can also apply to unlabeled structures, and the necessity of labeling only arises at the interface.

There are three cases to consider:

- (3)i. H – H Merge
- ii. H – Phrase Merge
- iii. Phrase – Phrase Merge

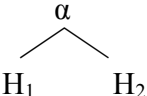
Chomsky (op. cit.) : labeling is straightforward in i and ii, but potentially problematic in iii.

8. A possible implementation

“Closeness” of a head may be computed in terms of c-command (NB: my definition; other definitions are imaginable):

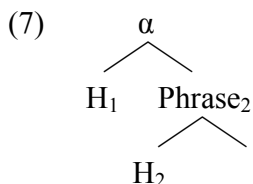
- (4) H_1 is the closest head to α iff
 - i. α contains H_1 , and
 - ii. there is no H_2 such that
 - i. α contains H_2 , and
 - ii. H_2 c-commands H_1 .

I. H – H Merge:

- (5) 

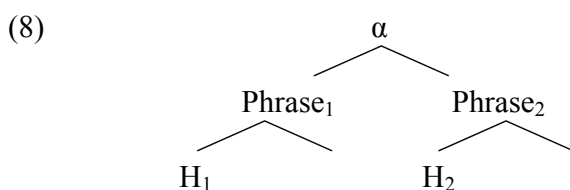
(6) Chomsky, op. cit.: if (external) H – H Merge only involves merger of a root not specified for category with a functional head expressing a categorial property (v, n, a, etc.) à la Marantz, the only category which can project is the one of the functional head because the root has no categorial label to project: [_n book + n]. So, “closest head” in (1) must be understood as “closest head with a label”.

II. H – Phrase Merge :



Here things are straightforward : H₁ is closer to α than H₂ (or any other lower head) hence α gets the label of H₁. So, for instance, in traditional X-bar notation, we have [_vV DP], [_TT VP], [_CC TP], etc.

III. Phrase – Phrase Merge:

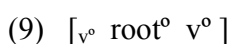


In case of Phrase Phrase Merge, the situation is ambiguous, as both H₁ nor H₂ qualify as the closest head to the new node created by Merge, so the algorithm gives inconsistent indications in (8), and α remains unlabeled. But this can only be a temporary state of affairs: under the assumption that nodes need labels for interpretation, α must receive a label before being passed on to the interpretive systems. So, something must happen here to make labeling possible.

9. A digressions: Head movement.

Suppose that Head movement (Head – Head internal Merge) exists, as distinct from phrasal movement. How can it be integrated in the labeling approach?

Let's first sharpen the assumptions on Head – Head external merge. I will assume that items drawn from the lexicon bear a feature (which I will continue to notate as “^o”, as in X-bar theory; but the current assumptions do not violate Inclusiveness). When the category undergoing merge with another category projects, this feature may disappear (in which case we get a phrasal projection) or remain (in which case we get a lexical projection, a category which still is a (complex) lexical item). So, external merge yields, for instance,



This is now a derived lexical item labeled v^o, a head (if we understand heads as elements bearing the “^o” feature). It can undergo Head – Phrase merge with a complement, e.g. to yield

(10) [[_{v°} root° v°] DP]

This category will now be labeled *v* under (1), yielding a verbal projection. The label of the new category can only be *v*, not *v°*, because it contains a phrase, and a lexical item cannot contain a phrase in the normal case.

Consider now head movement. For instance, T° (or, more plausibly, some lower inflectional head), attracts *v°* in (11)a, yielding (11)b:

(11)a T° [[_{v°} root° v°] DP]

b [_β [[_{v°} root° v°] T°] [<[_{v°} root° v°]> DP]

How is the complex head β , created by movement, labeled here? Perhaps, as both the simple head T and the complex head [root v] satisfy the definition of “closest head”, the system goes for the simple option, and labels the newly created head as T (alternatively, one could exploit the segment/category distinction à la Barriers, or the new head could have a complex label). The complex head thus created can further be head-moved to C, and then the new complex head will be labeled as C, etc., with the familiar properties of head movement (Mirror Principle, etc.).

