

The Linguistic Relevance of Tree Adjoining Grammar*

Anthony Kroch and Aravind Joshi
University of Pennsylvania
MS-CIS-85-16

June 1985

Contents

1	Introduction	2
1.1	Motivation	2
1.2	Factoring recursion and co-occurrence restrictions	4
1.3	The plan of the paper	5
2	An introduction to the Tree Adjoining Grammar formalism	6
2.1	Tree Adjoining Grammar–TAG	6
2.2	TAG’s with ‘links’.	13
2.3	TAG’s with local constraints on adjoining	16
2.4	Some formal properties of TAG’s	23
3	Some linguistic examples	24
4	Raising and equi constructions in a TAG	39
4.1	The basic issues	39
4.2	The problem of nominalizations	44
4.3	Further considerations	47
5	The Passive in a TAG	49
5.1	The link between raising and the passive	49
5.2	A TAG analysis of the passive	52
5.3	Impersonal and raising passives	55

*This work was supported in part by NSF grants MCS8219196-CER, MCS8207294 and a grant from the Alfred P. Sloan Foundation for a Program in Cognitive Science, Grant No. 84-4-15.

6	<i>Wh</i>- movement in a TAG	57
6.1	Subjacency in a TAG	57
6.2	The <i>that</i> -trace effect in a TAG	64
6.3	A further example	71
7	Conclusions	74
8	References	77
9	Notes	82

1 Introduction

1.1 Motivation

In this paper we apply a new notation for the writing of natural language grammars to some classical problems in the description of English.¹ The formalism is the Tree Adjoining Grammar (TAG) of Joshi, Levy and Takahashi 1975, which was studied initially only for its mathematical properties but which now turns out to be an interesting candidate for the proper notation of meta-grammar; that is for the universal grammar of contemporary linguistics. Interest in the application of the TAG formalism to the writing of natural language grammars arises out of recent work on the possibility of writing grammars for natural languages in a metatheory of restricted generative capacity (for example, Gazdar 1982 and Gazdar et al. 1985). There have been also several recent attempts to examine the linguistic metatheory of restricted grammatical formalisms, in particular, context-free grammars. The inadequacies of context-free grammars have been discussed both from the point of view of strong generative capacity (Bresnan et al. 1982) and weak generative capacity (Shieber 1984, Postal and Langendoen 1984, Higginbotham 1984, the empirical claims of the last two having been disputed by Pullum (Pullum 1984)). At this point the TAG formalism becomes interesting because while it is more powerful than context-free grammar, it is only “mildly” so. This extra power of TAG is a direct corollary of the way TAG factors recursion and dependencies, and it can provide reasonable structural descriptions for constructions like Dutch verb raising where context-free grammar apparently fails. These properties of TAG and some of its mathematical properties were discussed by Joshi 1983.

It is our hope that the presentation below will support the claim, currently controversial, that the exploration of restrictive mathematical formalisms as meta-languages for natural language grammars can produce results of value in empirical linguistics. In spite of the fact that the syntactic theory of natural languages and mathematical linguistics share a common origin, the relevance of the latter to the former is a matter of contention. Linguists agree that an explanatory theory of language requires a restrictive specification of universal grammar since that notion defines the space of possible human languages. Without a restrictive universal grammar the problem of language acquisition becomes intractable for the child, who has to entertain so many hypotheses as to the correct grammar for his language that the limited primary data of his experience will not allow him to choose among them. Linguists also agree that the formalization of transformational-generative grammar (TG) as it is used in the Aspects and related models is far too permissive. In the search for a more restrictive theory, however, different researchers have taken very different tacks. Some have claimed that progress can best be made by reanalyzing the syntax of English and other languages that have received extensive treatment in TG grammar within systems of grammar that are provably less powerful in generative capacity than transformational grammars are. Only in this way can researchers be sure that the grammars they construct will not only be learnable but also usable in the real-time linguistic activities of parsing and generation that grammati-

cal knowledge underlies. Under this approach transformational grammars must be excluded by the theory of universal grammar because, since they can generate non-recursive sets, the languages they generate cannot be expected to be parsable within reasonable (i.e., polynomial) time bounds (Generalized Phrase Structure Grammar (GPSG) of Gazdar takes this approach). Other linguists, most notably Chomsky, have argued that a restrictive theory of universal grammar can and should be developed by the empirically driven discovery of constraints on rules and representations and that these constraints cannot be expected to restrict the generative capacity of grammars in any interesting way. Chomsky appears to believe that the effect of the constraints that comprise universal grammar on the mathematically defined generative capacity of possible human language grammars has little linguistic relevance (Chomsky 1977, 1980). For him the learnability problem is the only one that should constrain universal grammar. The parsability of languages is not a goal that universal grammar should aim to account for because it is doubtful that the complete set of grammatical sentences of a human language is necessarily parsable and it may even be that parsing is not algorithmic.

The thrust of GPSG and similar approaches is to elaborate a formal theory that constrains the generative capacity of possible grammars and then to show that, in spite of its mathematical restrictiveness, grammars admitted by this theory give empirically satisfactory analyses of various syntactic phenomena previously analyzed in transformational terms. One may consider alternative linguistic analyses for a certain phenomenon in these frameworks, but the formal power of the theory is always well understood. Chomsky's Government Binding theory (GB), on the other hand, fails to restrict universal grammar in any significant way because GB theory is insufficiently formalized to be mathematically investigable. One result of this insufficient formalization is that modifications can be made to the theory without it being clear what effect they have on the theory's restrictiveness or coverage. In consequence of these considerations, it should be a question of considerable interest whether the theory of grammar can be stated in a notation which simultaneously fulfills the criterion of mathematical restrictiveness cum explicitness and allows the formulation of syntactic analyses linguistically similar to those current in GB work. Given the extensive reformulation of transformational grammar represented by Government Binding theory, it would not be surprising to discover that this was possible. For example, Berwick and Weinberg (1984) claim that the mechanisms responsible for much of the excessive weak generative capacity of Aspects-style transformational grammars are no longer used in GB. The existence of a provably restrictive notation for a grammatical theory of the GB sort would support the hypothesis that the generative capacity of natural language is mathematically characterizable in an interesting way. And this result would support the psychological interest of current research on algorithmic parsing, which depends on being able to assume reasonable time bounds for sentence recognition. Furthermore, it would demonstrate that there was no need to choose between parsability and learnability as the basis for universal grammar, since even analyses constructed without concern for the former could be closely modeled in a formalism of restricted mathematical power for which reasonable parsing algorithms are known.

It may be important at this juncture to stress that discovering the generative capacity of natural languages to be limited has value independently of whether there are a finite or an infinite number of possible core grammars. The finite number hypothesis has been claimed to have important implications for learnability (Chomsky 1981a, Osherson et al. 1985, Chomsky 1984); and, if true, it invalidates certain arguments against the psychological reality of grammars that generate nonrecursive sets. However, the hypothesis does not dictate that universal grammar not restrict the generative capacity of natural language grammars. Indeed, the finiteness hypothesis, whatever its status, has no bearing on the linguistic interest of investigations into formalisms with restrictive generative capacity.

1.2 Factoring recursion and co-occurrence restrictions

In previous work we have investigated the mathematical properties of the TAG formalism (Joshi, Levy and Takahashi 1975, Joshi 1983). In this paper we attempt to demonstrate that this same formalism is one in which linguistically attractive syntactic analyses can be stated in natural ways. In particular, we will show that the TAG formalism can instantiate transformational style analyses of a number of important syntactic constructions. Thus, the distinction between ‘raising’ verbs like *seem* and ‘equi’ (i.e., control) verbs like *try*, which transformational grammar has always treated configurationally, is easily formulable in a TAG. Moreover, this formulation allows one to express formally the appropriate syntactic generalization that relates the raising construction to the passive. Similarly, the principle of subadjacency, which Chomsky claims accounts for some of Ross’s island constraints (e.g., wh- island and complex NP), can be made to follow from the way that embedding is formalized in a TAG; and the empty category principle can be given an entirely local formulation without need of referring to intermediate traces. As we shall see, these results all follow from the way in which the TAG formalism factors recursion and co-occurrence restrictions. Unlike a context-free grammar, which writes recursion into the rules that generate phrase structure, a TAG defines a finite set of simple sentence elementary trees and an adjunction operation that produces complex sentences through the combination of simple sentences. This approach is reminiscent of the use of generalized embedding transformations in Chomsky’s early work, and it shares with that formalism the attractive feature of allowing many co-occurrence facts to be stated on simple sentences. At the same time it promises to allow the implementation of transformational style linguistic analyses in TAG even though the system does not have transformational power.

Our aims in this paper are necessarily limited by considerations of space and the preliminary stage of our investigations. Some of the limitations of the paper are worth explicit acknowledgment. First, in working out how a grammar of English might be stated within the TAG formalism, we have chosen to develop in detail only two areas, raising (including the passive) and wh- movement. We have not attempted to resolve all the syntactic issues that our example sentences are relevant to but have limited our attention to central features of the two constructions

mentioned. Thus, we have not considered such admittedly important issues as: the proper representation of the auxiliary, subject-aux inversion, extraposition, etc. Second, in the course of our presentation we will, of course, compare our analyses to those of others working in various frameworks, but we do not have as a goal giving a systematic comparison of TAG with other formalisms. We want rather to show the substantial extent to which accepted syntactic facts can be captured in a TAG analysis, often in a way that sheds new light on the linguistic generalizations involved. Finally, the analyses presented here do not exhaust those which are possible within the constraints of a TAG. We have made certain choices on the basis of our judgments of linguistic plausibility which were not dictated by the TAG formalism. Thus, even if the reader rejects particular features of our linguistic analyses he/she should not conclude that the putatively superior analyses are necessarily unstateable within the TAG system. One example of this point is the absence of NP trace in our discussion of raising and passive. We decided to do without NP trace because it did not seem necessary to assume its existence in order to capture the generalizations we were concerned with, but it is possible to write a TAG which incorporates NP trace. Conversely, we did incorporate PRO into our grammar because we felt it was convenient to do so, but the TAG formalism itself does not require that subjectless infinitives have abstract PRO subjects.

Of course, the flexibility of the TAG notation has an important consequence for its status in linguistic theory. Since it allows incompatible analyses for central constructions, it cannot by itself be a theory of universal grammar. Indeed, in our discussion we will assume the existence of various modules that constrain the set of possible elementary structures beyond the constraints imposed by the TAG formalism itself. Thus, we here assume an X-bar theory that constrains the phrase structure of elementary trees, a case theory that constrains the appearance of lexical NPs, a theory of thematic relations, and a principle of proper government that constrains the appearance of empty categories. The choice of these constraints was, of course, not dictated by the TAG formalism itself. Rather we have implemented in our presentation those which made it easiest for us to instantiate the configurational analysis of grammatical relations that we want to demonstrate. What the TAG notation itself contributes to the theory of grammar is a constrained theory of syntactic embedding, one requiring that embedded structures be composed out of elementary structures in a fixed way and one which forces co-occurrence relations between elements that are separated in surface constituent structure to be stated locally as constraints on elementary trees in which those elements are co-present. The extra generative power of TAG beyond context-free grammar emerges as a corollary of factoring recursion and co-occurrence relations.

