Unprobed movement and labeling: tree trimming as a key to understand successive cyclicity

Carlo Cecchetto, Caterina Donati

Unprobed movement and labeling: tree trimming as a key to understand successive cyclicity[1]

Carlo Cecchetto & Caterina Donati

1. Introduction: the status of unprobed movement

The status of unmotivated movement in syntax is a delicate issue. On the one hand, very general methodological considerations invite us to adopt a restrictive theory in which movement should always have a clear motivation, since in the absence of constraints no theory can make predictions that can be supported or falsified. On the other hand, assuming that movement must be motivated presupposes that it is inherently costly and this assumption is questionably. For example, Chomsky (2008: 137) claims that Merge must be costless, otherwise production of isolated lexical items should be preferred to complex linguistic structures. If movement is defined as Internal Merge, namely an ordinary case of Merge in which the merged category is taken from the derivation itself rather than from the lexicon, movement should also be costless, and in principle should be able to apply freely. This claim is repeated by Chomsky (2019). If movement is really free, then why does it quite systematically require a trigger?

Although ontological discussions on the status of basic syntactic operations are interesting and even necessary, we believe that they can be productive only if they are implemented in a concrete framework that ultimately allows the consequences of the opposing views to be evaluated empirically.

This is what we will try to do in this paper. We will investigate the notion of unmotivated movement, which, following current minimalist terminology, we will call 'unprobed movement'. We will show that unprobed movement is allowed but severely constrained. Importantly, this approach to unprobed movement will offer a new perspective on two classical topics in syntactic theory: successive cyclic movement and, to the extent that we can deal with this issue in this paper, head movement.

This paper is organized as follows. Section 2 introduces the framework, namely a theory of labels which underlines their role as *syntactic* devices, and an algorithm for determining labels at each step of the derivation when such a step is probed. Consistently with this system centered on syntactic labels, unprobed movement, which by definition yields an unlabeled syntactic object, is predicted to be very restricted and possess very specific properties. In section 3 a simplification of the Phase Impenetrability Condition (PIC) where only the label remains visible once a phase is completed is defended. This simplified definition can dispense with the spurious notion of "edge" and is consistent with the central role of labels in defining connectedness phenomena. The intermediate steps of successive cyclic movements are reinterpreted as instances of unprobed movement, which allow the

[1] We thank an anonymous reviewer for very useful comments. And…happy birthday Léa!

relevant phrase to get "tree-trimmed" out of the phase and go back into the working space, ready to be probed by some further probe. Section 4 highlights a consequence of the proposed analysis for the operation by which V exits the *v*P phase and gets to be probed by T. A general analysis of head movement in terms of unprobed movement is sketched. Section 5 concludes the paper.

2. Labels and how they are determined

In this section we summarize our assumptions about labeling. Although labels are not always defined, we believe that an explicit definition is useful and we adopt the following definition from Cecchetto & Donati's (2015a):

(1) *Labels*. When two objects α and β are merged, a subset of the features of either α or β become the label of the syntactic object {α, β}. A label:
(i) can trigger further computation
(ii) is visible from outside the syntactic object {α, β}.

Under (1), a label plays a crucial role *inside* the derivation. A label defines which of the features that are merged during a given derivation "stay alive", namely remain active (in Minimalist terms: can act as Probes) and visible (in Minimalist terms: are Goals that can be targeted by a Probe). We call the definition of label in (1) 'internal', since it stresses the role of labels inside the derivation. An important property of the definition in (1) is that labels are not assigned any special role at the syntax/semantics interface. This differs from the view of labels according to which they provide instructions to the semantic component on how to interpret syntactic objects (see recently Chomsky 2019). We call this the 'external' definition of label, since it stresses the role of labels at the interface, mostly at the C-I interface. In fact, core syntax would be label-free according to the supporters of the external definition (these include Collins 2002; Seely 2006; Narita 2014; Chomsky 20015; 2019, cf. Stockwell 2016 for a state-of-the-art article on this debate).

We believe that the main reason to reject an external definition of labels is that syntactic categories with the same label can correspond to different semantic types (a DP can be an individual, a generalized quantifier or a property; vice versa, a proposition can be a CP or a DP; a predicate can be a VP, an NP, an AP, and so on and so forth). This makes the role of labels as instructions for the semantic interface dubious at best. On the contrary, there are good reasons to assume an internal definition: first labels do play an important role in determining derivations in dimensions such as locality and intervention (see Cecchetto & Donati 2015a for a more extensive discussion and Frampton & Gutman 2002). Second, as we will explain in Section 3.1, the external definition of label inevitably implies a non-negligible computational cost.