10. Two possible solutions for unlabeled structures.

10.1. Movement

Phrase₁ moves further from [α Phrase₁ Phrase₂] in (8). At that point we get

(12) Phrase₁ ... [α <Phrase₁> Phrase₂]

“the intuitive idea is that the lower XP [Phrase₁] copy is invisible to LA [the labeling algorithm], since it is part of a discontinuous element, so therefore α will receive the label of YP [Phrase₂]” (Chomsky, op. cit., p. 22)

One possible implementation would be to understand the labeling algorithm (1) as stating “ α inherits the label of the closest head which has all of its occurrences internal to α ” (in fact the specification may be redundant if we properly understand the notions “internal”, “element”, and “occurrence of an element”); so H₁, head of Phrase₁, is both internal and external to α (it has internal and external occurrences), hence it is disregarded, and α receives the label of H₂, as desired.

So, for instance, the thematic subject of a transitive structure is merged with vP, which yields an [Phrase Phrase] structure:

(13) [α DP vP]

At this point the subject must vacate the position and raise, in order to allow proper labeling of the structure α as vP: DP (and D) are invisible (they are both internal and external to α), hence the closest head to the new node is *v*, unambiguously.

10.1.1. Digression 2: Labeling and locality.

The assumption that movement of one element in XP-YP makes it invisible for computation may seem ad hoc, and inconsistent with the copy theory of traces, in which traces have a full internal structure. But Chomsky manages to interestingly connect this assumption with the particular way of functioning of RM in structures like multiple questions (and, possibly, many other cases of “ordering preservations” with multiple movements, such as multiple scrambling in WF: Haegeman 1993).

- (14)a. Koj kakvo pravi? (Bulgarian: Rudin 1988,481-2)
who what does ‘Who is doing what?’
b. *Kakvo koj pravi?
what who does ‘What is who doing?’

- (15)a. Cine ce a văzut?
‘Who what saw?’
b * Ce cine a văzut?
‘What who saw ?’ (Rumanian: Soare 2009)

(16) Krapova & Cinque (2006)’s interpretation of Relativized Minimality:
in ... X ... Z ... Y ... ,
Z counts as an intervener between X and Y only if all the occurrences of Z intervene::

- (17)a. Cine ce a < cine > văzut < ce >?
‘Who what saw?’
b * Ce cine a < cine > văzut < ce >?
‘What who saw ?’

So, here too, movement makes a position “invisible” for the computation (of locality, in this case).

10.2. The creation of a criterial configuration.

At some point movement must stop. This happens when it reaches a criterial position (Rizzi 1996, 1997). Criteria are defined as configurations in which Spec and head share a major interpretable feature, e.g. Q in questions:

- (18) [_α [which_Q book] [did_Q you read]]

Chomsky’s idea is that the Criterial configuration permits labeling of the whole structure: Both heads in XP-YP share the most prominent feature relevant for labeling, Q in this case, so search of both XP and YP provides a non-ambiguous indication, Q, which can label the whole structure:

- (19) [Q [which_Q book] [did_Q you read]]

So, what characterizes a criterial configuration is that it receives the label of the criterial feature (and we get, in traditional X-bar notation, QP, TopP, FocP, RelP, etc.)

In short, in this system, the problem raised by XP-YP for labeling can be resolved either by moving one of the two elements, in which case the label of the unmoved element projects, or by creating a criterial configuration, in which case the shared label projects.

11. The “Halting Problem” for wh movement.

Wh-movement proceeds stepwise. But in certain environments it cannot stop, while in other environments it can (and in fact must) stop:

- (20)a You think [C [Bill read [which_Q book]]]
 b * You think [_α [which_Q book] [C [Bill read ____]]]
 c [_β [which_Q book] [Q [you think [_α ____ C [Bill read]]]]]

The system captures the fact that the wh phrase cannot stop in the embedded C in (20)b: for selectional reasons C cannot be Q (*think* does not select an indirect question), hence after-wh movement there is no way to label the XP-YP structure α , and the structure is rejected as unlabeled (see (25) below).

In (20)c the XP-YP structure β can be labeled as Q (it’s a criterial configuration, so both XP and YP are headed by Q), and this is fine. And α can now be labeled as C (or whatever more refined category we have here, presumably Decl (or Declarative Force, etc.) because the wh phrase has moved out, and there is only a trace (an occurrence of the wh phrase) in the Spec of C, which can be disregarded for labeling, according to the approach in 10.1.

