Missing Labels

Barbara Citko (University of Washington)

Abstract: This paper examines the nature of labels created by *External* and *Internal Merge* operations. The standard minimalist assumption is that in *External Merge* structures, one of the merged elements projects as the label, and in *Internal Merge* structures, the Probe does. However, these two options do exhaust the logical range of possibilities. For *External Merge*, the options are (*i*) Project α, (*ii*) Project β, (*iii*) *Project Both* α and β, and (*iv*) *Project Neither* α nor β. For *Internal Merge*, the options are: (*i*) *Project Probe*, (*ii*) *Project Goal*, (*iii*) *Project Both Probe and Goal*, and (iv) *Project Neither Probe nor Goal*. The proposal I defend in this paper, both on theoretical and empirical grounds, is that all the logical possibilities are in fact attested. I focus on the following 'missing labels': *Project Both* in *External Merge*, and *Project Goal* and *Project Both* in *Internal Merge* structures.

1.   Issue

My main goal in this paper is to examine the labels created by Merge, the basic structure building operation within the Minimalist Program. Following Chomsky (2001), I assume there are two kinds of Merge. One is familiar from earlier minimalist literature; it takes two distinct objects and combines them into one bigger one. This type of Merge, illustrated schematically in (1), is referred to as *External Merge*.[1]

1)   *External Merge* of α and β
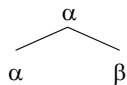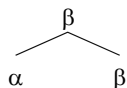
K

α β
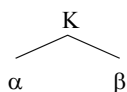
The second type of Merge, called *Internal Merge*, is like *External Merge* in that it also combines two objects. The only difference is that one of the objects is a subpart of the other, as shown in (2). *Internal Merge* is thus a minimalist way to capture syntactic displacement.

---

[1] Chomsky in Beyond Explanatory Adequacy also makes a distinction between *Set Merge*, illustrated in (1) above, and *Pair Merge*, which he argues is involved in the derivation of adjuncts. I will come back to this distinction in Section 3.1.1.

2)   *Internal Merge β and L*

```
              K
            /   \
          β       L
          ▲     /   \
              α       β̶
          |_____|
```
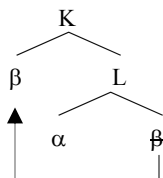
The question that I would like to focus on here is what determines the nature of K, which is the label of the object created by both *Internal* and *External Merge*. The standard minimalist assumption is that in *External Merge* structures either α or β projects as the label, whereas in *Internal Merge* structures the target, or to use the more current terminology, the Probe does.

These two options, i.e. Project either α or β in *External Merge* and Project Probe in *Internal Merge*, do not exhaust the logical range of possibilities. For *External Merge*, the options are (*i*) Project α, (*ii*) Project β, (*iii*) *Project Both* α and β, and (*iv*) Project neither α nor β. These options are illustrated in (3a-d), respectively.

3)   a.   Project α

```
              α
            /   \
          α       β
```

        b.   Project β

```
              β
            /   \
          α       β
```

        c.   *Project Both* α and β

```
            {α,β}
            /   \
          α       β
```
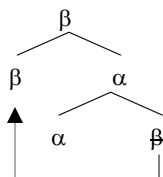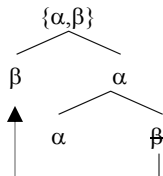
        d.   Project neither α nor β

```
            /   \
          α       β
```
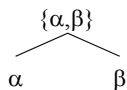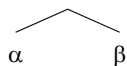
The logical possibilities for *Internal Merge* structures are illustrated in (4a-d). For α the Probe and β the Goal, these are: (*i*) Project Probe, (*ii*) Project Goal, (*iii*) *Project Both* Probe and Goal, and Project neither Probe nor Goal.
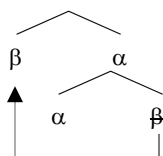
4)   a.   Project Probe
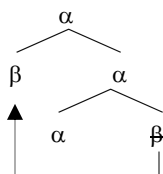
α
β   α
α   β̶

b.   Project Goal

β
β   α
α   β̶

c.   *Project Both* Probe and Goal

{α,β}
β   α
α   β̶

d.   Project neither Probe nor Goal

β   α
α   β̶

Before we move on, let me briefly comment on the convention used in this paper to represent *Project Both*. In (3c) and (4c), the label is represented as an unordered set {α, β} consisting of two constituent labels α and β. I assume labels to be shorthand descriptions for a set of features of the relevant constituent. To say that α is the label of the object given in (3a) or (4a) is nothing more than to say that the features of the lexical item α have projected to the next level. In the case of *Project Both*, the idea is to capture the intuition that the features of both constituents have projected to the next level. The hypothesis I would like to put to test here is that ALL logical possibilities illustrated in (3) and (4) are in fact attested. This is my understanding of the logic behind Chomsky's (2005) suggestion that 'the labeling algorithms apply freely, sometimes producing deviant expressions (Chomsky 2005:11).

        I will focus on the 'missing labels,' illustrated in (3c-d) and (4b-d) above, which are assumed to be impossible on standard minimalist assumptions. I will proceed as follows. In section 2, I will outline what I take to be standard minimalist assumptions concerning labels in both Internal and *External Merge* structures, on which only a subset of logically possible labels actually exists. From the diagrams given in (1) and (2) above, one might infer that the two are fundamentally different. This is somewhat misleading; in Section 2.3, I will review proposals that the same labeling algorithm operates in both *External* and

External Merge structures. The main contribution of this paper, however, is the examination of the missing labels, i.e. the ones departing from standard assumptions. In section 3, I will examine missing labels in *External Merge* structures. I will argue that *Project Both* captures otherwise puzzling properties of comparative correlatives and extended projections in the sense of Grimshaw 1991/2005. In Section 4, I turn to the missing labels in *Internal Merge* structures. I offer evidence in favor of the idea, going back to Larson 1998 (see also Bury 2003, Donati 2006, Iautridou, Anagnostopoulou and Izvorski 2001), that free relatives involve a *Project Goal* derivation. I further propose that *Project Both* can solve some of the long standing problems concerning head movement. I conclude the paper with a comparison of the present proposal with alternatives that eliminate labels altogether, such as that of Collins (2002) and Seely (2006).

2. Standard Minimalist View on Labels

2.1 Labels in *External Merge* Structures

Merge is defined as an operation that applies to two objects (α and β), and forms a new object out of them.[2] The result of Merge in the simplest case would be a set $\{\alpha, \beta\}$. Chomsky (1995) argues that this is not sufficient due to output conditions, and that labels are necessary because 'verbal and nominal elements are interpreted differently at LF and behave differently in the phonological complement' (Chomsky 1995:243). The object resulting from Merge thus must also include a label, represented as γ in (5).

5) $K = \{\gamma, \{\alpha, \beta\}$

Chomsky (2005) further notes that the label γ 'must be identifiable with minimal search, by simple algorithm' (Chomsky 2005:10). In order to avoid a violation of inclusiveness, the label γ *must* be constructed from the two constituents α and β. Logically, there are three possibilities: γ could be the intersection of α and β, the union of α and β, or one or the other of α and β. Chomsky excludes the first two options on the following grounds: 'the intersection of α, β will generally be irrelevant to output conditions, often null; and the union will be not only irrelevant but 'contradictory if α and β differ in value for some feature' the normal case. The only choice is then for either α or β to project. This opens up the possibility that *Project Both* could be possible as long as the two projecting elements do not conflict in relevant features.[3] I will come back to this possibility and explore its empirical consequences in Section 3.1 for *External Merge*, and in Section 4.3 for *Internal Merge*.

---

[2] A separate question is why *Merge* targets two elements (rather than one, zero, or more than two). Merging less than two objects is vacuous, whereas merging more than two is less economical than merging two (Collins 1997).
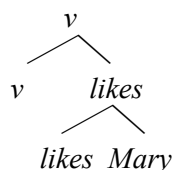[3] Pereltsvaig (2001) argues this is precisely what happens in small clauses. See Citko (2006), however, for empirical arguments against the existence of bare small clauses.

This reasoning raises an interesting question of which of the two elements projects as the label. The standard assumption is that in a given derivation at a given derivational step, only one choice leads to a convergent derivation. Consider in this light the derivation of a simple transitive clause such as 'John likes Mary', schematized (6a-e). The labels are underlined for the sake of clarity.
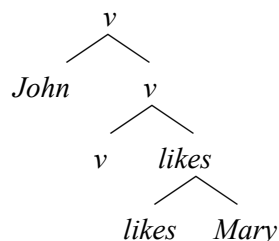
6)   a.   N = {John, Mary, likes, v, T}

        b.   Merge *likes* and *Mary* → {<u>likes</u>, {likes, Mary}}

```
        likes
       /    \
   likes    Mary
```
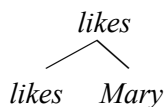
        c.   Merge v and {<u>likes</u>, {likes, Mary}} → {<u>v</u>, {v, {<u>likes</u>, {likes, Mary}}

```
        v
       / \
      v   likes
          /   \
       likes  Mary
```

        d.   Merge *John* and *{<u>v</u>,{v,{<u>likes</u>, {likes, Mary}}}* → *{<u>v</u>, {John,{<u>v</u>, {v, {<u>likes</u>, {likes, Mary}}}*

```
        v
       / \
   John   v
         / \
        v   likes
            /   \
         likes   Mary
```

        e.   Merge T and *{<u>v</u>, {John, {<u>v</u>,{v,{<u>likes</u>, {likes, Mary}}}}* → *{<u>T</u>, {T, {<u>v</u>, {John, {<u>v</u>,{v,{<u>likes</u>, {likes,Mary}}}}}*

```
         T
        / \
       T   v
          / \
      John   v
            / \
           v   likes
               /   \
            likes   Mary
```
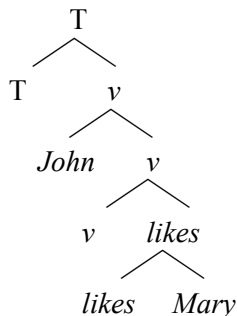
Furthermore, let us compare this convergent derivation with a partial nonconvergent one given in (7a-b) below. The two involve the same numerations, and merge the same elements; the only difference between them lies in which of the two merged elements projects. The first step in (7) involves projecting

the noun instead of the verb. Projecting the noun is consistent with the system given in Chomsky (1995), since the noun *Mary* is one of the two merged elements.

7)   a.   N = {John, Mary, likes, *v*, T}

        b.   Merge {*likes, Mary*} → {*<u>Mary</u>, {likes, Mary}*}

<div align="center">

Mary

likes   Mary

</div>
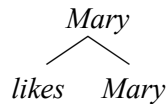
Intuitively, if *Mary* is the label of the object formed by this Merge operation, the derivation is going to be cancelled at the next step. This is essentially the intuition behind Chomsky's (2005) labeling algorithm, given in (8).

8)   In {H, α}, H an LI, H is the label (Chomsky 2005:**X**)

However, as Chomsky (2005) himself points out, there are some 'exceptional' cases that do not fall clearly under this algorithm. First, at the initial stage of every derivation, either of the two merged elements can project. The question then reduces to the question of what determines the head (as opposed to what determines the label), which is a mere restatement of the question I posed above. Instead of asking what determines the label of the objects formed by *External Merge*, we are now asking what determines the head of the *External Merge* structure. Another case that does not fall under the principle given in (8) concerns the merge of two XPs, in which case neither is the head, which is what takes place if the external argument merges with a non-minimal *v*P.[4]

What seems to be the crucial difference between the convergent (6b) and the nonconvergent (7b) is that in the former the constituent that projected was the one that had a feature or a set of features satisfied by the Merge operation. The feature(s) in question could be selectional features, EPP features, phi-features etc. This is the conclusion reached by Pesetsky and Torrego (2006), whose goal is to provide a unified motivation for both *Internal* and *External Merge*. They formulate it in terms of the Vehicle Requirement for Merge, given in (9) below:

9)   *Vehicle Requirement on Merge*   (Pesetsky and Torrego 2006:1)

        If α and β merge, some feature F of α must probe F on β.