Recently Chomsky (2019) has proposed a new conceptual argument against labeling in syntax. He argues that the proper definition of Merge should obey a 'third principle' requirement (Restrict Resources) which dictates that the number of objects in the Working Space should be augmented by at

most one by any application of Merge[2]. The Restrict Resources principle would be violated if when α and β are merged, the output is the object {α, β} plus the label of this object. The reason is that before Merge there are two objects in the Working Space (α and β) while after Merge there are four, namely α, β ,{α, β}) and the label of {α, β}. We are sensitive to the conceptual argument but we believe that it does not apply to the definition of label in (1). What (1) says is that each application of Merge is a case of 'feature stripping', namely (1) states that only a subset of the features of one of the two merged objects remains active at the next level up. We believe that, as a process of reduction of available information, labeling obeys the Restrict Resources principle and can even be seen as a response to it.

No matter whether labels are external or internal to syntax, any theory of labeling must be integrated with the formulation of the proper algorithm(s) that determines a label for each syntactic object that is being formed.

There are two main configurations to consider, namely Merge of an atom of the syntactic derivation (call it Lexical Item, LI) with an already formed syntactic object (call it a phrase), as in (2), and Merge of two phrases, as in (3):[3]

(2) {LI, XP}
(3) {XP, YP}

For the case in (2), although the technical implementations may vary, there is a general consensus that the LI provides the label. One simple way to capture this is Chomsky's (2013) Minimal Search principle. When LI and XP are inspected for label determination, LI is immediately detectable, being simpler.

The case in (3) is more complex because Minimal Search principle cannot return a label, as both XP and YP are equally complex in the relevant sense. In the labeling theory literature, there are two families of approaches to handle the problem raised by the configuration in (3). The first approach builds on the fact that Merge is not the only operation that takes place in core syntax. The other operation, namely Agree, is inherently asymmetric, since a Probe looks for a Goal that can value some of the unvalued features of the Probe (a paradigmatic example is T which probes the phi-features of the external argument). Therefore, the proposal is that when a Probe and a Goal are merged, it is the Probe that provides the label. If selection, qua asymmetric operation, is reduced to a special case of Probing, this approach can also derive the fact that the selector provides the label. Proponents of this approach include Adger 2003; Pesetsky & Torrego 2006; Cecchetto & Donati 2015 a.o.). Of course, when Merge is dissociated from probing, which is possible if Merge can apply freely, the object created by Merge is predicted to remain without a label.

---

[2] Chomsky refers in his argumentation to MERGE (in capital letter), which is an operation mapping a workspace into another (slightly bigger) workspace. MERGE replaces 'traditional' Merge in the framework sketched by Chomsky (2019). However, the conceptual argument made in the text is general enough to apply to different conceptualizations of Merge.

[3] A third configuration corresponds to step one of any derivation, namely Merge of two lexical items. For this, see Chomsky (2013) and Cecchetto and Donati (2015: 2.9.1).

The second approach to deal with the labeling problem associated to the configuration in (3) is the one developed by Chomsky (2013, 2015). It can be illustrated with the case of the external argument (EA) of a transitive verb. In this configuration no structural asymmetry can lead Minimal Search into finding a label. As Chomsky assumes that an object can remain without a label until it accesses the C-I interface, he considers unproblematic that the {EA, *v*} object remains without a label when it is created. However, the label-less configuration {EA, *v*} is destroyed when the external argument is internally merged with T (incidentally, this requires the non-innocent assumption that the copy of EA is not visible for labeling). As for the {EA, T} object created by Internal Merge, it will be labeled by a mechanism of feature sharing: T and EA share phi-features, therefore the {EA, T} object will have a label <ϕ, ϕ>. This approach can extend to all cases in which the moved category and the target share features (these include Criterial features like Q, Topic, Focus etc., as developed by Rizzi 2015; 2016).

In this paper, we will adopt the first approach, the one that builds on the asymmetric character of Probing, because we believe that assuming a syntactic object can freely remain label-less as long as it will be eventually destroyed later on in the derivation introduces a significant computational burden and is in sharp contrast with Third Factor principles that aim at minimizing computational costs (we go back to this in Section 3.1).

3. Unprobed movement and successive cyclic movement

The internal and external definitions of label briefly discussed above make opposite predictions concerning the type of syntactic objects that can remain label-less (if any). If labels are needed internally for the computation to proceed (as we will assume here), the expectation is that every syntactic object will need a label, except the root, where the computation ends. On the other hand, under the external definition, labels are needed at the interface, so a label is certainly needed at the end of the cycle (the root), when the structure is sent to interpretation. Let's us keep this in mind while discussing successive cyclic movement.

In the generative tradition, there is a conceptual motivation for why long movement must proceed by intermediate steps: the computational load is alleviated, as desirable, if, once the domain that contains a sub-link is completed, its internal details can be "forgotten".

To be concrete, consider (4), in which the *wh*-phrase must move from the embedded VP to the matrix CP. Assuming no successive cyclic movement, all the information about the entire structure must be retained until the very end of the derivation. However, if movement proceeds as shown in (4'), each time the *wh*-phrase escapes from a given domain by moving to its edge, the details concerning that domain exit the working space and stop being accessible (they can be "forgotten"). Supposedly, this alleviates the computation burden. In (4') we hypothesize, as is common practice, that successive cyclic movement of the *wh*-phrase targets the positions Spec,*v*P and Spec,CP on the path from its base position to its final landing site. Notice however that the conceptual argument is independent from the specifics

of the way successive cyclic movement is implemented within a more fine-grained analysis of the CP system[4].