1.3 The plan of the paper

We have organized our presentation as follows. In section 2 we present the TAG formalism itself and illustrate its workings with formal language examples. We also introduce the local constraints on tree adjunction, defined somewhat differently than

in earlier work, and ‘links’ between one node and another. Both of these devices will be used extensively in the linguistic discussion. The section ends with a brief discussion of the formal properties that can be proven to hold of TAG’s. Section 3 is an introduction to the application of the TAG formalism to linguistic data. We show how embedding structures, including both complement structures and relative clauses, are generated; and we show how the application of local constraints will prevent overgeneration. In section 4 we discuss in detail the (subject-to-subject) raising construction and demonstrate how a TAG analysis allows us to preserve a syntactic and configurational account of the construction. We also show that the non-existence of raising nominalizations follows from the nature of tree adjunction. Section 5 discusses the passive. Here we implement a lexical analysis of the passive and show how, under this analysis, exceptional case-marking verbs like *believe* become raising predicates as an immediate consequence of undergoing the lexical rule of passivization. In section 6 we discuss *wh*-movement from the perspective of capturing the constraints on its application that have been worked out by transformational grammarians over the past twenty years. We show that subjacency is a necessary consequence of the TAG formalism itself and that the Empty Category Principle (ECP) of Government Binding theory can be stated in a TAG in a straightforward way. We also point out that the TAG formalism has the advantage of not appealing to that troublesome entity the ‘intermediate trace’ in its treatment of unbounded dependencies. Finally, section 7 briefly summarizes our conclusions.

2 An introduction to the Tree Adjoining Grammar formalism

2.1 Tree Adjoining Grammar–TAG

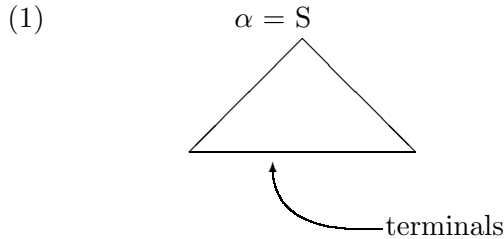
We will introduce the tree adjoining grammar (TAG) by first describing an alternate way of looking at the derivation of the strings and the corresponding derivation trees of a context-free grammar (CFG). Later we will introduce TAG’s in their own right. (This section is primarily based on the work reported in Joshi 1983, except for the section on local constraints, which is based on Vijay Shankar and Joshi 1985.) TAG’s are more powerful than CFG’s both weakly and strongly.² Consider the CFG G' defined in Example 2.1.1.

EXAMPLE 2.1.1: Let G' be a context-free grammar with the following productions.

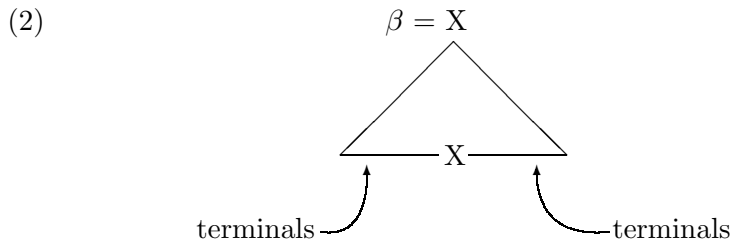
- $S \rightarrow a T b$
- $T \rightarrow S a$
- $T \rightarrow T T$
- $T \rightarrow a b$

S is the start symbol, S and T are the non-terminals and a and b are the terminal symbols.

We can now define a TREE ADJOINING GRAMMAR (TAG) G which is both weakly and strongly equivalent to G' . Let $G = (I, A)$ where I and A are finite sets of ELEMENTARY TREES. The trees in I will be called the INITIAL TREES and the trees in A , the AUXILIARY TREES. A tree α is an initial tree if it is of the form in (1):



That is, the root node of α is labeled S and the frontier nodes are all terminal symbols. The internal nodes are non-terminals. A tree β is an auxiliary tree if it is of the form in (2):

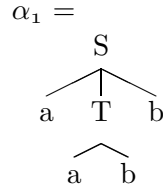


That is, the root node of β is labeled X where X is a non-terminal and the frontier nodes are all terminals except one which is labeled X the same label as that of the root. The node labeled X on the frontier will be called the foot node of β . The internal nodes are non-terminals. The initial and the auxiliary trees are not constrained in any manner other than as indicated above. The idea, however, is that both the initial and auxiliary trees will be ‘minimal’ in some sense. An initial tree will correspond to a ‘minimal’ sentential tree (i.e., without recursing on any non-terminal) and an auxiliary tree, with root and foot node labeled X will correspond to a ‘minimal’ recursive structure that must be brought into the derivation, if one recurses on X .

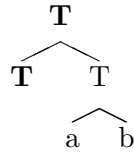
For the grammar in Example 2.1.1 above, we define our equivalent TAG, $G=(I,A)$ as in Example 2.1.2.

EXAMPLE 2.1.2:

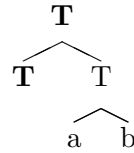
I:



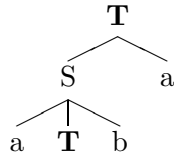
A : $\beta_1 =$



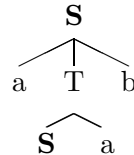
$\beta_2 =$



$\beta_3 =$



$\beta_4 =$



(The root and foot nodes of the auxiliary trees in this example have been set in boldface for the reader's convenience.)

We will now define a composition operation called adjoining (or adjunction), which composes an auxiliary tree β with a tree γ . Let γ be a tree containing a node \mathbf{n} bearing the label X and let β be an auxiliary tree whose root node is also labeled X . (Note that β must have, by definition, a node (and only one such) labeled X on the frontier.) Then the adjunction of β to γ at node \mathbf{n} will be the tree γ' that results when the following complex operation is carried out:

1. The sub-tree of γ dominated by \mathbf{n} , call it \mathbf{t} , is excised, leaving a copy of \mathbf{n} behind.
2. The auxiliary tree β is attached at \mathbf{n} and its root node is identified with \mathbf{n} .
3. The sub-tree \mathbf{t} is attached to the foot node of β and the root node \mathbf{n} of \mathbf{t} is identified with the foot node of β .

Figure 1 illustrates this operation.

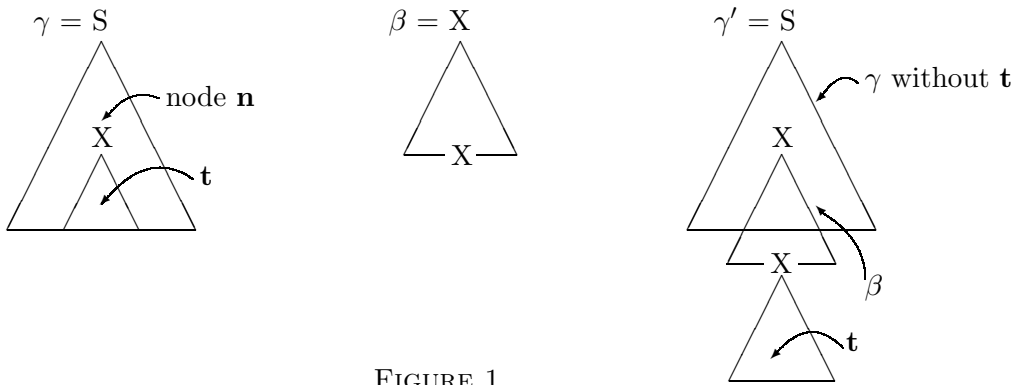


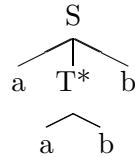
FIGURE 1

The intuition underlying the adjoining operation is a simple one but the operation is distinct from other operations on trees that have been discussed in the literature. In particular, we want to emphasize that adjoining is not a substitution operation.³

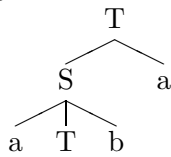
Let us now look at some derivations in the TAG, $G=(I,A)$ of Example 2.1.2, which we give as Example 2.1.3.

EXAMPLE 2.1.3:

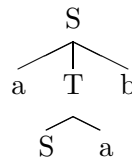
$$\gamma_0 = \alpha_1 =$$



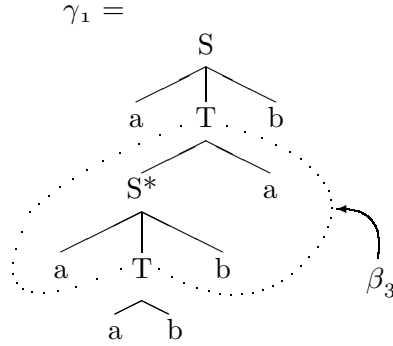
$$\beta_3 =$$



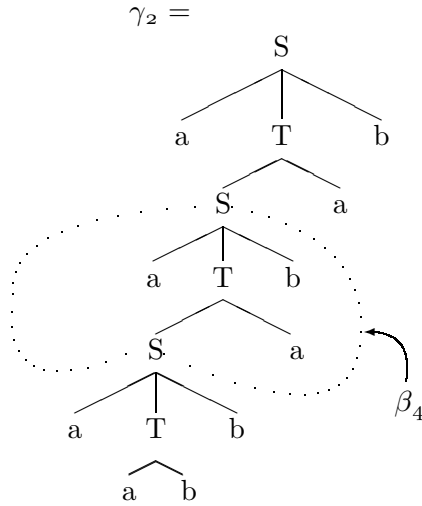
$$\beta_4 =$$



β_3 will be adjoined to γ_0 at T as indicated in γ_0 . (We will use * to indicate the node to which adjunction is made). The resulting tree γ_1 is then



We can continue the derivation by adjoining, say, β_4 , at S as indicated in γ_1 . The resulting tree γ_2 is then



Note that γ_0 is an initial tree, a sentential tree. The derived trees γ_1 and γ_2 are also sentential trees. It is clear in this example that the TAG, G will derive all and only the sentential trees of the CFG, G , starting from the initial tree of G . Thus G will also generate the string language $L(G)$ of G .

Let us now define two auxiliary notions, the tree set of a TAG and the string language of a TAG. Suppose $G=(I,A)$ is a TAG with a finite set of initial trees, a finite set of auxiliary trees, and the ADJOINING operation, as above. Then we define the TREE SET of a TAG G , $T(G)$, to be the set of all trees derived in G starting from initial trees in I . We further define the STRING LANGUAGE (or LANGUAGE) of G to be the set of all terminal strings of the trees in $T(G)$. The relationship between TAG's, context-free grammars, and the corresponding string languages can then be summarized in the following theorems (Joshi, Levy, and Takahashi 1975, Joshi 1983):

THEOREM 2.1.1: For every context-free grammar, G' , there is a TAG, G , which is both weakly and strongly equivalent to G' . In other words, $L(G)=(G')$ and

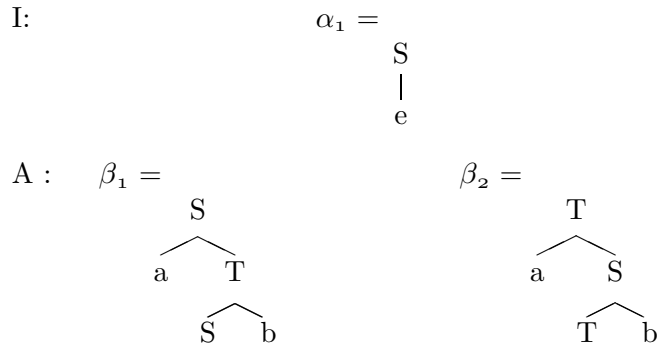
$T(G)=T(G')$. Furthermore, it can be shown that there is a general algorithm for obtaining the equivalent TAG for any CFG.

THEOREM 2.1.2: There exists a non-empty set of TAGs \mathcal{G}_1 such that for every $G \in \mathcal{G}_1$, $L(G)$ is context-free but there is no CFG G' such that $T(G')=T(G)$.