Even though their primary concern was with the motivation for Merge, not with the labels created by Merge, it seems natural to interpret their proposal as saying that *the label of the Probe is the one that projects*. The Vehicle Requirement on Merge captures perhaps one of the most fundamental syntactic intuitions, which is that nouns and verbs differ in their complementation (and modification patterns). In

---

[4] Chomsky (2005) hints at the possibility of neither label projecting in small clauses, which might explain why one of the two elements has to move (Chomsky 2005:11, *Ft* 11).

both cases, unvalued features act as Probes, which are valued via an Agree operation. Consider in this light the well-known pattern given in (10a-b), showing that the omission of *that* in a CP complement of a noun is not possible.

10)   a.   your proof *(that) Mary could not have committed the crime

       b.   the demonstration *(that) John was insane   (Pesetsky and Torrego 2006:5)

Pesetsky and Torrego (2006) derive the obligatory presence of *that* in (10ab) from the Vehicle Requirement on Merge. First, nouns have unvalued T features, which in Pesetsky and Torrego's (2001) system, are structural case features. The nominal complements therefore be headed by valued T features. This can only be accomplished if the CP has a T moved to its head. Bare CPs, i.e. CPs not headed by 'tensed' C heads, are not acceptable. This account relies on an independent proposal, also due to Pesetsky and Torrego (2001) that the complementizer *that* is a realization of a T head raised to C.

The combinatorial properties of verbs are also shown to follow from the VRM. Verbs have unvalued phi-features, and the complement of V has to value these features. Thus, a complement of V must be headed by an element with valued phi-features, such as a DP.

An interesting consequence of Pesetsky and Torrego's (2006) theory is that in both *External* and *Internal Merge* the constituent that projects has had its unvalued features valued through the operation. Thus, in a given derivational step involving Merge of α and β, only projecting one of them as the label will be possible; projecting the other one will result in ungrammaticality.

The question that I would like to address here is whether the computational system is built in such a way that for all derivations and all derivational stages, only projection of one element leads to a convergent result. If so, we might have an interesting argument in favor of a label-free system (see Preston 2006 for a discussion of this point). If not, what emerges instead is an interesting case in favor of the existence of labels. If Pesetsky and Torrego (2006) are right, cases of true *Project Either* should not exist.

A potential case of *Project Either* comes from Rooryck's (2000) analysis of French complementizers and wh-pronouns. The focus of Rooryck's paper is on the well-known *qui/que* alternation and the morphological identity between complementizers and wh-pronouns. [5] For the purposes of this paper, I will focus on the identity between complementizers and wh-pronouns, illustrated in (11) and (12).

11)   *qui/que* as interrogative pronouns

---

[5] Rooryck (2000) also departs from the canonical view regarding the *que/qui* alternation, which is to treat it as an agreeing complementizer. See also Taraldsen (2002) for arguments against treating *qui* as an agreeing complementizer. First, he points out that it lacks 'the characteristic morphosyntactic properties of agreeing complementizers in Germanic'. It lacks the person/number features that we would expect of an agreeing complementizer. Furthermore, it only 'agrees' if the subject has been wh-moved.

a. *Qui* est arrivé? [French]

who is   arrived

'Who arrived?'

b. *Qu*'as-tu   vu?

what have-you seen

'What did you see?'

12) *qui/que* as complementizers

a. Le paquet *que* j'ai vu.

the package that I've seen

'The package that I have seen'

b. Le paquet *qui* est venu.

the package that is   arrived

'The package that arrived.'

Rooryck (2000) derives this parallelism between wh-words and complementizers from the idea that the complementizer/wh-word *qui* or *que* is not a primitive lexical item. *Qui*, for example, is what he refers to as a functional complex, consisting of a C head and a D head. What yields a contrast between a complementizer and a wh-word is the option to project either the C or the D head. If C projects, the result is a complementizer. If D projects, the result is a wh-pronoun, as shown schematically in (13a-b):

13) a. Merge C and D, Project C

```
        C
       / \
      C   D
      |   |
     qu-  i
```

b. Merge C and D, Project D

```
       D
      / \
     C   D
     |   |
    qu-  i
```
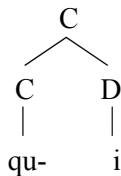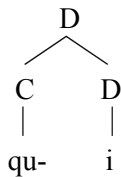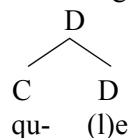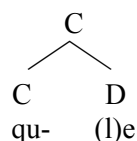
The non-subject complementizer/wh-word *que* can be derived in an analogous way. The only difference is that the D is the clitic *le* 'him/ it' rather than i(l) he/it, as shown in (14).

14)   a.   Merge C and D, Project D

```
           D
          /\
        C    D
       qu-  (l)e
```

      b.   Merge C and D, Project C

```
           C
          /\
        C    D
       qu-  (l)e
```

Rooryck's analysis of the French complementizer/wh-pronoun alternation can also shed some light on the issue of why it is relatively common for complementizers to be homophonous to wh-elements, and why diachronically complementizers often derive from wh-elements.[6] (15a-d) give some examples of complementizers that are related to wh-pronouns in a small selection of Indo-European languages.[7]

15)   a.   Russian   čto   (also *what*)

      b.   Polish   co   (also *what*)

      c.   Marathi   *ki*   (originally *what*)   (Bayer 1999)

      d.   Dakkhini-Hindi *ki*   (originally *what*)

The following examples from Polish, one of the languages given in (15), illustrate the morphological identity of wh-pronouns and complementizers.

16)   a.   *Co* czytasz?   [Polish]

             what read-2sg

             'What are you reading?'

      b.   Ten szef *co*   mu się   nic   nie podoba

             this boss COMP him REFL nothing not please

             'The boss that does not like anything…'

The free system of labeling proposed here sheds an interesting light on the historical change involving complementizers, in particular on the development of wh-words into complementizers. The

---

[6] The two observations are obviously related. Complementizers and wh-words are homophonous because they are historically linked.

[7] Bayer (1999), based on the data from Dravidian languages, also observes that final complementizers tend to be different from initial complementizers. Clause final complementizers quite often derive from verbs of saying, whereas clause initial complementizers are 'degenerate operators'.

change of wh-words into complementizers could simply be thought of as a change in which of the two merging elements projects.

The last question I want to address before turning to labels in *Internal Merge* structures is whether the option discussed here is compatible with Pesetsky and Torrego's Vehicle Requirement on Merge, according to which it is the element that had some unvalued feature valued through the Merge operation. In the case of complementizer/wh-pronoun alternation, the difference could be attributed to the difference of which of the two has an unvalued feature. D will project if it probes some features on C, and C will project if it is probing some feature on C. The possibility of Project Either discussed in this section thus does not seem to conflict with the Vehicle Requirement on Merge.

2.2   Labels in *External Merge* Structures

The standard minimalist assumption regarding labels in *External Merge* structures is that it is always the target of movement that projects. In current terms, it is always the Probe rather than the Goal that projects, as shown in (17a). Chomsky (1995) excludes the other logical possibility, *Project Goal* (illustrated in (17b)) on the grounds that no feature checking could be established in such a configuration.

17)   a.   Project Probe
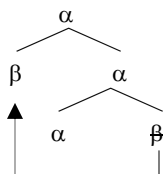
       b.   Project Goal
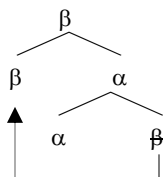
On the minimalist program, movement is assumed to be morphologically driven, and take place only as last resort. Chomsky (1995) considers three versions of Last Resort, given below, and argues that *Project Goal* is incompatible with all three of them.

18) α can target K only if   (Chomsky 1995:257)

       a.   a feature of α is checked by the operation

       b.   a feature of either α or K is checked by the operation

       c.   the operation is a necessary step toward some later operation in which a feature of α will be checked.

If α raises to target K, and projects as its label, the result is the object in (19).

19)   L = {H(α), {α, K} }

As a result of this operation, α ends up as a head of the new object, with K as its complement:

20)
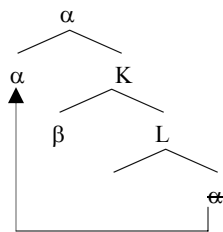
Chomsky (1995) excludes this possibility due to the fact that this movement does not establish a standard feature checking spec-head configuration. If β in (20) is the Probe, α is not a specifier of β after movement takes place. This objection, however, is not valid from the perspective of more current versions of the minimalist program, in which spec-head feature checking is replaced by a Probe-Goal Agree relationship. In (20), α is in the c-command domain of β; therefore an Agree relationship can be established between them (as long as β has some uninterpretable feature that can be valued by a matching feature on α).

Another possibility Chomsky (1995) considers is that after raising of α to target K, K becomes the specifier of α. In this case, α is maximal before movement and non-maximal after movement, which he takes to be impossible. This, however, is exactly the situation found in clitic movement cases, in which the moving clitic is maximal before movement and non-maximal after movement.

Interestingly, Chomsky (2005) departs slightly from Chomsky (1995) in a discussion of how labels are formed. In particular, Chomsky (2005) hints at the possibilities that were explicitly excluded by Chomsky (1995). Let us look again at the labeling algorithm given in Chomsky (2005). (21b) is what is crucial for *Internal Merge* cases.

21)   a.   In {H, α}, H an LI, H is the label

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

If α is a linguistic item, and β is not a linguistic item, the labeling algorithm in (21a-b) makes conflicting predictions. According to (21a), α should be the label, but according to (21b), β should be the label. This opens up the possibility, which I will explore in Section 4.2 of this paper, that sometimes the moved element can project.

Having outlined what I take to be standard minimalist assumptions about labels in *External* and *Internal Merge* structures, I turn to the main proposal, which is that the computational system does not impose any restrictions on what can project as the label. In the remainder of this paper, I will argue for the existence of the 'missing labels', illustrated in (22) for *External Merge*, and in (23) for *Internal Merge*.

22)   *Project Both* α and β

$$\{\alpha,\beta\}$$

α   β

23)   a.   Project Goal

β

β   α

α   β
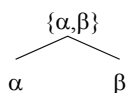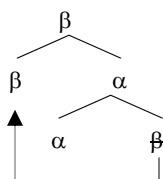
b.   *Project Both* Probe and Goal

$$\{\alpha,\beta\}$$

β   α

α   β
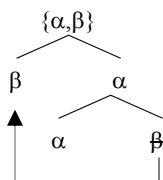
Note that the range of logical possibilities also includes the ones given in (24a-b), in which neither of the two input elements projects. I will set these aside till Section 5.

24)   a.   Project Neither in *External Merge* structures

α   β
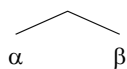
b.   Project Neither in *External Merge* structures

β   α

α   β
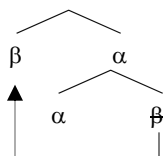
3.   Missing Labels in *External Merge*

3.1   *Project Both*

The labeling algorithm given in Chomsky (1995) leaves open the possibility that *Merge* can combine two objects and project both of them as long as there is no feature conflict between them. A natural question that arises here is what the relevant features are. I take them to be categorical features. More concretely, I propose that the grammar allows the structure in (25) as long as α and β are of the same category.

25)   *Project Both* α and β

$$\{\alpha,\beta\}$$

α   β
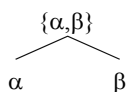
In the following two sections, I present two case studies to argue that this possibility is indeed attested. One involves comparative conditionals and the other one extended projections in the sense of Grimshaw 1991/2005.