12. Deriving Criterial Freezing from Labeling.

Consider now the complement of a verb selecting Q:

- (21)a John wonders [Q [Bill read [which_Q book]]]
 b John wonders [_α [which_Q book] [Q [Bill read ____]]]
 c * [_β [which_Q book] [Q [John wonders [_α ____ C [Bill read]]]]]

The wh phrase moves to the embedded C-system where a criterial configuration is created, and α can be properly labeled as Q.

Why is (21)c excluded? This is a violation of Criterial Freezing (Rizzi 2006, Rizzi & Shlonsky 2007): movement cannot undo a criterial configuration.

(22) Criterial Freezing: A phrase meeting a Criterion is frozen in place

Can Criterial Freezing be related to Chomsky’s labeling algorithm? As the algorithm accounts in a natural manner for the cases in which movement must continue, the possibility is worth exploring that labeling may also account for the cases in which movement **must stop**, thus providing a comprehensive solution for the “halting problem”. The point is not addressed in Chomsky (2012), but there is a natural possibility to consider.

Movement can only involve minimal or maximal projections: minimal projections, heads, in head movement (if indeed this option is allowed by UG) and maximal projections in phrasal movement. I.e.

given the traditional X-bar schema, X and XP can be moved, but the non-maximal, non-minimal projection X' is inert for movement.

(23) Movement can only involve minimal and maximal projections.

Minimal projections are heads, LI's extracted from the lexicon and complex heads formed by head movement (I will not try to extend the system to this case here).

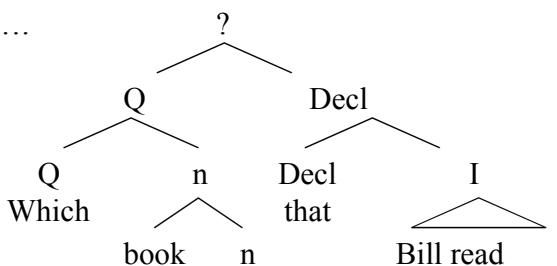
Under bare phrase structure, being a “**maximal projection**” is not a rigid inherent property of a node, as XP nodes in standard X-bar notation, but is a **dynamic** notion in the following obvious sense:

(24) α is a maximal projection if the node immediately dominating it does not have the same label.

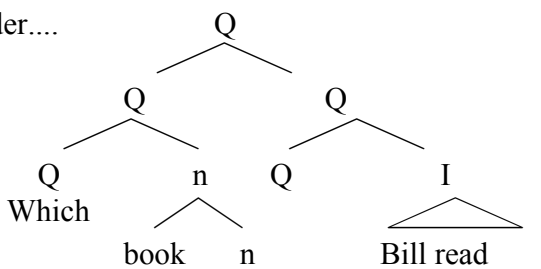
Then in the criterial configuration [XP YP], if the label is inherited from both XP and YP, neither is maximal, in the sense just defined: only the whole category [XP YP] is maximal; so, further movement of either XP or YP alone is excluded by the ban on movement of a non-maximal (non-minimal) projection (23) (see (26)).

So, both the necessary continuation of movement in intermediate C-systems ((20)b), and the halting in the criterial configuration ((21)c) can be made to follow from Chomsky's approach to labeling, under natural auxiliary assumptions. Here are configurations requiring continuation of movement (25) and determining freezing (26):

(25) think....



(26) wonder....



Notice that this approach accounts for simple cases of violation of Criterial freezing like (23)c, in which the same feature Q in *which book* is attracted twice (and for which alternative approaches in terms of “inactivation” could be considered), but it also accounts for the complex cases discussed in Rizzi (2006), in which two distinct criterial features are involved, i.e., Q on the determiner and Foc on the lexical restriction of a nominal expression:

(27) [quale_Q LIBRO_{Foc}]
‘which BOOK’

The lexical restriction can be focalized in situ in the embedded C system where the phrase satisfies the Q criterion, as in (28)a, but it cannot be Focus moved to the main C system, as in (28)b, as this would undo the criterial configuration. The whole indirect question can be marginally pied-piped through focus movement to the main C-system, as in (28)c:

(28)a Non sono riuscito a capire [[quale_Q LIBRO_{Foc}] [Q avesse letto]] ... (non quale articolo)
 ‘I havent managed to understand which BOOK he had read, not which article’

b * [quale_Q LIBRO_{Foc}] Foc non sono riuscito a capire [____ [Q avesse letto]] (non quale articolo)
 ‘Which BOOK I haven’t managed to understand he had read, not which article’

c ? [[quale_Q LIBRO_{Foc}] Q [avesse letto]] Foc [non sono riuscito a capire ____] (non quale articolo)
 ‘Which BOOK he had read, I didn’t manage to understand, not which article.’

In (28)b [quale_Q LIBRO_{Foc}] is extracted from the criterial configuration

(29) [_α [quale_Q LIBRO_{Foc}] [Q avesse letto]]

But, given the labeling algorithm, α is now labeled Q (we may assume that labeling takes place as soon as the conditions are met, as per Pesetsky’s Earliness Principle), hence [quale_Q LIBRO_{Foc}] is non maximal, and therefore it cannot be extracted from (29). In (28)c, the whole criterial configuration (29) is pied-piped, so the maximal phrase labeled Q is moved, and this is fine.

13. Digression: Successive cyclicity, “dangling preposition”, floating quantifiers.

Postal gave the following argument against Chomsky’s (1973) theory of successive cyclic wh movement: if wh movement goes through the intermediate C-system, why can’t it strand a preposition there? (the “dangling preposition” argument)

- (30)a Who do you think [_α t C [we should talk [to t]]]?
 b * Who do you think [_α [to t] C [we should talk t]]?
 c To whom do you think [_α t C [we should talk t]]

The impossibility of (30)b can now be made to follow from labeling: *to* is visible here because it’s entirely internal to the embedded clause, it competes with C for labeling (neither one c-commands the other, so they both qualify as “closest” to α), hence the embedded clause α cannot be labeled, and the structure is ill-formed. When the preposition is not stranded in the embedded C-system, as in (30)a or c, no problem arises, as the trace is not visible and C (presumably, Decl Force) wins the competition for labeling.

McCloskey (2000) argues that in certain varieties of Irish English a floating quantifier can be stranded by a wh element, apparently also in the intermediate C-system, thus providing straightforward evidence for successive cyclic wh movement:

- (31)a What *all* did he say (that) he wanted?
 b What did he say *all* (that) he wanted?
 c What did he say (that) he wanted *all*? (West Ulster English, McCloskey 2000)

This seems to be in direct contradiction with (our interpretation of) Postal's argument. If *all* is stranded in Spec C in (31)b, the structure should incur the same labeling problem as (30)b, under Sportiche's (1988) analysis of Q-float.

But perhaps floated quantifiers never remain in the position in which they are stranded, and move further to an adverbial position in the low IP space. So *all* could move to such a position in (31)b, thus vacating Spec C entirely, hence no labeling problem would arise.

The same conclusion holds for the classical case of Q-float from subjects:

- (32) Les amis on tous (bien) mange
'The friends have all (well) eaten'

Tous could not be stranded in Spec v in (32) because otherwise a competition would arise for labeling the vP, which would give rise to ill-formedness:

- (33) [tous t] [v VP]

So *tous* presumably moves out to an adverbial position vacating the Spec v position completely, and permitting proper labeling of vP. This is independently shown by the fact that *tous* is higher than the manner adverbial *bien* in (32), which suggests that *tous* cannot remain in Spec vP, and must move further, as the labeling approach would predict.

14. The status of subjects.

The canonical subject position is a fundamental halting point of movement, the final landing site of core cases of A-movement (unaccusatives, passive, raising, and in fact in any sentence under the vP-internal subject hypothesis). What does this imply for the labeling approach under consideration?

- (34) There is a Subject Criterion.

Otherwise the subject position would not be a possible halting point for phrasal movement: in order to label [Phrase₁ Phrase₂] in which Phrase₁ is the subject, we must be in a criterial configuration, otherwise labeling would fail.