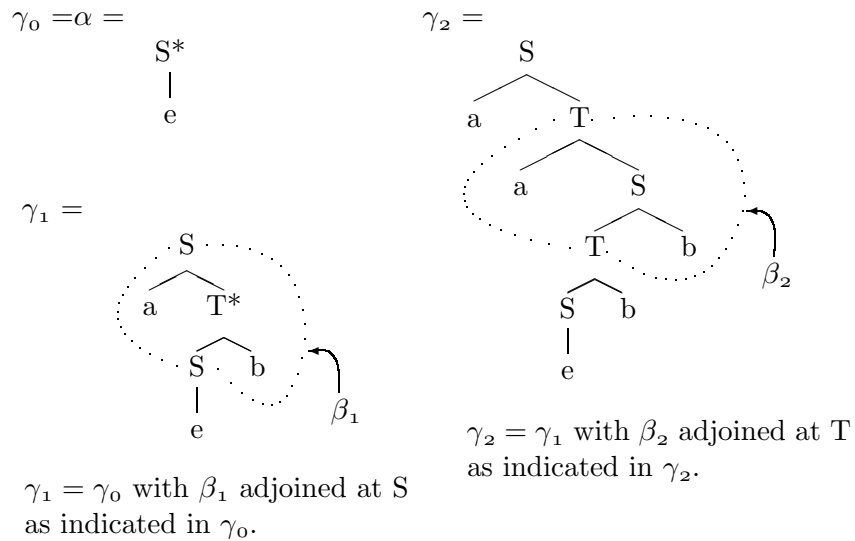
THEOREM 2.1.3: There exists a non-empty set of TAGs \mathcal{G}_2 such that for every $G \in \mathcal{G}_2$, $L(G)$ is strictly context sensitive; that is, there is no CFG grammar G' such that $L(G)=L(G')$.

Theorems 2.1.1 and 2.1.3 appear in Joshi, Levy, and Takahashi 1975. Theorem 2.1.2 is implicit in that paper, but we make it explicit here because of its importance in current linguistic discussions. Example 2.1 above illustrates theorem 2.1.1, and we illustrate the other theorems in the examples that follow.

EXAMPLE 2.2: Let $G = (I,A)$ where

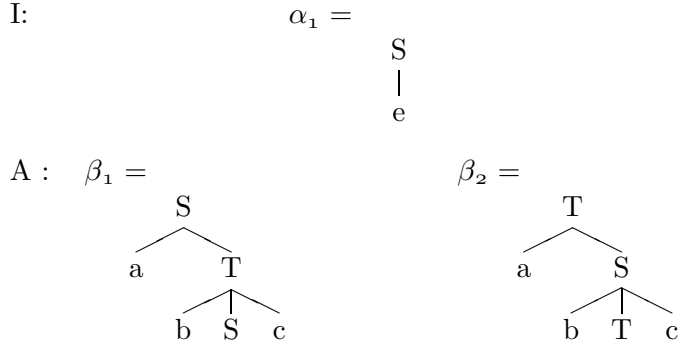


The language generated by G is context-free; but there is no CFG that is strongly equivalent to G . We can see this if we examine some derivations in G . Thus, consider the following trees:



Clearly, $L(G)$ is $\{a^n e b^n / n \geq 0\}$, which is a context-free language. Thus, there must exist a context-free grammar, G' , which is at least weakly equivalent to G . It can be shown, however, that there is no context-free grammar G' which is strongly equivalent to G ; i.e., for which $T(G) = T(G')$. This follows from the fact that $T(G)$ is NON-RECOGNIZABLE; i.e., there is no finite state bottom-up tree automaton that can recognize precisely $T(G')$. *Thus a TAG may generate a context-free language yet assign structural descriptions to the strings that cannot be assigned by any context-free grammar.*

EXAMPLE 2.3: Let $G = (I, A)$ where



The precise definition of $L(G)$ is as follows:

$$L(G) = L_1 = \{w e c^n / n \geq 0, w \text{ is a string of a's and b's such that}$$

1. the number of a's = the number of b's = n, and
2. for any initial substring of w, the number of a's \geq the number of b's.}

L_1 is a strictly context-sensitive language (i.e., a context-sensitive language that is not context-free). This can be shown as follows. Intersecting L_1 with the finite state language $a^* b^* e c^*$ results in the language

$$L_2 = \{a^n b^n e c^n / n \geq 0\} = L_1 \cap a^* b^* e c^*$$

L_2 is a well-known strictly context-sensitive language. The result of intersecting a context-free language with a finite state language is always a context-free language; hence, L_1 is not a context-free language. It is thus a strictly context-sensitive language. Example 2.3 thus illustrates Theorem 2.1.3.

We have seen that TAG's have more power than CFG's, but the extra power is quite limited. Joshi 1983 characterizes this limitation in detail, but the above example gives some indication of its nature. The language L_1 has an equal number of a's, b's and c's; however, the a's and b's are mixed in a certain way. The language

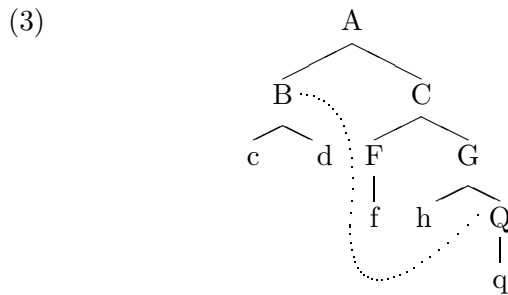
L_2 is similar to L_1 , except that all a's come before all b's. TAG's as defined so far are not powerful enough to generate L_2 . This can be seen as follows. Clearly, for any TAG for L_2 , each initial tree must contain an equal number of a's, b's and c's (including zero), and each auxiliary tree must also contain an equal number of a's, b's and c's. Further, in each case the a's must precede the b's. Then it is easy to see from the grammar of Example 2.3, that it will not be possible to avoid getting the a's and b's mixed. However, L_2 can be generated by a TAG with local constraints (see Section 2.3). The so-called copy language

$$L = \{ w e w / w \in \{a,b\}^* \}$$

also cannot be generated by a simple TAG but can be by a TAG with local constraints. Furthermore, it can be shown that TAG's, even with local constraints, cannot generate all context-sensitive languages (Joshi, 1983). Although TAG's are more powerful than CFG's, this extra power is highly constrained and apparently it is just the right kind for characterizing certain structural descriptions. TAG's share almost all the formal properties of CFG's (more precisely, the corresponding classes of languages). The string languages of TAG's can also be parsed in polynomial time, in particular in time Kn^6 , or less, where K is a constant that depends on the grammar and n is the length of the string (see Vijay Shankar and Joshi 1985 for further details).

2.2 TAG's with 'links'.

The elementary trees (initial and auxiliary trees) are the appropriate domains for characterizing certain dependencies (e.g., subcategorization dependencies and filler-gap dependencies). The characterization of certain of these dependencies can be achieved by introducing a special relationship between certain specified pairs of nodes of an elementary tree. This relationship, which we shall call 'linking', is pictorially exhibited by an arc (a dotted line) from one node to the other. For example, in the tree in (3) below, the nodes labeled B and Q are linked.



We will require the following conditions to hold for a link in an elementary tree:

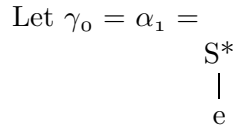
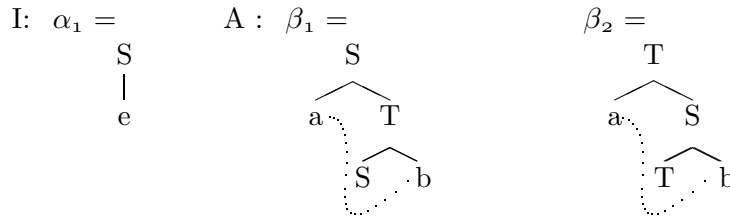
If a node n_1 is linked to a node n_2 then

1. \mathbf{n}_2 c-commands \mathbf{n}_1 (i.e., \mathbf{n}_2 precedes \mathbf{n}_1 and there exists a node \mathbf{m} which immediately dominates \mathbf{n}_2 and also dominates \mathbf{n}_1).
2. \mathbf{n}_1 dominates a null string (or a terminal symbol in the nonlinguistic formal grammar examples).

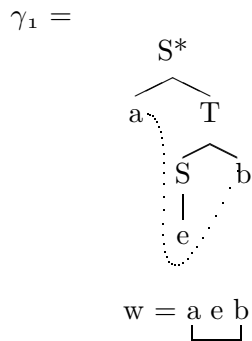
Linking is an asymmetric relation; and in the linguistic context both \mathbf{n}_1 and \mathbf{n}_2 will be of the same category, with \mathbf{n}_1 dominating the null string. The notion of linking thus defined is related to the one discussed in Peters and Ritchie 1982. A TAG with links is a TAG in which some of the elementary trees may have links as defined above. Henceforth, we may refer to a TAG with links just as a TAG.

Links are defined on the elementary trees. However, the important point is that the composition operation of adjoining will preserve the links. Links defined on the elementary trees may become stretched as the derivation proceeds. Example 2.4 will illustrate this point.

EXAMPLE 2.4: Let $G = (I, A)$ where

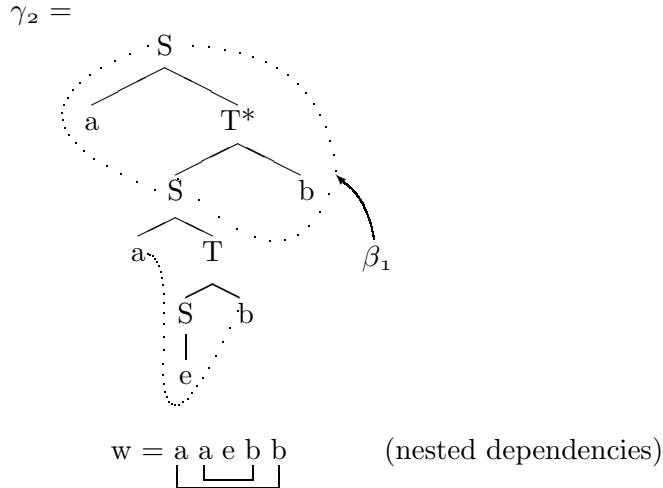


Adjoining β_1 at S as indicated in γ_0 , we have

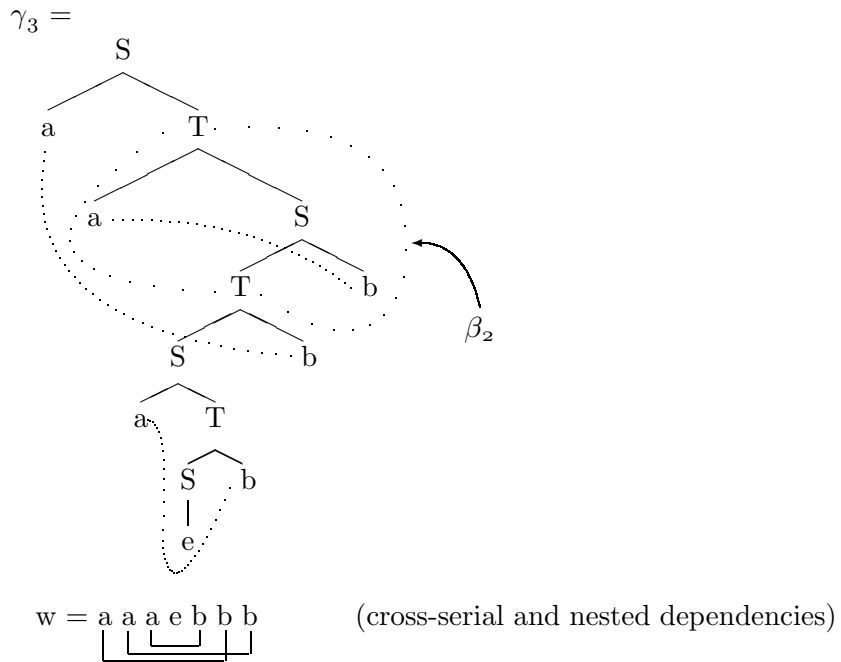


The terminal string corresponding to γ_1 is a e b, where the dependency is indicated by the solid line.

