Asymmetries in Long-Distance Extraction in a Tree-Adjoining Grammar

Anthony Kroch
University of Pennsylvania

March 5, 2001

In recent papers (Kroch and Joshi 1985, Kroch 1987) we claimed that, if one adopts the Tree Adjoining Grammar (TAG) formalism of Joshi, Levi, and Takahashi (1975) as the formal language of syntax, the ungrammaticality of extractions from *wh-* islands can be made to follow in a straightforward way from the nonexistence of multiple *wh-* fronting in simple questions. The analysis we gave was oversimplified, however, because it wrongly predicted all *wh-* island extractions to be ungrammatical, and we know that certain of them are well-formed, not only in languages like Swedish or Italian, but also in English (Chomsky 1986, Grimshaw 1986). Nevertheless, the analysis we gave had the attraction of providing a simple structural explanation for the *wh-* island effect, and it generalized directly to such other manifestations of subjacency as the Complex Noun Phrase Constraint (CNPC). In this paper we show that the analysis presented in our earlier papers can be extended in a reasonable way to several cases that were unaccounted for in the original discussion. In particular, we discuss such well-known examples as the following:

(1) a. * Who$_i$ does he think that [e$_i$ left]?
b. When$_i$ does he think that [we left e$_i$]?

(2) On Thursday$_i$, what$_j$ will you buy e$_j$ e$_i$?

(3) a. ? On that shelf$_i$, how many books$_j$ can you fit e$_j$ e$_i$?
b. * That many books$_i$, on what shelf$_j$ e$_i$ can fit e$_j$?

(4) a. ? What$_i$ were you wondering [how$_j$ to say e$_i$ e$_j$]?
b. * How$_i$ were you wondering [what$_j$ to say e$_j$ e$_i$]?

Accounting for some of these cases will require an extension of the TAG formalism as presented in our 1985 paper, but we hope to show that under this extension the principles governing all of the above contrasts receive a natural formulation.

In a recent squib, Baltin (1986) has pointed out that topicalized sentences like (5) are grammatical, although the corresponding *wh-* question in (6) is ungrammatical:

(5) After the party$_i$, I wonder who$_j$ e$_j$ will stay e$_i$.

(6) * When$_i$ do you wonder who$_j$ e$_j$ will stay e$_i$?

Baltin uses this contrast to argue that preposed adverbs may not leave traces, since under standard Government-Binding (GB) theory assumptions the trace of the adverb in (5) would not be properly governed. We, on the contrary, would take the significance of the above contrast to lie in the fact that it recapitulates the contrast between (7) and (8):

(7) After the party$_i$, who$_j$ e$_j$ will stay e$_i$?

(8) a. * When who will stay?

b. * I know when who will stay.

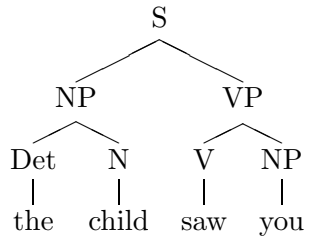
[1]

Since, as will become evident, our TAG analysis derives the ungrammaticality of (6) from the ungrammaticality of the simple sentence (8), it follows that we predict the grammaticality of (5), since the simple sentence (7) is grammatical.

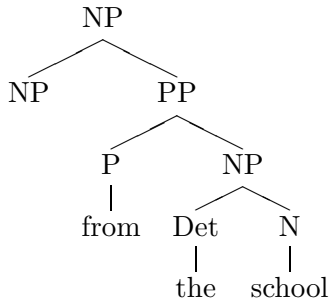
1 A sketch of the TAG formalism

The analysis that makes the above predictions is based on the formal theory of TAG grammar in Kroch and Joshi (1985). The reader is referred to that paper and others that have appeared since for a detailed introduction to the formalism (Joshi 1985, Joshi 1987, Vijay-Shankar, Weir, and Joshi 1986). To review briefly, the TAG formalism derives complex sentences by composing simple structures. These structures are phrase-structure trees, called in the theory ELEMENTARY TREES, and they come in two varieties, INITIAL TREES, which are representations of simple sentences and AUXILIARY TREES, which are the recursive structures of the language. Auxiliary trees are composed with other trees, both elementary and derived, by a tree combining operation called ADJUNCTION, which inserts the auxiliary tree into another tree. A TAG grammar consists simply of a finite set of elementary trees with the adjunction operation defined on it. The trees below illustrate how a derivation proceeds in a TAG. In these trees and in those to follow only those details relevant to the points under discussion are given.

(9) Initial tree:

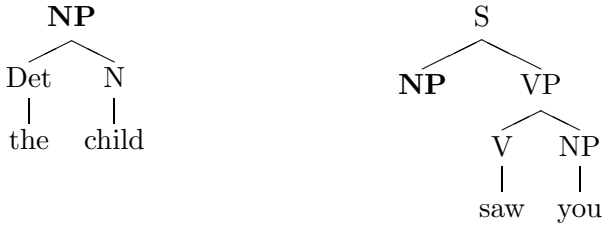


(10) Auxiliary tree:



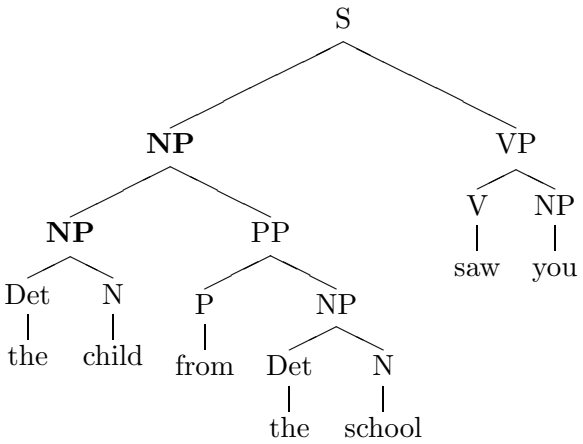
As the tree in (10) illustrates, auxiliary trees have a root node which may be any phrasal category; and on their frontier, all of their nodes are expanded to terminal symbols except one, called the FOOT node, which is identical in category to the root node. Adjunction works by first breaking an elementary tree at a phrasal node in such a way that the two pieces each contain a copy of the node at the break point. In (11) we illustrate this by breaking the tree in (9) at the boldface NP:

(11)



Then an auxiliary tree, in this case the tree in (10), whose root node is identical in category to this doubled node, is inserted at the broken node. This insertion occurs by the identification of the root and foot node of the auxiliary tree with the two instantiations of the doubled node. In our example the result of this insertion is the tree in (12), which corresponds to a sentence whose subject NP is modified by a prepositional phrase:

(12) Tree resulting from adjunction:



In the simplest case of adjunction, an auxiliary tree whose root node is of category XP may be adjoined at any node in another tree whose category label is also XP, but it is possible to constrain this adjunction further by associating with each node of an elementary tree a set of LOCAL CONSTRAINTS to specify which auxiliary trees may or must be adjoined at that node when the elementary tree appears in a derivation. In the application of TAG grammar to linguistics, these local constraints should, of course, reflect linguistic generalizations and not be used in an ad hoc way to force otherwise inadequate analyses not to misgenerate.

As should be evident even from this simple example, the structures of complex sentences generated by any TAG grammar are completely fixed by the inventory of elementary trees (initial and auxiliary). Since these trees are finite in number, it is

possible to construct a TAG grammar by simply listing the elementary trees it is to contain. Such a grammar would not be of linguistic interest, however, since it would contain no statement of the constraints governing the well-formedness of simple syntactic units. Hence, the TAG formalism must be supplemented with a substantive linguistic theory that defines well-formedness constraints for elementary structures, including the local constraints that they may contain. This theory must account for such linguistic constraints as those on phrase-structure well-formedness ($\overline{\text{X}}$ theory), on subcategorization and thematic role assignment, and on the appearance of empty categories. For reasons of space we cannot fully elaborate such a theory in this paper; but, as the reader will see, we largely assume, either directly or in modified form, the principles of GB theory for this purpose.

We also require, however, some constraints on the well-formedness of elementary structures that are particular to TAG grammar; and we will assume without argument the following constraints: First, we will require that the elementary trees be minimal; namely, that all initial trees be simple sentences and that auxiliary trees be minimal recursive structures. This minimality is expressed as a constraint that auxiliary trees rooted in a node of type XP contain no occurrences of nodes of that type on the path from the root to the required frontier node of type XP, except that Chomsky-adjunction structures in which the root node immediately dominates a node of its own type will not be considered to violate this constraint. Second, we will restrict auxiliary trees to being of two types, which we designate as ATHEMATIC TREES and COMPLEMENT TREES, respectively. The former have the structure of (13), with the order of the constituents variable:

(13) $\quad \text{X}^n \longrightarrow \text{X}^n \ \ldots \ (\text{Y}^{max}) \ \ldots$

An athematic tree introduces a modifying, complement or dislocated phrase, and the crucial feature of such a tree is that its foot node is the head of the phrase dominated by the root node, giving the tree the Chomsky-adjunction structure illustrated in (12). In previous TAG papers we have used athematic auxiliary trees to introduce such constructions as relative clauses and extraposed constituents. In the analysis below examples like (40a) are licensed as athematic trees.

A complement tree serves to introduce a predicate that subcategorizes for and assigns a thematic role to a phrase of the category of its foot node. This type of tree has the structure in (14):

(14) $\quad \text{X}^{max} \longrightarrow \ldots \ \text{Y}^0 \ \ldots \ \text{X}^{max} \ \ldots$, where Y^0 governs X^{max}.

We require that the foot node of such a complement tree be lexically governed. As in the case of athematic trees the order of constituents in complement trees is variable. But while the order in athematic trees is variable within languages, the order in complement trees will reflect the directionality of government and so will vary across but not within languages. Examples of complement trees used in the analysis in this paper are (21) and (40b), among others.

For reasons that will become apparent later, we require further that the foot node of a complement auxiliary tree have a certain government relationship to its root node. Let us define government as in (15):

(15) A minimal Y^0 GOVERNS a node X^{max} if and only if Y^0 is a sister of X^{max} and Y^{max}, the maximal projection of Y^0, dominates X^{max}. We call Y^{max} the GOVERNMENT DOMAIN of X^{max}. If the category of Y^0 is a lexical category (i.e., N, V, Adj, or P), then Y^0 LEXICALLY GOVERNS X^{max}, and Y^{max} is the LEXICAL GOVERNMENT DOMAIN of X^{max}.

Then we can define the transitive closure of the government domain relation as in (16):

(16) Let us say that the relation **G**(Y, X) holds of a pair of nodes if and only if Y is the government domain of X. We can then define **TG**, the transitive closure of **G**, as follows:

1. If **G**(Y, X), then **TG**(Y, X).
2. If **TG**(Y, X) and if $\overline{Y}$, the node immediately dominating Y, is a projection of Y, then **TG**($\overline{Y}$, X).
3. For any W and Z, if **TG**(W, X) and **TG**(Z,W) then **TG**(Z, X).