(4) Which student did Mary say that John thinks that Bill met?

(4') Which student did Mary [$_{v}$ [~~which student~~] say [$_{C}$ [~~which student~~] that John [$_{v}$ [~~which student~~] thinks [$_{C}$ [~~which student~~] that Bill [$_{v}$ [~~which student~~] met [~~which student~~]]]]]]?

3.1. The Phase Impenetrability Condition: the problem with intermediate steps

In current minimalist frameworks, the principle that forces long movement to take place in a successive cyclic fashion is the Phase Impenetrability Condition in (5).

(5) *Phase Impenetrability Condition*
The complement of a strong phase α is not accessible to operations at the level of the next highest strong phase β, but only the head and the edge of α are (Chomsky 2001).

The definition of Phase Impenetrability Condition must be complemented by the list of domains that act as phases. These include *v*Ps and CPs (the status of DPs and PPs is more controversial, see Matushansky 2006; Abels 2012 a.o. for a discussion). Phase Impenetrability Condition explains why movement must be successive cyclic by postulating a "time bomb mechanism", where the "explosion" moment is when a given phase stops being accessible. For example, in (6) the complement of *v*P (namely VP) stops being accessible by the moment C enters the derivation. Therefore, 'which student' must leave the VP before that moment, or remain otherwise buried within the phase. This is schematically shown in (6'), where we indicate in grey the structure that becomes inaccessible.

(6) Which student did you meet?

(6') a. [$_{T}$ you [$_{T}$ T [$_{v}$ [$_{V}$ meet which student]]]]
b. [$_{C}$ did [$_{T}$ you T [$_{v}$ which student [$_{v}$ [$_{V}$ meet [~~which student~~]]]]]]
c. [$_{C}$ [which student] [$_{C}$ did [$_{T}$ you T [$_{v}$ [~~which student~~ [$_{v}$ [$_{V}$ meet [~~which student~~]]]]]]]

The same logic applies to the case in (7), but this time there are three copies of 'which student' in intermediate positions, since three phases have been crossed (embedded *v*P, embedded CP and matrix *v*P).

[4] The second type of motivation for successive cyclic movement is empirical: in various languages there are overt reflexes of the intermediate traces of the moved *wh*-category in the relevant intermediate positions. We go back to some of these data later in the text (in particular Section 3.3).

(7) Which student do you think that John met?

(7') [$_C$ [which student] do [$_T$ you [$_v$ [~~which student~~] [$_V$ think [$_C$ [~~which student~~] [$_C$ that [$_T$ John [$_v$ [~~which student~~][$_V$ met [~~which student~~]]]]]]]]]

While the definition of Phase Impenetrability Condition in (5) successfully explains why (long) movement is successive cyclic, it suffers from some drawbacks, all of them widely discussed in the literature (cf. Lahne 2008 a.o.). The first concerns the identification of the categories that behave like phases. The initial motivation was that CPs and *v*Ps are phases because they have a special status at the interface(s), for example they are 'propositional' (sentences are CPs and *v*Ps are the domains where theta roles are assigned). However, under the definition of Phase Impenetrability Condition in (5) what exits the derivation and gets to the interface is not the phase (the object with a special status), but its complement (hence, by definition, an object without any special status at the interface). A second reason makes the definition of Phase Impenetrability Condition in (5) cumbersome: the category that is allowed to escape from the phase is spurious. In fact, what is taken to survive inaccessibility under (5) is two quite different objects: in traditional terms, they can be identified as the specifier and the head of the projection that counts as a phase. While there is an intuitive sense in which the Spec is at the edge, this is much less obvious for the head. In general, what is hidden under the concept of 'accessibility' in (5) is really a negative definition: what remains accessible is whatever is not a complement. Negative definitions are problematic because they reveal the incapacity to define a theoretical object by identifying (some of) its core features. There is a reason for this loose definition of 'edge': it is precisely designed to allow head movement and A'-movement out of phase (we go back to this in Section 4).

However, the biggest problem with the definition of Phase Impenetrability Condition in (5) is that it does not explain what motivates the movement to the edge of the phase. Certainly, this movement is ultimately beneficial, because it allows a *wh*-phrase (or any other category that undergoes long distance movement) to escape the phase that becomes inaccessible. However, *unless unprobed movement is allowed,* a motivation (a Probe) is necessary for each movement step. For certain languages and for certain types of movement a motivation for each step may be available. This holds for languages like Chamorro, or Irish or Asante Twi, where intermediate steps have a morpho-syntactic reflex, and it is thus reasonable to assume that they are morpho-phonologically triggered, hence probed (cf. Henry 1995 in Belfast English; Chung 1998 and Lahne 2008 for Chamorro; and McCloskey 1979; 2000 for Irish; Korsah & Murphy 2019 for Asante Twi). But in many (most?) other languages, embedded CPs and *v*Ps do not show any morpho-phonological feature associated to intermediate steps of successive cyclic movement.