Adjoining β_2 again at S as indicated in γ_1 , we have



Adjoining β_2 at T as indicated in γ_2 , we have



In this example β_1 and β_2 each have one link, and the composed trees γ_2 and γ_3 show how linking is preserved under adjunction. In γ_3 , for example, one of the links is stretched. It should be clear now how, in general, the links will be preserved during the derivation and we shall not give a formal treatment of this property here.

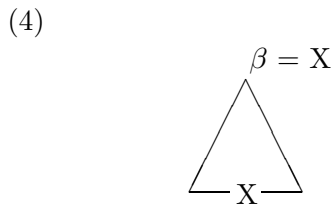
We should also note with regard to the above example that the dependencies in γ_2 between the a's and b's, as reflected in the terminal string, are properly nested, while in γ_3 two of them are properly nested, and the third one is cross-serial. The cross-serial one is crossed with respect to the nested ones (not, of course, a unique description). The two elementary trees β_1 and β_2 have one link each so that the nestings and crossings in γ_2 and γ_3 are the result of adjoining. There are two points of importance here:

1. TAG's with links can characterize certain cross-serial dependencies (as well as, of course, nested dependencies, which is not a surprise).
2. The cross-serial dependencies (as well as the nested dependencies) arise as a result of adjoining. But this is not the only way they can arise. It is possible to have two links in an elementary tree which represent cross-serial or nested dependencies, which will then be preserved during the derivation. The cross-serial dependencies, as well as nested dependencies, will arise in two distinct ways – either by adjoining or by being present in some elementary trees to start with.

It is clear from our example that the string language of TAG with links is not affected by the links; that is, links do not affect weak generative capacity. However, they make certain aspects of the structural description explicit which are implicit in a TAG without links.

2.3 TAG's with local constraints on adjoining

The adjoining operation as defined in Section 2.1 is 'context-free'. An auxiliary tree with the form in (4)



is adjoinable to a tree t at a node n if the label of node n is X independently of the (tree) context around n . In this sense, adjoining is context-free. In Joshi (1983), local constraints on adjoining similar to those investigated by Joshi and Levy (1978) were used. These are a generalization of the context-sensitive constraints studied by Peters and Ritchie (1969). It was soon recognized, however, that the full power of these constraints was never exploited, either in the linguistic context or in the formal languages cases. The so-called proper analysis contexts and domination contexts (as defined in Joshi and Levy (1978)) that were actually used in Joshi (1983) always

turned out to be such that the context elements were within a single elementary tree; that is, they were far more localized than the definitions required. Based on this observation and a suggestion in Joshi, Levy and Takahashi (1975), we will describe a new way of introducing local constraints. This approach not only captures the insight stated above, but it is more in the spirit of the TAG formalism, with its emphasis on locality. (For further details, see (Vijay Shankar and Joshi 1985)) The earlier approach was less so, although it was certainly adequate for the investigation in Joshi (1983). A precise characterization of the original approach remains an open problem.

Let $G = (I, A)$ be a TAG with local constraints if for each elementary tree $\mathbf{t} \in I \cup A$, and for each node, \mathbf{n} , in \mathbf{t} , we specify the set β of auxiliary trees that can be adjoined at the node \mathbf{n} . Note that if there is no constraint then any auxiliary tree whose root has the same label as the label of the node \mathbf{n} is adjoinable at \mathbf{n} . Thus, in general, β is a subset of the set of all auxiliary trees structurally adjoinable at \mathbf{n} .

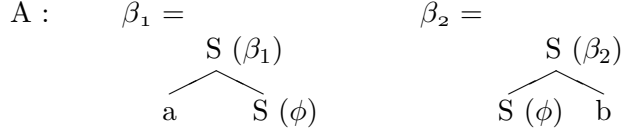
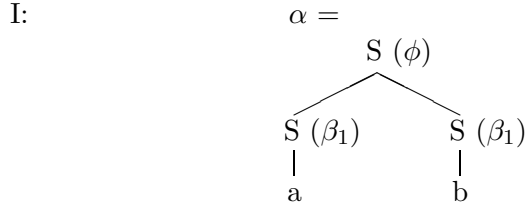
We adopt the following conventions for the statement of local constraints:

1. Since, by definition, no auxiliary trees are adjoinable to a node labeled by a terminal symbol, no constraint has to be stated for node labeled by a terminal.
2. If there is no constraint, i.e., all auxiliary trees with the appropriate root label are adjoinable at a node \mathbf{n} , then we will not state this explicitly, as this is the case we have discussed in Section 2.1.
3. If no auxiliary trees are adjoinable at a node \mathbf{n} , then we will write the constraint as (ϕ) , where ϕ denotes the null set.
4. We will also allow for the possibility that for a given node at least one adjoining is obligatory, of course, from the set of all possible auxiliary trees adjoinable at that node.

Hence, a TAG WITH LOCAL CONSTRAINTS is defined as follows, $G = (I, A)$ is a TAG with local constraints if for each node \mathbf{n} , in each tree \mathbf{t} , one (and only one) of the following constraints is specified:

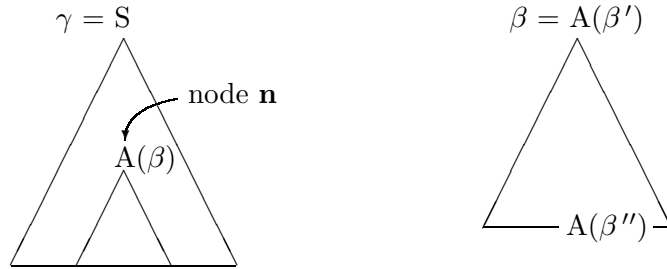
1. **SELECTIVE ADJOINING (SA)**: Only a specified subset of the set of all auxiliary trees are adjoinable at \mathbf{n} . SA is written as (β) , where β is a subset of the set of all auxiliary trees structurally adjoinable at \mathbf{n} .
If β equals the set of all auxiliary trees adjoinable at \mathbf{n} , then we do not explicitly state this at the node \mathbf{n} .
2. **NULL ADJOINING (NA)**: No auxiliary tree is adjoinable at the node \mathbf{n} . NA will be written as (ϕ) .
3. **OBLIGATORY ADJOINING (OA)**: At least one (out of all the auxiliary trees adjoinable at \mathbf{n}) must be adjoined at \mathbf{n} . OA is written as $O(\beta)$ where β is a subset of the set of all auxiliary trees adjoinable at \mathbf{n} .

EXAMPLE 2.5: Let $G = (I,A)$ be a TAG with local constraints where

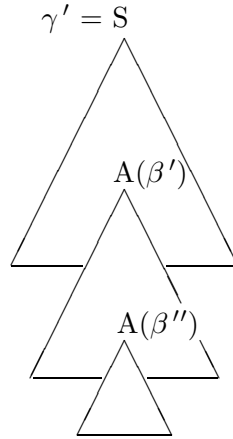


In α_1 no auxiliary trees can be adjoined to the root node. Only β_1 is adjoinable to the left S node at depth 1 and only β_2 is adjoinable to the right S node at depth 1. In β_1 only β_1 is adjoinable at the root node and no auxiliary trees are adjoinable at the foot node. Similarly for β_2 .

We must now modify our definition of adjoining to take care of the local constraints. Given a tree γ with a node \mathbf{n} labeled A and given an auxiliary tree α with the root node labeled A , we shall modify our definition of adjoining as follows: α is adjoinable to γ at node \mathbf{n} if $\alpha \in \beta$, where β is the constraint associated with node \mathbf{n} in γ . The result of adjoining α to γ will be as defined in Section 2.1, except that the constraint β associated with \mathbf{n} will be replaced by β' , the constraint associated with the root node of α and by β^\diamond , the constraint associated with the foot node of α . Thus, given

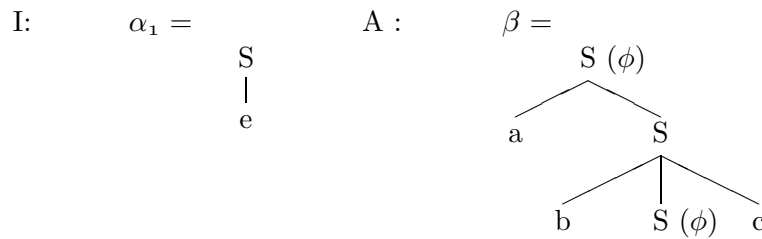


the resultant tree γ' is



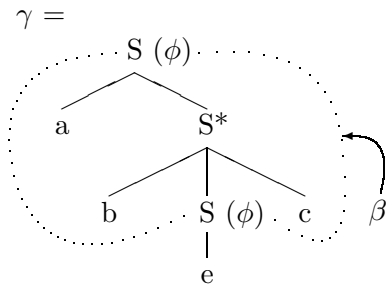
We also adopt the convention that any derived tree with a node which has an OA constraint associated with it will not be included in the tree set associated with a TAG, G . The string language L of G is then defined as the set of all terminal strings of all trees derived in G (starting with initial trees) which have no OA constraints left in them.

EXAMPLE 2.6: Let $G = (I, A)$ be a TAG with local constraints where

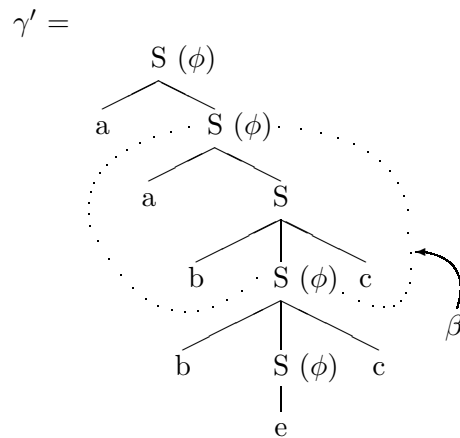


There are no constraints in α_1 . In β no auxiliary trees are adjoinable at the root node or the foot node and for the center S node there are no constraints.

Starting with α_1 and adjoining β to α_1 at the root node we obtain



Adjoining β to the center S node (the only node at which adjunction can occur) we have



It is easy to see that G generates the string language

$$L = \{ a^n b^n e c^n / n \geq 0 \}$$

EXAMPLE 2.7: Let G' be a TAG similar to G in Example 2.6, except that in G' there are no constraints in β . G' generates

$$L = \{ w e c^n / n \geq 0, \# \text{ a's in } w = \# \text{ b's in } w = n, \\ \text{and for any proper initial string } u \text{ of } w, \\ \# \text{ a's in } u \geq \# \text{ b's in } u. \}$$

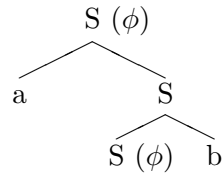
This is the same language as in Example 2.3. This language is closely related to the context-sensitive language discussed in Higginbotham 1984, which can also be shown to be a TAG language.

