Available online at www.sciencedirect.com
ScienceDirect
Lingua 118 (2008) 907–944
Lingua
www.elsevier.com/locate/lingua

Missing labels

Barbara Citko

Department of Linguistics, University of Washington, Box 354340, Seattle, WA 98195-4340, USA

Received 25 November 2006; received in revised form 7 January 2008; accepted 8 January 2008
Available online 9 April 2008

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 not exhaust all the possibilities. For *External Merge* of α and β, the options that do not violate Inclusiveness 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 these possibilities are in fact attested. I focus on the following, previously unattested, ones: *Project Both* in *External Merge*, and *Project Goal* and *Project Both* in *Internal Merge* structures.

Keywords: *Internal Merge*; *External Merge*; Labels; Comparative conditionals; Extended projections; Free relatives; Head movement

1. Issue

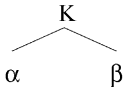
My main goal in this paper is to examine the labels created by Merge, the basic structure building operation within the Minimalist Program. Following Chomsky (2001a, 2001b), I assume there are two kinds of Merge. One is familiar from earlier minimalist literature; it takes two distinct objects and combines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

E-mail address: bcitko@u.washington.edu.

[1] Chomsky (2001a, 2001b) 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 discuss this distinction in section 3.1.1.

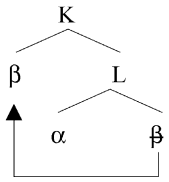
0024-3841/$ – see front matter
doi:10.1016/j.lingua.2008.01.001

(1) *External Merge* of α and β



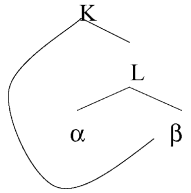
The second type of Merge, called *Internal Merge*, is like *External Merge* in that it also combines two objects. The only difference is that one of these two objects is a subpart of the other, as shown in (2).

(2) *Internal Merge* of β and L



Alternatively, the result of *Internal Merge* can be represented as (3). For the purposes of this paper, I will assume that the two are simply notational variants of one another.

(3)



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 either *Internal* or *External Merge*. This question raises three even more fundamental questions: (i) what are labels?, (ii) what kind of information do labels contain?, and (iii) are labels even necessary? The last question is particularly relevant in the context of recent work by Collins (2002) and Seely (2006), who argue that labels should be eliminated from the syntax altogether. I will discuss this possibility in section 5, where I show that it is indeed a valid option, but only under certain very restricted circumstances.

Merge is an asymmetric operation, projecting one of the objects to which it applies as the label of the newly formed object. The head of the projecting object becomes the label of the newly formed object (Chomsky, 1994:11, Epstein, 1999:341). On this view, a label is simply a set of features of one of the merged objects. The simplest assumption is that all features project (even though not all of them might be syntactically relevant). This is slightly different from what the X-bar theoretical notation might suggest labels are. In X-bar theoretic terms, labels were thought to represent two types of information: the object's phrase theoretical status (whether it is an X^0, X', or XP) and its category (whether it is a V or an N, for example). In Bare Phrase Structure theory, the object's phrase theoretical status is not a given, but can be derived from the structure: a category that does not project any further is a maximal category,

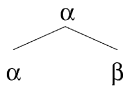
and a category that itself is not a projection of anything is a minimal projection (Chomsky, 1995:246).

In set theoretical terms, a label is represented as the first member of the set whose second member consists of the two objects that were input to Merge. In (4), this is γ.

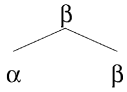
(4) $K = \{\gamma, \{\alpha, \beta\}\}$

The standard assumption is that in *External Merge* either α or β projects as the label γ, 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* structures, however, are not the only possibilities. For *External Merge*, the options are: (*i*) *Project* α, (*ii*) *Project* β, (*iii*) *Project Both* α and β, and (*iv*) *Project None*. These options are illustrated in (5a–d), respectively.[2]

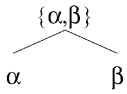
(5) a. *Project α*



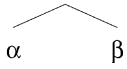
b. *Project β*



c. *Project Both* α and β



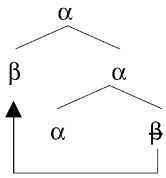
d. *Project None*



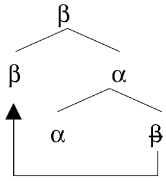
The logical possibilities for *Internal Merge* structures are illustrated in (6a–d). If α is a Probe and β a Goal, these are: (*i*) *Project Probe*, (*ii*) *Project Goal*, (*iii*) *Project Both Probe and Goal*, and *Project None*.

[2] Saying that α is the label of the object given in (5a) or (6a) is equivalent to saying that the features of α have projected to the next level. What this means is that in *Project Both* cases the features of both merged constituents have projected to the next level.

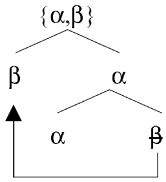
(6) a. *Project Probe*



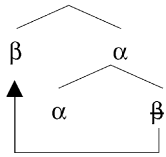
b. *Project Goal*



c. *Project Both Probe and Goal*



d. *Project None*



The hypothesis I would like to put to test here is that all the possibilities illustrated in (5) and (6) above are in fact attested. This is my understanding of the logic behind Chomsky's suggestion that 'the labeling algorithms apply freely, sometimes producing deviant expressions (Chomsky, 2005:11). I will focus on the 'missing' labels illustrated in (5c–d) and (6b–d) above. Crucially, I am using the term 'missing label' to refer to previously unattested labels, not to label-less projections. I will proceed as follows. In section 2, I will review what I take to be standard minimalist assumptions concerning labels, on which only a subset of the possibilities considered here exists. In section 3, I will discuss labels in *External Merge* structures, and show that *Project Both* captures some otherwise puzzling properties of comparative correlatives and Grimshaw's (1991) extended projections. In section 4, I will turn to labels in *Internal Merge* structures, and offer new evidence in favor of the idea, going back to Larson, 1998 (see also Bury, 2003; Donati, 2006; Iatridou et al., 2001), that free relatives involve a *Project Goal* derivation. I will also show that *Project Both* can solve some of the long-standing problems concerning head movement. I will conclude the paper with some stipulations concerning the *Project None* option, and a comparison of my proposal with alternatives that eliminate labels altogether, such as that of Collins (2002) and Seely (2006).

2. Standard minimalist approach to labels

2.1. Labels in External Merge structures

Merge is an operation that applies to two objects (α and β), and forms a new object out of them.[3] The result of Merge should thus be a simple set consisting of α and β. Chomsky in Bare Phrase Structure, however, argues that this is not sufficient due to the output conditions, and that the result of Merge also needs to include a label. Labels, he says, are necessary because 'verbal and nominal elements are interpreted differently at LF and behave differently in the phonological component' (Chomsky, 1994:396). The result of Merge thus has to be a more complex object, given in (4), in which γ represents a label.

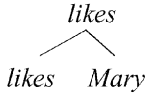
(7) $K = \{\gamma, \{\alpha, \beta\}\}$

In order to avoid a violation of the Inclusiveness Condition, which prohibits introduction of new elements in the course of the derivation, 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 and 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', which is often the case (Chomsky, 1995:244). The only choice is then for either α or β to project as the label. This opens up a possibility that *Project Both* could be possible as long as the two projecting elements do not conflict in relevant features.[4] I will explore the empirical consequences of this possibility in section 3.1 for *External Merge*, and in section 4.3 for *Internal Merge*.[5]

This reasoning raises the question of which of the two elements projects as the label. At a given derivational step, only one choice will result in a derivation that converges at the interfaces. To illustrate, let us consider a partial derivation of a simple transitive clause, given in (8a–e). For the sake of clarity, the labels are underlined and many details (such as cyclic Spell-Out, feature valuation, or movement to Spec, TP) are omitted.

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

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

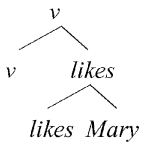


[3] A separate question is why *Merge* targets two elements (rather than one, zero, or more than two). For some relevant discussion, see Collins (1997), who argues that merging fewer objects than two is vacuous, whereas merging more than two is less economical than merging two.

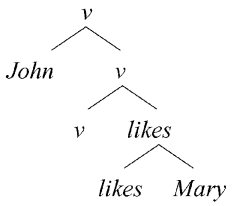
[4] Pereltsvaig (2007) argues this is precisely what happens in small clauses. See Citko (2008), however, for empirical arguments against treating small clauses in this way.

[5] Another possibility (or set of possibilities), brought to my attention by an anonymous reviewer, is that the label could also be the intersection of some subset of the properties of α and β. This is a very intriguing possibility; however, it raises the question of how to determine what the relevant subsets of features might be.