We will define as the MAXIMAL GOVERNMENT DOMAIN of X in a given elementary tree that node W such that **TG**(W, X) and such that there is no Z for which **TG**(Z,W).

Given the definition in (16), we require, as a first approximation, that complement auxiliary trees meet the condition in (17).[2] Note that we are assuming, with Chomsky (1986), that S is a projection of Infl and that $\overline{S}$ is a projection of Comp.

(17) The foot node of a complement auxiliary tree must be lexically governed, and its maximal government domain must be the root node of the tree.

Note that the relation **TG** is similar to the notion of "extraction domain" of Huang (1982) and related to the concept of "g-projection" of Kayne (1983). In a sense, therefore, the foot nodes of complement auxiliary trees can be thought of as empty categories that must be "properly governed" in the sense of Chomsky (1981). If we define proper government as in (18), we can state a TAG version of the GB theory Empty Category Principle (ECP) as in (19):

(18) For any node X in an elementary tree α, initial or auxiliary, X is properly governed if and only if one of the following conditions is satisfied:

i) The maximal government domain of X is the root node of α;

ii) X is coindexed with a "local" c-commanding antecedent in α.

(19) For any node X in an elementary tree α, initial or auxiliary, if X is empty, then it must either be properly governed or be the head of an athematic auxiliary tree.

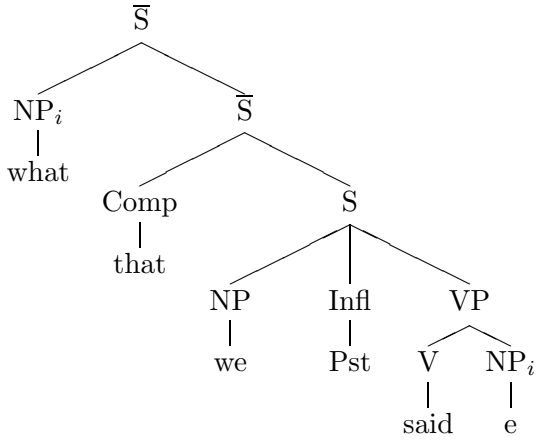
The utility of the requirement in (19) will become apparent in the course of the following section, where we discuss the mechanisms for capturing constraints on long-distance dependencies in a TAG.

2 *wh*- movement in a TAG

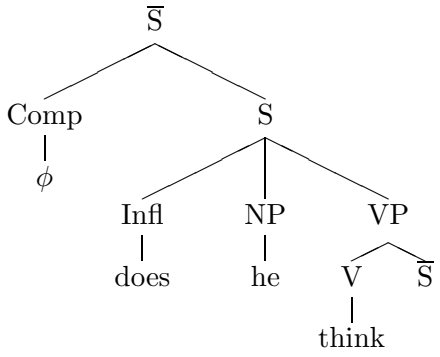
2.1 Generating unbounded *wh*- dependencies

The sets of string languages and tree languages generated by TAG grammars have many interesting formal properties that have been explored by Joshi and his mathematical collaborators (Joshi, Levy, and Takahashi 1975, Vijay-Shankar and Joshi 1985, Vijay-Shankar, Weir, and Joshi 1986); and some of these formal properties make the formalism an interesting candidate for the formal syntactic apparatus of linguistic theory, a point we discuss in detail elsewhere (Kroch and Joshi 1985, Kroch 1987). It is sufficient for present purposes to note that the TAG formalism allows a strict separation of the statement of local co-occurrence restrictions from the statement of recursion. The result of this factoring is that elements linked by unbounded dependencies can be made to originate in single elementary trees and then to move apart as a result of the adjunction of auxiliary trees. The following trees illustrate how unbounded dependencies are created:[3]

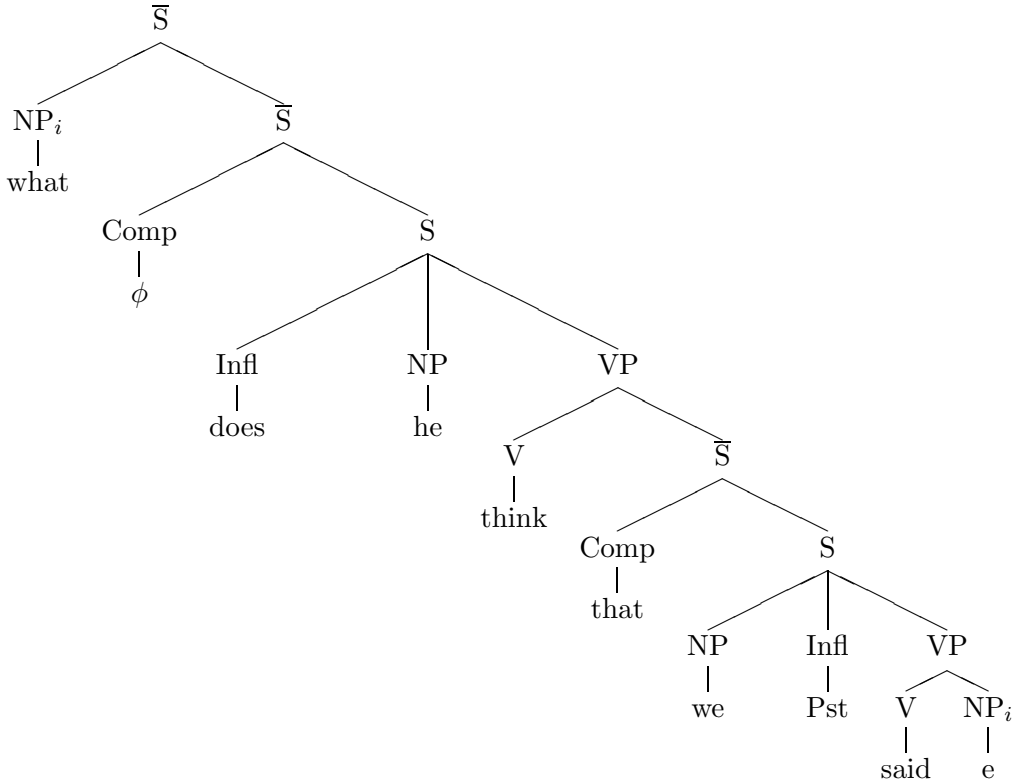
(20) Initial tree:



(21) Auxiliary tree:



(22) Output of adjunction:



The *wh-* pronoun in (20) is in the same tree as the verb with which it is construed, and its interpretation as the object of the verb *say* is guaranteed by that fact. Following standard conventions, we represent the relationship between the fronted constituent and the position in which phrases with its grammatical role normally appear by coindexing the fronted *wh-* with an empty category. The relationship between an indexed empty category and the categorially identical, c-commanding node with which it is coindexed, we call LINKING. The adjunction of the auxiliary tree (21) at the boldface $\overline{S}$ in (20) produces the tree (22), in which the *wh-* word is now initial in the matrix sentence.

2.2 ECP effects

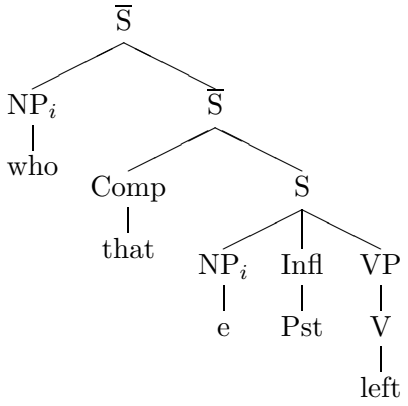
From the perspective of this analysis, let us now consider the cases in (1), repeated as (23):

(23) a. * Who$_i$ does he think that [e$_i$ left]?

b. When$_i$ does he think that [we left e$_i$]?

The initial tree needed to derive sentence (23a) will have to be as in (24), which would combine with the auxiliary tree in (21):

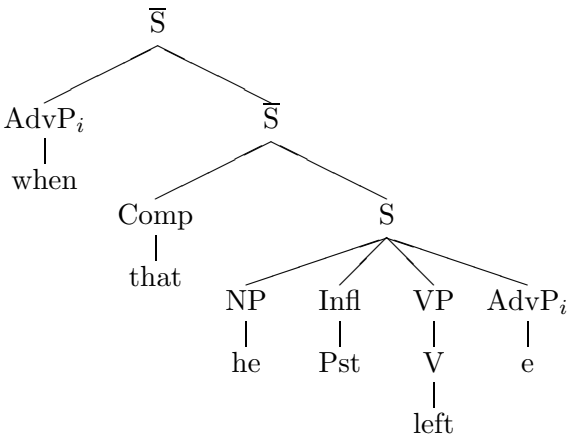
(24)



In this tree, however, the empty category in subject position is, under the standard GB assumptions we have adopted, not "properly governed." Being a subject, it is not "lexically governed" by a verb that subcategorizes it; and because the complementizer *that* intervenes between the fronted *wh-* and its trace in subject position, it is also not "antecedent-governed," as its c-commanding antecedent is not sufficiently local. Hence, since we have imposed the ECP as a constraint on the well-formedness of elementary trees, (24) is not a permissible initial tree and sentence (23a) is not derivable.

Unfortunately, this analysis runs into trouble in the case of a sentence like (23b). As Huang (1982) and Lasnik and Saito (1984) have pointed out, although adjuncts are thought to be properly governed in the same way as subjects (i.e., by antecedent government), they can be extracted out of subordinate clauses, as the grammaticality of (23b) shows. Under our derivation, the initial tree needed to derive (23b) would be as in (25):

(25)



In this tree, as in (24), the trace linked to the preposed *wh-* is not properly governed; hence the analysis fails. This difficulty, we should note, is not particular to our TAG version of the analysis but rather has a counterpart in all treatments of these cases that are based on GB assumptions. Thus, in Lasnik and Saito (1984), the contrast between (23a) and (23b) leads those authors to stipulate a difference in the licensing conditions on subject traces and the traces of adverbial adjuncts. Because

their analysis uses the GB level of representation known as "Logical Form" in a crucial way, we do not wish to incorporate it in our TAG treatment, one of whose strengths is that it seems able to state constraints on *wh-* extraction using only a single level of representation. If we consider additional data, moreover, we can find an alternative constraint on the licensing conditions on subject traces that produces the desired result. As the sentences in (2) and (3) (repeated as (26) and (27)) show, fronted PPs can occur adjoined to $\overline{\text{S}}$ to the left of a fronted *wh-* phrase:[4]

(26) On Thursday$_i$, what$_j$ will you buy e$_j$ e$_i$?

(27) a. ? On that shelf$_i$, how many books$_j$ can you fit e$_j$ e$_i$?

b. * That many books$_i$, on what shelf$_j$ e$_i$ can fit e$_j$?