3.1.1 Comparative Correlatives

In this section, I show that the option of projecting the labels of both merged constituents can explain an otherwise mysterious behavior of comparative correlatives. Comparative correlatives, illustrated in (26a-b), have correlative syntax and conditional semantics.[8]

26)   a.   The more you smile, the happier you get.

        b.   *Jiitnaa*   suuraj chamk-aa, *utnii*(-hii)   ThanD baRh-ii   [Hin]

               how-much sun   shine   that-much(-only) cold.F increase

               'The more the sun shone, the colder it got.'

        c.   *Minél*   többet olvasol,   *annál*   többet   [Hun]

               what-ADESS more   you.read   that-ADDESS more-ACC

               {megértesz/   értesz meg}.

               VM-you.understand you.understand-VM

               'The more you read, the more you understand.'

        d.   *Chem*   bol'she vina,   *tem*   veseleye.   [Rus]

               what-INSTR more   wine-GEN that-INSTR merrier

               'The more wine, the merrier.'   (Den Dikken 2005:499-500)

Both standard and comparative correlatives are typically analyzed as involving an adjunction structure of the kind given in (27) (see Beck 1997, Den Dikken 2005, McCawley 1988, 1998, among others).

27)   *Adjunction Structure for Correlative Clauses*[9]

---

[8] The crosslinguistic distribution of comparative correlatives is wider than that of standard ones. English, for example, does not allow standard correlatives, but allows comparative correlatives.
(i)   * What you like, this you eat.   *standard correlative*
(ii)   The more you smile, the happier you get.   *comparative correlative*
In languages that allow both, which are the languages I will focus on in this paper, the syntactic parallels between standard and comparative correlatives are more evident. As can be seen by comparing the examples in (26b-d) to the ones (i-iii), in both types the first clause contains a relative or interrogative pronoun and the second one a corresponding demonstrative one (Den Dikken 2005):
(i)   *Jo   CD sale par-hai*, Maya us CD-ko khari:d-egi: [Hin]
        which CD sale on is   Maya this CD   buy-will
        'Maya will buy the CD that is on sale.'
(ii)   *Aki   korán jött,* azt   ingyen beengedték.   [Hun]
        REL-who early came   that-ACC freely PV-admitted-3PL
        'Those who come early were admitted for free.'
(iii)   *Kogo ty predložiš,* togo   my vyberem.   [Rus]
        whom you suggest that-one we will-appoint
        'We will appoint whom you suggest.'

$$\text{CP} \rightarrow \text{CP}_{rel} \quad \text{CP}$$
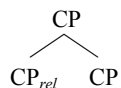
Such an adjunction structure makes a very straightforward prediction regarding extraction from correlatives. In particular, it predicts that extraction should be possible from the matrix clause but impossible from the adjunct clause. In what follows, I test this prediction in three languages that allow both standard and comparative correlatives: Polish, Hindi, and Hungarian.[10] As shown in (28-30), standard correlatives allow extraction from the matrix clause but not from the adjunct one.

28) a. * *Kogo_i* Jan myśli, że [kto lubi *t_i*], [ten powinien przeczytać Anne Kareninę]?   [Pol]
          who Jan thinks that who likes DEM should read Anna Karenina
          'Who do you think that the one that likes (him) should read Anna Karenina?'

     b. ? *Co_i* Jan myśli, że [kto lubi Tołstoja], [ten powinien przeczytać *t_i*]?
          what Jan thinks that who likes Tolstoy DEM should read
          'What does Jan thin that the one who likes Tolstoy should read?'

29) a. ???/* *Kaun-sii kitaab_i* Mary soch-tii hai ki [jis-ko *t_i* pasand hai]   [Hin]
          which book Mary think-HAB.F be.PRS that REL-DAT like is
          [us-ko Anna Karenina paRh-nii chaahiye]
          he-DAT Anna Karenina read-INF should
          'Which book does Mary think that the one who likes (it) should read Anna Karenina?'

     b. ?*Kaun-sii kitaab_i* Mary soch-tii hai ki [jis-ko War and Peace
          which book Mary think-HAB.F is that REL-DAT War and Peace
          pasand hai] [us-ko *t_i* paRh-nii chaahiye]
          like is he-DAT read-INF should
          'Which book does Mary think that the one who likes War and Peace should read?'

30) a. * *Kiket_i* hallott Mari, [hogy aki szeret *t_i*], [annak el kell olvasni az Anna [Hun]
          who heard Mari that REL-who likes, that-DAT PV need read-INF the Anna
          Kareninát]?
          Karenina-ACC
          'Who did Mary hear that the one who likes (him) should read Anna Karenina?'

     b. ? *Miket_i* hallott Mari, hogy [aki szereti Tolstoyt], [annak el kell olvasni *t_i*]? [Hun]
          what heard Mari that REL-who likes Tolstoy-ACC that-DAT PV need read-INF

---

[9] Srivastav (1991), Dayal (1995), and Bhatt (2003) analyze the matrix clause in Hindi correlatives as an IP rather than a CP. Izvorski (1996), however, shows that in Slavic languages demonstrative pronouns in the matrix clause undergo the same kind of movement as relative or wh-pronouns in questions, which suggests that at least in some languages the matrix clause has to be a CP.

[10] The Hungarian judgments are due to Aniko Liptak, and the Hindi ones to Rajesh Bhatt, both of whom I am very thankful to. The Polish judgments reported in this paper are mine unless otherwise indicated.

'What did Mary hear that the one who likes Tolstoy needs to read?'

What is unexpected given the structure in (27) is the fact that comparative correlatives do not show the expected adjunct/matrix asymmetry with respect to wh-movement. Either both clauses allow extraction, or both disallow it. Culicover and Jackendoff (1999) were the first ones to note this fact for English comparative correlatives:[11]

31)   a.   a person *who$_i$* [the more you meet t_i], [the more you hate him]

       b.   a person *who$_i$* [the more you meet him], [the more you hate t_i]

Den Dikken (2005) further notes that ATB-style movement is also possible:

32)   a person *who$_i$* [the more you meet t_i], [the more you hate t_i]   (Den Dikken 2005:504)

The three languages that we looked at above, Polish, Hindi, and Hungarian, also do not show the expected adjunct/matrix asymmetry. However, there is an interesting contrast between Polish and Hindi on the one hand, and Hungarian on the other. Polish and Hindi pattern with English in that they allow extraction from either clause (or both clauses simultaneously), as shown in (33a-c) and (34a-c), respectively. Hungarian, on the other hand, disallows movement completely as shown in (35a-c).[12]

33)   a.   *Kogo$_i$* myślisz,   że [im lepiej poznasz t_i], [tym bardziej go   polubisz]?[13]   [Pol]

             whom think.2sg that the better know.2sg   the more   him like.2sg

             'Whom do you think that the more you get to know, the more you will like him?'

       b.   *Kogo$_i$* myślisz,   że [im lepiej go   poznasz], [tym bardziej   polubisz t_i]?

             whom think.2sg that the better him know.2sg   the better   like.2sg

---

[11] Culicover and Jackendoff (1999) analyzed such examples as involving parasitic gaps in which either gap is independently possible, thus assimilating them to examples of the following kind:

(i)   *Who$_i$* did you tell t_i that you would pay a call on t_i?

(ii)   *Who$_i$* did you give pictures of t_i to friends of t_i?

Such parasitic gaps, however, seem to be quite rare, and not all languages that allow extraction from both clauses in a comparative correlative allow such parasitic gaps. Polish, for example, is like English in that it allows wh-movement out of comparative correlatives. However, it does not allow parasitic gaps of the kind given in (i-ii):

(iii) ?? *Kogo$_i$*   ostrzegłeś t_i że odwiedzisz t_i?   [Pol]

         who-ACC warn.PST.2SG that visit.2SG

         'Who did you warn that you will visit (him)?'

(iv)   * *Kogo$_i$*   dałeś   zdjęcia t_i przyjaciołom t_i?

         who-ACC give.PST.2SG pictures-ACC   friends-DAT

         'Who did you give pictures of to friends of?'

[12] Den Dikken (2005) also notes that German and Dutch correlatives behave similarly to Hungarian with respect to extraction out of comparative correlatives.

[13] Replacing the pronouns in (33a-b) with names does not affect grammaticality, as shown in (i) and (ii).

(i)   *Kogo$_i$* myślisz,   że [im lepiej poznasz t_i], [tym bardziej   lubisz Jana]?   [Pol]

         whom think.2sg   that the better know.2sg   the more   like.2sg Jan

         'Whom do you think that the more you get to know, the more you will like John?'

(ii)   ?*Kogo$_i$* myślisz,   że   [im lepiej poznasz Jana], [tym bardziej   lubisz t_i]?

         whom think.2sg that the better   know.2sg Jan   the better   like.2sg

         'Who do you think that the more you get to know John, the more you will like?'

'Who do you think that the more you get to know him, the more you will like?'

c. *Kogo_i* myslisz, że [im lepiej poznasz *t_i*], [tym bardziej polubisz *t_i*]?

who think.sg that the better know.2sg the more like.sg

'Who do you think that the more you get to know, the more you will like?'

34) a. ?*Kis-ko_i* Mary soch-tii hai ki [tum *t_i* jitnaa zyaadaa jaanoge] [Hin]

who-ACC Mary think-HAB.F be.PRS that you how.much more know-FUT

[tum us-ko utnaa-hii zyaadaa pasand karoge]?

you he-ACC that.much-only more like do-FUT

'Who does Mary think that the more you get to know, the more you like him?'

b. ? *Kis-ko_i* Mary soch-tii hai ki [tum us-ko jitnaa zyaadaa jaanoge]

who-ACC Mary think-HAB.F be.PRS that you he-ACC how.much more know-FUT

[tum *t_i* utnaa-hii zyaadaa pasand karoge]?

you that.much-only more like do-FUT

'Whom does Mary think that the more you get to know him, the more you will like (him)?'

b. *Kis-ko_i* Mary soch-tii hai ki [tum *t_i* jitnaa zyaadaa jaanoge]

who-ACC Mary think-HAB.F be.PRS that you how.much more know-FUT

[tum *t_i* utnaa-hii zyaadaa pasand karoge]

you that.much-only more like do-FUT

'Who does Mary think that the more you get to know (him), the more you will like (him)?'

35) a. * *Kiket_i* hallott Mari, hogy [minél jobban ismerted o"ket], [annál [Hun]

who-PL-ACC heard Mari that what-at better knew-2sg them, that-at

jobban szeretted *t_i*]?

better liked-2sg

'Who did Mari hear that the better you knew (them), the better you liked (them)?'

b. ???*Kiket_i* hallott Mari, hogy [minél jobban ismerted *t_i*], [annál jobban

who-PL-ACC heard Mari that what-at better knew-2sg, that-at better

szeretted öket]?

liked-2sg they-ACC

'Who did Mari hear that the better you knew (them), the better you liked (them)?'

c. * *Kiket_i* hallotta Mari, [hogy minél jobban ismerted *t_i*], [annál jobban szeretted *t_i*] ?

who-PL-ACC heard Mari that what-at better knew-2sg, that-at better liked-2sg

'Who did Mari hear that the better you knew (them), the better you liked (them)?'