EXAMPLE 2.8: Let $G = (I,A)$ be a TAG with local constraints where

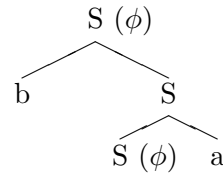
I: $\alpha_1 =$



A : $\beta_1 =$



$\beta_2 =$



G generates the language

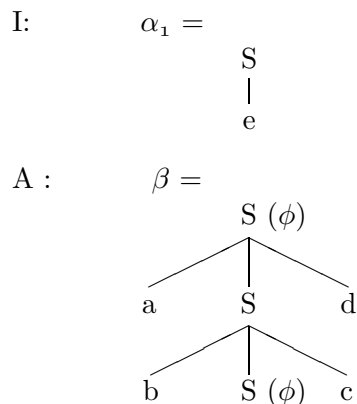
$$L = \{ w e w / w \in \{a,b\}^* \}$$

EXAMPLE 2.9: Let G' be a TAG which is the same as G in Example 2.8 but without any local constraints. The corresponding language is

$$L = \{ w e w' / w, w' \in \{a,b\}^*, w = w' = 2n, \\ \# \text{ a's in } w = \# \text{ a's in } w' = \\ \# \text{ b's in } w = \# \text{ b's in } w' = n. \}$$

This language is related to the Swiss-German example in Shieber (1984).

EXAMPLE 2.10: Let $G = (I,A)$ be a TAG with local constraints where



G generates

$$L = \{ a^n b^n e c^n d^n / n \geq 0 \}$$

Note that it can be shown that languages

$$L^1 = \{ a^n b^n e c^n d^n e^n / n \geq 1 \}$$

and

$$L^2 = \{ w w w / w \in \{a,b\}^* \}$$

cannot be generated by TAG's either with or without local constraints (Joshi 1983). Other languages, such as $L' = \{ a^{n^2} / n \geq 1 \}$, also cannot be generated by TAG's. This is because the strings of a TAG grow linearly (for a detailed definition of this property, called the 'constant growth' property, see (Joshi, 1983). L' does not satisfy this property.

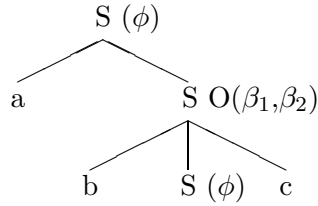
For those familiar with Joshi (1983), it is worth pointing out that the SA constraint is only abbreviating i.e., it does not affect the power of TAG's. The NA and OA constraints, however, do affect the generative power of TAG's. Thus, NA is needed to generate the languages in Examples 2.6, 2.7, and 2.8. OA is needed to generate the language in Example 2.11 below:

EXAMPLE 2.11: Let $G=(I,A)$ be a TAG with local constraints

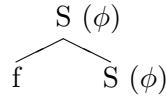
I: $\alpha_1 =$

S
|
e

A: $\beta_1 =$



$\beta_2 =$



G generates

$$L = \{ a^n f b^n e c^n / n \geq 1 \}$$

Repeated adjoining of β_1 generates

$$\{ a^n b^n e c^n / n \geq 1 \}$$

The resulting trees each have one node still with an OA constraint, which can be removed by adjoining β_2 , generating L.

In contrast to Joshi 1883, where we stated for each auxiliary tree the constraints on its adjoinability, we have here stated, for each node in each elementary tree, the constraints on what auxiliary trees can be adjoined there. This way of looking at local constraints has not only greatly simplified their statement, but it has also allowed us to capture the insight that the ‘locality’ of the constraint is storable in terms of the elementary trees themselves!

2.4 Some formal properties of TAG’s

Here we state without proof the most important formal properties of TAG’s. Almost all of the computationally significant properties of context-free grammars are carried over to TAG’s. For instance, it can be shown that Tree Adjoining Languages (TAL) are closed under union, concatenation, substitution, Kleene-star, and intersection with regular sets. That is, given two TAL’s L and L’ a regular set (finite state language) R, and an empty language λ :

1. $L \cup L'$ is a TAL, where \cup denotes set union,
2. $L \cdot L'$ is a TAL, where \cdot denotes concatenation,
3. $L^* = \lambda \cup L \cup LL \cup LLL \cup LLL \cup \dots$ is a TAL, where $*$ denotes the Kleene star operator,

4. the language obtained by substituting TAL's for each terminal symbol of L is a TAL,
5. $L \cap R$, the language obtained by intersecting L with a regular set is a TAL.

A pumping lemma similar to that for context-free languages can be established for TAL's also. This lemma allows one to establish that certain languages are not TAL's, for example L^1 and L^2 in Example 2.10.

The tabular parsing algorithm for context-free grammars (the so-called CKY algorithm) can be extended in a natural fashion for parsing TAL's, although the extension is not immediate because adjoining is not a substitution operation. The time bound of the parsing algorithm is proportional to n^6 (where n is length of the string to be parsed) as compared to the n^3 bound that has been established for context-free grammars.

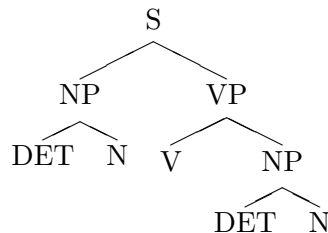
Recently, Pollard (1984) has introduced a class of grammars called Head Grammars (HG). HG's are remarkably similar to TAG's. All formal properties proved so far for TAG's also hold for HG's (except for substitution, which has been not been established yet). It has been shown that TAG's are contained in HG's. The question of whether HG's are entirely equivalent formally to TAG's remains open. For further details of the formal properties of TAG's see Joshi 1983 and Vijay Shankar and Joshi 1985; for HG's see Roach 1984.

3 Some linguistic examples

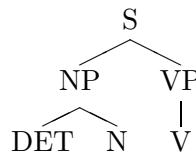
Our purpose in this section is to give some simple linguistic examples that illustrate the applicability of the TAG formalism to the description of natural language phenomena. Many details which do not serve the purpose of illustration have been ignored or simplified, and we do not intend here to offer detailed justification of the analyses presented. Rather we hope that these examples will familiarize the reader with TAG derivations of natural language sentences, will answer some obvious questions that may arise in the reader's mind, and will help the reader to follow the discussion in subsequent sections.

Let $G = (I,A)$ be a TAG where I is the set of initial trees and A is the set of auxiliary trees. We will list only some of the trees in I and A, those relevant to the derivation of our illustrative sentences. Rather than introduce all these trees at once, we shall introduce them as necessary.

I (initial trees): $\alpha_1 =$



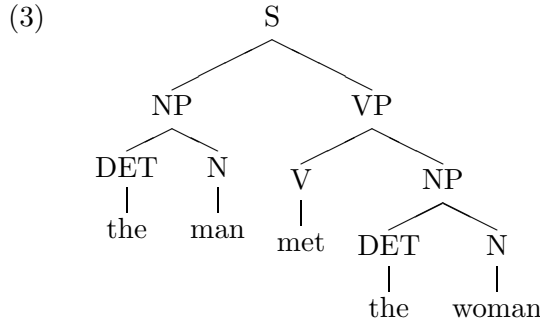
$\alpha_2 =$



Tree α_1 corresponds to a minimal sentence with a transitive verb, as in (1); and α_2 corresponds to a minimal sentence with an intransitive verb, as in (2):

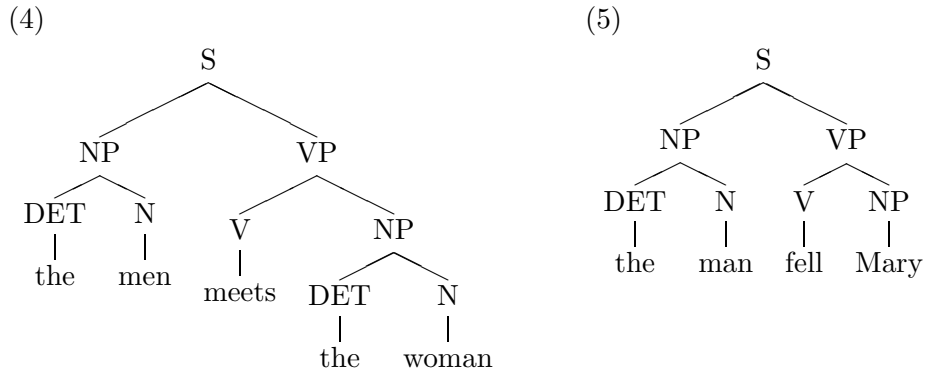
- (1) The man met the woman.
- (2) The man fell.

Initial trees as we have defined them require terminal symbols on the frontier. In the linguistic context, the nodes on the frontier will be preterminal lexical category symbols such as N, V, A, P, DET, etc. The lexical items are inserted for each of the preterminal symbols as each elementary tree enters the derivation. Thus, we generate the sentence in (1) by performing lexical insertion on α_1 , yielding (3):



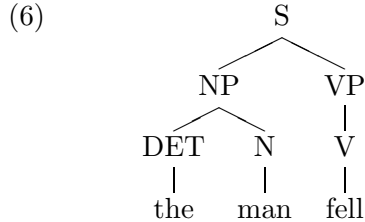
As we continue the derivation by selecting auxiliary trees and adjoining them appropriately, we follow the same convention, i.e., as each elementary tree is chosen, we make the lexical insertions. Thus in a derivation in a TAG, lexical insertion goes hand in hand with the derivation. Each step in the derivation selects an elementary tree together with a set of appropriate lexical items.

Note that as we select the lexical items for each elementary tree we can check a variety of constraints, e.g., agreement and subcategorization constraints on the set of lexical items. Thus, for example, the following choices of lexical items will not be permitted:



In (4) number agreement has been violated and in (5) the intransitive verb fell has been inserted into a transitive verb phrase. These constraints can be checked easily

because the entire elementary tree that is the domain of the constraints is available as a single unit at each step in the derivation. If we had started with the initial tree α_2 then the choice of the intransitive verb would be permitted, yielding the well-formed tree in (6):

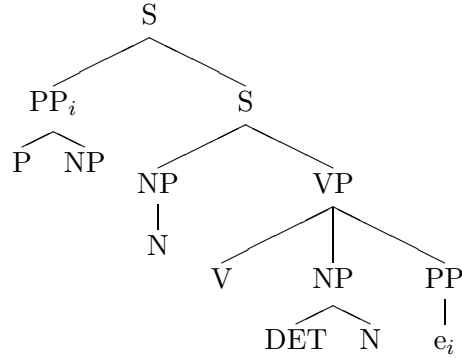


When an auxiliary tree enters the derivation, similar considerations hold. In addition, further constraints, both contextual and lexical, can be checked by means of local constraints, which will be illustrated later.

As the reader will have noted, we require different initial trees for the sentences “John fell” and “the man fell” because the expansion of NP is different in the two cases. Since the structure of these two sentences is otherwise identical, we cannot be content with a linguistic theory that treats the two sentences as unrelated. In a fully articulated theory of grammar employing the TAG formalism, the relationships among initial trees would be expressed in an independent module of the grammar that specified the constraints on possible elementary (initial or auxiliary) trees. These constraints might be expressed in a number of alternative ways; for example as a set of rules for the projection of syntactic structure from lexical heads, perhaps incorporating features of categorial grammar. We might even provide schemata or rules for obtaining some elementary structures from others. Some of these rules will have the effect of conventional transformational rules. For example, we might obtain the trees for constituent questions by movement of a *wh*- phrase into COMP. But in a TAG these rules need not be formulated as classical transformations were nor need they be subject to the sorts of constraints that prevent a rule like “MOVE α ” from overgenerating. Since both the domains and codomains of these rules would be finite, we can formulate them directly as tree rewriting rules. In any case, the rules will be abbreviatory, in the sense that they will generate only a finite set of trees and so will not affect the formal power of the TAG. The most important point regarding the source of elementary trees, as we emphasized in our introduction, is that using the TAG formalism allows us to treat as orthogonal the principles governing the construction of minimal syntactic units and those governing the composition of these units into complex structures.