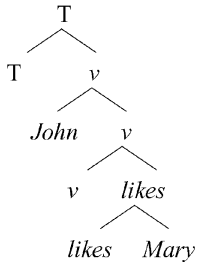
c. Merge *v* and {likes, {likes, Mary}}, Project *v* → {*v*, {*v*, {likes, {likes, Mary}}



d. Merge *John* and {*v*, {*v*, {likes, {likes, Mary}}}, Project *v* → {*v*, {John, {*v*, {*v*, {likes, {likes, Mary}}}}

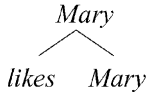


e. Merge T and {*v*, {John, {*v*, {*v*, {likes, {likes, Mary}}}}}, Project *T* → {*T*, {T, {*v*, {John, {*v*, {*v*, {likes, {likes, Mary}}}}}}



Let us focus on the first instance of Merge. If instead of the verb *likes*, the noun *Mary* projected as the label, resulting in (9), there would be no way for this derivation to continue, since the next element to be merged, *i.e. v*, selects a verbal not a nominal complement.

(9) Merge {*likes, Mary*} → {*Mary*, {*likes, Mary*}}



Chomsky (2005) suggests a more specific algorithm to determine labels, which is given in (10a–b). (10a) deals with *External Merge*, and (10b) with *Internal Merge*. For now, I will focus on (10a). I will come back to (10b) in the next section.

(10) a. In {H, α}, H an LI, H is the label (Chomsky, 2005:10)
b. If α is internally merged to β, forming {α, β}, then the label of β is the label of {α, β}

According to (10a), if one of the merged elements is a lexical item, it is going to be the element that projects as the label. A lexical item is an item that is selected directly from the Numeration (as

opposed to being a result of a previous Merge operation). From this perspective, both lexical and functional elements are considered to be lexical items. This part of the algorithm makes correct predictions for (8c) and (8e). In (8c), *v* is selected directly from the Numeration, and in (8e) T is; both thus count as lexical items. However, the algorithm makes incorrect predictions for (8b) and (8d). In (8b) both elements are selected directly from the Numeration, therefore either one should be able to project. In (8d) *John* is a lexical item, and thus is incorrectly predicted to project as the label.

Chomsky himself points out some of these 'exceptional' cases, which do not clearly follow from this algorithm. The initial stage of every derivation is problem, as we just saw, since either of the two merged elements is predicted to be able to project. Another problematic case involves merging two XPs, in which case neither is a lexical item. This is what takes place, for example, when a non-minimal external argument merges with *v*P.[6] This is a variant of the issue raised above concerning (8d).

Coming back to the choice between (8b) and (9), what distinguishes them is that in (8b) the projected constituent, the verb *likes*, had a selectional feature satisfied by the Merge operation. So perhaps the constituent that determines the label is the one that has some kind of a feature (or a set of features) satisfied via the Merge operation. The feature(s) in question could be selectional features, EPP features, phi-features, case features, wh-features, or any other features that drive syntactic computations.[7] This is similar in spirit to 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 as a Vehicle Requirement on Merge, given in (11):

(11) *Vehicle Requirement on Merge*
If α and β merge, some feature F of α must probe F on β.

Even though their primary focus was on the motivation behind (*External*) *Merge*, not on the labels created by *Merge*, it seems reasonable to interpret their proposal as saying that *the element that probes is the one that projects as the label.* In both *External* and *Internal Merge* cases, the element that projects as the label needs to have its unvalued features valued through the operation. While conceptually appealing, the Vehicle Requirement on Merge is not going to work for all cases of Merge. For example, it is not going to work for adjuncts. The result of adjoining a PP to VP is a VP, even though neither the VP nor the PP has had a feature valued through this operation.[8] Furthermore, the Vehicle Requirement on Merge seems to require a somewhat non-standard view of the *Agree* relationship, which is the structural relationship between Probes and Goals required for feature valuation. *Agree* is possible only in structures that have already been built. Here, it seems that *Agree* must happen before two elements merge, since it is *Agree* that determines whether they can merge, and which one of them will project.

[6] Chomsky (2005:11, *Ft* 11) suggests that maybe in such cases neither element projects. He further suggests, referring to Moro's (2000) work, that this is what might happen in small clauses, which involve merge of two XPs. It is not clear to me, however, how label-less constituents might be further integrated into the structure.

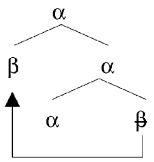
[7] Note that this crucially relies on the assumption that theta roles are not features but reflexes of syntactic configurations (as argued by Hale and Keyser (1993), for example). If *Mary* has a theta feature that gets checked via merge with a verb, both *likes* and *Mary* have had a feature satisfied. Thus, according to the Vehicle Requirement on Merge both should be able to project.

[8] This is not a problem if adjuncts are specifiers of designated functional projections (as argued by Cinque, 1999).

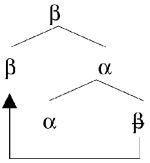
2.2. Labels in internal merge structures

The standard minimalist assumption regarding labels in *Internal Merge* structures is that the target of movement, or to use the more current term, the Probe always projects. This is illustrated in (12a). Chomsky (1995) excludes the other possibility, *Project Goal* (illustrated in (12b)) on the grounds that no feature checking could be established in such a configuration.

(12) a. Project Probe



b. Project Goal

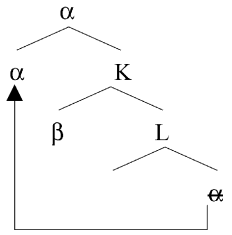


Movement is assumed to be morphologically driven, and to be possible only as a Last Resort operation. Chomsky considers three versions of Last Resort, given in (13a–c), and argues that *Project Goal* is incompatible with all three of them.

(13) 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

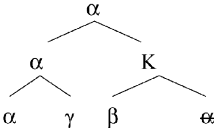
If α raises to target K, and projects as its label, the result is the object in (14).[9]

(14)



[9] There are two possible results here, depending on the status of α. If α is minimal, K ends up as its complement, as shown in (14) above. If α is not minimal, K will end up as its specifier, as shown in (i). In both cases, the phrase-theoretic status of α changes throughout the derivation.

(i)



Chomsky (1995) excludes the possibility of *Project Goal* due to the fact that the movement in (14) does not establish a specifier-head feature checking configuration. If β in (14) is the Probe, α is not a specifier of β after movement takes place. This objection, however, is not valid from the perspective of current minimalism, in which feature checking is done in situ, and a spec-head feature checking is eliminated and replaced by a Probe-Goal Agree relationship. Prior to movement in (14), α is in a 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).

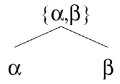
Interestingly, the labeling algorithm of Chomsky, 2005 seems to allow some of the labeling possibilities that were explicitly excluded on earlier minimalist assumptions, in particular the *Project Goal* option. The relevant algorithm is repeated in (15a and b); (15b) deals with *Internal Merge* cases.

(15) 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}

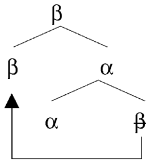
If α is a linguistic item, and β is not a linguistic item, the labeling algorithm makes conflicting predictions. According to (15a), α should be the label, but according to (15b), β should be the label. If the grammar can choose in such conflicting cases, choosing (15b) will yield *Project Goal*.

Having outlined what I take to be standard minimalist assumptions about labels in *External* and *Internal Merge* structures and having pointed out some problems with these standard assumptions, I turn to the main proposal of this paper, which is that all the labeling possibilities compatible with the Inclusiveness Condition are attested. In the remainder of this paper, I will provide empirical evidence in favor of the following ones: *Project Both* in *External Merge* structures (section 3), *Project Goal* and *Project Both* in *Internal Merge* structures (sections 4.2 and 4.3). These options are illustrated in (16a–c).[10]

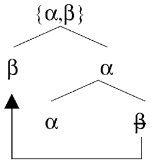
(16) a. *Project Both* α and β



b. *Project Goal*



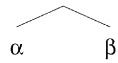
c. *Project Both Probe and Goal*



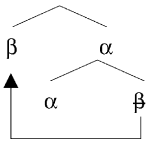
[10] As pointed out to me by one of the reviewers, *Project Both* in cases of total identity follows from the Axiom of Extensionality, which defines a set by its membership. According to this axiom, the set {α, α} is equivalent to {α}.

Note that the range of logical possibilities also includes the ones given in (17a–b), in which neither of the two input elements projects. I will set these aside till section 5.

(17) a. *Project None* in *External Merge* structures



b. *Project None* in *External Merge* structures



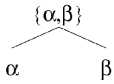
An immediate question that arises here is what determines what particular labeling option is taken at any given stage for any given derivation. I assume that labeling is essentially free, as long as it does not violate the Inclusiveness Condition. However, only a subset of possible choices will yield convergent derivations.

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. This opens up the possibility that the grammar allows the structure in (18) as long as α and β are of the same category.[11]

(18) *Project Both* α and β



In the following two subsections, 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).

3.1.1. Comparative correlatives

This section shows that the option of projecting the labels of both merged constituents can explain an otherwise mysterious extraction behavior of comparative correlatives. Examples of comparative correlatives from English, Hindi, Hungarian, and Polish are given in (19a–d).

[11] An anonymous reviewer raises the issue of how we can determine which of the two α's has projected in {α, {α α}}. This might not be possible to determine by simply looking at the resulting representation, and thus might serve as an interesting argument in favor of a derivational approach to syntax.