16

The proposal I would like to make here is that the possibility of projecting both labels can capture the differences in extraction possibilities between standard and comparative correlatives. In particular, I propose that standard correlatives involve a standard asymmetric adjunction structure, in which the label of the matrix CP projects, whereas comparative correlatives involve a more 'symmetric' structure, in which the labels of both the matrix CP and the relative CP project. The contrast between the two is illustrated in (36a-b):

36) a. Merge the relative and the matrix CP, Project the matrix CP   *standard correlatives*

$$CP$$
$$CP_{rel} \quad CP$$
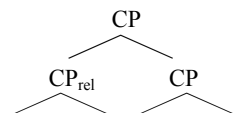
 b. Merge the relative CP and the matrix CP, *Project Both*   *comparative correlatives*

$$\{CP,CP\}$$
$$CP \quad CP$$
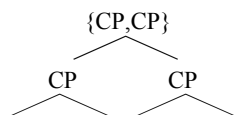
Crucially, the assumption that the labels of both CPs project in (36b) does not imply that there are no asymmetries whatsoever between the two constituent CPs. With respect to many diagnostics, the two clauses in a comparative conditional construction behave asymmetrically with respect to each other.[14] I follow the suggestion made by Chomsky (2001) to handle a similar issue that his *Pair Merge* structures face.[15] The *Project Both* option proposed here is similar in spirit to Chomsky's (2001) *Pair Merge*, which also projects the labels of two merged constituents. In both types of Merge, two merged objects project as the label of the new object. Chomsky (2001) suggests that symmetric Pair Merge structures are converted to a more standard asymmetric Set Merge structures at the level of Transfer, which is the point at which the derivation is shipped to the conceptual-intentional and perceptual-articulatory interfaces. At the level relevant for the purposes of linearization and interpretation, what we have is a standard asymmetric structure. For the purposes of narrow syntax processes, which I take wh-movement to be, however, there is no asymmetry between the two CPs in a symmetric Merge structure.

  Given this structural difference between the two types of correlatives, we can account for the differences between the two with respect to wh-movement by appealing to any version of the Adjunct

---

[14] Culicover and Jackendoff (1999) discuss two such diagnostics: question tags and subjunctive mood. As shown below, with respect to both of them, there is an asymmetry between the first and the second CP:
(i) I demand that the more John eats, the more he pay.
(ii) * I demand that the more John eat, the more he pays.
(iii) The more we eat, the angrier you get, don't you?
(iv) * The more we eat, the angrier you get, don't we?   (Culicover and Jackendoff 1999:549)
[15] The two, Chomsky's *Pair Merge* and my *Project Both Merge*, are not fully equivalent though. Chomsky's Pair Merge creates an ordered set, my *Project Both Merge* creates an unordered set. Furthermore, for Chomsky (2001), *Pair Merge* is limited to adjuncts.

Condition. For the sake of concreteness, I will adopt an independently motivated Agree-based version of Rackowski and Richards (2005). The crucial innovation in their proposal lies in the idea that the CP out of which the wh-phrase moves has to (independently) agree with the matrix v. It is this agreement that makes the CP transparent for extraction. The ungrammaticality of extraction from the relative clause in standard correlatives can be accounted for in the following manner. Following the general consensus in the field, I assume wh-movement is successive cyclic and proceeds through the edge of the vP phase. In a Probe-Goal system, this means that the v embedded inside the relative clause first agrees with the wh-pronoun, which subsequently moves to its specifier.

37) $[_{CP}$ C $[_{vP} \dots v \dots$ $[_{CPmtx} [_{CPrel} \dots [_{vP} wh [_{v'}$ v $[_{VP} \dots t_{wh} \dots]]]] [_{CPmtx} [_{vP} \dots]]]]$?

Next, the highest v first agrees with the embedded CP, which is what makes this CP transparent for future Agree relationships:

38) $[_{CP}$ C $[_{vP} \dots v \dots [_{CPmtx} [_{CPrel} \dots [_{vP} wh [_{v'} v [_{VP} \dots t_{wh} \dots]]]] [_{CPmtx} [_{vP} \dots]]]]$?

Now, the matrix v agrees with the next closest phase containing the wh-pronoun. In this particular case, it is going to be the relative CP *not* the relative vP. The wh-pronoun is thus not going to be movable, not being able to undergo Agree with the higher v.

39) $[_{CP}$ C $[_{vP} \dots v \dots [_{CPmtx} [_{CPrel} \dots [_{vP} wh [_{v'} v [_{VP} \dots t_{wh} \dots]]]] [_{CPmtx} [_{vP} \dots]]]]$?

By contrast, in comparative correlatives, the matrix v will agree with both the matrix and the relative CP due to the fact that both of them have projected as the label of the newly formed constituent. Consequently, both the matrix and the relative CP are going to be transparent for extraction.

40) $[_{CP}$ C $[_{vP} \dots v \dots [_{CPmtx,CPrel} [_{CPrel} [_{vP} wh [_{v'} v [_{VP} t_{wh}]]] [_{CPmtx} [_{vP} wh [_{v'} v [_{VP} t_{wh}] \dots]]$?

The next *Agree* relationship can skip both the matrix and the relative CP. The next phase containing the wh-phrase is going to be either of the two vPs. Thus, extraction from both is going to be possible.

41) $[_{CP}$ C $[_{vP}$ $v \dots [_{CPmtx,CPrel}$ $[_{CPrel} [_{vP} wh [_{v'} v [_{VP} t_{wh}]]] [_{CPmtx} [_{vP} wh [_{v'} v [_{VP} t_{wh}] \dots]]$?

In summary, the analysis I have outlined above explains why extraction from both CPs is possible in comparative correlatives. This is the pattern found in languages like Hindi, Polish, or English. However, it requires extra assumptions to explain the extraction pattern found in languages like Hungarian, Dutch or German, where extraction from both clauses in comparative conditionals is blocked. At this point I can only speculate that there is something in these languages that blocks agreement with a multi-labeled constituent. What is significant though is the fact that even in these languages, the two clauses in a comparative correlative construction behave in a symmetric fashion with respect to extraction; an unexpected fact given the standard adjunction analysis.

3.1.2 Extended Projections as Merge and *Project Both*

Allowing Merge to project the features of both merged constituents as the label of the new object can also provide a fairly straightforward way to recast Grimshaw's (1991)/(2005) extended projection theory.[16] The main idea behind extended projections is to account for the combinatorial properties of lexical and functional heads; and in particular, to provide answers to the following two questions. One is why the combinatorial properties of functional heads are much more limited than those of lexical heads. The attested combinations of lexical and functional heads are only a subset of the logically possible ones. A given functional head can only select a couple, typically just one, head as its complement. For example, a C head selects a T head (more accurately, a phrase headed by a T head), not the other way round. The second question is why selectional restrictions are encoded as restrictions between lexical heads, with intervening functional heads ignored. For example, a verb like 'kill' requires an animate object. Animacy, however, is a property of the head noun embedded within the complement DP. At least in English, determiners are not specified for animacy. Further evidence showing that there is no direct selection between verbs and determiners, due to Sportiche 1997, comes from the fact that there are no idioms consisting of verbs and determiners.[17]

More concretely, Grimshaw notes that the patterns given in (42) and (43) are possible ones, whereas the logically possible ones in (44) are not (Grimshaw 2006:9).[18]

42)　　Possible complements of functional heads

　　　C-TP　P-DP

　　　T-VP　D-NP

43)　　Possible complements of lexical heads

　　　V-PP　V-DP　V-NP　C-CP　V-TP　V-VP

　　　N-NP　N-DP　N-NP　N-CP　N-TP　N-VP

44)　　Impossible complements of functional heads

　　　T-NP　T-DP　T-PP　T-CP

　　　D-VP　D- TP　D-CP　D-PP

　　　C-NP　C-DP　C-VP

　　　P-VP　P-PP

　　　T-TP　D-DP

---

[16] This is what Hornstein, Nunes and Grohmann (2005) pose as an open ended question. The following section can be seen as an extended answer to this question.

[17] Such considerations led Sportiche (1997) to propose that determiners and their complement NPs do not start out as constituents. At the point of first (External) Merge, the verb selects directly for a noun, and the determiner is merged higher. Only later the noun moves and attaches to its corresponding determiner. This idea has come to be known as a DP-partitioning proposal.

[18] Grimshaw uses I's instead of T's in her analysis.

Grimshaw derives the impossibility of the patterns in (44) from a simple assumption that the functional heads in (44) do not form extended projections with their complements. An extended projection consists of a lexical head and all the functional heads that share its category specification. Furthermore, syntactic categories consist of category specification (+/-V, +/-N) and F (functional) feature specification. The specifications for common lexical and functional categories are given in (45). For Grimshaw (1991)/(2005), F values can range from 0 to 2, but nothing prevents the range to be wider in a system with a larger number of functional projections.

45)   *category specification   functional specification*

V	[+V, -N]	F0	(Grimshaw 2005:4)
T	[+V, -N]	F1	
C	[+V, -N]	F2	
N	[-V, +N]	F0	
D	[-V, +N]	F1	
P	[-V, +N]	F2	

Since V, T, C heads are non-distinct in terms of categorial status, they form one extended projection. So do N, D, and P heads.[19] The combinatorial properties illustrated in (42-44) follow from the assumption that heads of higher F values can only take projections with lower F values as their complements. In other words, functional heads take only complements that they form extended projections with. In all the ungrammatical combinations in (44) above, a functional head takes as its complement a category it cannot form an extended projection with. Complements of lexical categories are not subject to this restriction.

The idea that the labels of both merged constituents can project captures the main insights behind Grimshaw's extended projections. If the two merged objects do not conflict in category specification, the features of both can project as the label of the newly formed object. The nominal and verbal extended projections are illustrated in (46) and (47), respectively.
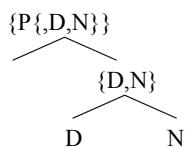
46)   a.   Merge D and NP, *Project Both*

                    {D,N}
                   /   \
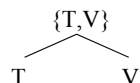   D   N

---
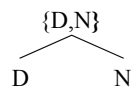
[19] Grimshaw's definition of a projection is given below (Grimshaw 2005:4):

(i)   X is the head of YP, and YP is a projection of X iff :
        a.   YP dominates X
        b.   The categorial features of YP and X are consistent
        c.   There is no inconsistency in the categorial features of all the nodes intervening between X and YP.
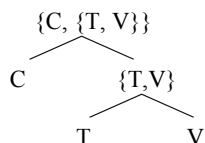
b.   Merge P with DP, *Project Both*

$$\{P\{,D,N\}\}$$

$$\{D,N\}$$

$$D \qquad N$$

47)   a.   Merge V and T, *Project Both*

$$\{T,V\}$$

$$T \qquad V$$

b.   Merge T and C, *Project Both*

$$\{C, \{T, V\}\}$$

$$C \qquad \{T,V\}$$

$$T \qquad V$$

Selectional restrictions now become a matter of straightforward sisterhood relationship. There is no need for a selecting head to skip intermediate nodes. If a DP object merges with a verb, the verb's selectional requirements can be met straightforwardly by the noun, since the features of the noun have also projected as the label of the entire phrase.

To sum up the discussion so far, I have motivated the option of projecting both labels in *External Merge* operations. In the next section, I turn to labels in *Internal Merge* structures.

4.   Labels in *Internal Merge* Structures

4.1   Logical Possibilities

The logical possibilities for determining labels in *Internal Merge* structures are given in (48).

48)   a.   Project Probe

$$\alpha$$

$$\beta \qquad \alpha$$

$$\alpha \qquad \beta$$
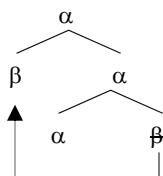
b.   Project Goal

$$\beta$$

$$\beta \qquad \alpha$$

$$\alpha \qquad \beta$$

c.   *Project Both* Probe and Goal

d.   *Project Neither* Probe nor Goal
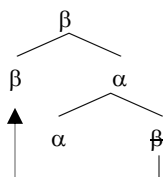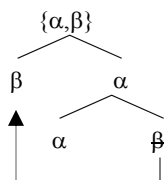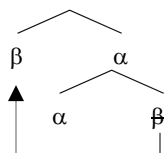

The standard minimalist assumption is that the Probe is the one that always projects. In the following two sections, I provide empirical illustration of the options given in (48b) and (48c): *Project Goal* and *Project Both*. I will argue that *Project Goal* is involved in the derivation of free relatives, and *Project Both* (Probe and Goal) is involved in the derivation of constructions involving head movement.