A subject criterion is made independently plausible by certain interpretive properties that go with the subject position (Rizzi 2006). The subject is the argument "about which" the event is presented. So, an active and a passive sentence (also in "all new" contexts) differ in "aboutness": the "hitting event" is presented as being about the truck in (35)a, and about the bus in (35)b:

- (35)a Un camion ha tamponato un autobus
'A truck hit a bus'
b Un autobus è stato tamponato da un camion
'A bus was hit by a truck'

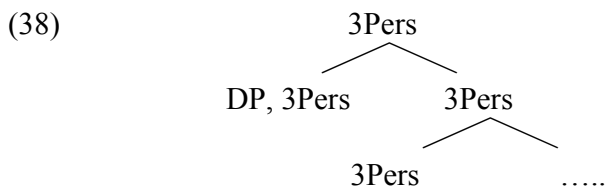
This has clear consequences for the overall interpretation and discourse articulation: for instance, in a Null Subject Language, *pro* in the following sentence in discourse can only pick up the “aboutness” subject (as observed in Calabrese (1986)):

- (36) Poi, *pro* è ripartito
 ‘Then, *pro* left’ (*pro* = truck after (35)a; *pro* = bus after (35)b)

In previous work (Rizzi 2006, Rizzi & Shlonsky 2007, building from Cardinaletti 2004), the criterial head (Subj) and the attracting feature (+N) were not fully identified, as in other cases (a Q head and a Q feature; a Foc head and a Foc feature, etc.). But perhaps this can be done, in the spirit of the overall criterial approach. Let us tentatively propose that the relevant attracting feature is Person, so SubjP is in fact PersonP (see also Shlonsky, in progress). Then, a Person head in the high functional structure of the clause attracts a DP endowed with person features, thus creating a Criterial configuration which allows movement to stop in that position. The “aboutness Subject – Predicate” interpretive routine is then triggered:

- (37) [Un camion 3pers, sing] [Person [ha [t tamponato un autobus]]]
 [“aboutness” subject] [predicate]]

Movement can stop here because the whole clause can be labeled as “person”, the criterial feature in common between XP and YP. So we get a subtree like the following:



In fact, Subject movement must stop in (38): neither XP (DP, 3Pers) nor YP (3Pers....) are maximal, in the intended sense, so the subject cannot move further, under (23) and (24). This gives a strong version of the “Fixed Subject Constraint” (Bresnan 1977). That – trace effects are thus derived from Criterial Freezing and now, ultimately, from labeling:

- (39) * Who do you think [that [t Person [will come]]]

Who satisfies the Subject (Person) Criterion in the embedded clause, and then it is frozen there because neither XP nor YP are maximal in the criterial configuration thus created:

- (40) ... that [_{3Pers} [who 3Pers] [3Person [will [t come t]]]]

Languages then may use “strategies of Subject extraction” (Rizzi & Shlonsky 2007) to circumvent the freezing effect and allow wh-extraction of a subject. For instance, Italian (and other Null Subject Languages) permit a “skipping strategy” consisting of the use of expletive *pro* to formally satisfy the Subject Criterion, which allows the thematic subject to skip the freezing position, so that it remains available for further movement (much as in the original ECP-based analysis in Rizzi (1982)).

- (41) Chi credi [che [_{3Pers} [*pro* 3*Pers*] [3*Pers* [t verrà t]]]]
 ‘Who do you think that pro will come?’

15. Not all agreement positions are criterial.

Not all positions of Spec of a head with matching phi features are stopping positions. For instance, movement can (and must) continue from the subject position of a small clause under a raising predicate:

- (42)a [Gli amici] sembrano [____ Num, Gen simpatici]
 ‘The friends seem nice NumPlur, GenMasc’
 b * Sembrano [[gli amici] Num, Gen simpatici]
 ‘Seem the friends nice’

Or, with compound tenses, the object clitic triggers participial agreement (Kayne 1989), but then it must continue to move to the clitic position:

- (43)a Gianni li ha [____ Num, Gen incontrati ____]
 ‘Gianni them has met NumPlur, GenMasc’
 b * Gianni ha [li Num, Gen incontrati ____]
 ‘Gianni has them met NumPlur, GenMasc’