(19) a. The more you read, the more you understand.
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 (Den Dikken, 2005:499–500)
'The more you read, the more you understand.'
d. Im więcej czytasz, tym więcej rozumiesz. [Pol]
rel more read.2SG DEM more understand.2SG
'The more you read, the more you understand.'

Descriptively speaking, comparative correlatives, also referred to in the literature as comparative conditionals, have been analyzed as having correlative syntax and conditional semantics.[12,13] For example, the example given in (19a) above can be paraphrased as (20).

(20) If you read more, you will understand more.

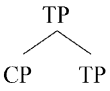
What distinguishes them from standard conditionals is the fact that they contain a comparative element in both clauses. Both standard and comparative correlatives are typically analyzed as

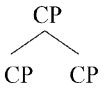
[12] The crosslinguistic distribution of comparative correlatives is wider than that of standard ones. English, for example, does not allow standard correlatives, but allows comparative correlatives.
(i) *What you like, this you eat. *standard correlative*
(ii) The more you smile, the happier you get. *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 (19b–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.'

[13] I am using the following abbreviations in the examples that follow:

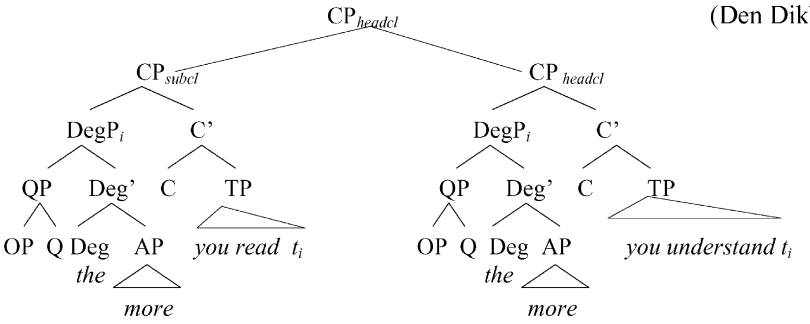
addess	addessive
instr	instrumental
f	feminine
acc	accusative
gen	genitive
vm	verbal marker
dat	dative
rel	relative
hab	habitual

involving the adjunction structure of the kind given in (21a) or (21b) (see Beck, 1997; Den Dikken, 2005; McCawley, 1988, 1998, among others).[14]

(21) a. 

b. 

More concretely, Den Dikken posits the following structure for the example given in (19a) above.

(22) 

(Den Dikken 2005:511)

The two CPs in (22), the main one (marked as CP_{headcl} in the diagram) and the subordinate one (marked as CP_{subcl} in the diagram) are distinguished by the featural make-up of their respective C heads. This might be hard to see in languages like English, in which the specifiers of both Cs contain the same string ('the more'), but the examples in (19b–d) above show that in languages such as Polish, Hindi, or Hungarian, the specifier of the subordinate C is filled by a relative pronoun, whereas the specifier of the matrix C is filled by a demonstrative pronoun. The featural make-up of the two C heads is thus different, since they require different elements in their specifiers.

What is relevant for our purposes is the prediction such an adjunction structure makes 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 have both standard and comparative correlatives: Polish, Hindi, and Hungarian.[15] I will first examine the extraction behavior of standard correlatives, and next I will turn to comparative correlatives.

The examples in (23–25) show that standard correlatives behave as expected with respect to wh-movement; they allow extraction from the matrix clause but not from the relative one. This follows straightforwardly from the structure in (21), in which the relative CP is adjoined to the matrix one. The ungrammaticality of the b examples can be reduced to the Adjunct Condition.

[14] Srivastav (1991), Dayal (1995), and Bhatt (2003) analyze the matrix clause in Hindi correlatives as a TP 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.

[15] 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.

(23) 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 does Jan 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 think that the one who likes Tolstoy should read?'

(24) 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?'

(25) 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]?
what heard Mari that REL-who likes Tolstoy-ACC that-DAT PV need read-INF
'What did Mary hear that the one who likes Tolstoy needs to read?'

Given the adjunction structure in (21), we would expect comparative correlatives to show a similar relative/matrix clause asymmetry with respect to wh-movement. Interestingly, this is not what happens; either both clauses allow extraction, or both disallow it. In other words, comparative correlatives show a symmetric extraction pattern. This is what Culicover and Jackendoff (1999) note with respect to English comparative correlatives[16]:

[16] Culicover and Jackendoff (1999) analyzed examples 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?
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?'

(26) 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 further notes that ATB-style movement out of comparative correlatives is also possible:

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

The three languages whose standard correlatives we have examined above, Polish, Hindi, and Hungarian, also do not show the expected adjunct/matrix asymmetry. There is an interesting contrast, however, 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 (28a–c) and (29a–c), respectively. Hungarian, on the other hand, disallows movement completely, as shown in (30a–c).[17]

(28) a. *Kogo*$_i$ myślisz, że [im lepiej poznasz t_i], [tym bardziej go polubisz]?[18] [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
'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.2SG that the better know.2SG the more like.2SG
'Who do you think that the more you get to know, the more you will like?'

(29) 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?'

[17] Den Dikken (2005) also notes that German and Dutch correlatives behave similarly to Hungarian with respect to extraction out of comparative correlatives as long as the comparative element in the matrix clause is fronted. If it is not fronted, extraction from the matrix clause becomes grammatical (Den Dikken, 2005:505). Abeillé et al. (2006) observe a similarly symmetric extraction pattern in Spanish and French comparative correlatives.

[18] Replacing the pronouns in (28a and b) with proper 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?'

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)?'

c. *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)?'

(30) a. * *Kiket*$_i$ hallott Mari, hogy [minél jobban ismerted öket], [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)?'

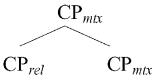
b. ???*Kiket*$_i$ hallott Mari, hogy [minél jobban ismerted t_i], [annál jobban
who-PL-ACC heard Mari that what-at better knew-2SG, that-at better
szeretted öket]?
liked-2SG they-ACC
'Who did Mari hear that the better you knew (them), the better you liked (them)?'

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)?'

The differences in extraction behavior between standard and comparative correlatives can be accounted for by allowing either the label of just the matrix CP or both the relative and the matrix CP to project. In particular, I propose that standard correlatives involve an 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 and the relative CPs project. This difference between the two is illustrated in (31a–b).

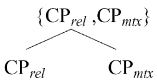
(31) a. Merge the relative and the matrix CP, Project the matrix CP [19]

standard correlatives



[19] The subscripts CP$_{mtx}$ and CP$_{rel}$ are used solely for expository purposes.

b. Merge the relative CP and the matrix CP, *Project Both* [20]



comparative correlatives

Crucially, the assumption that both CPs project does not imply that there are no asymmetries whatsoever between the two CPs in a comparative correlative. For example, the two clauses in a comparative correlative construction behave asymmetrically with respect to diagnostics such as question tag formation, illustrated in (32a–b).

(32) a. The more we eat, the angrier you get, don't you?
b. *The more we eat, the angrier you get, don't we?
(Culicover and Jackendoff, 1999:549)

The two clauses also behave asymmetrically semantically; the first clause modifies the second one. To tackle the issue of why comparative correlative exhibit both symmetric and asymmetric properties, I follow a suggestion made by Chomsky (2001a, 2001b) to handle a similar issue faced by *Pair Merge* structures.[21] *Project Both* proposed here is similar in spirit to Chomsky's *Pair Merge*, which also projects the labels of both merged constituents. Chomsky suggests that symmetric *Pair Merge* structures are converted to more standard asymmetric structures, which he refers to as *Set Merge* structures, at the point of *Transfer*, which is the point at which the derivation is 'shipped' to the Conceptual-Intentional and Perceptual-Articulatory interfaces. At the levels relevant for the purposes of linearization and semantic interpretation (PF and LF, respectively), the result is a standard asymmetric structure. However, for the purposes of narrow syntax processes, which I take wh-movement to be, there is no asymmetry between the two clauses.