4.2   Project Goal

To the best of my knowledge, Larson (1998) was the first one to show that a number of properties of free relatives follow straightforwardly from a derivation in which the Goal rather than the Probe projects.[20] Free relatives, also known as headless relatives, are constructions that appear to be headed by the wh-pronouns themselves or lack heads altogether, as shown in (49a-b).

49)   a.   Bill reads *what(ever)* Tom reads.

         b.   Bill reads *whatever books* Tom reads.

Project Goal allows for a derivation in which the wh-pronoun undergoes *Internal Merge* with C and projects as the label of the newly formed constituent.

50)   *Project Goal* derivation of free relatives
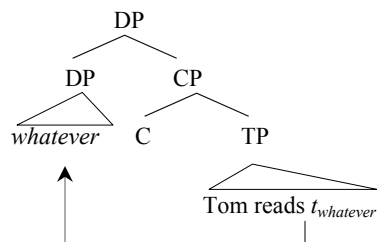
What is interesting about such a *Project Goal* account of free relatives is that it can reconcile the two seemingly incompatible accounts, the so-called Comp Account and the Head Account. Proponents of

---

[20] Other proponents of the *Project Goal* type accounts of free relatives include Iatridou, Anagnostopoulou, and Izvorski (2001), Bury (2003), and Donati (2006). They differ from each other in details of implementation, but the basic insight remains the same.

both accounts agree that a free relative is like a headed relative with some element missing. The two disagree, however, on what exactly is missing. On the Comp Account, the head is missing and the wh-phrase is in Spec,CP, as shown in (51a) (Groos and van Riemsdijk (1981), Grosu (1998), Gračanin-Yuksek (2005), among many others). On the Head Account, the [Spec, CP] position is empty and the head position is empty, as shown in (51b) (Bresnan and Grimshaw (1978), Larson (1987), Citko (2002).

51)   a.   John plays [$_{DP}$ Ø [$_{CP}$ *whatever* [$_{TP}$ he likes t(whatever)]]]   (Comp Account)

    b.   John plays [$_{DP}$ *whatever* [$_{CP}$ Ø [$_{TP}$ he likes __]   (Head   Account)

Convincing arguments in favor of the Comp Account come from locality effects; in particular from the fact that free relatives show the same locality restrictions on movement as wh-questions. This parallelism between free relatives and wh-questions is illustrated in (52-54) with respect to the Complex Noun Phrase Constraint, the Wh-Island Constraint, and the Adjunct Condition, respectively.

52)   a.   * John plays *whatever$_i$* he hears the claim that Mary likes t_i.

    b.   * *What$_i$* did John hear the claim that Mary likes t_i?

53)   a.   * John plays *whatever$_i$* he wonders why Mary plays t_i.

    b.   * *What$_i$* does John wonder why Mary plays t_i?

54)   a.   * John did *whatever$_i$* Mary left because John did t_i?

    b.   * *What$_i$* did John leave because Mary did t_i?

The presence of the so-called category and case matching effects, on the other hand, seems to favor the Head Account. Category matching, illustrated in (55a-b) refers to the requirement that the category of the wh-pronoun heading the free relative satisfy (or match) the requirements of the embedding predicate.

55)   a.   I will listen to whatever you listen to.

    b.   * I will listen [$_{DP}$ [$_{DP}$ whatever] Mary plays t.]

Case matching effects are best illustrated in languages with richer morphological systems (than English). As shown by the following examples from German, the case of the wh-pronoun inside the free relative has to be simultaneously satisfy the requirements of the matrix verb and the embedded verb. Citko (2000) discusses analogous data from Polish, a language whose case system is richer than German.

56)   a.   Ich nehme [wen$_{ACC}$ du   mir empfiehlst]$_{ACC}$.   (Groos and van Riemsdijk 1981:177)

        I take   whom   you me recommend

        'I take whom you recommended to me.'

    b.   * Ich nehme, [wem$_{DAT}$ du vertraust]$_{ACC}$

        I   take   whom   you trust

        'I take who you trust.'

In the grammatical example given in (56a), the verb *nehme* 'take' requires an Accusative object. Since the   wh-pronoun heading the relative is also Accusative, the result is grammatical. By contrast, in the

ungrammatical example given in (56b), the wh-pronoun heading the relative is Dative, which conflicts with the Accusative requirement of the matrix verb.

The *Project Goal* account can explain both locality and matching effects, thus combining the insights of both the Head and the Comp Account. The movement of the wh-pronoun in (50) above is forced by the C head. Thus, as far as the Probe-Goal relationship is concerned, this movement is the same as the movement that takes place in wh-questions. The parallelism in locality between the two follows naturally. The fact that the moved element projects, on the other hand, explains the category and case matching effects illustrated above.[21]

As mentioned above, there are many different implementations of the Project Goal account of free relatives.[22] In Larson 1998, the movement of the wh-phrase is akin to QR rather than wh-movement, in that it targets a TP rather than a CP domain, as shown in (57). While this derivation can explain the lack of overt complementizers in free relatives, it raises questions concerning the motivation for QR (or the movement of the wh-pronoun *whatever*). Furthermore, the evidence shows that the movement of wh-pronouns in free relatives can be successive cyclic, whereas QR typically assumed to be clause-bound.

---

[21] The situation with case matching is actually a little more complicated than the examples given in the text might imply. As shown in (i-ii), case matching effects are sensitive to morphological case syncretism. In other words, mismatches in Abstract case are fine as long as there exists an appropriate case syncretic form. The example in (i) is ungrammatical because the Genitive wh-pronoun cannot satisfy the case requirements of the matrix verb 'lubić' (which requires an Accusative object). The example in (ii), on the other hand, is grammatical because the wh-pronoun 'kogokolwiek' can be both Accusative and Genitive. See Citko (2000) for further discussion of case syncretism in free relatives, and Citko (2005) for a discussion of analogous case syncretism effects in ATB wh-questions.

(i)   * Lubię   [czegokolwiek$_{GEN}$ Maria nienawidzi t_{GEN}]$_{ACC}$ [Polish]
         like-1SG   whatever   Maria hate-3SG
         'I like whatever Maria hates.'
(ii)   Lubię   [kogokolwiek$_{ACC/GEN}$ Maria nienawidzi t_{GEN}]$_{ACC}$
         like-1SG whoever   Maria hate-3SG
         'I like whoever Maria hates.'

[22] Pesetsky and Torrego (2006), following the insights of Bhatt (1999) extend the idea of *Project Goal* to headed relatives. This allows them to extend the *Vehicle Requirement on Merge* beyond complementation, to cover cases of modification as well. In (ii), the head of the relative, the noun *person* is the projecting Goal:

(i)
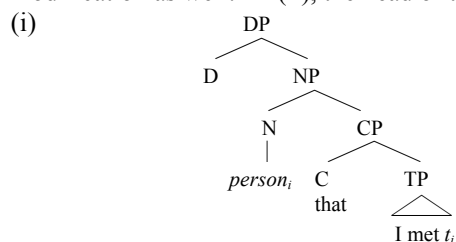
The *Project Goal* derivation of headed relatives captures all the facts that motivated the head raising/promotion analyses, at the same time avoiding the constituency problems that Kayne's analysis faces. For Kayne (1994), the headed relatives involve the structure given in (i) below. What distinguishes it from the structure in (i) is the fact nominal head is in [Spec,CP], and the CP itself is the complement of a determiner.
(ii)   [$_{DP}$ [$_{D'}$ the [$_{CP}$ person$_i$ [$_{C'}$ that [$_{TP}$ I met t$_i$]]
See Borsley (1997) for empirical arguments against this implementation of head raising account of headed relatives.

57)

$$DP$$

$$DP \quad TP$$

whatever   Bill reads $t_{whatever}$
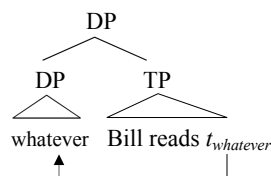
Donati (2006) proposes a somewhat more restricted version of the *Project Goal* derivation for free relatives. It is more restricted in a sense that it derives only a subset of what is t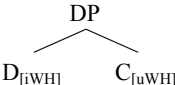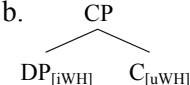raditionally thought as free relatives via a *Project Goal* mechanism. In particular, Donati (2006) proposes that only free relatives headed by simple wh-pronouns, such as the one given in (58a) involve a *Project Goal* derivation.[23] Relatives headed by more complex wh-phrases, such as the one in (58b), have to be reanalyzed as headed (rather than headless) ones. This might be true for some languages, but seems somewhat counterintuitive for languages like English.

58)   a.   John reads what Mary recommends.

   b.   John reads whatever books Mary recommends.

Donati (2006) argues that Project Goal is only possible if the projecting Goal is a head. Analyzing head movement as involving *Project Goal* is conceptually plausible and can solve a number of well-known problems with head movement. I will come back to head movement in the next section of this paper. However, restricting *Project Goal* to head movement strikes me as conceptually unmotivated. Donati derives this restriction in the following manner. First, she takes head movement to be universally preferred to phrasal movement. This preference for head over phrasal movement can be thought of as following from economy conditions, since head movement pied-pipes less material than phrasal movement. Thus, phrasal movement takes place only if head movement instead would result in a non-convergent derivation. The fact that the wh-word can project in free relatives follows from the interaction of economy constraints with the interface principles determining LF convergence. Consider in this light the structures given in (65a-b); in both of them, C acts as a Probe attracting a wh-feature:

59)   a. DP b. CP

   $D_{[iWH]}$ $C_{[uWH]}$ $DP_{[iWH]}$ $C_{[uWH]}$

Economy favors the structure in (59a). Donati further argues that the Chain Uniformity Condition requires the head to project in this case. However, the existence of movements that do violate the Chain Uniformity Condition, such as the often discussed case of clitic movement, weakens Donati's point. What excludes projection of the Goal in wh-questions is the fact that the result would be a nominal constituent, which cannot be interpreted as an interrogative element at the interface (Donati 2006:31).

---

[23] Chomsky endorsed Donati's idea that the goal projects in free relatives, and that free relatives headed by complex wh-phrases like 'I read what book you wrote.' He further hints that free relatives might be the case of Project Both.

However, a moved head can project if it is embedded in a nominal selecting environment. This, Donati, argues is exactly what happens in free relatives. Donati's proposal has three undesirable consequences, in my view. First, the explanation given above for why the Goal cannot project in wh-questions seems to ignore the existence of concealed questions.   These are clearly DPs that allow an interrogative interpretation. Second, the movement of the wh-head in free relatives that Donati proposes would have to exempt from the effects of the Head Movement Constraint. And third, the relatives given in (60a-b), despite appearances, have to be treated as headed relatives rather than headless ones.

60)   a.   I shall visit what towns you will visit.

        b.   I will dig with whatever shovel John digs.   (Larson 1987)

Donati's suggestion that only heads can move and project receives some support from languages like Italian or Polish, in which free relatives headed by complex wh-phrases such as 'how many cookies' or 'what towns' are indeed ungrammatical:

61)   a.   * Ho   mangiano *quanti   biscotti* hai   preparato.   [Italian]

                have-1sg eaten   how-many cookies have-2sg prepared

                'I have eaten what cookies you have prepared.'   (Donati 2006:32)

        b.   * Odwiedzę *które miasta* ty też odwiedzisz.   [Polish]

                visit-1sg which towns you also visit-2sg

                'I will visit which towns you will also visit.'