Linked trees can be used to represent so-called ‘unbounded’ dependencies like topicalization and *wh*- constructions. In (7) we give a possible topicalized structure, and in (8)/(9) we give two *wh*- questions:

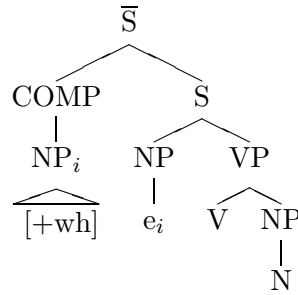
(7) $\alpha_3 =$



to Mary John gave a book

(8)

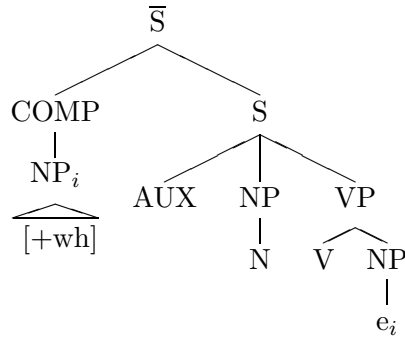
$\alpha_4 =$



who met Mary

(9)

$\alpha_5 =$



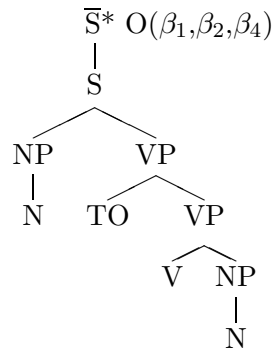
who did John meet

(Note that for convenience we will, in the rest of this paper, replace the dotted line links with coindexing of the linked nodes. Since in the linguistic case the link mother node is always an empty category, the coindexing notation is adequate to specify the linking relation.) In (7) we have shown a link from the lower PP node to a higher PP node. When the lexical items below the tree are inserted for the preterminal nodes in α_3 , we can check not only that a verb requiring a PP object has been inserted, but also that the preposition P is *to* as required by the verb *give*.

Thus far all of the initial trees that we have defined correspond to minimal root sentences. We now introduce, in (10) below, some initial trees which are minimal but do not give root sentences. The motivation for introducing these trees will be clear from the examples and the subsequent use of these trees in derivations. Since these trees are not possible root sentences, it is necessary that they undergo at least one adjunction (of a specific type) and the resulting tree becomes a possible independent sentence. This requirement can be very easily stated as a local constraint. The effects of the local constraints associated with these trees will be illustrated later in this section.

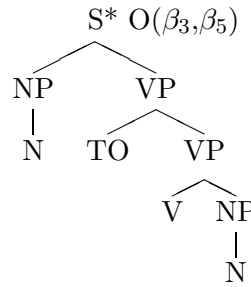
(10)

$\alpha_6 =$



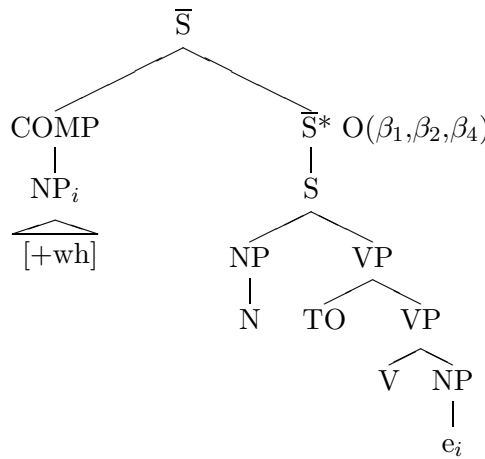
PRO to invite Mary

$\alpha_7 =$



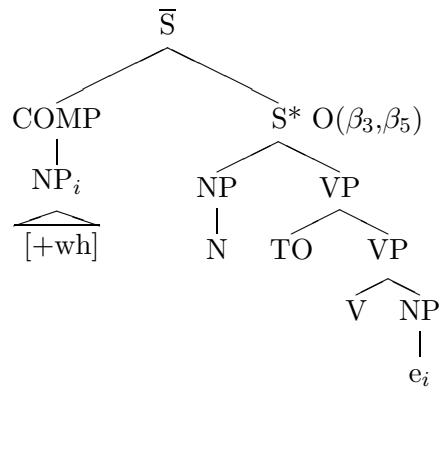
John to like Mary

$\alpha_8 =$



who PRO to invite

$\alpha_9 =$



who Bill to invite

Tree α_6 and α_7 are similar, except that in the first case the subject NP will be un-governed when adjoined to a matrix predicate. Tree α_6 will be used in the derivation of sentences like (11) and (12):

(11) John persuaded Bill to invite Mary.

(12) John tried to invite Mary.

Tree α_7 will be used in deriving sentences like (13):

(13) John expected Bill to like Mary.

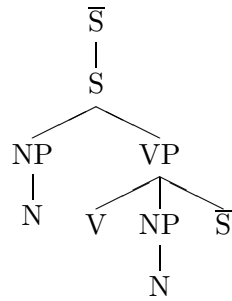
Trees α_8 and α_9 differ in the same way as α_6 and α_7 ; α_8 will be used in deriving sentences like (14) while α_9 will be used in deriving sentences like (15):

(14) Who did John try to invite?

(15) Who did John expect Bill to invite?

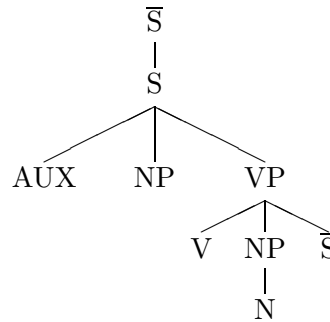
Now we introduce auxiliary trees that will adjoin to the above infinitival initial trees to produce complete independent sentences:

(16) $\beta_1 =$



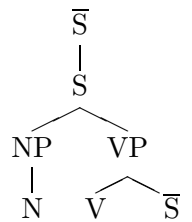
John persuaded Bill \bar{S}

$\beta_4 =$



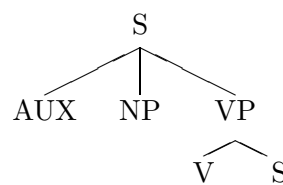
Did John persuade Bill \bar{S}

$\beta_2 =$



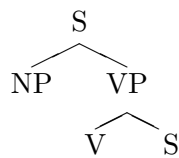
John tried \bar{S}

$\beta_5 =$



Did John expect S

$\beta_3 =$

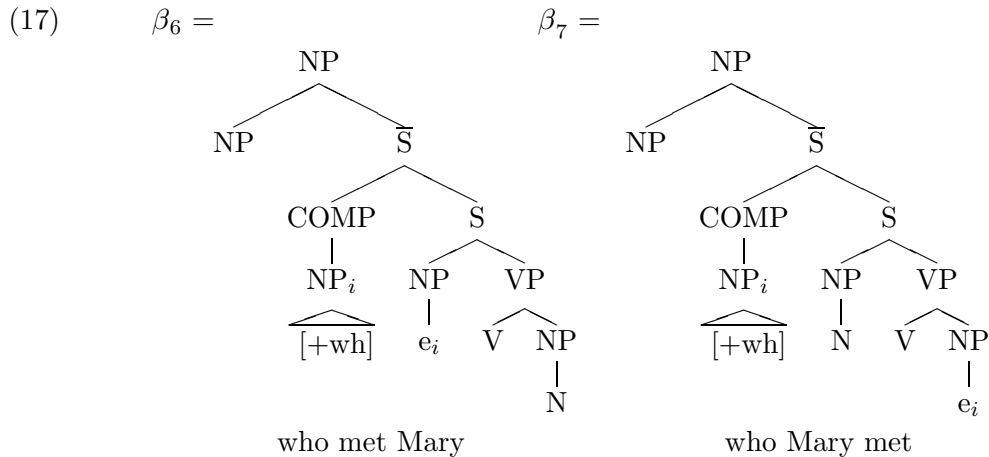


John expected S

As the reader can check for him or herself, the sentences of (11) – (15) will be derived if the appropriate auxiliary trees listed as satisfying the obligatory adjoining constraint are adjoined at the starred nodes of the initial trees in (10). We shall go through some of these derivations in detail in Examples 3.2, 3.3, and 3.4 below. The one difficulty implicit in these examples is that of ensuring that PRO will appear only in ungoverned positions and that only PRO will appear in such positions. This can be accomplished in more than one way. Thus, we might have treated PRO as a terminal constant, thereby explicitly distinguishing those elementary trees in which it occurs from those in which it does not. We prefer, however, to adopt a version of GB case theory (Chomsky 1981a) to produce the same result. Let us assume that lexical insertion is obligatory and that it occurs in a derivation whenever the

conditions governing it are met. Normally, this will guarantee that lexical insertion occurs in each elementary tree as it enters a derivation. Let us further stipulate that NPs can be lexically filled out only when governed by a case assigner. This requirement will prevent the subjects of infinitives in trees like those of (10) from being inserted until and unless an adjunction produces a governor for the subject position. If and only if the position remains ungoverned, PRO is inserted in place of lexical material.⁴

Now let us introduce some auxiliary trees that will allow us to generate sentences with relative clauses:



Tree β_6 can be used to build sentences with subject relatives, as in (18); and β_7 can be used to build sentences with object relatives, as in (19):

(18) The boy who met Mary left.

(19) The boy who Mary met left.

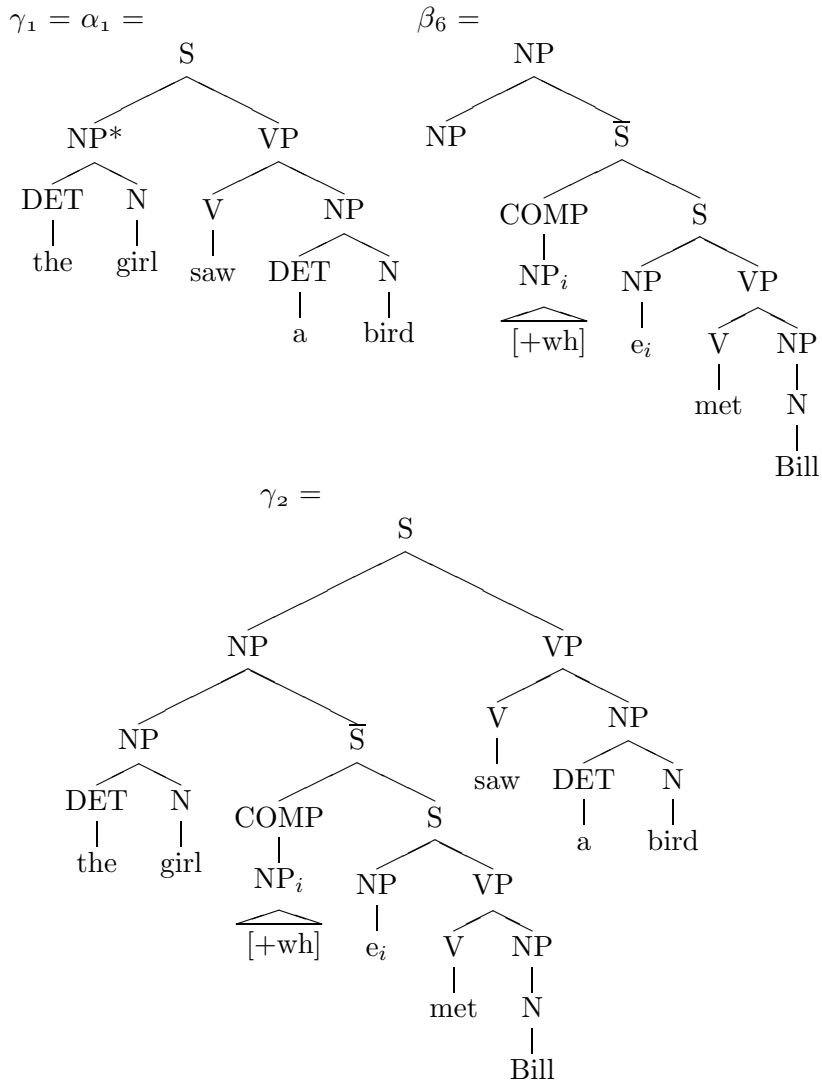
As in the case of initial trees, we see here explicitly that some auxiliary trees will have to be related to simpler structures by a theory that constrains the form of elementary trees. In the case of relative clauses this theory will have to specify, first, that an \bar{S} functioning as a relative clause has essentially the same structure as any other \bar{S} and second, that the relative \bar{S} be of the appropriate syntactic type to allow it to function in the predication relationship between head and clause that constitutes relative clause semantics. We shall leave the formulation of these constraints to another occasion.