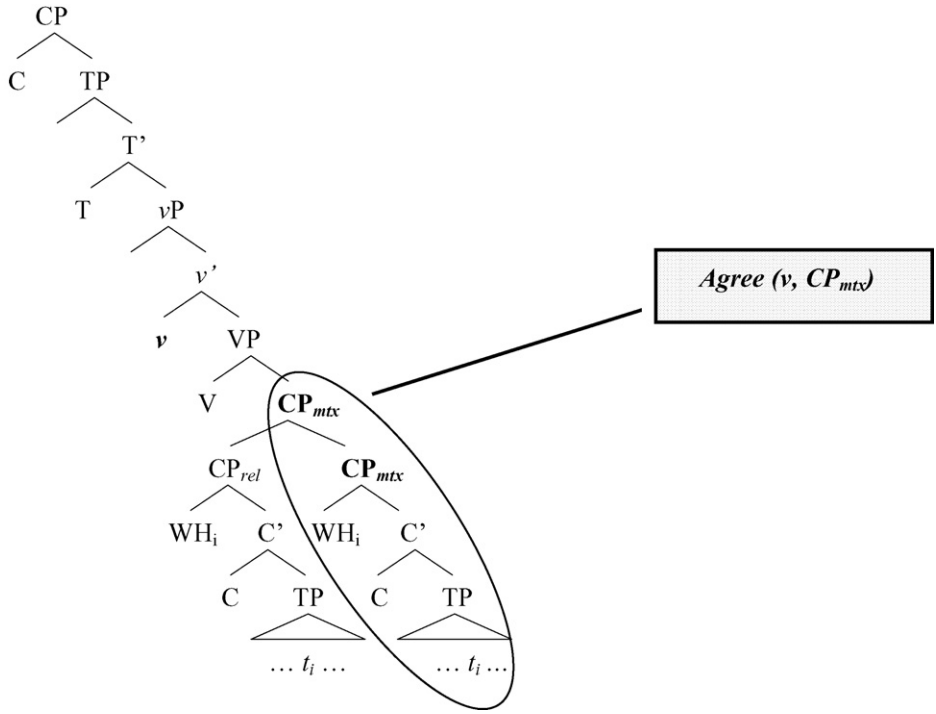
Given the difference in derivation between standard and comparative correlatives (*Project One* versus *Project Both*), we can account for the differences between them with respect to wh-movement. To do so, I adopt an independently motivated *Agree*-based version of the Adjunct Condition developed by Rackowski and Richards (2005). The crucial innovation in Rackowski and Richards's (2005) proposal lies in the idea that the CP out of which the wh-phrase moves has to (independently) undergo Agree with the matrix little *v*. It is this Agree relationship that makes the CP transparent for extraction. In standard correlatives, the matrix *v* undergoes Agree with the matrix CP (represented in (33) as CP$_{mtx}$). This makes the matrix CP (but not the relative one) transparent for extraction.[22]

[20] A reviewer raises the issue of whether the set {CP, CP} is distinct from a set consisting of just one CP. The two members of the set are categorically non-distinct, however, they are headed by different types of C heads. One is headed by a 'relative' C, *i.e.* a C that has a relative pronoun in its specifier. The two CPs are not non-distinct with respect to all of their features.

[21] Chomsky's *Pair Merge* and *Project Both Merge* employed here are not fully equivalent. Chomsky's *Pair Merge* creates an ordered set, whereas *Project Both Merge* proposed here creates an unordered set.

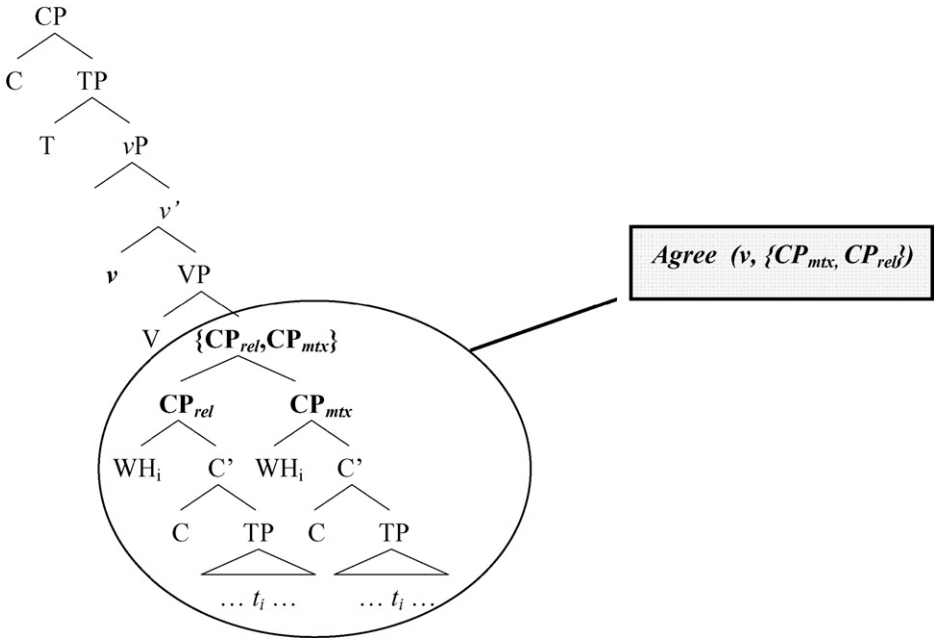
[22] I depart from Rackowski and Richards in one respect. For them, wh-movement proceeds directly from the specifier of the embedded *v*P to the specifier of the matrix *v*P, skipping the specifier of CP. Since nothing in my analysis hinges on this assumption, I will maintain a more 'traditional' view, on which wh-movement proceeds through the specifier of CP as well.

(33)



By contrast, in comparative correlatives, the matrix *v* undergoes *Agree* with both the matrix and the relative CP. This makes both of them transparent for extraction.

(34)



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, in which extraction from both clauses in comparative conditionals is blocked. What is significant is the fact that even in such 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. An idea I would like to suggest is that languages differ in whether they allow Agree with multi-labeled objects. If a language does not allow such more complex forms of Agree, extraction will be banned.

Wh-movement behavior of comparative correlatives is not the only property that does not follow from the standard adjunction structure. Another puzzle for the adjunction structure involves the optionality, or lack thereof, of the subordinate CP. The contrast in (35a–b) shows that in comparative correlatives, both CPs are required.

(35) a. The more you smile, the happier you get.
b. *The happier you get.

In standard correlatives (as well as standard conditionals), the subordinate clause can be omitted without affecting the grammaticality of the sentence. This is illustrated in (36a–b) for Polish correlatives, and in (37a–b) for English conditionals.

(36) a. Jak sobie pościelisz, tak się wyśpisz. [Pol]
how SELF make-bed DEM self sleep
'You've made your bed, now lie in it.'
b. Tak się wyśpisz.
dem self sleep
'You will sleep this way.'

(37) a. If you smile more, you will get happier.
b. You will get happier.

The lack of optionality also seems to suggest that correlatives do not involve a standard adjunction structure.

3.1.2. Extended projections as merge and project both

Allowing *Project Both* as one of the labeling options can also provide a fairly straightforward way to capture the main insights behind Grimshaw's (1991/(2005) theory of extended projections.[23] There are two issues that Grimshaw's extended projection theory addresses. One is why the combinatorial properties of functional heads are much more limited than those of lexical heads. In particular, why are the patterns given in (38) and (39) possible, whereas the logically possible ones in (40) are not (Grimshaw, 2005:9)?[24]

(38) Possible complements of functional heads

C-TP	P-DP
T-VP	D-NP

[23] This is what Hornstein et al. (2005) pose as an open-ended question. This section can be seen as an extended answer to this question.

[24] I have updated somewhat Grimshaw's notation, by replacing I's with T's.

(39) 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

(40) 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		

What (38–40) show is that 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, not the other way round.

The second issue 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 noun (not the determiner). The fact that there are no idioms consisting of verbs and determiners provides further evidence that there is no direct selection between verbs and determiners.[25]

In Grimshaw's theory, the impossibility of the patterns in (40) follows from a simple assumption that the functional heads 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 (41). 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.

(41)

	Category specification	*Functional specification*
V	[+V, −N]	F0
T	[+V, −N]	F1
C	[+V, −N]	F2
N	[−V, +N]	F0
D	[−V, +N]	F1
P	[−V, +N]	F2

(Grimshaw, 2005:4)

Since V, T, C heads are non-distinct in terms of their categorial status, they form one extended projection. So do N, D, and P heads.[26] The combinatorial properties illustrated in (38–40) thus follow from the assumption that heads of higher F values can only take projections with lower F values as their complements. In all the ungrammatical combinations in (40), a functional head

[25] Facts of this sort led Sportiche (1997) to develop a DP Partitioning Hypothesis, which states that determiners and the nouns they select do not enter the derivation as constituents. Nouns are selected directly by the verbs, and become constituents with their determiners only as a result of movement.

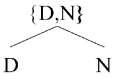
[26] Grimshaw's definition of a projection is given below (Grimshaw, 2005:4):

(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.

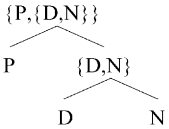
takes as its complement a category it cannot form an extended projection with. Complements of lexical categories are not subject to this restriction.

The option of projecting both labels in External Merge structures can capture the main insights behind Grimshaw's theory of extended projections in the following way. If the two merged objects are non-distinct in a sense to be made precise shortly, the features of both can project as the label of the newly formed object, as shown in (42a–c) for nominal projections, and in (43a–c) for verbal ones. The notion of non-distinctness that I am appealing to here relies on category specification, as defined in (41) above. The elements that can undergo *Project Both* are categorically non-distinct; they differ only with respect to their functional specification.