In criterial terms this amounts to saying that number and gender features are not criterial in the clausal structure, hence movement doesn’t stop there, and in fact it must continue. Consider the following possible implementation:

Number and gender features do not define an independent head in the functional structure of the clause (possibly, a consequence of the fact that they are “uninterpretable”, in the sense of Chomsky 1995), but are merely attached to other interpretable heads to “register” the application of movement and other structural relations. So, what we have in fact is

- (44)a ... [_{α} [gli amici] [_{a} NumPlur, GendMasc simpatici]]
 b ... [_{β} [li] [_{Asp} NumPlur, GendMasc incontrati]]

where ” a ” in (44)a is the functional head defining the category Adjective, and Asp in (44)b defines the aspectual interpretation expressed by the past participial construction.

Here the only possibility to label α and β arises if the Spec moves further, so that the label of YP can project.

The difference between Person and Number – Gender would then be that the former defines an autonomous head in the clausal spine, while the latter do not, a distinction possibly connected to the interpretable – uninterpretable divide. So we may assume that

- (45) Features that project are categorial features, which can define an independent head (Q, Top, Foc, the features of the Cinque hierarchy, T, M, Asp, Voice..., n, v, a,..., but also Pers,...)

Possible independent morphological manifestations of Person as an autonomous head could be the system of subject clitics in the Northern Italian dialects:

- (46) Le ragazze le son venute (Brandi & Cordin 1989, Poletto 2000, Manzini & Savoia
 ‘The girls Scl have+3pl come’

The element in Spec cannot stop there in (44)a-b because the features entering into the agreement process are not categorial here, hence they cannot define a criterial configuration which would permit the labeling of the phrases. Therefore the nominal expression in Spec must move further.

16. Halting, complements, and specifiers.

Can a phrase ever halt and be spelled out in a non-criterial position? On the basis of the labeling approach it can surface in a complement position (say, an object position), because there H – Phrase Merge, or X-YP Merge, straightforwardly permits labeling of the new category as XP.

Objects (and complements in general) can remain *in situ* (as far as labeling is concerned), or move if other properties require movement. Specifiers, on the other hand, are halting positions, or position from which further movement is compulsory, depending on whether they give rise to criterial configurations or not.

A potential problem for this view of the halting problem is raised by the subject position of small clauses: in the complement of some verbs, the equivalent of (44)a is a possible spell-out configuration, and further movement is not required (but it is possible, as shown in (48)a-b):

- (47) Considero [α [i tuoi amici] [simpatici]]
 ‘I consider your friends nice’
- (48)a I tuoi amici sono considerati [β ____ [simpatici]]
 ‘Your friends are considered nice’
- b Gli amici che considero [β ____ [simpatici]]
 ‘The friends that I consider nice’
- c Li considero [β ____ [simpatici]]
 ‘I them consider nice’

One possibility is to assume that $\alpha = \beta$, and modify the system so as to permit Spec positions which are consistent both with halting and continuation of movement. Another possibility is to assume that $\alpha \neq \beta$, and continue to assume a rigid complementarity between “halting” Specs and Specs requiring further movement. Then, the subject of the small clause would be criterial in (47), but not in (48).

A possible indication in favor of the second solution is that bare plurals in Italian are possible in object position, but not as subjects of small clauses (Belletti 1988); but the bare plural can apparently be moved and become the head of a relative:

- (49) Gianni frequenta amici
 ‘Gianni sees friends’

(50)a * Gianni considera [[amici] [simpatici]]
'Gianni considers friends nice'

b Gianni frequenta amici [che considera [____ [simpatici]]]
'Gianni sees friends that he considers nice'

So, it may be the case that the small clause optionally allows a criterial position in its Spec, whose interpretive import is incompatible with bare plurals. So, α in (47) can be labeled, (50)a is excluded by the interpretive incompatibility, while (48) and (50)b do not involve the criterial position, hence no semantic incompatibility in (50)b, but movement must proceed to a higher criterial destination.

(51) If something along these lines is tenable, we can stick to a simple picture on the "halting problem", as far as labeling is concerned:

- complements can stay where they are, or move;
- specifiers can (and must) stay if they are in a criterial configuration, otherwise they must move.