Baltin (1982) argues on the basis of examples like (26) that adverb phrases can appear in this position, and example (27a) shows that even argument PPs can appear there. Subject NPs, on the other hand, cannot occur, as the ungrammaticality of (27b) shows. Sentences (26) and (27a) raise the question of how the traces of the fronted PPs are to be governed. In the case of the argument PP in (27a), one might treat the trace as lexically governed; but the adjunct PP in (26) cannot be so analyzed. We propose, therefore, the following difference between antecedent government of the traces of PPs or adverbs and subject NPs: Let us define local c-command (see Aoun and Sportiche 1983) so that the configuration in (24) and (25) does permit antecedent government of the trace bound by the moved *wh-*. This can be accomplished in a TAG simply by limiting the locality requirement to co-presence in the same elementary tree of an empty category and its antecedent governor. Then the extractability of adjuncts is predicted. To block the extraction of subjects, we require that the traces of subject NPs obey a stricter locality condition; namely, we require that they be adjacent to their antecedent governor. This stipulation will rule out the tree in (24) and so block the derivation of sentence (23a).

While we cannot claim to explain the source of the requirement that the traces of subject NPs, when antecedent-governed, be adjacent to their governors, we do think that it can be independently motivated. For one thing, it explains why English does not allow heavy-NP shift of subjects. Thus, we find the contrast between (28) and (29):

(28) a. I saw an old St. Bernard with a limp yesterday.

b. I saw e$_i$ yesterday [an old St. Bernard with a limp]$_i$.

(29) a. An old St. Bernard with a limp came by yesterday.

b. * e$_i$ came by yesterday [an old St. Bernard with a limp]$_i$.

Given our expectation that extractions should be general, we might expect (29b) to be grammatical, as it is in many languages. From a structural point of view, moreover, the dislocated NP in (29b) is just as much a local c-commanding antecedent as the dislocated subject in a simple question like the following:

(30) [Which old St. Bernard]$_i$ e$_i$ came by yesterday?

Apparently, the only relevant difference between the two cases is that in the latter the trace and its antecedent are adjacent.[5]

Consider next the fact, pointed out by Koopman (1984), that *do*-support cannot occur in subject questions. Thus, sentence (31) is ungrammatical, unless *do* is stressed and so interpreted emphatically:

(31) * Who *did* come to town?

Koopman assigns the structure in (32) to this sentence and argues that its ungrammaticality follows from the ECP, since the configuration in Comp is identical to that in the case of an extraction from the subject of an object complement clause with a lexical complementizer:

(32) $[_{\overline{S}}\ [_{Comp}\ Who_i\ did_j]\ [_S\ e_i\ [_{Infl}\ e_j]\ [_{VP}\ come\ to\ town]]]$

The difficulty with this analysis is that Baltin (1982) has given arguments that subject-aux inversion must be a local permutation of NP and Aux rather than a movement of Aux to Comp. If he is correct, as we believe him to be, then Koopman's explanation for the ungrammaticality of (31) is no longer available since its structure will now be as in (33):

(33) $[_{\overline{S}}\ [_{Comp}\ Who_i\ ]\ [_S\ [_{Infl}\ did]\ e_i\ [_{VP}\ come\ to\ town]]]$

If, on the other hand, we require that subject traces be adjacent to their antecedents, then (31) is ruled out as desired under the structural description in (33).

Finally, if we assume with Baltin (1982) and Lasnik and Saito (1984) that topicalized constituents are adjoined to S rather than to $\overline{S}$, the adjacency requirement will explain the contrasts in (34) and (35):

(34) a. ? Who_i e_i thinks $[_{\overline{S}}$ $beans_j$ John will never eat $e_j]$?
b. ? Why_i do you suppose $[_{\overline{S}}$ $beans_j$ John never eats e_j $e_i]$?
c. * Who_i do you think $[_{\overline{S}}$ $beans_j$ e_i will never eat $e_j]$?

(35) a. * Who_i did they say $[_{\overline{S}}$ e_i would eat beans] and eat beans, e_i did?
b. ? Who_i did they say e_i would eat beans and eat beans, he did?

The sentences of (34) differ unexpectedly in acceptability. While (34a) and (34b) are quite awkward, as topicalizations inside subordinate clauses generally tend to be,[6] they are more acceptable than (34c), which is completely out. Why this distinction should exist is unclear, unless it is due to the fact that the topicalized NP *beans* lies between the subject position and the original position of the fronted *wh*-, thereby preventing the antecedent governor of the subject from being adjacent to the subject position.[7] If we can assume that extraction from coordinate structures is constrained by some version of the "Across the Board Constraint" of Williams (1978), the sentences of (35) make the same point. Sentence (35a) is completely unacceptable, on our account because adjacency between the subject position and its governor is disturbed by the fronted VP *eat beans*. That this configuration produces

an ECP violation is suggested by the fact that inserting a resumptive pronoun, as in (35b), improves the acceptability of the sentence, as is generally the case with such violations.[8]

3 Extraction from NP

3.1 Subject/object asymmetries in extraction from NP

The treatment of *wh-* movement we have so far presented will not incorporate one important case of an unbounded *wh-* dependency, that known as "extraction from NP." Consider, for example, the following sentences:

(36) Which painting$_i$ did you see e$_i$?

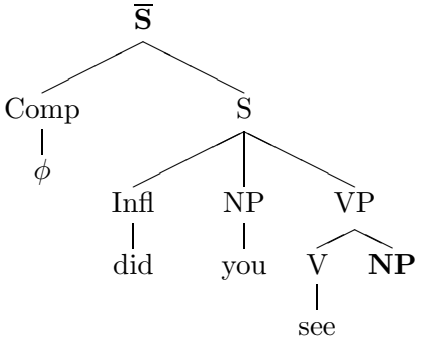
(37) Which painting$_i$ did you see a copy of e$_i$?

(38) Which painting$_i$ did you see a photograph of a copy of e$_i$?

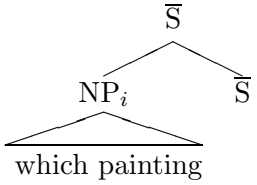
It is clear from examples like these that *wh-* phrases can be linked to empty categories embedded under any number of noun complements and hence that extraction from NP yields an unbounded dependency. If we wish to preserve a straightforward compositional semantics for our TAG, we cannot derive sentences like (37) and (38) from (36) by the adjunction of auxiliary trees rooted in NP to the NP node dominating the empty category. The problem is that under such a derivation, the preposed *wh-* phrase changes its thematic role with each adjunction and the interpretation of the derived tree is not a simple function of the interpretations of its component elementary trees. Indeed, we never want to allow derivations under which thematic roles, once established, are altered by further adjunctions, and we will block such derivations by, in every tree, placing a particular local constraint on every node that is assigned a thematic role by a governor. This constraint limits adjunctions at that node to athematic auxiliary trees in the sense of Section 1 (see Kroch and Joshi forthcoming for further details).[9] This general well-formedness constraint can be thought of as a TAG version of the GB constraint known as the "Projection Principle" (Chomsky 1981).[10] In order to permit the derivation of cases of extraction from NP in a way consistent with this constraint, it is necessary to extend the formalism beyond what we have been assuming. The extension we have in mind is one which is already defined in the original mathematical paper on TAGs and which we have made use of for linguistic purposes in our paper on extraposition (Kroch and Joshi 1987). In that paper we extended the definition of an auxiliary tree to include sets of such trees. An auxiliary tree set is simply a set of elementary trees, each of which meets the definition of an auxiliary tree. When an auxiliary tree set is adjoined to another tree, each of its components is adjoined to a distinct node in the other tree. The only requirement on the adjunction of an auxiliary set is that all of its components be adjoined into the same elementary tree. So long as the adjunction of auxiliary sets is defined in this way, all the mathematical properties of TAG are preserved (Joshi 1987).

Once we allow auxiliary sets, we can define the linking relationship between nodes in different component trees of the same set. When two components contain nodes that are linked, however, we will require that, upon adjunction, the empty category wind up in a position where its coindexed antecedent c-commands it. With this machinery we can now derive sentences like (37) and (38). The initial and auxiliary trees needed to derive (37) are as follows:

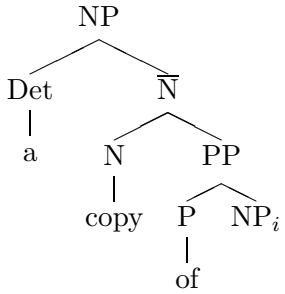
(39)



(40) a.



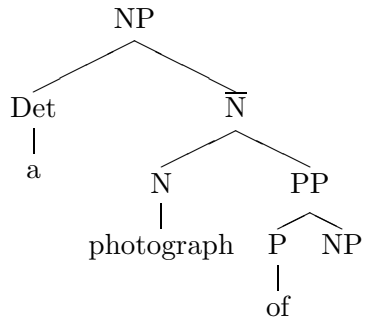
b.



When the two components of the auxiliary set in (40) are adjoined to the boldface nodes in (39), the result is the proper tree for sentence (37). One interesting feature of these trees is the occurrence of an empty category in the initial tree not linked to any antecedent. The linking here is between nodes in the two components of the auxiliary set, those bearing the subscript i; and it is only when adjunction takes place that the empty category receives an interpretation by virtue of the index that adjunction supplies. It is because the empty category is uninterpreted until adjunction occurs that the semantic interpretation of the complex structure remains compositional. Since the empty category is uninterpreted in the initial tree, it has no thematic role; and so, rather than changing its thematic role as a result of adjunction, it merely acquires one for the first time.

The unbounded character of extraction from NP follows straightforwardly under the analysis we have given. Thus, if we add the following auxiliary tree to our grammar, we can derive sentence (38):

(41)



The derivation proceeds by simply adjoining the tree in (41) to the root node of (40b). Then the resulting derived auxiliary set, which consists of (40a) and the complex structure composed of (41) and (40b), is adjoined as before to the boldface nodes in (39). Clearly, the adjunction of structures like (41) into an auxiliary tree can be repeated any number of times before the set is adjoined to (39), ending the derivation. Thus, our analysis correctly allows extraction from NP to be unbounded.

Let us consider now the contrast between (37) and the following, ungrammatical, sentence:

(42) * Which painting$_i$ was a copy of e$_i$ seen by many people?