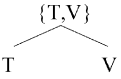
(42) a. Merge D and NP, *Project Both*



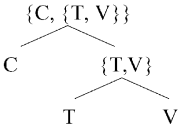
b. Merge P with DP, *Project Both*



(43) a. Merge V and T, *Project Both*



b. Merge T and C, *Project Both*



Selectional restrictions now reduce to a straightforward sisterhood relationship. Since the features of the noun are present in the label, the verb can directly select these features. For example, if a DP object merges with a verb, the verb's selectional requirements can be satisfied by its sister, since the sister's label contains features of the noun, in addition to those of the determiner.[27]

[27] The idea that sisterhood is the crucial relationship is similar in spirit to Epstein's (1999) and Epstein et al's (1998) proposals that the only legitimate syntactic relationship is derivational c-command, which in turn is defined in terms of sisterhood.

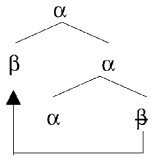
(i) *Derivational definition of c-command* (Epstein, 1999:329)
X c-commands all and only the terms of the category Y with which X was paired (by Merge or by Move) in the course of the derivation.

4. Labels in *Internal Merge* structures

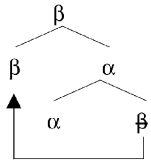
4.1. Labeling options

The possibilities for determining labels in *Internal Merge* structures are repeated in (44a–d).

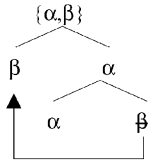
(44) a. Project Probe



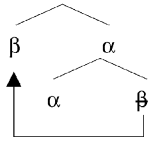
b. Project Goal



c. *Project Both* Probe and Goal



d. *Project None*



In the following two sections, I provide evidence that, in addition to the *Project Probe* option, the grammar also allows the options illustrated in (44b) and (44c), *i.e. Project Goal* and *Project Both*. More specifically, 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.[28] Free relatives, also known as headless relatives, are constructions that

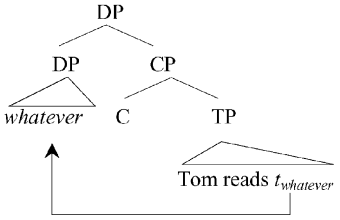
[28] Other proponents of *Project Goal* accounts of free relatives include Iatridou et al. (2001), Bury (2003), and Donati (2006). They differ from each other in details of implementation, but the basic insight remains the same. I will comment on the differences between these proposals and mine below.

appear to be headed by the wh-pronouns themselves or appear to lack heads altogether, depending on how they are analyzed. Examples of free relatives are given in (45a–b).

(45) a. Bill reads *what(ever)* Tom reads.
b. Bill reads *whatever books* Tom reads.

Project Goal allows for a derivation illustrated in (46), in which the wh-pronoun undergoes *Internal Merge* with C and projects as the label of the newly formed constituent.

(46) *Project Goal* derivation of free relatives



What is interesting about such a *Project Goal* derivation for free relatives is that it can reconcile the two seemingly incompatible accounts, the so-called Comp Account and the Head Account. The proponents of 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 (47a) (Groos and van Riemsdijk (1981), Grosu (1998), Gracanin-Yuksek (2005), among many others). On the Head Account, the [Spec,CP] position is empty and the head position is filled by the wh-phrase, as shown in (47b) (Bresnan and Grimshaw (1978), Larson (1987), Citko (2002)).

(47) a. John plays $[_{DP}$ Ø $[_{CP}$ *whatever*$_i$ $[_{TP}$ he likes $t_i]]]$ (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 (48–50).

(48) 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?
Complex Noun Phrase Constraint

(49) a. *John plays *whatever*$_i$ he wonders why Mary plays t_i.
b. **What*$_i$ does John wonder why Mary plays t_i? *Wh-Island Constraint*

(50) a. *John did *whatever*$_i$ Mary left because John did t_i?
b. **What*$_i$ did John leave because Mary did t_i? *Adjunct Condition*

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 (51a–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.

(51) a. I will listen to whatever you listen to.
b. *I will listen [$_{DP}$ [$_{DP}$ whatever$_i$] Mary plays t_i.]

Case matching effects are best illustrated in languages that have 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 simultaneously satisfy the requirements of the matrix verb and the embedded verb. In Citko, 2000, I discuss analogous data from Polish, a language whose case system is richer than German.

(52) 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 (52a), the verb *nehme* 'take' requires an Accusative object. Since the wh-pronoun heading the free relative is also Accusative, the result is grammatical. By contrast, in the ungrammatical example given in (52b), the wh-pronoun heading the relative is Dative, which conflicts with the matrix verb's need for an Accusative object.

The *Project Goal* account can explain both locality and matching effects, thus combining the insights of both the Head and the Comp Accounts. The movement of the wh-pronoun in (46) above is driven by features on 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 thus follows naturally. The fact that the moved element projects, on the other hand, explains the category and case matching effects illustrated in (51) and (52).[29]

As I mentioned above, there are many different implementations of the *Project Goal* account of free relatives. In the remainder of this section, I will briefly comment on some of the most significant differences between my proposal and two other *Project Goal* proposals: Larson's (1998) and Donati's (2006).

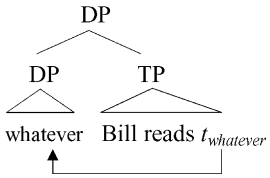
Larson (1998) analyses movement of the wh-phrase in free relatives as an overt counterpart of QR. The moved wh-phrase adjoins to TP, rather than moves to the specifier of CP, as shown in (53).

[29] 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 either Accusative or 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}]$_{ACC}$ [Pol]
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.'

(53)



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*), especially in languages with covert QR. It also predicts that movement of wh-pronouns in free relatives should be subject to the same locality restrictions as QR. QR, however, is well-known to be clause-bound, as shown by the lack of wide scope for the universal quantifier *everyone* in (54a). Movement of wh-phrases in free relatives, on the other hand, is not clause-bound, as shown by the grammaticality of (54b).[30]

(54) a. Someone thinks that everyone likes syntax. (∃>∀, *∀>∃)
b. John reads *whatever*$_i$ Bill thinks [$_{CP}$ t_i that he should read t_i]

Donati (2006) proposes a more restricted version of the *Project Goal* derivation for free relatives. It is more restricted in the sense that it derives only a subset of what is traditionally thought as free relatives via a *Project Goal* mechanism. In particular, she proposes that *Project Goal* is only possible if the projecting Goal is a head. What this means is that only free relatives headed by simple wh-pronouns, such as the one in (55a), can involve a *Project Goal* derivation. Free relatives headed by complex wh-phrases, such as the one in (55b), have to be reanalyzed as headed (rather than headless) ones.[31]

[30] This contrast could be accounted for on the assumption that overt movement is less restricted than covert movement. There does not seem to be any independent support for this assumption. For example, covert wh-movement in wh in situ languages is less restricted than overt movement.

[31] There are languages, such as Italian or Polish, in which free relatives headed by complex wh-phrases such as 'how many cookies' or 'what towns' are indeed ungrammatical:

(i) * Ho mangiano *quanti biscotti* hai preparato. [Ita]
have-1SG eaten how-many cookies have-2SG prepared
'I have eaten what cookies yo u have prepared.' (Donati 2006:32)

(ii) * Odwiedzę *które miasta* ty też odwiedzisz. [Pol]
visit-1SG which towns you also visit-2SG
'I will visit which towns you will also visit.'

However, both become grammatical if the morpheme *ever* is added.

(iii) Mangerò qualunque bicotto preparerai. [Ita]
will-eat-1SG whatever cookie will-prepare-2SG
'I will eat whatever cookie you will prepare.' (Donati 2006:41, Ft 10)

(iv) Odwiedzę którekolwiek miasta ty też odwiedzisz. [Pol]
visit-1SG whichever towns you also visit-2SG
'I will visit which towns you will also visit.'

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

(v) [$_{DP}$ [$_{D}$ 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. While the 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.

(55) a. John reads *what* Mary recommends. *simple free relative*
b. John reads *whatever books* Mary recommends. *complex free relative*

This is problematic, given that they pattern with headless (rather than headed relatives) with respect to a number of diagnostics that distinguish between headed and headless ones. In what follows, I will briefly discuss three phenomena that distinguish headed from headless relatives, and show that with respect to all three of them, free relatives headed by complex wh-phrases behave like simple free relatives (and not headed relatives). The three phenomena are: compatibility with overt complementizers, matching, and extraposition in German.

First, simple free relatives differ from headed relatives in that they disallow the complementizer *that.* In this respect, complex free relatives pattern with simple free relatives rather than headed ones.

(56) a. We'll hire whom (*that) you recommended to us.
b. We'll hire the man (that) you recommended to us.
c. We'll hire whichever man (*that) you recommended to us.

Second, free relatives differ from headed ones with respect to matching.[32] What is crucial (and well-documented in a wide range of languages) is that headed relatives do not show matching effects. The following data from Polish provide an illustration: the ungrammatical example in (57a) is a non-matching free relative; the grammatical one in (57b) is a corresponding non-matching headed relative.[33]

(57) 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.'

Crucially, complex free relatives pattern with simple ones in that they also require matching. If (58b) were a headed relative, we would expect it to be grammatical.

(58) 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.'

[32] 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).

[33] 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.

And third, extraposition in German provides a useful diagnostic to establish the position of the wh-phrase in a free relative, and indirectly distinguish between headed and free relatives (Groos and van Riemsdijk, 1981). As shown by the contrast in (59a-b), German allows extraposition of CPs but not DPs.