In order to deal with this problem, Chomsky (2001) assumes that intermediate Cs (and *v*s) can be endowed with an abstract (EPP) feature which probes the category that undergoes successive-cyclic movement. However, this move is far from being satisfactory. First, this EPP feature suspiciously looks like a *deus ex machina*. Furthermore, as stressed by Bošković (2008), we are forced to assume that an intermediate C has the relevant EPP feature if and only if C happens to be on the path of a category that undergoes long movement. In other cases, the intermediate C cannot have the EPP feature, otherwise sentences like (8) should be grammatical, with 'which student' moved to the intermediate C endowed with a probing EPP feature.

(8) *Who thinks which student that John met?

Therefore, the EPP feature solution is totally ad hoc.

A way to handle the problem of the lack of motivation for successive cyclic movement is simply dropping the requirement that Internal Merge is motivated and that every syntactic object must be labeled (Chomsky 2013: 44). As we say in Section 2, the idea defended by Chomsky (2013, 2015) is that structures of the form {XP, YP} that are not associated with any type of feature sharing (including the intermediate step of successive-cyclic-movement) have no label and are tolerated, as long as they are destroyed by the end of the derivation. For example, the structure in (9) is tolerated if by the end of the derivation (the step illustrated in 9') 'which student' and the dedicated head in the COMP area where it eventually sits share a Q feature and hence make up a labeled syntactic object.

(9) [[$_D$ which student] [$_C$ that John met]]
(9') [$_Q$ [$_D$ Which student] [$_C$ do you think [[$_D$ which student] [$_C$ that John met]]]]?

This analysis requires adopting the external definition of label, discussed above, according to which labels are needed only at the interface. But we believe that this analysis suffers from a major drawback. For the sake of the argument, assume that that the structure in (9) has been created and can be temporarily tolerated. Still, the derivation must proceed in order to reach the resolutive step in (9'). The intended derivation is one in which the label-less object is selected by 'think' but it is not clear how selection can take place in absence of a label. Selection requires a label, but a label is not (yet) present when it is needed. It seems that the no-labeling approach to successive-cyclic-movement requires that any sort of merging and probing operations be allowed freely, at least until the interface is accessed. This amounts to assuming that multiple alternative derivations are allowed, including those that will be filtered out at the interface. Allowing 'ungrammatical' sub-derivations to run freely, although temporarily, seems to us a clear case of computational *in*efficiency.

This problem is recognized by proponents of the idea that label-less objects are not problematic for the derivation. For example, Ott (2010) stress that if access to the C-I interface is done cyclically the deviance will detected quickly (at the phase level). However, at most this minimizes the problem, it does not solve it. Even more importantly, there is another issue pointed out by Stockwell (2016:139-140):

> "if Merge is entirely free, there is nothing that could force a phase head to be merged. This means that a derivation could continue through an infinite number of derivational steps without ever being transferred to the interfaces. Rather than ensuring that a fatal problem is spotted reasonably early, a problem may never be spotted at all….
> Second, even merging a phase head would never necessarily resolve the issue. The specifier of the phase head is not transferred to the interfaces in cyclic spell out; only its complement is. It seems that on the free-merge approach Internal Merge can operate as 'interface bounceback'. The edge feature allows any SO with an unvalued feature that is rejected by the interfaces to be merged as a specifier of the phase head. This would allow any deviant or crashing element to escape from the phase, allowing the derivation to continue through infinitely many phases without the deviance being noticed. Far from identifying deviance early on, phases could in fact prolong the agony."

Given this consideration and in agreement with so-called crash-proof syntax (Frampton & Gutmann 2002), we maintain that labels play a pivotal during the derivation, also by determining selection.

However, we do think that the intuition that the intermediate steps of successive cyclic movement involve the creation of an unlabeled node is correct. In the next section we propose a novel implementation of this idea that is fully compatible with the internal definition of label and does not suffer from the drawbacks we just mentioned.

3.3 The tree trimming mechanism

As anticipated, in order to fix the problem of the absence of a motivation for intermediate steps of successive cyclic movement, we take the bull by the horns and assume that the intermediate step of successive cyclic movement in languages like English, where it has no overt morphological reflex, is what it looks like: an instance of unprobed merge. Before getting there, we also propose the version of Phase Impenetrability Condition in (10), which is much simpler than the canonical cumbersome definition in (9) and is consistent with the central role that labels have in syntax in the system we are defending here.

(10) *Simplified Phase Impenetrability Condition*.
When a phase is concluded, only its label remains accessible to further syntactic computation.