Donati acknowledges in a footnote that this prohibition is not universal, as shown by the grammatical status of the English examples in (60a-b) above. Furthermore, in both Italian and Polish, adding the particle *-ever* makes the ungrammatical examples in (61a-b) grammatical:

62)   a.   Odwiedzę którekolwiek miasta ty też odwiedzisz.   [Polish]

                visit-1sg whichever   towns you also visit-2sg

                'I will visit which towns you will also visit.'

        b.   Mangerò   qualunque bicotto preparerai. (Donati 2006:41, Ft 10)

                will-eat-1sg whatever   cookie will-prepare-2sg

                'I will eat whatever cookie you will prepare.'

The structure Donati (2006) assigns to such relatives is given in (63). Following Kayne (1994), it capitalizes on the near homophony between the English determiner/quantifier *every* and the particle *ever*.

63)   [DP [D ever/unque [CP [DP what town] […]]]

There are two problems with the treatment of such relatives as headed relatives. First, this homophony is not universal; the Polish particle *–kolwiek* is not morphologically related to any determiner.[24] While the

---

[24] According to Błaszczak (1999), -kolwiek is etymologically related to –koli, which is related to the alternative particle *li ... li* 'either or'.

issue of why languages like Polish or Italian only allow free relatives headed by simple wh-words is certainly worth further inquiry, the conclusion that these are headed relatives rather than free ones strikes me as very theory internal. It is hard to evaluate it without independent evidence that such relatives pattern with headed rather than headless relatives, and differ from free relatives headed by simple wh-pronouns. In order to do so, we need to establish clear diagnostics that would distinguish headed relatives from free relatives. The diagnostic established by Groos and van Riemsdijk (1981) to argue that the wh-pronoun in a free relative is in [Spec,CP] rather than the head position is a useful one.[25] This diagnostic is somewhat language specific, since it involves extraposition in German. As shown by the contrast in (64), in German the extraposition of DPs is impossible, whereas the extraposition of CPs is possible:

64)   a.   Der Hans hat das Geld t_i zurückgegeben, [$_{CP}$ das er gestohlen hat]

the Hans has das Geld returned   that he stolen   has

'Hans has returned the money that he has stolen.'

b.   * Der Hans hat t_i zurückgegeben, [$_{DP}$ das Geld das er gestohlen hat]

the Hans has returned the money that he stolen has

'Hans has returned the money that he has stolen.'

Furthermore, the contrast in (65a-b) shows that the wh-pronoun in a free relative clause construction behaves differently with respect than the head in a headed relative.

65)   a.   * Der Hans hat was t_i zurückgegeben, [er gestohlen hat]$_i$

the Hans has what returned   he stolen   has

'Hans has returned the money that he has stolen.'

b.   Der Hans hat t_i zurückgegeben, [was er gestohlen hat]$_i$

the Hans has   returned   what he stolen has

'Hans has returned the money that he has stolen.'

If Donati is right, free relatives headed by complex wh-phrases are predicted to pattern with headed rather than headless relatives. This prediction is not confirmed, as shown by the following examples (Klaus Abels, personal communication).[26]

---

[25] A potential question that might arise here is whether the German extraposition data are compatible with the *Project Goal* account advocated here. Groos and van Riemsdijk's (1981) extraposition test was designed to discriminate between the Head and the Comp Account of free relatives. Since the *Project Goal* account is in a sense a combination of both accounts, it does not seem to be incompatible with these facts.

[26] The following examples illustrate the same point (Dirk Bury, personal communication):

(i)   Ich werde   dir welches Auto auch immer du (dir)   aussuchst kaufen.
         I   will   you which car   also ever   you you-DAT choose   buy
         'I'll buy you whichever car you choose (for yourself).'

(ii)   Ich werde dir   kaufen welches Auto auch immer du (dir)   aussuchst.
         I   will you buy   which   car also ever   you you-DAT choose
         'I'll buy you whichever car you choose (for yourself).'

(iii)   * Ich werde dir   welches Auto kaufen auch immer du (dir)   aussuchst.

66)   a.   * Der Hans hat welches Geld t_i auch immer zurückgegeben [er gestohlen hat]$_i$

the Hans has which   money also always returned   he stolen   has

'Hans has returned whichever money he has stolen.'

    b.   Der Hans hat t_i zurückgegeben [welches Geld auch immer er gestohlen hat]$_i$

the Hans has returned   whichever money also always he stolen has

'Hans has returned whatever money he has stolen.'

There are two other diagnostics that can distinguish headed from free relatives, and with respect to both of these diagnostics free relatives with *ever* pattern with free relatives without *ever* rather than headed relatives. These are the presence (or absence of the so-called matching effects) and the presence or absence of overt complementizers.   Let us consider them in turn.   As shown by the contrast between (67a) and (67b), free relatives differ from headed ones with respect to matching.[27] What is crucial (and well-documented in a wide range of languages), headed relatives do not show matching effects. The ungrammatical example in (67a) is a non-matching free relative; the grammatical one in (67b) is its headed counterpart, in which the head of the relative and the relative pronoun bear different cases.[28]

67)   a.   * Zatrudnimy [komu   ufamy t_{DAT}]$_{ACC}$.   [Pol]

hire-1pl   who-DAT   trust-1pl

'We'll hire whom(ever) we trust.'

    b.   Zatrudnimy tego   człowieka, któremu ufamy t_{DAT}.

hire-1pl   this-ACC man-ACC   who-DAT trust-1pl

'We'll hire the man that we trust.'

As shown by the contrast in (68a-b), free relatives with *ever* parallel simple free relatives in that they also require matching.

68)   a.   Zatrudnimy [któregokolwiek studenta   nam polecisz t_{ACC}]$_{ACC}$.   [Pol]

hire-1pl   whichever-ACC student-ACC us   recommend-2sg

'We'll hire whichever student you recommend to us.'

    b.   * Zatrudnimy [któremukolwiek studentowi   ufamy t_{DAT}]$_{ACC}$.

hire-1pl   whichever-DAT   student-DAT trust-1pl

'We'll hire whichever student we trust.'

I   will   you which   car   buy   also ever   you you-DIRE choose
'I'll buy you whichever car you choose (for yourself).'

[27] Even though non-matching free relatives are possible in some languages, I take matching in free relatives to be the norm (see, for example, Vogel (2002) and Pittner (1991) for a discussion of non-matching free relatives in some German dialects, and Daskalaki (2006) for a discussion of non-matching free relatives in Greek).

[28] The lack of matching effects in headed relatives is something that does not follow straightforwardly from a head promotion account without extra assumptions or stipulations. See Kayne (1994), Bianchi (1999) for ways to explain the lack of matching effects in headed relatives on the head promotion account.

With respect to compatibility with overt complementizers, free relatives with *ever* also parallel free relatives without *ever* rather than headed ones. The contrast between (68a) and (68b) shows that free relatives do not allow overt complementizers, whereas headed ones freely allow them. The ungrammaticality of the example given in (69c) would be unexpected if it involved a headed (rather than a free) relative.

69)   a.   * We'll hire whom that you recommended to us.

       b.   We'll hire the man that/whom you recommended to us.

       c.   * We'll hire whichever man that/whom you recommended to us.

These diagnostics show that free relatives headed by simple wh-pronouns and the ones headed by complex wh-pronouns behave similarly. Therefore, treating the former as true free relatives (involving *Project Goal* derivations) and the latter as headed relatives does not seem to be on the right track.

4.3   *Project Both* Probe and Goal

In this section, I propose that *Project Both* can solve a number of well-known and often discussed issues with head movement. Let us start by outlining some of these issues. First, head movement is unlike phrasal movement in that it violates the Extension Condition of Chomsky 1995. Second, the moved head does not c-command its trace. This is illustrated by the movement of V to *v* in (70); the movement of the V to *v* does not extend the tree, and the moved V does not c-command its trace:[29]

70)
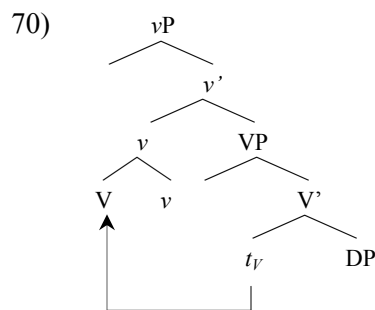
There are a number of proposals designed specifically to remedy these two (obviously related) problems with head movement. They fall into three general categories. One is to relegate head movement to the PF component of the grammar (Chomsky (2000), Boeckx and Stjepanovic (2001)). The other is to reanalyze it as phrasal movement (Koopman and Szabolcsi (2000)). And the third one is to reconceptualize verb movement in a way that makes it compatible with the Extension Condition (Matushansky (2006), Bury (2003), Bobaljik and Brown (1997), Suranyi (2005)). The proposal I would like to develop here falls into the third category and is close in spirit to Bury's (2003), Matushansky's

---

[29] Head movement is not the only case of countercyclic movement: Richards' (2001) tucking-in type movements and most cases of covert movements are also countercyclic.

(2006), and Suranyi's (2005) proposals. I will comment on the other two types of proposals after presenting the gist of my proposal.

There are two crucial ingredients to the analysis of verb movement I would like to propose here. One is that verb movement involves *Project Both* (Probe and Goal), and the other one is that head movement does in fact obey the Extension Condition by always targeting the root of the tree. This is illustrated in (68); I assume that *v* has an uninterpretable feature and V has a corresponding interpretable feature. Agree provides values to these features. In addition, *v* has an EPP feature which is what drives V movement.[30] The movement illustrated in (71b) departs from the 'standard' countercyclic views of head movement; note that the moving head targets the root node.

71)   a.

v

$v_{[uF,EPP]}$ V

V$_{[iF]}$ D

   b.

{V,v}

V v

v V

t_V D
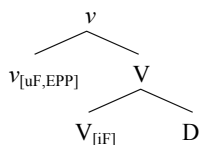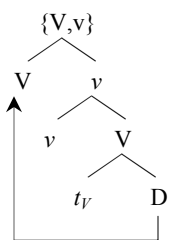
I further assume that lexical items enter the derivation fully inflected, following Chomsky (1995), which means that a head adjunction configuration is not a necessary input to the Spell-Out of individual words. What is relevant is that the moved head enters in a local relationship with a projection of V, not that it forms a head adjunction structure with it. Subsequent verb movement to T yields the following structure:

72)   {V,T}
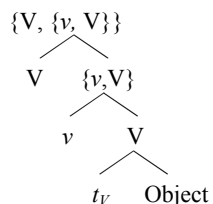
V   T

T   {V,v}

V v

v V

t_V D

---
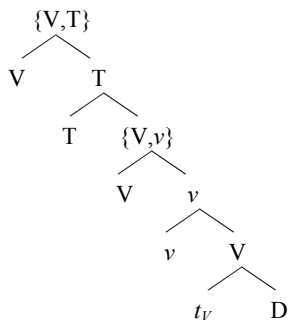
[30] This is following Alexiadou and Anagnostopoulou (1998) and (2001), who argue that EPP features can be checked either by head or phrasal movement. Furthermore, I assume a generalized approach to EPP, on which other heads besides T can have an optional EPP feature.