(59) a. Der Hans hat das Geld t_i zurückgegeben, [$_{CP}$ das er gestohlen hat] [Ger]
the Hans has the money 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 (60a–b) shows that the wh-pronoun in a free relative clause construction behaves differently with respect to extraposition than the head in a headed relative.

(60) a. * Der Hans hat was t_i zurückgegeben, [er gestohlen hat]$_i$
the Hans has what returned he stolen has
'Hans has returned what 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 what 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, Dirk Bury, personal communication).

(61) 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 whatever 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.'

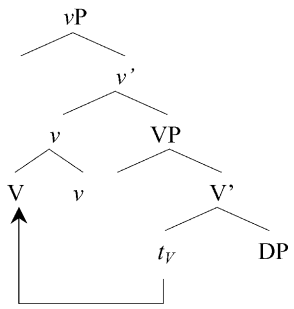
These diagnostics show that free relatives headed by simple wh-pronouns and the ones headed by complex wh-pronouns behave similarly. Therefore, treating the former as true free relatives (involving *Project Goal* derivations) and the latter as headed relatives does not seem to be on the right track.

4.3. Project Both Probe and Goal

This section explores the possibility that *Project Both* can solve a number of well-known and often-discussed problems with the standard analysis of head movement. Let me start by

reviewing these problems. First, head movement is unlike phrasal movement in that it violates the Extension Condition, which requires all syntactic operations (that is, *Merge* and *Move*) to extend the tree by targeting its root. Second, the moved head does not c-command its trace. This is illustrated by the movement of V to *v* in (62); the movement of the V to *v* does not extend the tree, and the moved V does not c-command its trace[34]:

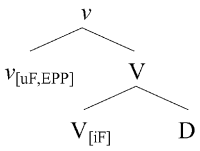
(62)



There are a number of proposals designed specifically to remedy these two (obviously related) problems. They fall into three general categories. The first one is to relegate head movement to the PF component of the grammar (Boeckx and Stjepanović, 2001; Chomsky, 2000; Harley, 2004). The second is to reanalyze it as phrasal movement (Bentzen et al., 2007; Mahajan, 2000; Koopman and Szabolcsi, 2000). And the third is to reconceptualize head movement in a way that makes it compatible with the Extension Condition (Bobaljik and Brown, 1997; Bury, 2003; Matushansky, 2006; Surányi, 2005). The analysis of head movement I will develop here falls into the third category. I will comment on the other two types of accounts after presenting the gist of my proposal.

There are two crucial ingredients to the analysis of head movement I would like to propose here. One is that head 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 (63a–c) with respect to V–*v*–T movement; I assume that *v* has an uninterpretable V feature and the verb has a corresponding interpretable feature. The uninterpretable one is valued via an *Agree* mechanism. In addition, *v* has an EPP feature which drives V movement.[35]

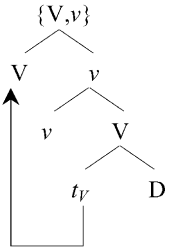
(63) a.



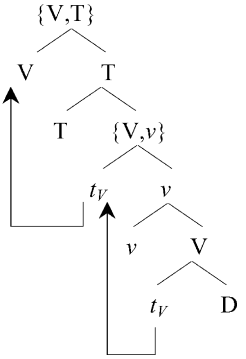
[34] Head movement is not the only example of a movement that does not extend the tree. Richards' (2001) tucking-in type movements and most cases of covert movements are also countercyclic. This is the reason why Groat and O'Neil (1996) reanalyze covert object shift as an overt movement in which the lower copy (rather than the higher one) is the one that is pronounced.

[35] This is following Alexiadou and Anagnostopoulou (1998, 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 EPP features.

b.



c.

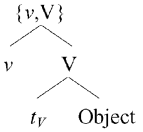


In (63b), the moved V becomes a sister of a *v*P. Since sisterhood is the most basic syntactic configuration, the EPP feature (which is the feature that gets checked via V movement) can be checked in this configuration just as well as the configuration obtained via the standard 'head adjunction' operation.

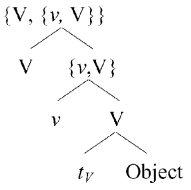
There are two immediate questions that the *Project Both* analysis of verb movement (head movement in general) raises. One involves its interaction with *(External) Merge* of (external) arguments, and the other one its compatibility with the idea that extended projections also involve *Project Both*, as I argued for in section 3.1.2 above.

Given the fact that the result of merging *v* and V involves projection of both labels, a more accurate representation of V to *v* movement is given in (64b).

(64) a. Merge *v* and V, *Project Both*

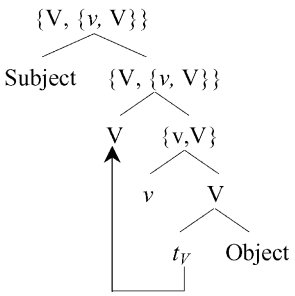


b. Move V to *v*, *Project Both*



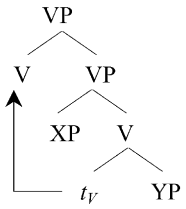
I furthermore assume that *Internal* and *External Merge* can interleave. What this means for the case at hand is that the subject gets externally merged *after* V to *v* movement. While this departs from the standard view that *Internal Merge* follows *External Merge*, I do not see anything that would exclude it on principled grounds.[36] Furthermore, it parallels the move from proposals that covert movement follows overt movement to proposals that the two can also interleave (see, for example, Fox and Nissenbaum (1999), for evidence to support this view).

(65)



My proposal bears some resemblance to Bury's (2003) and Matushansky's (2006) analysis of head movement. Bury (2003), for examples, analyzes verb movement as *Project Goal,* along the lines shown in (67) (see also Ackema et al., 1993).[37]

(66)



(Bury 2003:12)

I share Bury's intuition that the Probe does not always project. We differ, however, with respect to the label of the resulting structure. For me, both the Probe and the Goal project, whereas for Bury, only the Goal does. Furthermore, for Bury, the result of verb movement is what he terms 'a derived VP', which is essentially equivalent to a Larsonian VP shell. The result of V to T to C movement thus is also going to be a VP. This predicts that in languages that allow embedded V2, the matrix verb should select a VP (rather than a CP). The *Project Both* account does not make

[36] One of the reviewers brings to my attention the well-known contrast between (i) and (ii), which is typically taken to show preference for External Merge over Internal Merge. Crucially, however, it does not show that External Merge precedes Internal Merge; it only shows that External Merge preempts Internal Merge.

(i) There seems to be a man in the garden.

(ii) *There seems a man to be in the garden.

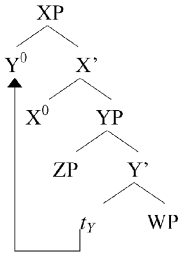
Furthermore, in (i–ii) we are dealing with two XPs (the expletive *there* and the DP 'a man') and in (65) we are dealing with a head and a phrase. This might also suggest that *External Merge* over *Internal Merge* preference is operative only if it involves elements of the same type (two heads or two phrases), giving it a Relativized Minimality flavor.

[37] One of the reviewers brings to my attention Epstein's (1998) analysis of verb movement as head-adjunction to a higher functional projection, followed by LF replacement of the target projection by a projection of the verb. This is essentially equivalent in spirit to a Project Goal account of head movement.

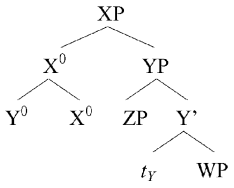
this prediction; the information about the target of head movement is preserved in the label of the new constituent.

My proposal also bears some resemblance to Matushansky's (2006) proposal, who analyzes head movement as a composite of two independently attested operations: movement to a specifier position and morphological merger, illustrated in (67a–b), respectively.

(67) a. Step 1: movement of Y to [Spec,XP] (Matushansky 2006:81)



b. Step 2: morphological merger



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 (treated by Matushansky as movement to a specifier position) have to be followed by m-merger. Otherwise, it would not be clear how to exclude excorporation.

To conclude this section, I would like to comment briefly on the differences between my account and two alternatives: PF movement and remnant movement accounts of head movement. As pointed out by Matushansky (2006), Surányi (2005), Zwart (2001) among others, relegating head movement to a PF component adds redundancy to the grammar 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, Surányi (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 would be 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

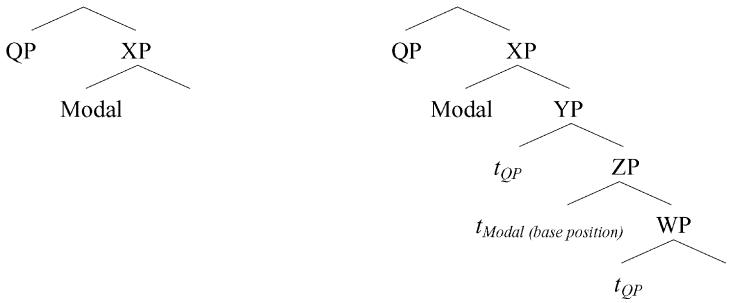
are no differences in interpretation between raised and non-raised verbs, it is not true that there are no semantically significant instances of head movement. Furthermore, as also pointed out by Matushansky (2006), if semantic vacuity were a sole diagnostic for PF movement, all movements resulting in reconstruction would have to reclassified as PF movements.