The definition in (10) has several advantages:
(i) it does not contain any reference to the spurious notion of edge;
(ii) it does not postulate a 'time bomb' mechanism, since a phase stops being accessible right when it is concluded and not at a designated point later on in the derivation; and
(iii) what has a special status at the interface is the phase, not its complement.

Apparently, there is a huge price to pay for this simplicity, namely (10) does not seem to allow any movement out of the phase. This is clearly wrong, as movement out of a phase does occur.
In this section we deal with this problem as far as Successive Cyclic Movement is concerned. In Section 4 we shall briefly address head movement, which in some cases also targets a position outside the phase.

We can illustrate our approach to successive cyclic movement based on (10) by considering once again the derivation of the sentence (7) above, repeated as (11). For simplicity we illustrate only the step involving the intermediate step to the embedded Spec,CP. However, the same logic applies to the intermediate step involving the edge of *v*P.

(11) Which student do you think that John met?

Step 1: The *wh*-phrase is internally merged to the embedded C as a case of unprobed merge.

(11') [[$_D$ Which student] [$_C$ John met ~~with student~~]]

If the label is determined by the probe, as we assume, no Probe means no label, so an unlabeled syntactic object is created. This is temporarily OK because derivationally speaking the structure to which 'which student' merges is a *root* clause. By definition, a root clause does not *need* to enter further derivation, hence does not need a label (under the internal definition we are defending here). However, a problem arises when the main verb 'think' enters the derivation because, under the definition of label in (1), an unlabeled syntactic object is not visible and cannot be selected. Let us see what happens in the next steps.

Step 2: Under the simplified Phase Impenetrability Condition in (10), only the label C remains in the computation, *not the syntactic object under the label* (here we indicate an object that becomes inaccessible by using a grey color). Remember that, under the definition (1), the label C is a subset of the features of the embedded C. This must include the categorial feature C.

(11'') [[$_D$ which student] [$_C$ John met ~~which student~~]]

At this point, 'which student' and the label C do not form a constituent because the syntactic object to which 'which student' has attached in the former cycle is not accessible any more (it has abandoned the derivation). The label C and 'which student' are hence both available for further derivation, but as disconnected elements. It is convenient to give a name to the mechanism according to which two objects that were previously connected become disconnected, because one of them becomes inaccessible under the simplified PIC in (10). Let us call this mechanism 'tree trimming'. We assume that simplified PIC (and tree trimming as a result) are a consequence of the Restrict Resources principle, which minimizes the number of objects allowed in the Working Space.

After tree trimming, the label C and 'which student' still belong to the working space of the derivation. Under the reasonable assumption that by the end of the computation any material in the working space needs to be merged into the structure, both these elements need to be integrated into the new structure under construction.

Step 3: The label of the (now inaccessible) embedded clause, namely C, is probed by the selecting verb 'think':

(11''') [$_V$ think [$_C$ _ _ _]]

The structure gets a label from the probe (matrix V), under the assumption that selection is a form of probing.

Step 4: 'which student' remains in the working space until a suitable probe kicks in, with which it can merge: this is the matrix C. When matrix C probes 'which student', it labels the resulting syntactic object, which ends up having label C:

(11’’’’) [$_C$ [$_D$ Which student] [$_C$ do you [$_V$ think [$_C$ _ _ _]]]?

The advantages of the tree trimming approach to successive cyclic movement should be self-evident. First, tree trimming is a direct consequence of the minimally simple definition of the Phase Impenetrability Condition in (10). Second, the tree trimming approach correctly restricts “intermediate steps” to phase edges only, without stipulating any edge or special EPP feature attached to them[5]. In any other location, an unprobed merging operation would yield an unlabeled syntactic object that would not be “trimmed” and would therefore stop the derivation. Third, it avoids the computational costs discussed above for other implementations of the idea that successive-cyclic-movement creates an unlabeled object[6].

An interesting question arises concerning those morpho-syntactic reflexes of intermediate steps attested in some languages, which we mentioned above. By the way tree trimming works, there is no proper multi-links chain in so-called successive cyclic movement, which would be composed by the copy in the landing site, the one(s) in the intermediate position(s) and the one in the base position. Only two copies are assumed to be represented through the derivation, the head and the foot, while the copy merged unprobed at the edge of the phase is immediately removed through tree trimming. But if there is no copy where the morpho-syntactic reflex of so-called successive cyclic movement arises, how can this reflex be explained?

Let us address this question by focusing on one easy case among those cited above, namely auxiliary inversion in Belfast English. The basic facts are the following: as in Standard English, auxiliary inversion is observed in direct questions (cf. 12a). Unlike Standard English, Belfast English requires auxiliary inversion in embedded questions (cf.12b) and even in the intermediate site of successive-cyclic movement (cf. 13). Crucially, no such inversion is possible in positions lower than the gap (14).