There are two immediate questions that the *Project Both* analysis of verb movement (head movement in general) raises. One involves its interaction with *(External) Merge* of (external) arguments, and the other one its compatibility with the idea that extended projections of lexical heads also involve *Project Both*, which is what I argued for in Section 3.1.2 above. Thus, given the fact that merging *v* with V involves projection of both labels, subsequent movement of V would result in the following multi-labeled structure.[31]

73) a. Merge *v* and V, *Project Both*

$$\{v,V\}$$
$$v \qquad V$$
$$t_V \qquad \text{Object}$$
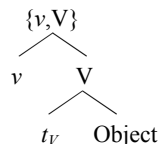
b. Move V to *v*, *Project Both*

$$\{V, \{v, V\}\}$$
$$V \qquad \{v,V\}$$
$$v \qquad V$$
$$t_V \qquad \text{Object}$$

I furthermore assume that *Internal* and *External Merge* can interweave. What this means for the case at hand is that the subject can get externally merged *after* V-*v* movement. While this departs from the standard view that *Internal Merge* follows *External Merge*, I do not see anything that would exclude it on principled grounds. Furthermore, it parallels the move from proposals that covert movement follows overt movement to proposals that the two can also interweave (see, for example, Fox and Nissenbaum 1999, for concrete arguments in favor of this view).

74)

$$\{V, \{v, V\}\}$$
$$\text{Subject} \qquad \{V, \{v, V\}\}$$
$$V \qquad \{v,V\}$$
$$v \qquad V$$
$$t_V \qquad \text{Object}$$
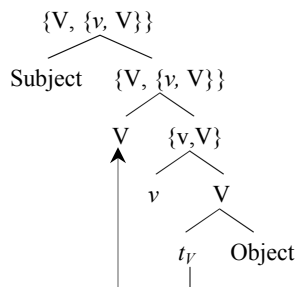
This proposal bears some resemblance to Bury's (2003) and Matushansky's (2006) analysis of head movement. Bury (2003), also working within the framework of Chomsky's Bare Phrase Structure, argues that verb movement involves *Project Goal* along the lines schematized in (80). I share Bury's intuition

---

[31] Note that the set $\{V, v\}$ is different from the set $\{v, \{v,V\}\}$.
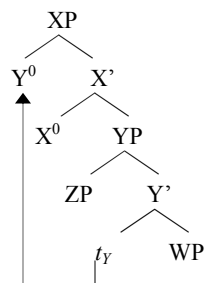
that the target does not always project in movement constructions. I depart from him, however, in what

the label of the resulting structure is. For Bury, the result of verb movement is a verb phrase projection,

as shown in (75), which is essentially equivalent to a Larsonian VP shell.

75)

```
        VP
       /  \
      V    vP
      ↑   /  \
         v    V
            /  \
          t_V   Object
```
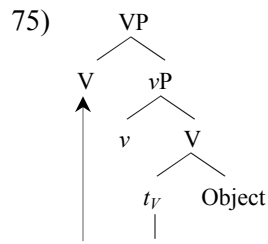
The result of subsequent verb movement to T and C is also going to be a VP, which seems somewhat

counterintuitive. On the current view, the information about the target of head movement is preserved in

the label of the new constituent. The advantage of the *Project Both* analysis is that the result of verb

movement (to T and to C) is a projection of T and C (and V).[32]

Matushansky (2006) proposal reanalyzes head movement as a composite of two independently

attested operations: movement to a specifier position and morphological merger, illustrated in (76a-b)

respectively (Matushansky 2006:81)

76)   a.   Step 1: movement of Y to [Spec,XP]

```
              XP
             /  \
           Y^0   X'
           ↑    /  \
              X^0   YP
                   /  \
                  ZP   Y'
                      /  \
                    t_Y   WP
```

    b.   Step 2: morphological merger

```
              XP
             /  \
           X^0   YP
          /  \   /  \
        Y^0  X^0 ZP  Y'
                   /  \
                 t_Y   WP
```
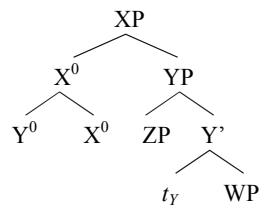
If head movement targets a specifier and if m-merger is an independent operation, it is not clear

why head movement should be subject to stricter locality conditions than phrasal movement (Head

Movement Constraint versus Island Constraints). Evidence seems to suggest that all cases of head

movement (now conceived as movement to a specifier position) have to be followed by m-merger.

Otherwise, it would not be clear how to exclude excorporation. Furthermore, in a system with early

---

[32] Suranyi's (2005) proposal avoids this problem.

vocabulary insertion, which is what I am adopting here, m-merger is not necessary, at least for the cases under consideration.

To conclude this section, I would like to comment briefly on the differences between my proposal and the two alternatives mentioned above, PF movement and remnant movement analyses of head movement.   As pointed out by Matushansky (2006), Zwart (2001), Suranyi (2005) among others, relegating head movement to a PF component adds redundancy to the system in that the same mechanisms of displacement are duplicated in both the syntactic and phonological component. Furthermore, if head movement were a purely phonological process, it should not have any syntactic or semantic consequences. With respect to syntactic consequences, Suranyi (2005) makes a convincing argument against PF treatments of head movement based on the fact that it participates in a number of syntactic correlations, which is unusual for PF movement.   The generalizations sensitive to head movement are the well-known Holmberg's Generalization and perhaps a somewhat lesser known Vikner's Generalization, which correlates the availability of Transitive Expletive Constructions with both V to T movement and V2.

One of Chomsky's main arguments in favor of treating head movement as phonological came from the fact that it does not have any semantic effects.   While it might be true that there are no differences in interpretation between raised and non-raised verbs, one cannot infer from it that there are no semantically significant instances of head movement.   Furthermore, as also pointed out by Matushansky (2006), if semantic vacuity were the sole diagnostic of PF movement, all movements resulting in reconstruction would have to reclassified as PF movements.

Lechner (2005) provides convincing arguments that there are cases of semantically active head movement (*S(*emantically) *A(*ctive) *H(*ead) *M(*ovement) Conjecture)).   His evidence comes from the so-called scope splitting cases, illustrated below, in which the modal has wide scope over the quantified subject, as shown by the paraphrases of the (a) examples given in (b).

77)   a.   Not every pearl can be above average size.   (Lechner 2005:3)

      b.   It is not possible that every pearl is above average size.

78)   a.   Not everyone can be an orphan.

      b.   It is not possible that everyone is an orphan.

79)   a.   Not every boy can make the basketball team.

      b.   It is not possible, that every boy makes the basketball team.

All the (a) examples in (77-79) all involve the following configuration (Lechner 2005:2):
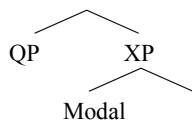
80)   Syntax: Interpretation:
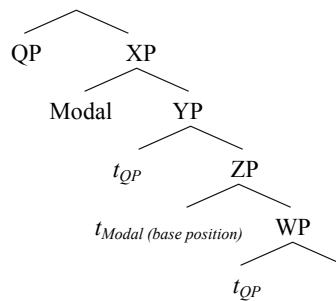
The modal is base-generated in a position that Lechner shows to be lower that the reconstructed position of the subject.   Furthermore, the wide scope interpretation for the modal cannot be a result of reconstructing the subject. This is shown by the examples in (81a-c), which show that strong DPs, including negative ones, do not reconstruct (Lasnik 1999:205).

81)   a.   No one is certain to solve the problem.

        b.   Every coin is 3% likely to land heads.

        c.   No large Mersenne number was proven to be prime.

For Lechner (2005), the reason why the modal has wide scope over the subject is that it undergoes LF movement to a position above the reconstructed position of the subject. This evidence in favor of LF movement is incompatible with the views that treat head movement as a purely phonological phenomenon.

An alternative solution would to reanalyze head movement as phrasal movement, or more precisely, to get the effects of head movement from phrasal movement. This is what Koopman and Szabolcsi (2000) do. While their proposal is developed primarily to deal with verbal complexes in Hungarian and Dutch, it should be generalizable to more paradigmatic cases of head movement. On the remnant movement approach, the derivation of the French word order *Subject V Adverb/Neg Obj* would have to involve movement of the object to some position below the adverb, as shown in (82b). If adverbs are VP adjuncts, it is not clear what that position would be. And the second step would have to involve remnant movement of the VP containing the trace of the object to some position above the Adverb, as shown in (82c).

82)   a.   Subject Adverb [$_{VP}$ V Object]

        b.   Subject Adverb Object [$_{VP}$ V t_{obj}]

        c.   Subject [$_{VP}$ V t_{obj}] Adv Object t_{VP}

This idea might be a good solution for some movement types (and word order alterations). However, it seems somewhat ad hoc for paradigmatic cases of head movement. As has been pointed out by many researchers, it results in unnecessary proliferation of unnecessary functional projections whose sole purpose seems to be to provide landing sites for various moved elements.

To sum up the discussion in this section, I have tried to show that while relegating head movement to PF or reanalyzing it as phrasal remnant movement avoids violations of the Extension Condition, it raises its own non-trivial issues. The *Project Both* account of head movement I have argued for in this section avoids these issues. Unlike in PF accounts, the movement is syntactic and thus is predicted to have syntactic (and semantic) consequences. Unlike in remnant movement accounts, the movement does not require any extra steps or empty projections whose sole purpose is to provide landing spots for moved constituents. A further confirmation that the movement posited here might be on the right track comes from the fact that it can explain why head movement is restricted to extended projections. Thus, V heads raise only as far as C (passing through *v* and T on their way up) and N heads raise to D. As pointed out by Grimshaw (1991)/(2005) in a different context, the highest projection in an extended projection of a given lexical head typically provides a limit of how far this head can move. This restriction follows the fact that *Project Both* is only possible if the two projecting elements do not conflict in categorial features. And we saw in Section 3.1 that this is only possible within an extended projection.

5.   Project None

The current proposal departs from Collins (2002) and Seely (2006), who argue in favor of a total elimination of labels. For them, the only option is *Project None.* The main conceptual argument against labels comes from the fact that they violate the Inclusiveness Condition. However, if a label is thought of as simply a copy of one of the merged elements, inclusiveness is not violated (no new information is added; copying is an independently motivated operation). Thus, labels are *not* a priori undesirable, and the system proposed here, which does not impose any labeling restrictions, is to be preferred on minimalist grounds.

Chomsky (1995) also argues that labels are necessary at the interface since 'verbal and nominal elements are interpreted differently at LF and behave differently in the phonological complement.' The work that labels did in a system with labels is done in a label-free system by the Locus Principle of Collins 2002:

83)   'Let X be a lexical item that has one or more probe/selectors. Suppose X is chosen from the lexical array and introduced into the derivation. Then the probe/selectors of X must be satisfied before any new unsaturated lexical items are chosen from the lexical array. Let us call X the locus of the derivation.' (Collins 2002:46)

The crucial difference between a locus and a label is that there is only one locus in a given structure/at a given derivational step, but there are many labels. The following elements can serve as probes (or selectors): theta roles, phi-features, Case features, EPP features, and subcategorization features.

Seely (2006) deduces the lack of labels from Chomsky's (1995) definition of a syntactic term. He argues that no operation makes reference to labels, no operation can change or rearrange or delete

labels because labels are syntactically inert.[33] Seely (2006) argues that labels are not terms, according to the definition of a term from Chomsky 1995, given below:

84)   The label X is never derivationally c-commanded (by any category Y)

      i.   For any structure K

      ii.   K is a term of K

      iii.   If L is a term of K, then the members of the members of L are terms of K (Chomsky 1995)

In (85) below, the entire object is a term. The first member of this set, *the*, is not a member of a member, thus is not a term. The second member of the set has two members, thus these two members are terms.

85)   {the, {the, picture}

Similarly, after we merge *see* with it, the entire object is a term. The members of the first and second objects now become terms. The label <u>see</u> is not a term.

86)   {<u>see,</u> {see, {<u>the,</u> {the, picture}}}

Only terms can be derivationally c-commanded (and for Seely this is the only kind of c-command), which gives Seely the welcome result that labels are syntactically inert. Epstein's (1999) definition of derivational c-command is given in (87).