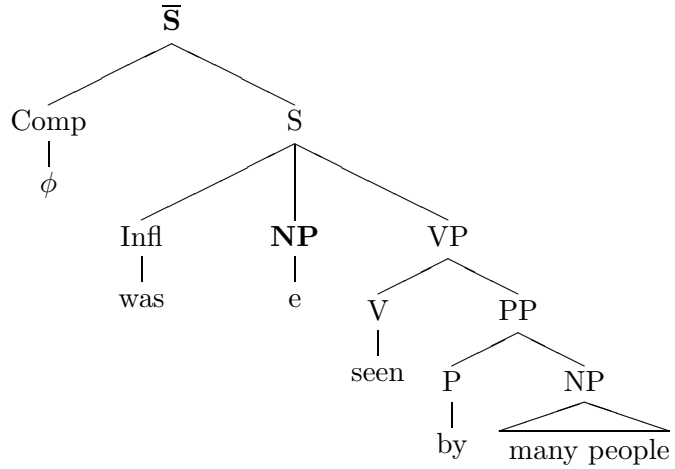
There have been many attempts to explain this contrast in the recent literature, and all of them in one way or another reduce it to the general contrast between subject and object extractions. The difficulty posed by this case is that, as standardly formulated, the ECP, which is usually responsible for subject/object asymmetries in GB analyses, appears not to rule out (42) (see Huang 1982). The trace in this sentence is lexically governed by a preposition and the preposition, which one might want to say can transmit lexical government but is not by itself sufficient for proper government, is governed by the lexical head of a major category. Therefore, if, as is usual, proper government is stated on the most local domain to which it might apply, the ECP is not violated in (42). It was for this reason that Huang (1982) proposed his Condition on Extraction Domains. According to the CED, syntactic extraction from a properly governed position is possible only if there is a CHAIN of lexically governed positions linking the position of the empty category with a matrix predicate. In the case of (37) the empty category is lexically governed by the preposition *of*, which is in turn governed by the noun *copy*, which is in turn governed by the matrix verb *see*; and this chain of government licenses the extraction of *which painting*. In (42), on the other hand, the empty category is governed by *of*, which is governed by *copy* as before; but *copy* is not lexically governed, so the chain is broken before it reaches the matrix predicate. Hence, the extraction is not licensed.

The CED looks very much like an extension of the ECP, and the phenomena it covers have been analyzed by Kayne (1981, 1983) as ECP effects. Although various considerations have been invoked to argue that the two constraints should not be collapsed, we believe that the arguments for maintaining a distinction between the two principles are weak.[11] Hence we find it interesting that it is easy in a TAG

to incorporate CED effects under the ECP in much the way that Kayne has, and it is such an analysis that we are developing in this paper. If it turns out that, for empirical reasons, the CED should appear as an independent stipulation in the theory of grammar, we can easily impose it in a TAG as a local constraint on the adjunction of the appropriate auxiliary trees. Even in this latter case, however, the effects of the CED would be linked to the ECP quite directly.

We can account in a natural way for the difference between (37) and (42) in our TAG analysis only if the initial trees for the two sentences differ appropriately, since the auxiliary set adjoined into the initial trees is the same in the two cases. Let us note then that the empty category in the initial tree for (37), namely (39), is lexically governed by the verb and that it is properly governed under the definition in (18) because its maximal government domain is the root of the tree in which it appears. On the other hand, in the initial tree for (42) (given as (43)) the empty category is in subject position, where it is not governed by the verb:

(43)



Since the initial tree (43) also does not contain an antecedent for the empty category, the empty category is neither antecedent-governed nor lexically governed, and hence it is not licensed by the constraints in Section 1. Once (43) is ruled out as a possible initial tree, (42) is excluded as ill-formed without further stipulation. Since (39) is licensed by the TAG version of the ECP, sentence (37) is correctly predicted to be grammatical.[12] This use of the ECP to rule out CED violations is made easy in a TAG grammar because the natural domain of locality in a TAG, the elementary tree, is larger than that normally assumed in other approaches and because the adjunction operation allows the recursive extension of the constraints imposed on a local domain.

3.2 Lexical government and extraction from NP

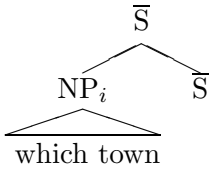
The above analysis leads to a natural explanation for the contrast in acceptability between extraction from complement PPs and modifier PPs. As is well known, extraction from the latter is not possible, leading to the contrast between (37) (repeated as (44)) and (45):

(44) Which painting$_i$ did you see a copy of e$_i$?

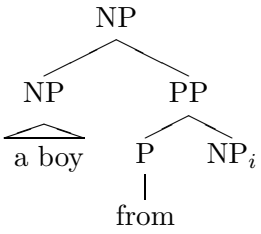
(45) * Which town$_i$ did you see a boy from e$_i$?

The reason for the unacceptability of (45), however, is apparent from the structure of the auxiliary tree set needed to derive it. The set would have to be as in (46):

(46) a.



b.



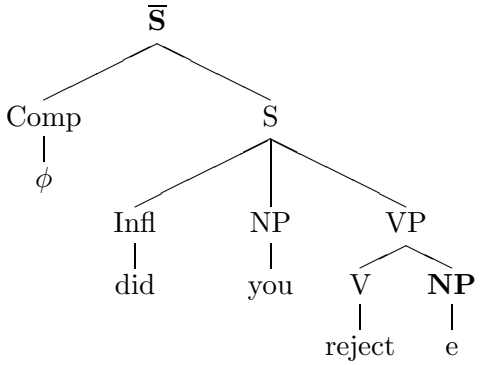
We see immediately that in tree (46b) the empty category of the extracted NP is not properly governed. It is lexically governed by the preposition *from*, but the preposition itself is not lexically governed since it is not c-commanded by a zero-bar-level category. Hence, by the definitions in (15) and (16), the maximal government domain of NP$_i$ is only PP and not the root of the auxiliary tree. Thus, NP$_i$ is not properly governed and so the configuration in (46b) is ruled out by the ECP.

Once we have introduced auxiliary sets, we must alter our analysis of the CNPC. We showed, in Kroch and Joshi (1985) and Kroch (1987), that in a TAG with only simple auxiliary trees, CNPC violations are not generable. In a TAG with auxiliary sets, however, it is possible to generate such violations unless the relevant tree sets are excluded. Thus, consider the following sentence:

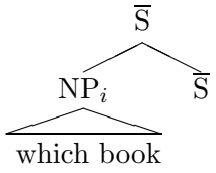
(47) * Which book$_i$ did you reject the report which$_j$ e$_j$ mentioned e$_i$?

Given the initial tree in (48), if we allow the tree set in (49), we can generate (47) by adjoining the tree set to the boldface nodes in the initial tree:

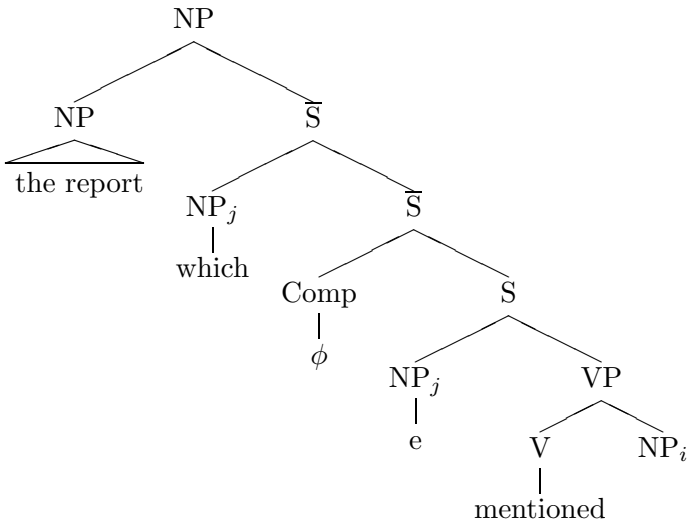
(48)



(49) a.



b.



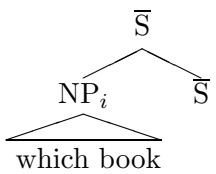
Since the tree in (49b) has the configuration that we have proposed for relative clauses (Kroch and Joshi 1985), the tree set (49) is basically well-formed. Like the tree in (46b), however, the maximal government domain of the foot node of (49b) is only a daughter of the root node and not the root node itself. Thus, extraction from relative clauses reduces to an ECP violation, as in the analysis of Kayne (1981).

As has been noted by more than one author (see for example Chomsky 1986), the CNPC does not apply with the same strictness to extraction from nominal complements as it does to extraction from relative clauses. The relative acceptability of (50) as compared to (47) is clear evidence of this difference:

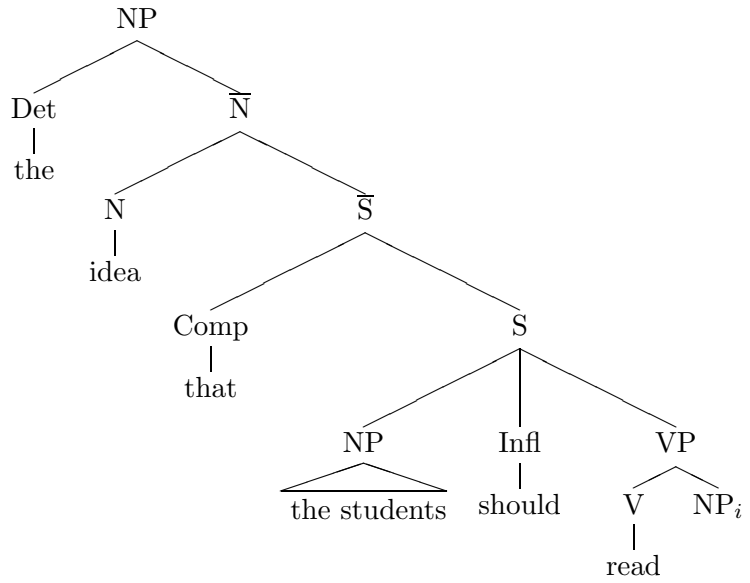
(50) ? Which book$_i$ did you reject the idea that students should read e$_i$?

The analysis we are proposing relates the difference between these two sentences to the difference in the internal structure of relative clause and nominal complement NPs. Since the N^o head of a nominal complement governs its complement clause, the maximal government domain of the foot node of the auxiliary tree in (51b) is the root node of the tree so that the tree is well-formed:

(51) a.



b.



When the tree set (51) is adjoined to the initial tree (48), the sentence (50) is generated.