(12) a. What have you done?
b. She asked [who had I seen].
(13) Who did John hope [~~who~~ would he see]?
(14) *Who do you think [~~who~~ did John tell ~~who~~ [did Mary leave]?
(Belfast English, Henry, 1995: 106)

In principle, this can be interpreted as evidence that it is the general mechanism of successive-cyclic-movement which triggers auxiliary inversion and that the latter is the mechanical reflex of the former. However, things are more complicated than this. In fact, other instances of successive-cyclic-movement do *not* trigger auxiliary inversion. This is shown by the ungrammaticality of (15) in which there is long

[5] Under this aspect the tree trimming approach we are defending, although it assumes no probe for intermediate steps, qualifies the path of successive cyclic movement as punctuated, not uniform, under the terminology proposed by Abels (2012).

[6] The account proposed in the text is similar but not identical to Cecchetto & Donati’s (2015). An important consequence of the tree trimming approach is that it can explain for free the occurrence of strong island effects. We refer to Cecchetto & Donati (2015: 4.4 - 4.7) for an elaboration of this important point.

distance movement of the relative pronoun 'who', which must have passed through the CP of the complement clause of the verb 'claim'. Still, embedded auxiliary inversion is not allowed.

(15) *This is the man who John claimed ~~who~~ did I see.
(Belfast English, Henry, 1995: 106)

As rightly stressed by Abels (2012), this means that auxiliary inversion is not a simple reflex of successive cyclic movement. Depending on the nature of the moving category (relative versus interrogative) and of the embedded complementizer, inversion takes place or not. This strongly suggests that the movement step that is responsible for auxiliary inversion in Belfast English is feature triggered, namely probed, hence challenging the approaches that take successive cyclic movement to involve an unprobed step (including the tree trimming analysis). However, notice that, under the tree trimming analysis, if the embedded complementizer is a suitable probe, nothing prevents it from attracting a given category, before this category further moves unprobed to the edge of this CP, ultimately tree trimming. We hypothesize that this is what happens in (13), as schematically illustrated below:

(13') Who did John hope [~~who~~ [C ~~who~~ would he see who]]

The question remains of what is the feature which is probed by the embedded C in (13), and whether this kind of explanation can be extended to all the other cases where intermediate steps of successive cyclic movement have morpho-syntactic effects that have been discussed in the literature. Since this paper is dedicated to the general mechanism of successive cyclicity, a fully worked out case by case identification of the features that can probe the step preceding unprobed movement is outside our scope. However, a point is important: this feature cannot be the same as the one that is responsible for the final probing, which will allow the moved and tree trimmed category to re-enter the derivation. The reason is that we make the standard assumption that a specific feature can be checked only once in the derivation. To be concrete: in the case of (13), if the interrogative wh-feature of 'who' were checked in the embedded CP, 'who' could not re-enter the derivation because it could not be probed by matrix CP. This is of course in line with the fact that in (13) the embedded C is not interrogative: if it probed an interrogative *wh*-feature, there would a mismatch between its interrogative nature in syntax and its non-interrogative interpretation. We stress this because it shows how the tree trimming mechanism can capture core cases of what Rizzi (2010) calls Criterial Freezing effects, namely the fact that once a phrase enters into a criterial configuration, it normally becomes unavailable to further movement. Given tree trimming, this translates into the fact that once a category α has entered into a checking configuration for a feature β, it can still move unprobed to the edge of a phase and go back to the working space. But that same feature β cannot be probed again, allowing α to re-enter the derivation (a different feature in α can be probed if available, though).

As for (15), we hypothesize that the embedded C does not probe the relative 'who', much like in Standard English, and therefore unprobed movement is not preceded by a probed movement step[7].

4. Perspectives on Head Movement

Still**,** it is uncontroversial that a head can move out of a phase, as in V to T in Romance, or Verb Second configurations and so-called "long head movement", instantiated for example by participle fronting in Bulgarian (cf. Lema & Rivero 1990 a.o.). This is to be expected if there are no rule-specific constraints. Still, prima facie at least, the Simplified Phase Impenetrability Condition in (10) does not allow head movement out of a phase, since the notion of edge of the phase (which, by stipulation, includes the head of the phase) is dispensed with. The system outlined in the paper allows a way by which a phase head can escape the phase, though. We illustrate this with movement of V out of the phase *v*P: V can get out of *v*P via unprobed movement to its edge (cf. 15).

(15) [V [$_v$ *v* [V ~~V~~]]]

unprobed movement

At this point however a difference emerges with respect to the successive cyclic movement case discussed in section 3: V is a lexical item and under minimal search it can provide a label. Therefore, V can label the structure formed after it has moved unprobed to the edge of the *v*P. Informally, we can say that *v*P is turned into a category with label V[8].

(16) [$_V$ V [$_v$ *v* [V ~~V~~]]]

It is only at this point that, under the simplified Phase Impenetrability Condition, the object under the label *v*, i.e. the phase, becomes inaccessible.