Relatives introduced by *that* instead of by *wh-* can be handled in our TAG in either of two ways. We might postulate an empty *wh-* type operator in the COMP of the relative clause, giving an analysis similar to that we have for overt *wh-* relatives. Alternatively, we might have no *wh-* element in our structure and instead establish a link between the head NP of the relative clause and the gap position. This latter analysis is statable, and easily so, in a TAG because the head of a relative clause

appears in the same elementary tree as the clause itself. Although we shall not pursue this non-*wh*- analysis of *that* relatives in this paper, it has linguistic advantages that make it worth further consideration.

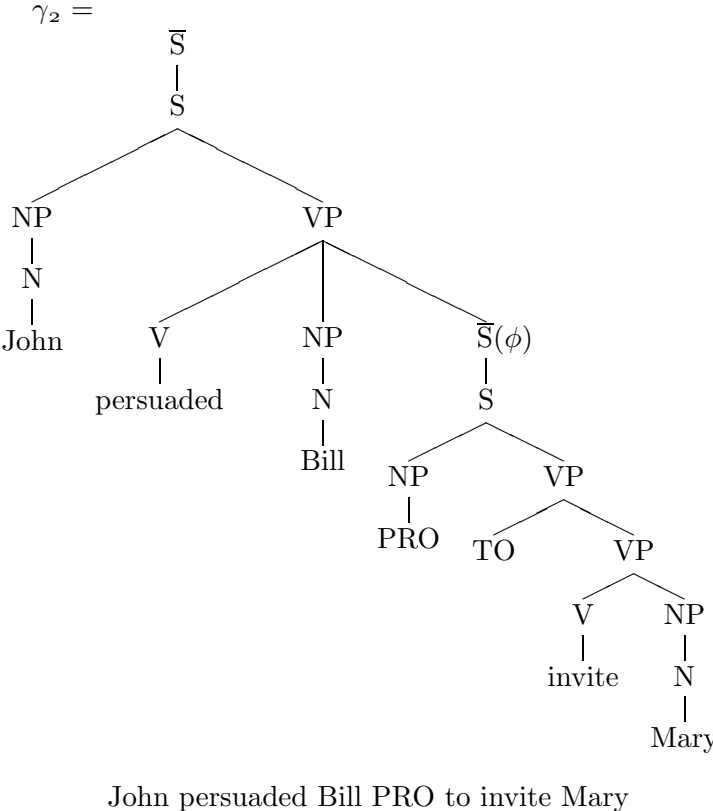
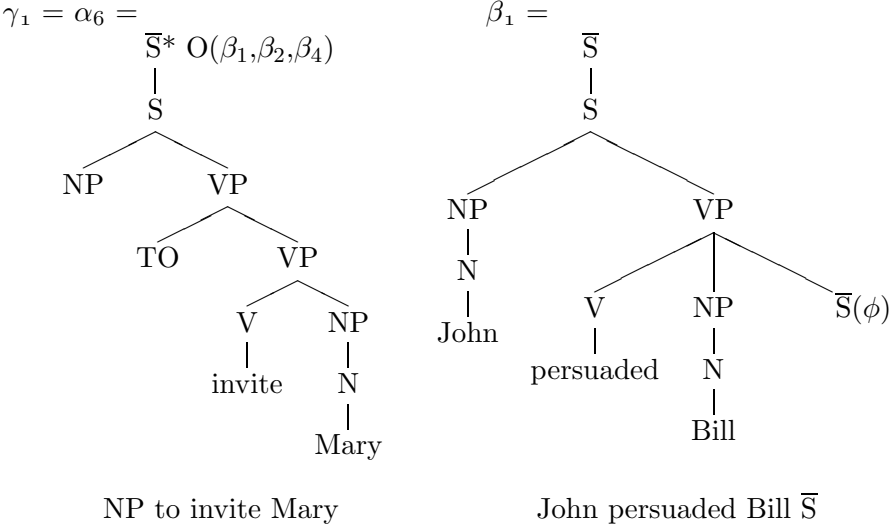
The examples below should serve to illustrate some of the derivations possible in a TAG containing the initial and auxiliary trees introduced in this section. As we go through the derivations we shall point out how local constraints function to prevent overgeneration.

EXAMPLE 3.1: Starting with the initial tree $\gamma_1 = \alpha_1$ and then adjoining β_6 (with appropriate lexical insertions) at the indicated node in α_1 , we obtain γ_2 .



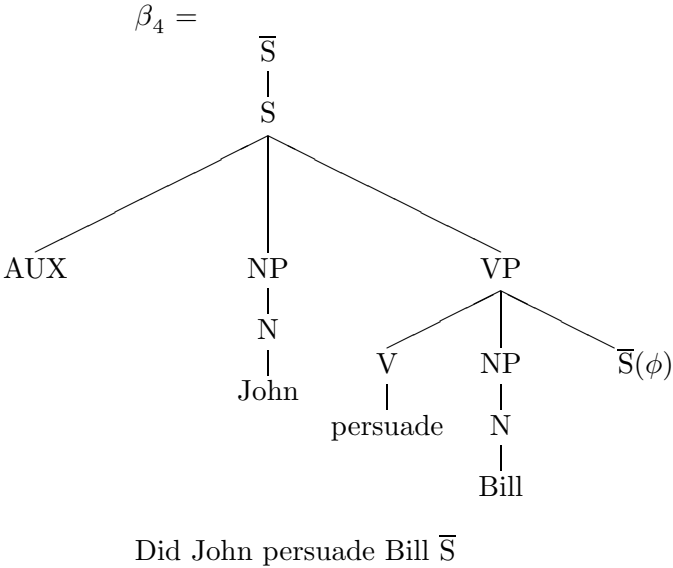
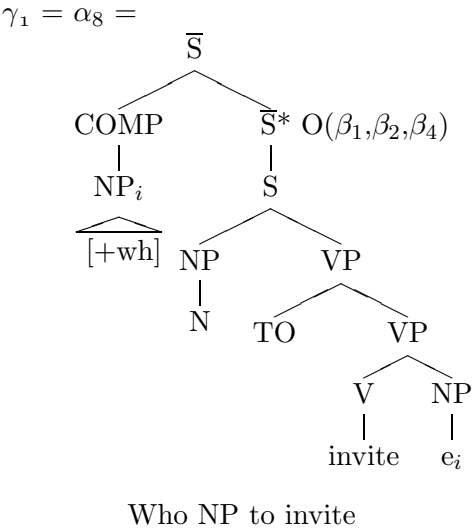
The girl who met Bill saw a bird

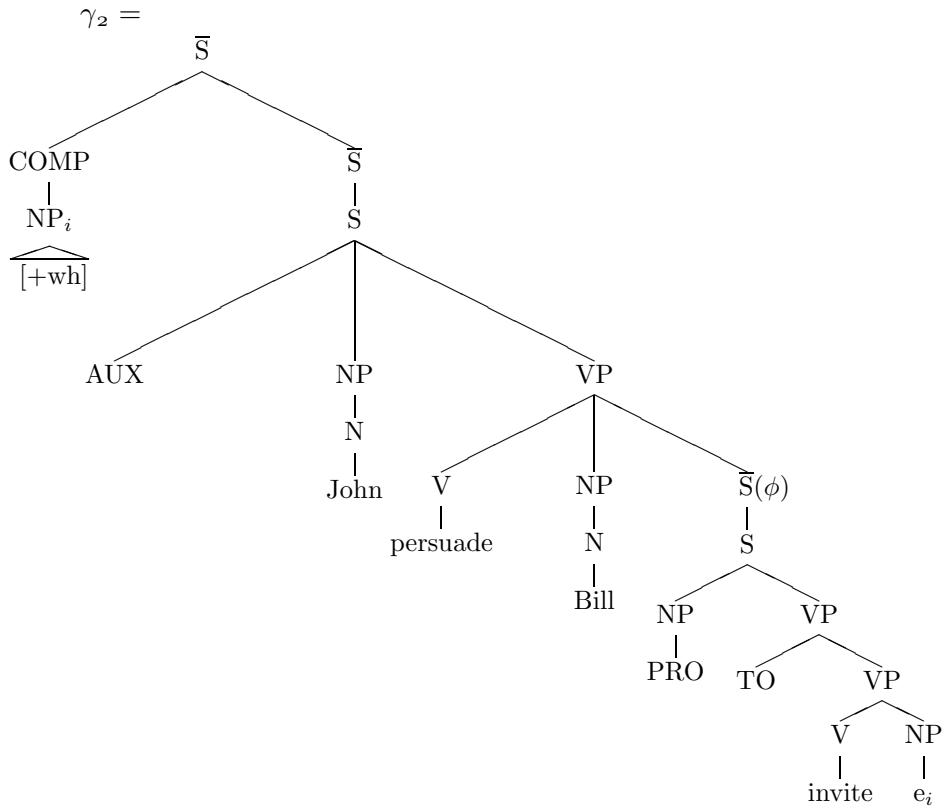
EXAMPLE 3.2: Starting with the initial tree $\gamma_1 = \alpha_6$ and adjoining β_1 at the indicated node in α_6 we obtain γ_2 . Lexical insertion proceeds as indicated with each tree as fully specified as possible at each stage of the derivation.



Since the initial tree α_6 is not a root sentence, it must undergo an adjunction at its root node, for example, by the auxiliary tree β_1 as shown above. Thus, for α_6 we have specified a local constraint $O(\beta_1, \beta_2, \beta_4)$ for the root node, indicating that α_6 must undergo an adjunction at the root node by an auxiliary tree β_1 . In a fuller grammar there might, of course, be many alternatives in the scope of $O()$. The local constraint (ϕ) at the foot node of β_1 prevents further adjoining at this node in γ_2 . Note that PRO first appears in γ_2 , as required by our formulation of case theory.

EXAMPLE 3.3: Starting with the initial tree $\gamma_1 = \alpha_8$ and adjoining β_4 to α_8 at the indicated node in α_8 , we obtain γ_2 .