4 *Wh-* islands

4.1 Blocking extraction from indirect questions

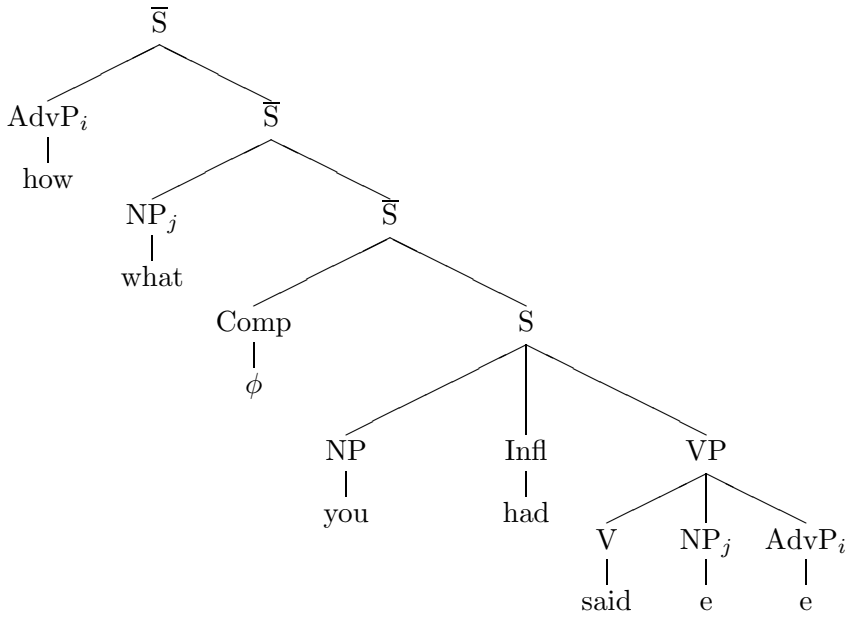
Restricting our attention for the moment to simple, one-part, auxiliary trees, we can insure the ungrammaticality of *wh-* island violations by imposing a well-formedness constraint on elementary trees to the effect that they can contain no more than one fronted *wh-*. This stipulation is needed in any case since multiple *wh-* fronting is impossible in English simple sentences. Since, by the definition of linking, the dislocated element in a *wh-* type dependency must originate in the same elementary tree as the empty category with which it is construed, there will be no way of generating *wh-* island violations. Thus, under TAG analysis using only one-part trees, the ungrammaticality of the sentences in (52) follows directly from the ill-formedness of the initial trees which would enter into their derivation:

(52) a. * How$_i$ did he wonder what$_j$ you had said e$_j$ e$_i$?

b. * Who$_i$ did he wonder what$_j$ e$_i$ had said e$_j$?

Thus, the initial tree for (52a) would be (53), in which both *how* and *what* are left adjoined to $\overline{\text{S}}$:

(53)



This analysis of the *wh-* island effect makes an interesting prediction that we are, fortunately, able to confirm. If we find a language which, like English, allows extraction freely from tensed clauses and in which multiple *wh-* fronting is possible, we expect extractions from indirect questions to occur freely. As it happens, there is at least one language of the sort needed to test this expection. The language in question is Rumanian, which shares free extractability from tensed clauses with its Romance sister languages and which has borrowed multiple *wh-* fronting from the Slavic languages that it has for so long been in contact with. Comorovski (1986) reports that, in Rumanian, the multiple *wh-* fronted sentence in (54) is perfectly acceptable, as are *wh-* island extractions like (55), whose English counterpart would be completely unacceptable:

(54) Cine$_i$ cui$_j$ ce$_k$ ziceai ca t$_i$ i$_j$-a promis t$_k$ t$_j$?[13]
'Who$_i$ did you say t$_i$ promised what$_k$ to whom$_j$?'

(55) Pentru care clauza$_i$ vrei sa afli cine$_j$ t$_j$ nu a decis inca ce$_k$ va vota t$_k$ t$_i$?
'For which paragraph$_i$ do you want to learn who$_j$ t$_j$ has not decided yet what$_k$ he will vote t$_k$ t$_i$?'

As sentence (55) shows, Rumanian allows extraction out of an unbounded number of *wh-* islands, a fact that correlates with the, in principle, unbounded number of *wh-* phrases that can appear at the front of a given clause. Comorovski points out that this correlation is just what is predicted by the subjacency constraint; and the fact that our TAG analysis makes just the same prediction illustrates how directly subjacency follows from the factoring of recursion and local dependencies that the TAG formalism enforces.

4.2 Asymmetries in extraction from indirect questions

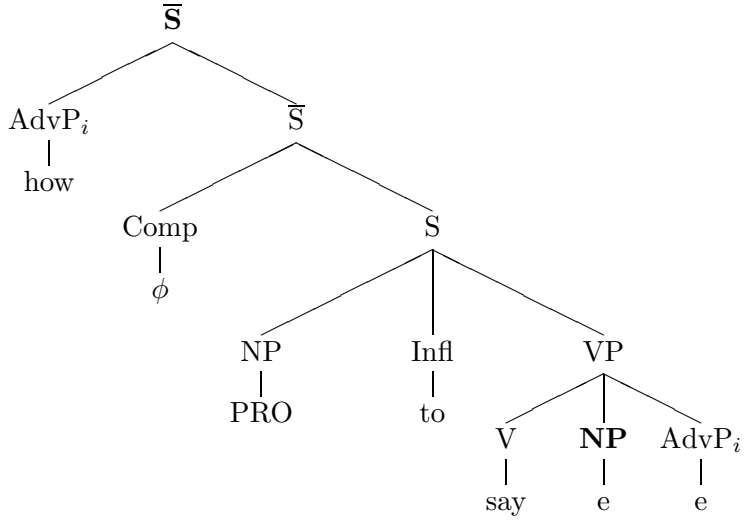
Let us now consider the contrast in acceptability between the sentences of (4), repeated as (56):

(56) a. ? What$_i$ were you wondering [how$_j$ to say e$_i$ e$_j$]?

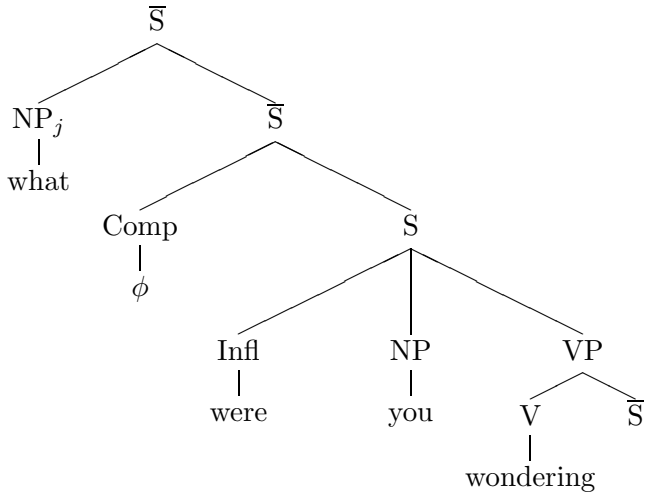
b. * How$_i$ were you wondering [what$_j$ to say e$_j$ e$_i$]?

Sentence (56a), while slightly unnatural, seems to us essentially grammatical while (56b) is entirely unacceptable. Under the analysis of *wh-* islands that we have sketched, this difference cannot be accommodated, since in the derivation of either sentence the initial tree would have to contain two preposed *wh-* constituents, and such trees are ill-formed. However, the extension of the TAG formalism to allow auxiliary sets gives us a straightforward way of generating the acceptable sentence (56a) without admitting sentences like (56b). The initial and auxiliary trees needed to derive (56a) are as follows:

(57)



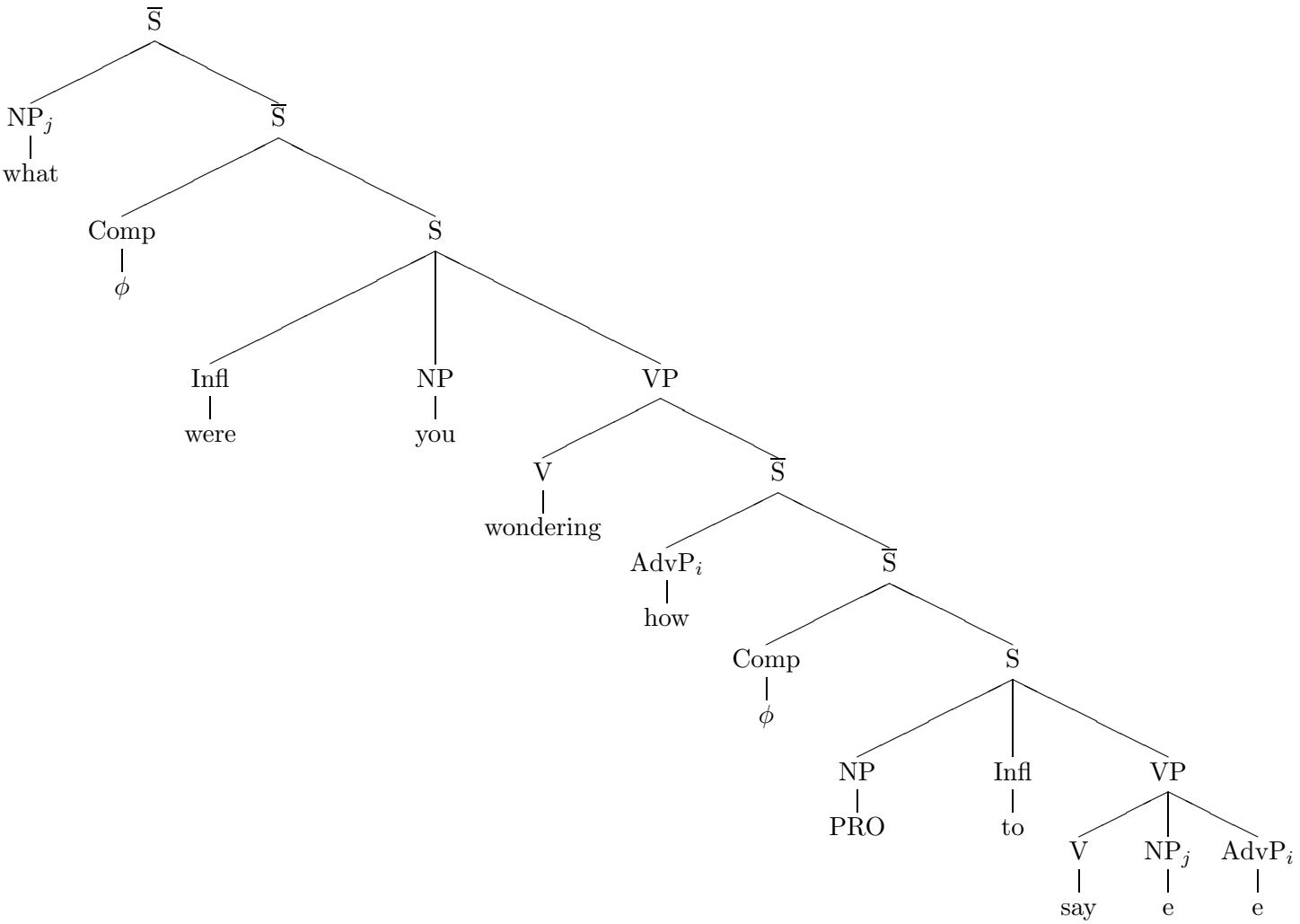
(58) a.



b. NP$_j$

The component (58b) of the tree set is in a sense degenerate since it consists of only one node; but it fits the definition of an athematic auxiliary tree and its use causes no formal difficulties. When it is adjoined to the NP node dominating the empty category in (57) its only effect is to add the index j to that node. The fact that it bears an index is enough to guarantee that it will only be adjoined to a node dominating an empty category, since otherwise the definition of linking would be violated. When the component in (58a) is adjoined to the boldface $\overline{\text{S}}$ node in (57) and the second component to the boldface NP node, the resultant tree (59) represents the sentence (56a).