87)   *Derivational definition of c-command* (Epstein 1999:329)

      X c-commands all and only the terms of the category Y with which X was paired (by Merge or by Move) in the course of the derivation.

However, there seems to be some circularity in Seely's argument. Seely notes a problem with subcategorization defined in a system without labels, which leads him to redefine the notion of a term. However, by this redefined notion of the term, labels would count as terms as well. To see the problem, consider the label-free representation given in (88).

88)   {see, {the, picture}}

The problem with (88) is that *see* doesn't derivationally c-command *the* and *picture*. Neither *the* nor *picture* are terms of {the, picture} according to the definition of term given in (84). Consequently, *see* does not (derivationally) c-command *the* and *picture*, since *the* and *picture* are not terms of the category with which *see* was merged. This leads Seely (2006) to propose a modified definition of a term, which is given in (89) (Seely 2006:201):

89)   For any structure K,

      i.   K is a term of K

      ii.   the members of K are terms of K

---

[33] If Seely (2006) is right in that labels are not terms, Hornstein and Uriagereka's (2002) reprojection theory, which allows labels of the same constituent to change during the course of the derivation, cannot be right.

      iii.   the members of the members of K are terms of K

According to this definition, *the* and *picture* in (85) are terms. However, according to the same definition, the label in a representation with labels would be a term as well.   Consider the following structure:

90)   {<u>see</u>, {see, {<u>the</u>, {the, picture}}}

According to Seely's revised definition, however, the labels ARE terms, which takes away the original argument in favor of eliminating labels.

      Another argument in favor of labels comes from the fact that they reduce the computational load. During the derivation, we only need 'access' to the label of the complement without having to look into its internal structure. This opens up an interesting possibility that *Project None* might be an option in the last instance of Merge, where the issue of accessing the label for future Merge operation does not arise. Since the empirical evidence of favor of allowing Project None in this one case, I will leave it as an open possibility, to be investigated further.

6.   Conclusion

To conclude briefly, this paper has provided both theoretical and empirical arguments in favor of new labeling possibilities. It has examined all the logical labeling possibilities, for both *External* and *Internal Merge*, and explored the consequences of the hypothesis that all of these possibilities are in fact attested. The tables in (91a-b) summarize the empirical results of the paper.

91)   a.   Missing Labels in *External Merge* structures

	Project Both	*Project None*
External Merge	• extended projections • comparative conditionals	• possibly final Merge

     b.   Missing Labels in *Internal Merge* structures

	Project Goal	*Project Both*	*Project None*
Internal Merge	• free relatives	• head movement	• possibly final Merge

The existence of the 'missing labels', I believe, is a welcome result from the perspective of the strong minimalist thesis. In an optimal system, there is no room for restrictions or constraints on what counts as a possible label and what does not.

      There is one issue that the unrestricted system of labeling argued for here might raise.   In conclusion, let me briefly address this issue. It concerns the potential complication of the computational system it introduces by allowing a lot of extra labeling possibilities. The standard labeling system, in which only one label can project is arguably more restrictive. However, from the perspective of the

strong minimalist thesis, a system that imposes restrictions on labeling (unless they can all be shown to be imposed by the interfaces) is less optimal than a system that does not.

References

Alexiadou, A. & E. Anagnostopoulou. 1998. 'Parametrizing AGR: Word Order, Verb-movement and EPP checking,' Natural Language and Linguistic Theory 16.3: 491-539.

Alexiadou, A. & E. Anagnostopoulou. 2001. 'The Subject-in-Situ Generalization and the Role of Case in Driving Computations,' Linguistic Inquiry 32.2, 193-231.

Bayer, J. 1999. Final complementizers in hybrid languages. Journal of Linguistics, 35, 233-271.

Beck, S. 1997. 'On the semantics of comparative conditionals,' Linguistics and Philosophy 20. 229–71.

Bhatt, R. 1999. Covert modality in non-finite contexts, unpublished doctoral dissertation, University of Pennsylvania.

Bhatt, R. 2003. 'The locality in correlativization,' Natural Language and Linguistic Theory 21:3, 485–541.

Bianchi, V. 2000. The raising analysis of relative clauses: reply to Borsley. Linguistic Inquiry 31:123-140.

Błaszczak, J. 1999. '-Kolwiek pronouns in Polish: negative polarity items or free choice items or both, In Böttger, K., M. Giger and B. Wiemer (eds). Beiträge der Europäischen Slavistischen Linguistik (POLYSLAV 2). [Die Welt der Slaven, 3]. Verlag Otto Sagner: München.

Bobaljik, J. and Brown. 1997. Interarboreal Operations: head Movement and the Extension Requirement. Linguistic Inquiry 28.2, 345-356.

Boeckx, C, and S. Stjepanović. 2001. 'Head-ing toward PF,' Linguistic Inquiry 32:345–355.

Borsley, R. 2004. The view from the Polish periphery. Annual Workshop on Formal Approaches to Slavic Languages: the Ottawa Meeting 2003. O. Arnaudova, W. Browne, M. L. Rivero, D. Stojanović (eds.), pp. 59-91, Michigan Slavic Publications.

Bury, D. 2003. Phrase Structure and derived heads. University College London. Doctoral dissertation.

Chomsky, N. 1995. The Minimalist Program. Cambridge, MA: MIT Press.

Chomsky, N. 2000. Minimalist inquiries. In R. Martin, D. Michaels and J. Uriagereka. (Eds.) Step by step: Essays on minimalist syntax in honor of Howard Lasnik. (pp. 89-156.) Cambridge, Mass.: MIT Press.

Chomsky, N. 2001. Beyond Explanatory Adequacy. MIT Occasional Papers in Linguistics 20.

Chomsky, N. 2001. Derivation by phase. In M. Kenstowicz. (Ed.) Ken Hale: A life in language. (pp. 1-52.) Cambridge, Mass.: MIT Press.

Chomsky, N. 2005. 'On Phases,' unpublished manuscript, MIT.

Chomsky, N. 2006. 'Approaching UG from below,' unpublished manuscript, MIT.

Citko, B. 2000. Parallel Merge and the Syntax of Free Relatives. Unpublished doctoral dissertation. Stony Brook University.

Citko, B. 2005. 'On the Nature of Merge: External Merge, External Merge, and Parallel Merge,' Linguistic Inquiry 36:475-497.

Citko, B. 2006. 'Small clauses: not so small and not all alike,' Unpublished manuscript, University of Washington.

Collins, C. 1997. Local Economy. MIT Press, Cambridge, MA.

Collins, C. 2002. Eliminating Labels. In Samuel Epstein and Daniel Seely (eds.), Derivation and Explanation in the Minimalist Program. Blackwell.

Culicover, P., and R. Jackendoff. 1999. The view from the periphery. Linguistic Inquiry 30: 543–571.

Dayal, V. 1995. 'Quantification in correlatives,' in E. Bach, E. Jelinek, A. Kratzer, and B. H. Partee, eds., Quantification in Natural Languages, Studies in Linguistics and Philosophy 54, Kluwer Academic Publishers, Dordrecht, Reidel, 179–205.

Den Dikken, M. 2005. 'Comparative Correlatives Comparatively,' Linguistic Inquiry 36:497-533.

Donati, V. 2006. 'On Wh-Head Movement,' In Wh-Movement. Moving On, ed by Lisa Lai-Shen Cheng and Norbert Corver, pp. 21-46. Cambridge, MA: MIT Press.

Epstein, S.D, E. Groat, R. Kawashima, and H. Kitahara. 1998. A derivational approach to syntactic relations. New York/Oxford: Oxford University Press.

Fox, D. and J. Nissenbaum. 1999. Extraposition and Scope: A case for overt QR. WCCFl 18.

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

Groos, A., and H. van Riemsdijk. (1981). Matching effects in free relatives: a parameter of core grammar,' In Theory of Markedness in Generative Grammar: Proceedings of the IVth GLOW Conference (A. Belleti, L. Brandi, and L. Rizzi, eds.), pp. 171-216, Scuola Normale Superiore di Pisa, Pisa.

Hornstein, N., J. Nunes, and K. Grohmann. 2005. Understanding Minimalism. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Hornstein, N. and J. Uriagereka. 2002. Reprojections. In Derivation and Explanation in the Minimalist Program. Ed by S. D. Epstein and T. Daniel Seely, pp. 106-132. Oxford/Malden: Blackwell Publishing.

Iatridou, S., E. Anagnostopoulou and R. Izvorski (2001) Observations about the form and meaning of the perfect. In M. Kenstowicz. (Ed.) Ken Hale:A life in language. (pp. 189-238.) Cambridge, Mass.: MIT Press.

Izvorski, R. 1996. 'The syntax and semantics of correlative proforms, In K. Kusumoto, ed., Proceedings of NELS 26, Amherst, Massachusetts, GLSA, 133–147.

Koopman, H. and A. Szabolcsi. 2000. Verbal Complexes. Cambridge, Mass.: MIT Press.

Larson, R.K. 1987. 'Missing prepositions' and the analysis of English free relative clauses. Linguistic Inquiry 16: 239-266.

Larson, R.K. 1998. Free relative clauses and missing P's: reply to Grosu. Unpublished manuscript, Stony Brook University.

Lasnik, H. 1999. Chain of arguments. In S. Epstein and N. Hornstein (eds.), Working Minimalism, pp. 189-215. Cambride, Mass: MIT Press.

Lechner, W. 2005. Interpretive effects of head movement. Ms., University of Tubingen.

Liptak, A. 2004. On the correlative nature of left-peripheral relatives. ZAS Working Papers in Linguistics 35: 287-313.

Matushansky, O. 2006. Head-movement in linguistic theory. Linguistic Inquiry 37:69-107.

McCawley, J. 1988. The comparative conditional constructions in English, German and Chinese. In Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society. Berkeley: BLS. 176–87.

McCawley, J. 1998. The syntactic phenomena of English (2nd edn). Chicago: University of Chicago Press.

Moro, A. 2000. Dynamic Antisymmetry. Cambridge: Mass: MIT Press.

Pereltsvaig, A. 2001. On the Nature of Intra-Clausal Relations. Unpublished doctoral dissertation. McGill University.

Pesetsky, D. and E. Torrego. 2001. "T-to-C Movement: Causes and Consequences (with Esther Torrego, In Michael Kenstowicz (ed.), Ken Hale: A Life on Language. Cambridge, MA: MIT Press.

Pesetsky, D. and E. Torrego, 2006. Probes, Goals and Syntactic Categories, to appear in Proceedings of the Seventh annual Tokyo Conference on Psycholinguistics (Y. Otsu, ed.)

Rackowski, A., and N Richards. 2005. Phase Edge and Extraction: A Tagalog Case Study. Linguistic Inquiry 36:565-599.

Richards, N. 2001. Movement in Language: Interactions and Architectures. Oxford University Press,

Rooryck, J. 2000. Configurations of sentential complementation. London: Routledge.

Seely, D. 2006. Merge, derivational C-command, and subcategorization. In C. Boeckx (ed). Minimalist Essays, pp. 182-217). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Sportiche, D. 1997. Reconstruction and constituent structure. Manuscript, UCLA.

Srivastav, V. 1991. 'The syntax and semantics of correlatives,' Natural Language and Linguistic Theory 9, 637–686.

Surányi, B., 2005. Head movement and reprojection. In: Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica. Tomus XXVI. Budapest: ELTE. 313-342.

Taraldsen, K. 2002. The Que/Qui alternation and the distribution of expletives. In Subjects, Expletives, and the EPP, Peter Svenonius (ed.). Oxford/New York: Oxford University Press.

Zwart, C. J-W. 2001. Syntactic and phonological verb movement. Syntax 4.1: 34-62.