---

[7] A rather different type of effects that are normally taken as empirical motivation for successive-cyclic-movement concern interpretation, cf. Barss (1986). For example, assuming that an anaphor in a *picture*-NP is subject to Principle A (something that, incidentally, is fairly controversial, cf. Reinhart & Reuland, 1993), the fact that the anaphor can be bound by 'John' in (i) is taken as evidence that a copy of the *picture*-NP sits at the edge of the most embedded CP, where 'himself' is locally bound by 'John'.
(i) Which picture of himself$_i$ does Mary believe that John$_i$ thinks that she is going to buy?
Assuming tree trimming, we would be forced to say that the operation that moves the anaphor close to its antecedent in (i) is s a topicalization-like movement. We assume that this movement is probed and leaves the copy that counts for binding purposes and that it is followed by unprobed movement and tree trimming. Cf. Abels 2012 for a survey of the many interpretation reflexes of successive movement that have been discussed in the literature.

[8] Notice that we are implicitly endorsing here an approach *à la* Matushansky (2006), who propose that canonical head movement configurations involve "normal" syntactic movement to a specifier position, followed by a morphological process, which she calls *m*-merger, by which the moved verb gets fused with its affixes. Cf. also Harizanov & Gribanova's (2019) for a different implementation of the same idea.

(17) [V V [v _ _ _ _ _ _]]

This means that there is a difference between what happens after unprobed movement of a head and after unprobed movement of a phrase. In the latter case, tree trimming takes place and both the label of the phase and the moved phrase re-enter the working space, where they remain until a suitable probe attracts each of them back into the derivation. However, we do not assume that tree trimming takes place after a head, say V, moves unprobed to the edge of a phase, say *v*P. The reason is that V, as a lexical item, can provide the label to the resulting structure, even if the movement by itself is unprobed. Therefore, for tree trimming to take place at that level would involve destroying the node labeled V created on the top of *v*P. We assume that this is blocked by the No-Tampering Condition, which says that you should not destroy a structure that has been created. Of course, in case of unprobed movement of a phrase, the No-Tampering Condition is irrelevant, since there is no label to destroy, given that unprobed movement has created no label to begin with.

There is some empirical evidence suggesting that V does not re-enter the working space. Consider the ungrammaticality of sentences like (18), traditionally explained as a violation of the Head Movement Constraint (Travis 1984). If V could go back to the working space after its unprobed movement to *v*, the sentence in (18) would be predicted to be OK, contrary to facts. This is so because V could be probed by C and merged there directly.

(18)* [C What [C buy] [T John [T will] [v ~~buy~~] bread]] ?

On the other hand, if the head V remains in the derivation after its unprobed movement to *v*P, obvious locality considerations make 'will' closer to C than 'buy'.

Notice an important prediction of the analysis of V-to-T that we are sketching: if the label of *v*P becomes V, T should select V, not *v*. We think that this prediction is borne out (and is at the basis of the notion of extended projection introduced by Grimshaw 1991). We offer the following piece of empirical evidence here: T is sensitive to the lexical semantics of the verb, which is expressed at the level V , not to argument structure which is expressed at the level *v*. We illustrate this with Italian. Consider the imperative. The inflection of imperative (expressed in C but present at T as well) selects specific lexical verbs, irrespective of their thematic properties. This is shown by the following examples. Unaccusative, unergative, and transitive verbs can be imperative (cf. 19 a–c) but only if the imperative is compatible with their lexical semantics.

(19)	a.	Mangia la pasta! Eat-IMP the pasta	(transitive)
	b.	Respira! Breath-IMP	(unergative)
	c.	Parti! Leave-IMP	(unaccusative)
	d.	*Conosci l'inglese! Know English!	(transitive)

e. *Suda! (unergative)
Sweat-IMP
f. *Svieni! (unaccusative)
Faint-IMP

The pattern in (19) can be described by saying that (imperative) T must see the label V while it is not sensitive to the label *v*.

5. Conclusion

This paper addressed the issue of the status of unprobed merge in syntax, which we have claimed to be possible but severely restricted since it cannot create a labeled syntactic object. While in general we have been assuming that the presence of an unlabeled object, which is syntactically invisible, blocks the derivation, this problematic configuration can be tolerated at phase edges because it is destroyed when the phase exits the derivation, according to the mechanism which we called tree trimming. This is the direct consequence of the simplified version of the Phase Impenetrability Condition that we have proposed, which assumes that, when a phase is completed, only its label survives. As a result, any material that has merged unprobed to the phase itself gets "detached" (tree trimmed) from it when this becomes inaccessible and this material is put back into the working space. Successive cyclic movement can be modeled this way, as we have discussed in details in this article.

We believe that the approach outlined in this article, which opens the way to considering successive cyclic movement and head movement as a new natural class, that of unprobed movements, can not only provide a new explanation to old problems, but indirectly makes a strong case in favor of a label-driven syntax in response to recent attempts (cf. Chomsky 2013, 2019) to expunge labels from syntax and relegate them to the interface.

References

Abels, Klaus. 2012. *Phases. An essay on cyclicity in syntax*. Berlin: De Gruyter.