(59)



Inspection of the elementary trees (57) and (58) makes it clear how to block the generation of (56b). In (57) there are two empty categories, one coindexed with *how* and the other not bearing an index. Both of these empty categories, however, are properly governed, since the indexed one is antecedent-governed and the other is lexically governed. In the trees needed to derive (56b), on the other hand, the

initial tree corresponding to (57) would have an index on the lexically governed node, which would thus also be antecedent-governed, and no index on the node which was not lexically governed. Hence the second empty category would not be properly governed. Since we require that empty categories be properly governed in the elementary tree in which they appear, the initial tree needed to generate (56b) is not well-formed. Thus, although we have had to use the extension of our formalism to auxiliary sets in order to generate (56a), we need make no changes in our linguistic constraints on the well-formedness of elementary trees to rule out (56b), a welcome result.[14] It should be obvious, moreover, that sentences like (52b), in which a subject is extracted from a *wh-* island, will be ruled out on an auxiliary set analysis just as (56b) is.

4.3 Extraction from indirect questions in Italian and English

In his well-known article on subjacency in Italian, Rizzi (1982) pointed out that Italian allowed extraction from indirect questions and so apparently differed from English. More recent work on English, however, has shown that the pattern of allowable extractions is quite similar in the two languages. Thus, the contrast between (56a) and (56b) discussed above is reproduced in Italian (Zanuttini, pers. comm.), and the pattern described by Rizzi can be reproduced in English (Grimshaw 1986). That pattern, whose analysis in a TAG we now propose to consider, is given in the sentences of (60) and (61). These sentences are slightly modified versions of Grimshaw's sentences (3), (4), (5a) and (5b), respectively.

(60) a. (I knew) which book$_i$ the students would forget who$_j$ e$_j$ wrote e$_i$.

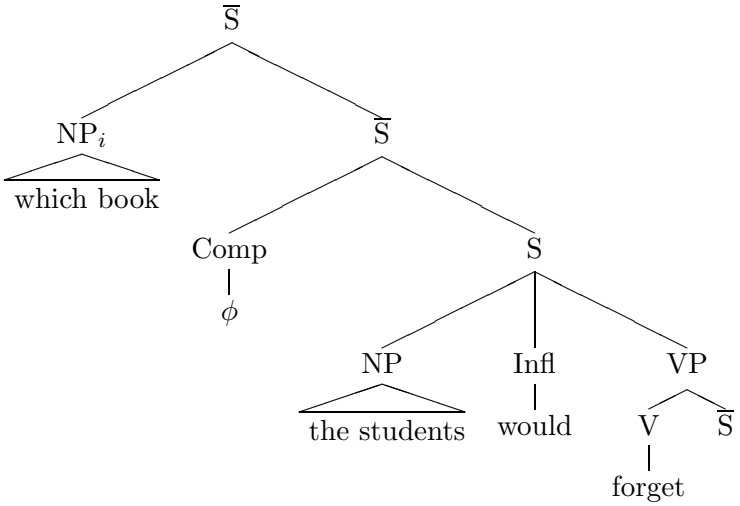
b. (I knew) which book$_i$ the TAs told us (that) the students would forget who$_j$ e$_j$ wrote e$_i$.

(61) a. * (I knew) which book$_i$ the students would forget who$_j$ e$_j$ told us (that) Dorothy Sayers wrote e$_i$.

b. * (I knew) which book$_i$ the students would forget who$_j$ e$_j$ told us who$_k$ e$_k$ wrote e$_i$.

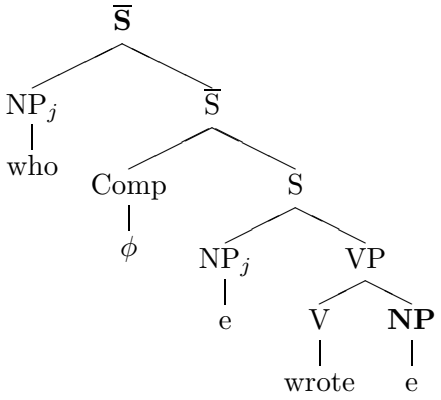
As Grimshaw points out, the sentences of (60) are awkward in English but they are a great deal more acceptable than those of (61). Assuming that we wish to capture this contrast in our TAG grammar, we can do so fairly simply. Sentence (60a) is exactly parallel to (56a) and can be generated by adjoining the auxiliary tree set in (62) to the boldface nodes in the initial tree (63). Note that we are ignoring, for the sake of brevity, the inversion of subject and auxiliary. The proper statement of the constraints of this inversion are discussed in detail in Moser (1987).

(62) a.



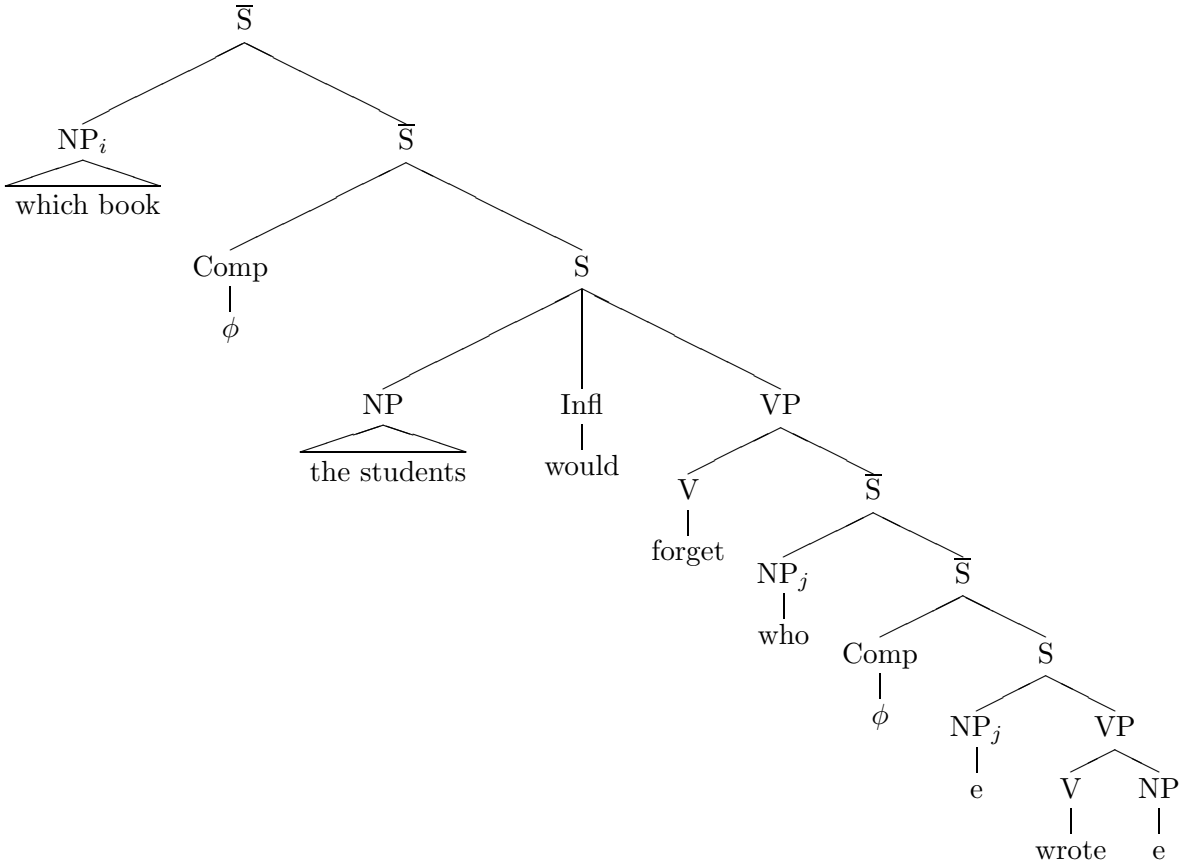
b. NP$_i$

(63)



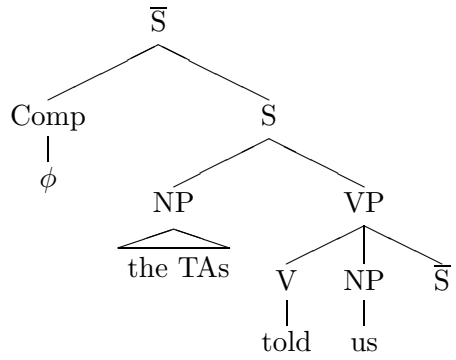
The resultant tree is (64):

(64)



Sentence (60b) can now be generated from (64) by the adjunction of the object complement auxiliary tree in (65) at the boldface node in the former:

(65)



The structure of the resultant tree is obvious.

To avoid generating the sentences of (61), we must add a stipulation to our grammar in the form of a local constraint on all elementary trees with a *wh-* in Comp. The constraint will have the effect of blocking the adjunction of any other *wh-* tree if the result would be to produce a configuration with two *wh-* elements

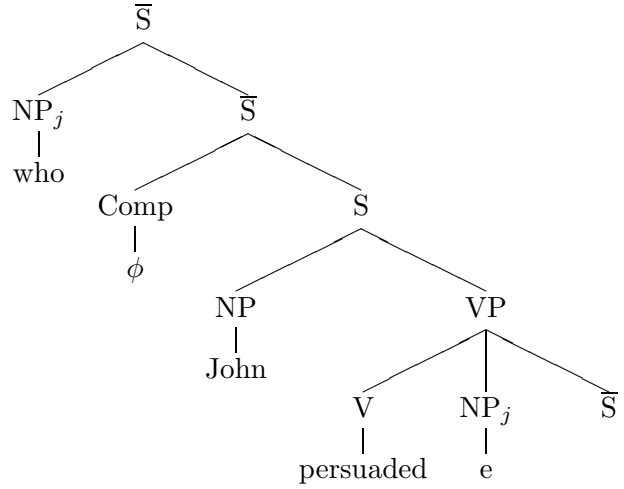
adjacent to one another. This stipulation is needed in any case to rule out sentences like the following:

(66) * (I know) which town$_j$ who$_i$ John persuaded e$_i$ to visit e$_j$.