Lechner (2005) provides convincing arguments that there are cases of semantically active head movement. His evidence comes from the so-called scope splitting cases, illustrated in (68–70), in which the modal has wide scope over the quantified subject, as shown by the paraphrases of the (a) examples given in (b).

(68) a. Not every pearl can be above average size. (Lechner, 2005:3)
b. It is not possible that every pearl is above average size.
(69) a. Not everyone can be an orphan.
b. It is not possible that everyone is an orphan.
(70) 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 (68–70) involve the following schematic configurations (Lechner, 2005:2):

(71) Syntax: Interpretation:



The modal is base-generated in a position that Lechner shows to be lower than the reconstructed position of the subject. Furthermore, wide scope interpretation for the modal cannot be a result of reconstructing the subject. This is shown by the examples in (72a–c), which show that strong DPs, including negative ones, do not reconstruct below T.

(72) 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. (Lasnik, 1999:205)

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

The proponents of XP movement approach to head movement include Koopman and Szabolcsi (2000), Mahajan (2000) among many others. They derive the effects of head movement from a sequence of phrasal movements, often involving remnant movement. However, it is not clear whether such proposals can be generalized to cover all cases of head

movement. To see why, let us consider Koopman and Szabolcsi's (2000) account of verbal complexes in Hungarian and Dutch, exemplified in (73a–c) for Hungarian.[38]

(73) a. Nem fogok akarni be menni. [Hun]
not will-1SG want-INF in go-INF
'I will not want to go in.'

b. Nem fogok be menni. akarni
not will-1SG in go-INF want-INF

c. * Nem fogot akarni menni be
not will want go in

On the assumption that (73c) is the underlying order, (73a) can be plausibly analyzed as involving head adjunction of *be* 'in' to *menni* 'go-inf', and (73b) as involving head adjunction of the entire complex *bemenni* 'go in' to *akarni* 'want'.

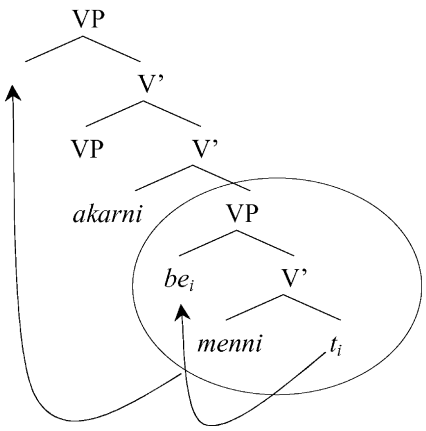
Even though the pattern suggests a head movement account, Koopman and Szabolcsi suggest that such an account is both insufficient and redundant. It is insufficient in that it does not capture the fact that not only separable prefixes undergo (obligatory) inversion with the verb. Other elements, which are undeniably phrasal, such as directional or locative PPs containing full DPs or predicative APs also undergo the inversion processes:

(74) a. [a szobaba]$_i$ menni t_i
the room-into go-inf

b. [ostobanak]$_i$ bizonyulni t_i
stupid-DAT prove-INF (Koopman and Szabolcsi 2000:19)

It is redundant in that independently motivated sequences of phrasal movement can account for the same facts. The verbs (or prefixes) undergo phrasal movement, landing in specifier positions. A schematic derivation of (73b) is given in (75).

(75)



[38] The string in (73c) is not a possible surface string, which suggests that the first instance of movement is obligatory.

Mahajan (2000), on the other hand, shows that a remnant movement analysis offers a way to dispense with rightward scrambling in the grammar of Hindi. The examples of the kind given in (76), instead of involving rightward scrambling, can be analyzed as involving *leftward* movement of the object followed by a further *leftward* movement of VP/IP.

(76) Raam-ne $[_{IP}[_{VP}\ t_i$ khaaye] the $t_{VP}]$ *saare phal*$_i$ t_{IP} [Hin]
Raam-ERG eat.PERF.MASC.PL be.MASC.PL.PST all fruits
'Raam had eaten all the fruits.'

While the idea of deriving the effects of head movement through a combination of phrasal movements might be a good solution for some movement types (and word order alternations) in some languages, it seems somewhat ad hoc for paradigmatic cases of head movement. As has been pointed out by many researchers, it often results in unnecessary proliferation of functional projections whose sole purpose seems to be to provide landing sites for various moved elements. For example, a 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 (77b). If adverbs are VP adjuncts, it is not clear what that position would be.

(77) a. Subject Adverb $[_{VP}$ V Object]
b. Subject Adverb Object $[_{VP}$ V $t_{obj}]$
c. Subject $[_{VP}$ V $t_{obj}]$ Adv Object t_{VP}

To sum up the discussion in this section, I have shown that while relegating head movement to PF or reanalyzing it as remnant movement avoids a violation 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 in the following manner. Unlike PF movement account, the *Project Both* account is syntactic and thus is predicted to have syntactic (and semantic) consequences. Unlike remnant movement, the *Project Both* movement does not require any extra steps or empty projections whose sole purpose is to provide landing spots for moved constituents.

Interesting further evidence in favor of the *Project Both* account of head movement comes from locality considerations. Typically, V heads raise only as far as C (passing through *v* and T on their way up) and N heads raise to D and P. As pointed out by Grimshaw (1991) in a different context, the highest head in an extended projection of a given lexical head typically provides a limit to how far this lexical head can move. In my account, this restriction follows the fact that *Project Both* is only possible if the two projecting elements do not conflict in categorial features. As we saw in section 3.1, this is only possible within an extended projection.

5. Project None

My proposal departs from Collins (2002) and Seely (2006), who argue in favor of a total elimination of labels. For them, the only option that the grammar allows is *Project None*. Their main conceptual argument against labels comes from the fact that labels violate the Inclusiveness Condition. However, if a label is thought of as a copy of one of the merged elements, inclusiveness is not violated. No new information is added, and copying is an independently motivated operation. Thus labels are not undesirable *a priori*, and the system proposed here, which employs all the labeling possibilities, is to be preferred on conceptual grounds.

In order to evaluate my proposal against a label-less alternative, let us examine the issue of why we needed labels to begin with. One of the main reasons for having labels is category selection. In a system without labels, how can we capture the fact that a verb selects a DP or a PP? Collins (2002) does this via the Locus Principle, given in (78) below.

(78) '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.'

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

Seely (2006), on the other hand, deduces the lack of labels from Chomsky's (1995) definition of a syntactic term. He argues that no syntactic operation can make reference to labels because labels are not syntactic terms, according to the following definition of a term[39]:

(79) For any structure K:
a. K is a term of K
b. If L is a term of K, then the members of the members of L are terms of K
(Chomsky, 1995:247)

In (82) 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.

(80) {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 see, however, is not a term.

(81) {see, {see, {the, {the, picture}}}

Since labels, not being terms, cannot be derivationally c-commanded (nor can they derivationally c-command anything), they become 'syntactically inert.' They are not in a syntactic relation to anything (since relations require derivational c-command), which is tantamount to the absence of labels. Derivational c-command, defined in (82) below, allows only terms to be c-commanded.

(82) *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.

[39] If Seely (2006) is right 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.

However, there seems to be some circularity in Seely's argument. More specifically, he notes a problem with subcategorization in a label-less grammar, 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 this problem, consider the label-free representation given in (83).

(83) {see, {the, picture}}

The problem with (83) 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 a term given in (79). 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 to propose a modified definition of a term, which is given in (84).

(84) For any structure K,
i. K is a term of K
ii. the members of K are terms of K
iii. the members of the members of K are terms of K. (Seely, 2006:201)

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 count as a term as well. Consider the following structure:

(85) {see, {see, {the, {the, picture}}}

According to Seely's revised definition, however, labels are terms, which takes away the original argument in favor of eliminating labels.[40]

Another argument in favor of labels comes from the fact that they, similarly to phases, 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 in favor of allowing *Project None* in this one case is hard to think of, I will leave it as an open possibility, to be investigated further.

6. Conclusion

To conclude briefly, I have provided in this paper both theoretical and empirical arguments in favor of a number of new labeling possibilities. I have examined the labeling possibilities for both *External* and *Internal Merge* structures, and explored the consequences of the hypothesis that these possibilities are in fact attested. The tables in (88a–b) summarize the empirical results of the paper.

[40] This problem arises only if we adopt Chomsky's original definition of a term in conjunction with label-less representations.

(86) a. Missing Labels in *External Merge* structures

	Project Both	***Project None***
External Merge	• extended projections • comparative correlatives	• final Merge

b. Missing Labels in *Internal Merge* structures

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

An interesting question that arises here is whether my proposal complicates the computational system 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 a minimalist perspective, a system that imposes restrictions on labels (unless these restrictions can be shown to be imposed by the interfaces) is less optimal than a system that uses all the possibilities, which is what I have argued for here. The existence of the missing labels is thus a welcome result from the perspective of the strong minimalist thesis.