Adger, David. 2003. *Core Syntax. A Minimalist Approach*. Oxford: Oxford University Press.

Barss, Andrew. 1986. *Chains and anaphoric dependence*. Cambridge, MA: MIT dissertation.

Blümel, Andreas. 2012. Successive cyclic movement as recursive symmetry-breaking. In Nathan Arnett & Ryan Bennett (eds.). *Proceedings of WCCFL 30.*

Bošković, Zeljko. 2008. On Successive Cyclic Movement and the Freezing Effect of Feature Checking. In Jutta Hartmann, Veronika Hegedus & Henk van Riemsdijk (eds.) *Sounds of Silence: Empty Elements in Syntax and Phonology*, 195-233, Amsterdam: Elsevier.

Cecchetto, Carlo & Caterina Donati. 2015a. *(Re)labeling*. Cambridge MA: MIT Press.

Cecchetto, Carlo & Caterina Donati. 2015b. Please do not Move That remnant too much! In Gunther Grewendorf (ed.), *Remnant Movement*, 179-200. Berlin: De Gruyter.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52, Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero & Maria Luisa Zubizarreta (eds.). *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press .

Chomsky, Noam. 2013. Problems of Projection. *Lingua*, 130. 33–49.

Chomsky, Noam. 2015. Problems of projection. Extensions. In Elisa Di Domenico, Cornelia Hamnan & Simona Matteini (eds.) *Structures, strategies and beyond. Studies in honor of Adriana Belletti*, 1-16. Amsterdam: Benjamins.

Chomsky, Noam 2019. The UCLA Lectures. lingbuzz/005485

Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*, Chicago: University of Chicago Press.

Collins, Chris. 2002. Eliminating labels. In Samuel D. Epstein & T. Daniel Seely (eds.). *Derivation and explanation in the minimalist program*, 45-61. Oxford: Blackwell.

Frampton, John. & Sam Gutmann. 2002. Crash-proof syntax. In Samuel D. Epstein & Daniel Seely (eds.), *Explanation and derivation in the minimalist program*, 90–105. Oxford: Blackwell.

Grimshaw, Joan. 1991. Extended Projections. Unpublished Manuscript, Brandeis University. [Reprinted as Grimshaw 2005. In *Words and Structure*. CSLI Lecture Notes Number 151, Stanford, CA: CSLI Publications]

Harizanov, Boris & Vera Gribanova. 2019. Whither head movement? *Natural Language and Linguist Theory* 37. 461–522.

Henry, Alison (1995). *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: Oxford University Press.

Korsah, Sampson & Andrew Murphy, A. 2019. Tonal reflexes of movement in Asante Twi. *Natural Language & Linguistic Theory* 24. 57–111.

Lahne, Antje. 2008. Where There is Fire There is Smoke. Local Modelling of Successive cyclic Movement. Universität Leipzig Dissertation.

Lema, José & María-Luisa Rivero. 1990. Long head movement: ECP vs. HMC. *Proceedings of Nels* 20. 333–347.
McCloskey, James. 1979. *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish*. Dordrecht: Reidel.
McCloskey, James. 2000. Quantifier Float and Wh-Movement in an Irish English. *Linguistic Inquiry*, 31. 57–84.
Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37. 69–109.
Narita, Hiroki. 2014. *Endocentric structuring of projection-free syntax.* Amsterdam: Benjamins.
Ott, Dennis. 2010. Grammaticality, interfaces, and UG. In Michael T. Putnam (ed.) *Exploring Crash-Proof Grammars* 89-104. Amsterdam: Benjamins.
Pesetsky, David & Esther Torrego. 2006. Probes, Goals and Syntactic Categories. In Yukio Otsu, (ed.) *Proceedings of the Seventh Tokyo Conference on Psycholinguistics*, 25-60. Tokyo: Hituzi Syobo.
Reinhart, Tanya & Eric Reuland, 1993. Reflexivity. *Linguistic Inquiry*. 24. 657-720.
Rizzi, Luigi. 2010. On some properties of criterial freezing. In Phoevos Panagiotidis (ed.) *The complementizer phase: Subjects and operators*, 17–32. Oxford: Oxford University Press.
Rizzi, Luigi. 2015. Cartography, criteria, and labeling. In Ur Shlonsky (ed.), *Beyond the functional sequence*, 314–338. New York: Oxford University Press.
Rizzi, Luigi. 2016. Labeling, maximality and the head – phrase distinction. *The Linguistic Review* 33. 103-127.
Seely, T. Daniel. 2006. Merge, derivational c-command and subcategorization. In Cédric Boeckx (ed.) *Mimimalist Essays*, 182-217. Amsterdam: Benjamins.
Stockwell, Richard. 2016. Labelling in Syntax. *Cambridge Occasional Papers in Linguistics* 9. 133-155.
Travis, Lisa. 1984. *Parameters and effects of word order variation*. Cambridge, MA: MIT dissertation.