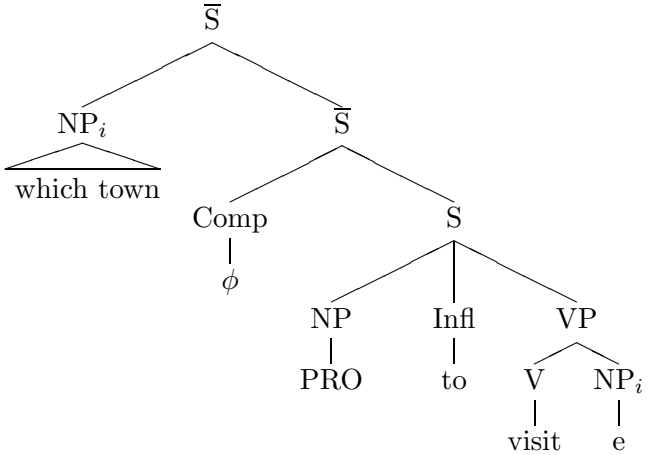
[15]

Without an appropriate local constraint to block their adjunction, sentence (66) will result from the adjunction of the auxiliary tree in (67) to the initial tree in (68):

(67)



(68)



These trees are, of course, well-formed and are needed to generate the grammatical sentences of (69) and (70) respectively:

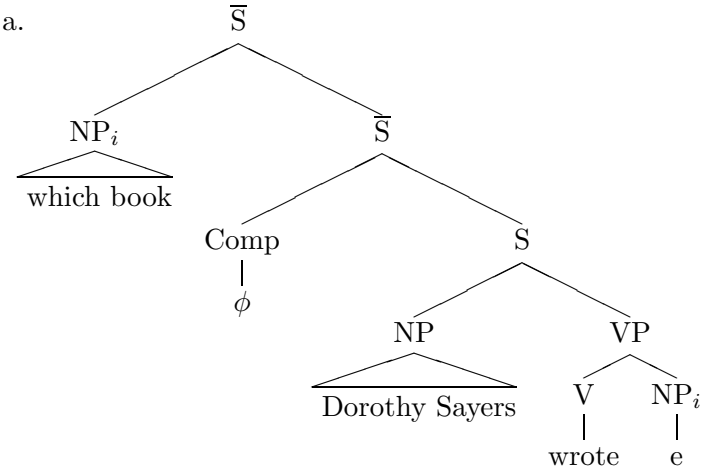
(69) (I know) who John persuaded to visit the town.

(70) (I know) which town John persuaded you to visit.

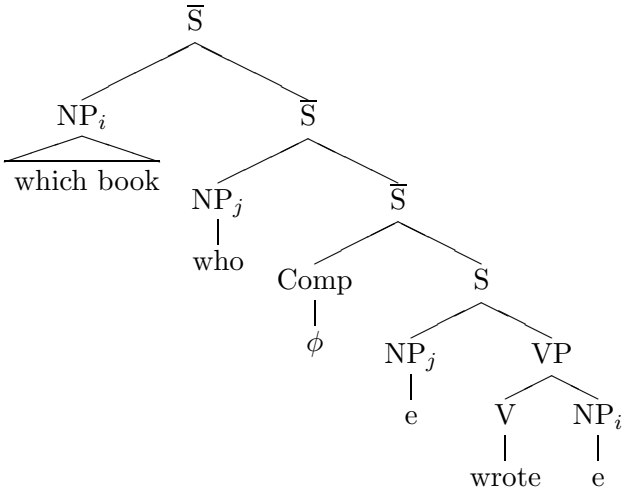
The local constraint we are proposing is also needed to block the generation with two-part trees of simplex sentences with multiple fronted *wh-* elements like those that would correspond to the tree in (53). We have ruled out such structures as possible elementary trees; but with the addition of auxiliary sets, they can now be generated by adjoining an appropriate auxiliary set to a simple sentence with one fronted *wh-* and one unbound empty category. Needless to say, the fact that we must rule out structures with multiple *wh-* fronting at two different levels suggests that our analysis has missed a linguistic generalization. Rectifying this problem is beyond the scope of the present discussion because it requires a substantial reformulation of the notion of local constraint. The required reformulation has been worked out by Vijay-Shankar (1987); and under it, the desired unification of the two constraints is possible (see Moser 1987). This unified constraint remains, however, an independent, hence language-particular, stipulation. That this result is correct is demonstrated by the fact, stressed by Comorovski, that Rumanian shows none of the limitations on extraction out of *wh-* islands found in Italian and English. Furthermore, the Rumanian counterpart of sentence is entirely (66) grammatical.

Once adjunctions are blocked which create structures with multiple fronted *wh-* elements, the sentences of (61) are ruled out whether simple auxiliary trees or auxiliary sets are used in their derivation. Thus, if we attempt to derive the sentences with simple auxiliary trees, the initial trees would have the structures in (71) and the auxiliary trees the structures of (72) and (73):

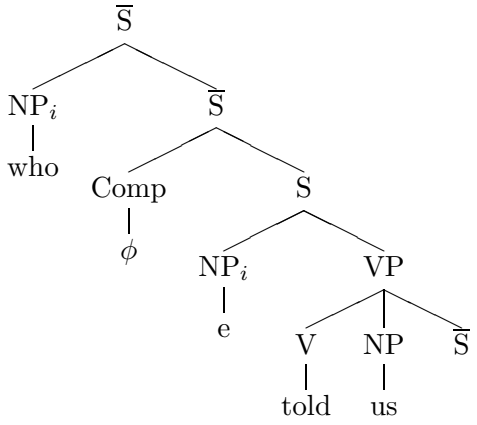
(71) a.



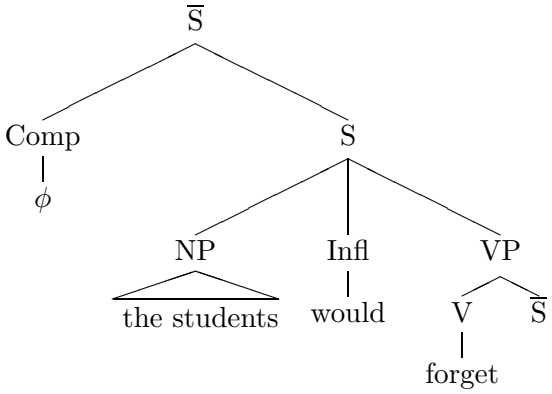
b.



(72)



(73)



From these structures it is immediately clear that (61b) is ruled out since its initial tree (71b) contains two fronted *wh-* elements. Of course, (61a) is also ruled out since the adjunction of (72) to (71a) produces the illicit configuration. Note that the derivation under which (73) is first adjoined to (71a) and then (72) is adjoined into the resulting structure is ruled out by the TAG version of the Projection Principle.

If we now try to derive the sentences of (61) using auxiliary sets, we find that the constraint against multiple *wh-* fronting continues to block the derivations. Because

of the requirement that all member trees of an auxiliary set be adjoined into the same elementary tree, the auxiliary set containing the preposed *wh-* phrase *which book* must be adjoined to the initial tree containing the verb *wrote* as the first step in the derivations of both (61a) and (61b). As readers can easily check for themselves, this requirement means that any otherwise licit derivation of either sentence will run afoul of the constraint against multiple *wh-* fronting. As before, derivations which avoid violating this constraint are blocked by the local constraint that enforces the Projection Principle.

5 Conclusions

We hope to have shown in this paper that a very wide range of constraints on *wh-* extraction can be accommodated within the TAG formalism in a natural way. A major attraction of the formalism for a generative linguist is the fact that the effects of subjacency follow from the adjunction operation itself once the form of elementary trees is fixed. In an earlier paper (Kroch and Joshi 1985) we presented this result, demonstrating that a TAG grammar would reflect such subjacency phenomena as the *wh-* island effect and the CNPC, although it contained no explicit reference to the bounding of long-distance dependencies. We were not able, however, to give a satisfactory account of the exceptions to subjacency that have been of concern in the recent transformational literature. In the present paper, by introducing the use of auxiliary tree sets as a supplement to simple auxiliary trees, we have been able to extend our earlier analysis to cover a wide range of apparently exceptional cases. Our earlier work also showed that ECP could be stated easily as a constraint on the well-formedness of elementary trees, thus accounting for the simplest cases of subject/object extraction asymmetries. The present discussion has shown that the analysis can be extended to the range of cases covered by the CED. Moreover, this extension is made possible by extending the ECP, the major well-formedness constraint on empty categories, to the foot nodes of auxiliary trees, thereby pointing toward an explanation of CED effects along lines similar to Kayne's recent work.

The naturalness with which the TAG formalism incorporates the empirical results of work on constraining *wh-* movement provides, in our opinion, further evidence that the TAG formalism should be considered a serious candidate for the proper notation of universal grammar. This evidence has been accumulating since our first attempts to apply the formalism to linguistic description. In other work, we have shown, among other results, that TAG allows a natural syntactic treatment of subject-to-subject raising (Kroch and Joshi 1985), that it allows an elegant statement of the constraints on extraposition (Kroch and Joshi 1987), and that its additional generative capacity beyond context-free grammar can be exploited to provide a sound syntactic analysis of West Germanic verb raising (Kroch and Santorini forthcoming). The present paper should be seen as a contribution to a larger ongoing project, an attempt to establish the extent and limitations of the utility of TAG to the task of fixing the form of universal grammar.

6 References

Aoun, J. and D. Sportiche. 1983. On the Formal Theory of Government. The Linguistic Review 2:211–236.

Baltin, M. 1982. A Landing Site Theory of Movement Rules. Linguistic Inquiry 13:1–38.

Baltin, M. 1986. Adverb Preposing and Extraction Constraints. Ms., New York University.

Chomsky, N. 1981. Lectures on Government and Binding. Dordrecht: Foris.

Chomsky, N. 1986. Barriers. Cambridge, MA: MIT Press.

Comorovski, I. 1986. Multiple *Wh* Movement in Romanian. Linguistic Inquiry 17:171–177.

Grimshaw, J. 1986. Subjacency and the S/$\overline{\text{S}}$ Parameter. Linguistic Inquiry 17:364–369.

Huang, J. 1982. Logical Relations in Chinese and the Theory of Grammar. Ph.D. dissertation, MIT.

Joshi, A. 1985. How much context-sensitivity is required to provide reasonable structural descriptions: Tree Adjoining Grammars. In D. Dowty, L. Karttunen, and A. Zwicky, eds., Natural Language Processing: Psycholinguistic, Computational and Theoretical Perspectives, 206–250. New York: Cambridge University Press.

Joshi, A. 1987. An Introduction to Tree Adjoining Grammars. In A. Manaster-Ramer, ed., The Mathematics of Language, 87–114. Philadelphia: Benjamins.

Joshi, A., L. Levy, and M. Takahashi. 1975. Tree Adjunct Grammars. Journal of the Computer and System Sciences 10:136–163.

Kayne, R. 1981. ECP Extensions. Linguistic Inquiry 12:93–133.

Kayne, R. 1983. Connectedness. Linguistic Inquiry 14:223–251.

Koopman, H. 1984. The Syntax of Verbs: From Verb Movement Rules in the Kru Languages to Universal Grammar. Dordrecht: Foris.

Kroch, A. 1982. A Quantitative Study of Resumptive Pronouns in English Relative Clauses. Paper presented at the 13th NWAVE conference, University of Pennsylvania.

Kroch, A. 1987. Unbounded Dependencies and Subjacency in a Tree Adjoining Grammar. In A. Manaster-Ramer, ed., The Mathematics of Language, 143–172. Philadelphia: Benjamins.