Acknowledgements

I am thankful to two anonymous *Lingua* reviewers for very useful comments and suggestions. I am alone responsible for any remaining errors and omissions.

References

Abeillé, A., Borsley, R., Espinal, M-T., 2006. The syntax of comparative correlatives in French and Spanish. In: Muller, S. (Ed.). Proceedings of the HPSG06 Conference, pp. 6–26.

Ackema, P., Neeleman, A., Weerman, F., 1993. Deriving functional projections. Proc. NELS 23, 17–31.

Alexiadou, A., Anagnostopoulou, E., 1998. Parametrizing AGR: word order, verb movement and EPP checking. Natural Language and Linguistic Theory 16, 491–539.

Alexiadou, A., Anagnostopoulou, E., 2001. The subject-in-situ generalization and the role of case in driving computations. Linguistic Inquiry 32, 193–231.

Beck, S., 1997. On the semantics of comparative conditionals. Linguistics and Philosophy 20, 229–271.

Bentzen, K., Hrafnbjargarson, G.F., Hroarsdottir, T., Wiklund, A-L., 2007. The Tromsø guide to the Force behind V2. Working Papers in Scandinavian Syntax 79, 93–118.

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

Bobaljik, J., Brown, S., 1997. Interarboreal operations: head movement and the extension requirement. Linguistic Inquiry 28, 345–356.

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

Bury, D., 2003. Phrase structure and derived heads. Doctoral Dissertation. University College London.

Chomsky, N., 1994. Bare phrase structure. MIT Occassional Papers in Linguistics. Department of Linguistics and Philosophy, MIT. Also published in Webelhuth, G. (Ed.), 1995. Government and Binding Theory and the Minimalist Program. Blackwell, Oxford.

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

Chomsky, N., 2000. Minimalist inquiries. In: Martin, R., Michaels, D., Uriagereka, J. (Eds.), Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik. MIT Press, Cambridge, MA, pp. 89–156.

Chomsky, N., 2001a. Beyond explanatory adequacy. MIT Occasional Papers in Linguistics 20.

Chomsky, N., 2001b. Derivation by phase. In: Kenstowicz, M. (Ed.), Ken Hale: A Life in Language. MIT Press, Cambridge, MA, pp. 1–52.

Chomsky, N., 2005. On phases. Unpublished Manuscript. MIT.
Citko, B., 2000. Parallel merge and the syntax of free relatives. 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., 2008. Small clauses: not so small and not all alike. Lingua 118, 261–295.
Collins, C., 1997. Local Economy. MIT Press, Cambridge, MA.
Collins, C., 2002. Eliminating labels. In: Epstein, S., Seely, D. (Eds.), Derivation and Explanation in the Minimalist Program. Blackwell Publishing, Oxford, pp. 42–64.
Culicover, P., Jackendoff, R., 1999. The view from the periphery. Linguistic Inquiry 30, 543–571.
Dayal, V., 1995. Quantification in correlatives. In: Bach, E., Jelinek, E., Kratzer, A., Partee, B.H. (Eds.), Quantification in Natural Languages, Studies in Linguistics and Philosophy 54. Kluwer Academic Publishers, Dordrecht, pp. 179–205.
Den Dikken, M., 2005. Comparative correlatives comparatively. Linguistic Inquiry 36, 497–533.
Donati, V., 2006. On wh-head movement. In: Cheng, L., Corver, N. (Eds.), Wh-Movement. Moving On. MIT Press, Cambridge, MA, pp. 21–46.
Epstein, S., Groat, E., Kawashima, R., Kitahara, H., 1998. A Derivational Approach to Syntactic Relations. Oxford University Press, New York/Oxford.
Epstein, S., 1998. Overt scope marking and covert verb second. Linguistic Inquiry 29, 181–227.
Fox, D., Nissenbaum, J., 1999. Extraposition and scope: a case for overt QR. In: Bird, S., Carnie, A., Haugen, J.D., Norquest, P. (Eds.), Proceedings of the 18th West Coast Conference on Formal Linguistics. Cascadilla Press, Somerville, MA.
Grimshaw, J. 1991. Extended Projections. Unpublished Manuscript. Brandeis University. [Published as Grimshaw 2005. Words and Structure. CSLI Lecture Notes Number 151, Stanford, CA: CSLI Publications].
Groat, E., O'Neil, J., 1996. Spell-out at the LF interface. In: Abraham, W., Epstein, S., Thráinsson, H., Zwart, J.W. (Eds.), Minimalist Ideas: Syntactic Studies in the Minimalist Framework. John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 189–198.
Groos, A., van Riemsdijk, H., 1981. Matching effects in free relatives: a parameter of core grammar. In: Belleti, L., Brandi, A., Rizzi, L. (Eds.), Theory of Markedness in Generative Grammar: Proceedings of the IVth GLOW Conference. Scuola Normale Superiore di Pisa, Pisa, pp. 171–216.
Hale, K., Keyser, J., 1993. On argument structure and the lexical representation of syntactic relations. In: Hale, K., Keyser, J. (Eds.), The View from Building 20. MIT Press, Cambridge, pp. 53–109.
Harley, H., 2004. Merge, conflation, and head movement: the first sister principle revisited. In: Moulton, K., Wolf, M. (Eds.), Proceedings of NELS, 34. GSLA, University of Massachusetts, Amherst.
Hornstein, N., Nunes, J., Grohmann, K., 2005. Understanding Minimalism. Cambridge University Press, Cambridge.
Hornstein, N., Uriagereka, J., 2002. Reprojections. In: Epstein, S.D., Seely, D. (Eds.), Derivation and Explanation in the Minimalist Program. Blackwell Publishing, Oxford/Malden, pp. 106–132.
Iatridou, S., Anagnostopoulou, E., Izvorski, R., 2001. Observations about the form and meaning of the perfect. In: Kenstowicz, M. (Ed.), Ken Hale: A Life in Language. MIT Press, Cambridge, MA, pp. 189–238.
Izvorski, R., 1996. The syntax and semantics of correlative proforms. In: Kusumoto, K. (Ed.), Proceedings of NELS, 26. GLSA, Amherst, MA, pp. 133–147.
Koopman, H., Szabolcsi, A., 2000. Verbal Complexes. MIT Press, Cambridge, MA.
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: Epstein, N., Hornstein, N. (Eds.), Working Minimalism. MIT Press, Cambridge, MA, pp. 189–215.
Lechner, W., 2005. Interpretive effects of head movement. Manuscript. University of Tubingen.
Mahajan, A., 2000. Eliminating head movement. The GLOW Newsletter # 44, 2000, 44–45.
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: The Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society, Berkeley Linguistics Society, University of California, Berkeley, pp. 176–187.
McCawley, J., 1998. The Syntactic Phenomena of English. University of Chicago Press, Chicago.
Moro, A., 2000. Dynamic Antisymmetry. MIT Press, Cambridge, MA.
Pereltsvaig, A., 2007. Copular Sentences in Russian: A Theory of Intra-Clausal Relations. Springer.

Pesetsky, D., Torrego, E., 2006. Probes, goals and syntactic categories. In: Otsu, Y. (Ed.), Proceedings of the Seventh Annual Tokyo Conference on Psycholinguistics. Hituzi Syobo Publishing Company, Tokyo, pp. 25–60.
Rackowski, A., Richards, N., 2005. Phase edge and extraction: a Tagalog case study. Linguist. Inquiry 36, 565–599.
Richards, N., 2001. Movement in Language: Interactions and Architectures. Oxford University Press, Oxford.
Seely, D., 2006. Merge, derivational c-command, and subcategorization. In: Boeckx, C. (Ed.), Minimalist Essays. John Benjamins Publishing Company, Amsterdam/Philadelphia, pp. 182–217.
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: In: Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae. Sectio Linguistica. Tomus XXVI. ELTE, Budapest, pp. 313–342.
Zwart, C.J.-W., 2001. Syntactic and phonological verb movement. Syntax 4, 34–62.

Further reading

Bianchi, V., 2000. The raising analysis of relative clauses: reply to Borsley. Linguistic Inquiry 31, 123–140.
Borsley, R., 2004. The view from the Polish periphery. In: Arnaudova, O., Browne, W., Rivero, M. L., Stojanović, D. (Eds.), Annual Workshop on Formal Approaches to Slavic Languages: the Ottawa Meeting 2003. Michigan Slavic Publications, Ann Arbor, pp. 59–91.
Chomsky, N., 2006. Approaching UG from below. Unpublished manuscript. MIT.
Liptak, A., 2004. On the correlative nature of left-peripheral relatives. ZAS Working Papers in Linguistics 35, 287–313.
Pesetsky, D., Torrego, E., 2001. T-to-C movement: causes and consequences. In: Kenstowicz, M. (Ed.), Ken Hale: A Life in Language. MIT Press, Cambridge, MA, pp. 355–426.