Kroch, A. and A. Joshi. 1985. The Linguistic Relevance of Tree Adjoining Grammar. University of Pennsylvania Department of Computer and Information Sciences Technical Report MS-CIS-85-16.

Kroch, A. and A. Joshi. 1987. Analyzing Extraposition in a Tree Adjoining Grammar. In G. Huck and A. Ojeda, eds., Syntax and Semantics: Discontinuous Constituents, 107–149. New York: Academic Press.

Kroch, A. and A. Joshi. Forthcoming. Some Aspects of the Linguistic Relevance of Tree Adjoining Grammar. To appear in Linguistics and Philosophy.

Kroch, A. and B. Santorini. Forthcoming. The Derived Constituent Structure of the West Germanic Verb Raising Construction. To appear in R. Freidin, ed.,

Proceedings of the Princeton Conference on Comparative Grammar. Cambridge, MA: MIT Press.

Lasnik, H. and M. Saito. 1984. On the Nature of Proper Government. Linguistic Inquiry 15:235–289.

Moser, M. 1987. Some Feature Assignments in TAGs. Ms., University of Pennsylvania.

Rizzi, L. 1982. Issues in Italian Syntax. Dordrecht: Foris.

Vijay-Shankar, K. 1987. A Study of Tree Adjoining Grammar. Ph.D. dissertation, University of Pennsylvania.

Vijay-Shankar, K. and A. Joshi. 1985. Some Computational Properties of Tree Adjoining Grammars. Proceedings of the Association of Computational Linguistics 11:82–93.

Vijay-Shankar, K., D. Weir, and A. Joshi. 1986. On the Progression from Context-Free to Tree Adjoining Languages. Ms., University of Pennsylvania.

Williams, E. 1978. Across the Board Application. Linguistic Inquiry. 9:31–45.

7 Notes

* This paper is based on work supported in part by NSF grants DCR 84-11726 and CER MCS-8219196, Aravind Joshi, principal investigator. The author expresses his thanks to the many people who have discussed the issues involved. Special thanks are due to the members of the TAG research group at the University of Pennsylvania: Sharon Cote, Caroline Heycock, Aravind Joshi, Megan Moser, Beatrice Santorini, K. Vijay-Shankar, and David Weir.

1. This sentence can be assigned two structural descriptions, one with two preposed *wh-* phrases and one with the subject *wh- in situ* . Both structures are ungrammatical but only the former is of interest for our discussion.

2. One of the advantages of stating the well-formedness constraints on auxiliary trees as we have is that they block extraction from sentential subjects. The type of auxiliary tree that would be required for such extractions would have a foot node $\overline{\text{S}}$ in subject position, and such a tree would not be a possible complement tree since subject position is not lexically governed. Since the tree would not have a Chomsky-adjunction structure, it would also not be a possible athematic tree.

3. We assume in this paper that preposed *wh-* phrases are adjoined to $\overline{\text{S}}$, as proposed in Baltin (1982).

4. Data from a certain marginal construction in English parallel the pattern in (26) and (27). In unplanned speech doubly filled Comps sometimes occur in indirect questions, as in the following example taken from a corpus that we have gathered through informal observation:

(i) That tells you [how many days]$_i$ that the car will be in the shop e$_i$.

Such examples occur from time to time when the *wh-* element introducing the indirect question is a phrase rather than a simple pronoun, though whether they should be considered grammatical in the colloquial language is difficult to judge. It is striking, however, that, while examples like (i) or cases with lexically governed traces

are possible, examples with subject gaps, like (ii), both fail to occur and are entirely unacceptable:

(ii) * I asked her [which car]$_i$ that e$_i$ was in the shop least often.

It is clear that the difference between (i) and (ii) is due to a difference in the conditions under which subject and adjunct extractions are licensed.

5. This analysis has interesting implications for the analysis of Italian (Chomsky 1981, Rizzi 1982) and other languages that do allow the free inversion of subjects. It has been noted that these languages also allow subject pro drop, have relatively rich subject-verb agreement, and do not display the *that*-trace effect in extraction from object complements. More than one attempt has been made to relate these features, and our adjacency requirement on the antecedent government of subject traces provides a new perspective on the problem of how to formulate this relationship. Suppose that the verbal agreement affix is sufficiently pronominal to be an antecedent governor of the subject position in a language like Italian but not in English. Then, since Infl is adjacent to the subject position, it will always function as such a governor in Italian, and so the appearance of inverted and extracted subjects will not be constrained by the adjacency requirement, which will always be met by the agreement affix. In English, on the other hand, movement of the subject will be highly constrained, since the moved subject itself must be the antecedent governor of the subject trace.

6. Sentences (34a) and (34b) would be more acceptable if the subordinate clause were introduced by the overt complementizer *that*. Unfortunately, the presence of an overt complementizer in (34c) would make it a violation of the ECP even if the topicalized NP had no effect.

7. In pre-*Barriers* GB analyses this position is the Comp of the subordinate clause. In the TAG analysis presented here, it is the Chomsky-adjoined fronted position in the initial tree of the subordinate clause.

8. A further piece of evidence that adjacency with an antecedent is required for the government of subject traces is provided by such examples as the following:

(i) a boss who$_i$ if you were late, e$_i$/he$_i$ used to yell at you

In a corpus of relative clauses that we collected for another study (Kroch 1982), we found that resumptive pronouns were extremely common and highly acceptable in relative clauses of this type. Indeed, in speech, the resumptive pronoun occurs much more often than not. Moreover, while in other constructions, resumptive pronouns on subject position are very rare, in particular rarer than resumptive pronouns in VP, in this case they are more common on subject than object position. Thus, resumptive pronouns were more common in sentences like (i) than in sentences like (ii):

(ii) a boss who$_i$ if you were late, you had to lie to e$_i$/him$_i$

This pattern makes this sort of relative clause seem similar to relatives in which a resumptive pronoun amnesties an ECP violation, as in (iii):

(iii) a problem that$_i$ I couldn't figure out how$_j$ it$_i$/*e$_i$ had been solved e$_j$

These cases are clearly ungrammatical without a resumptive pronoun and are quite acceptable in colloquial speech with one. To see how these facts are related to the point under discussion here, note the following pair of sentences:

(iv) When it rains, Mary likes to go for walks.

(v) Mary, when it rains, likes to go for walks.

Both of these sentences are grammatical. In Baltin's (1982) terms, the adverbial clause in (iv) is adjoined to S and the one in (v) to VP. Assuming this description for the sake of concreteness, we can note that a relative clause on subject position based on the structure of (iv) will violate the adjacency requirement on the subject trace-antecedent link, as the following bracketing illustrates:

(vi) a girl [$_{\overline{S}}$ who$_i$ [$_S$ when it rains [$_S$ e$_i$ likes to go for walks]]]

On the other hand, if the relative clause is based on the structure of (v), the adverbial clause, being adjoined to VP, will not intervene between the antecedent and trace. From these facts, the adjacency requirement predicts that relative clauses containing preposed adverbials should be possible both with and without resumptive pronouns, which is indeed the case. Furthermore, it is not surprising, given the highly marked character of the word order in (v), that the resumptive pronoun relative is the favored option in colloquial speech.

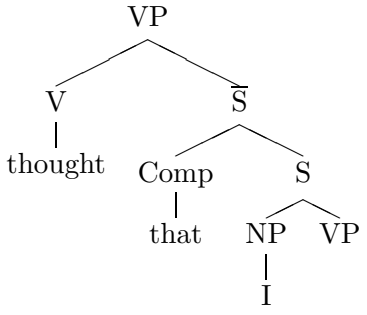
9. This constraint is not easily statable in the TAG formalism as we have presented it here. The addition of feature structures to nodes worked out in Vijay-Shankar (1987) makes possible the natural statement of this constraint as well as the other local constraints mentioned in the paper.

10. A TAG version of the Projection Principle is needed, among other reasons to rule out an illicit derivation of a sentence like (i) from the initial tree in (9) (repeated here as the labelled bracketing in (ii)) and the auxiliary tree in (iii):

(i) The child thought that I saw you.

(ii) [$_S$ [$_{NP}$ the child] [$_{VP}$ saw you]]

(iii)



The tree in (iii) does not violate any of the well-formedness constraints on elementarty trees that we have proposed, but its use in any derivation would lead to the

sort of change in thematic relations that we want to rule out in principle. We thank Ken Safir for pointing out the problem posed by such unwanted structures.

11. Huang's argument for distinguishing the ECP from the CED is based on facts from Chinese, a language without syntactic *wh-* movement. Huang gives evidence that Chinese obeys the ECP but not the CED, and so he concludes that the ECP applies at the level of Logical Form while the CED is a constraint on syntactic movement. This argument depends on assumptions about the nature of Logical Form that seem to us not to be well established.

12. A TAG analysis of the sort we are giving agrees with Kayne (1983) in treating the subjects of the infinitive complements of exceptional case marking (ECM) verbs as being antecedent-governed rather than lexically governed even though these subjects are case-marked by the matrix verb. The correctness of this approach is supported by the ungrammaticality of extractions from such subject NPs. Thus, note the contrast between (i) and (ii):

(i) Who$_i$ do you expect the parents of e$_i$?

(ii) * Who$_i$ do you expect the parents of e$_i$ to visit Rome?

Given the discussion of antecedent government in Section 2.2, our analysis also predicts that heavy-NP shift of the subjects of complements to ECM verbs should be ungrammatical. Here the facts are less clear, but there seems to be a contrast between the object control case in (iii) and the ECM case in (iv):

(iii) I persuaded e$_i$ to tackle the problem [a friend of mine from school]$_i$.

(iv) ? I believe e$_i$ to have made the discovery [a friend of mine from school]$_i$.

To the extent that (iv) is better than expected, this may be due to a superficial process of reanalysis under which the phrase *believe to have made the discovery* is treated as a complex verb.

13. The dative pronoun in this sentence is, as Comorovski shows, not a resumptive pronoun but an ordinary Romance doubled clitic.

14. For reasons that we do not understand, extractions that require the use of two-part auxiliary tree sets in a TAG share the feature that they are much more acceptable when the subject of the initial tree and the subject of the auxiliary tree are coreferential. Thus, the examples in the text are more acceptable than the following:

(i) ?? Which lines were you wondering how the actor said?

(ii) ?? Which city did you buy John's picture of?

(iii) ?? Which novel were you wondering who wrote?

Similarly, extractions from *wh-* islands, at least in English, are more acceptable when the embedded clause contains an infinitive or modal than when the embedded verb is tensed. Since the temporal interpretation of infinitive and modal verbs in indirect questions is usually controlled by the tense of the matrix sentence, this may also indicate that binding relations are relevant to extraction in two-part-tree cases.

15. We thank Polly Jacobson for pointing out the difficulties posed for us by sentences like (66).