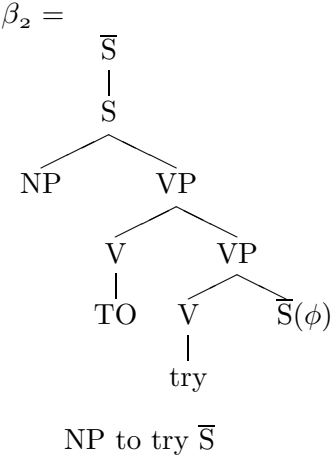
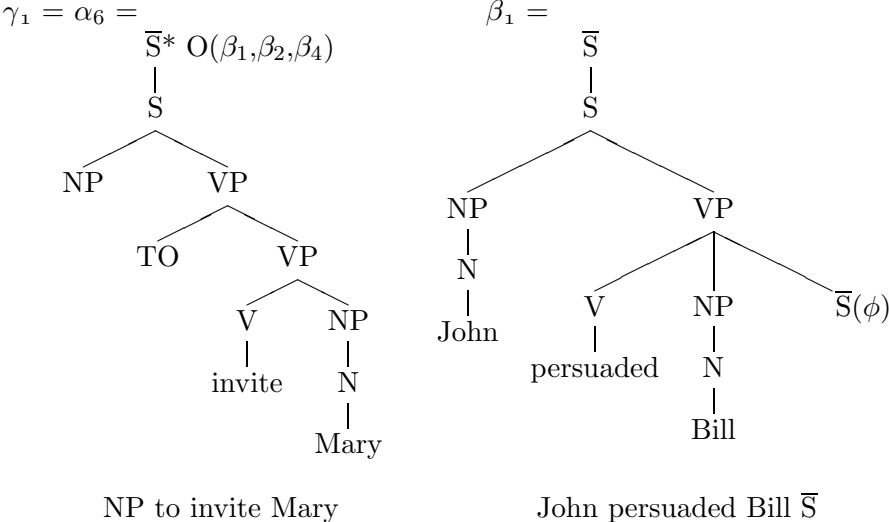
Who did John persuade Bill PRO to invite

Note that the link in γ_1 is preserved in γ_2 ; it is stretched, resulting in a so-called unbounded dependency. Also note that, as in previous examples, α_8 is an initial tree that cannot serve as a root sentence and the obligatory adjunction possibilities are as indicated. Again the local constraint (ϕ) at the foot node of β_4 prevents further adjoining at this node in γ_2 .

EXAMPLE 3.4: Suppose that we want to derive the sentence in (20):

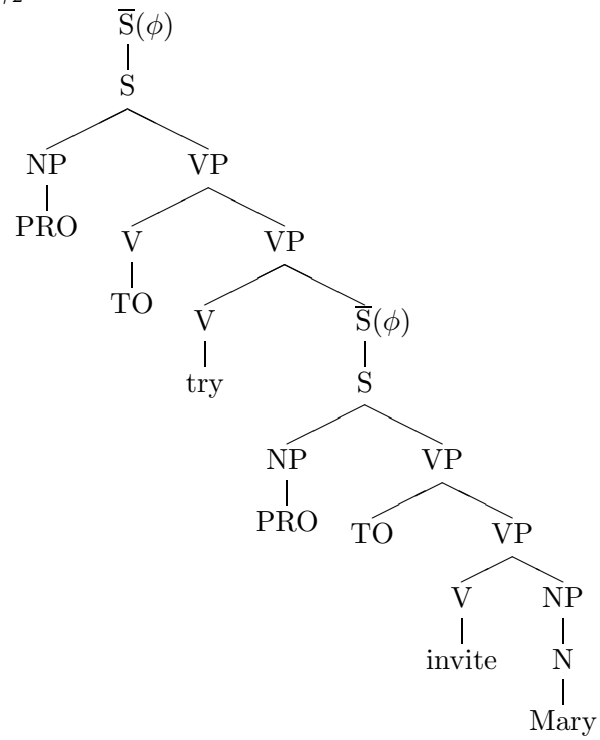
(20) John persuaded Bill PRO to try PRO to invite Mary.

The initial tree in the derivation will again be α_6 and the auxiliary trees will be β_1 and β_2 , as given below:



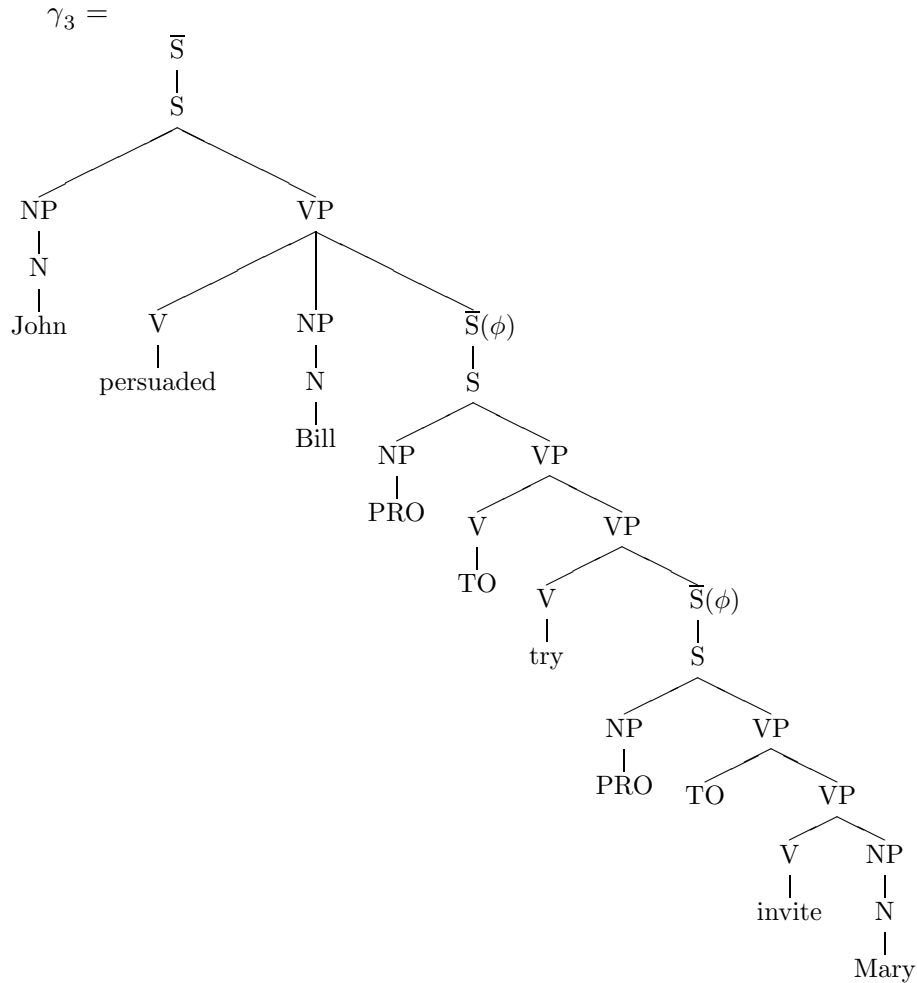
If we adjoin β_2 to α_6 , we get γ_2 :

$\gamma_2 =$



NP to try PRO to invite Mary

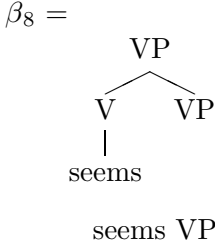
If we then adjoin β_1 to γ_2 , we get γ_3 , the tree we want:



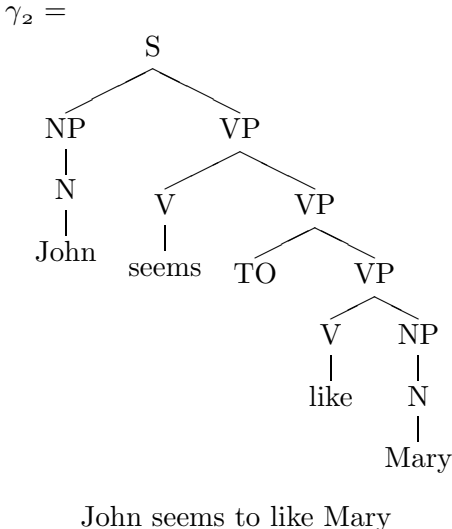
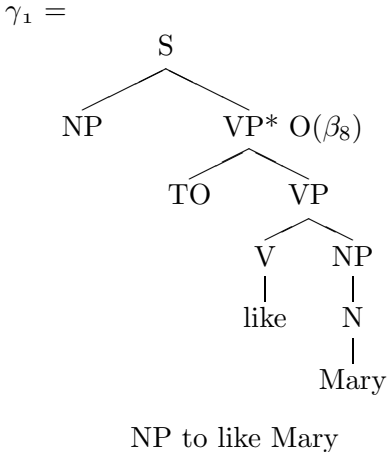
John persuaded Bill PRO to try PRO to invite Mary

Note that it would be possible in principle to adjoin β_1 to α_1 and then to adjoin β_2 to the result. This derivation is, however, blocked by the null adjoining local constraint we have placed on the foot node of β_1 (and other auxiliary trees). We do not want this derivation to succeed because it would prevent us from integrating our TAG derivations with a step by step compositional semantics. While we shall have little to say about the character of a semantics for a TAG, it is possible to define in an interesting way a formal semantics for the “derivation structures” of a TAG. The details of such a semantics will be worked out in future papers.

EXAMPLE 3.5 Let us now add the tree in β_8 to our inventory of auxiliary trees:



If we create a modified version of α_6 with an obligatory adjunction site as indicated below, we can adjoin β_8 to it, producing the raising structure in γ_2 :



4 Raising and equi constructions in a TAG

4.1 The basic issues

One of the attractive features of the TAG formalism as a notation for writing grammars is that it allows us to capture the distinction between raising and control predicates in the syntactic component of the grammar. It was a substantial achievement of the standard theory transformational grammar of the 1960's to be able to distinguish between these two kinds of predicates while generating identical surface structures for them. Thus, the sentences of (1) and (2) look alike, but they are distinguished by several clear diagnostics:

- (1) a. John seems to be happy.
- b. John is likely to return.
- (2) a. John wants to be happy.
- b. John is eager to return.

Semantically, the distinction between predicates of the *want* type and those of the *seem* type is easily defined: the former take two arguments, a term and a proposition, while the latter take only a single, propositional argument. This semantic difference has systematic syntactic correlates and the standard theory succeeded in accounting for both the semantic and the syntactic characteristics of these predicates by deriving the sentences in which they appeared from distinct underlying forms, as in (3) and (4).

- (3) --- seem [s John to be happy]
- (4) John want [s John to be happy]

The surface forms of the sentences in (1) and (2) are then produced by applying the transformations of SUBJECT-TO-SUBJECT RAISING and EQUI-NP-DELETION to (3) and (4) respectively.

These derivations allow the underlying forms of the sentences to express in a direct way the semantic differences between *seem*-type and *want*-type predicates and they also provide a clear way to express the syntactic differences between the two predicate types. These syntactic differences are several in number:

I. *seem*-type predicates may have expletive there subjects but *want*-type predicates may not.

- (5) a. There seems to be something missing.
- b. * There wants to be something missing.

II. *seem*-type but not *want*-type predicates may have surface subjects that are parts of idiomatic constructions in the complement infinitive.

- (6) a. Tabs seem to have been kept on Mary.
- b. *Tabs want to have been kept on Mary.

III. *seem*-type predicates may have “weather” *it* subjects but *want*-type predicates may not.⁵

- (7) a. It seems to be dark.
- b. *It wants to be dark.

IV. *seem*-type predicates appear with expletive *it* subjects but *want*-type predicates do not:

- (8) a. It seems to bother Bill that the problem has solved itself.
- b. *It wants to bother Bill that the problem has solved itself.

All of these differences amount to one thing: The subject of a *seem*-type predicate with an infinitive complement bears no thematic relation to the matrix verb and it may appear as the matrix subject whenever it is acceptable as the subject of the complement predicate. It makes no difference whether the subject of the complement predicate is an underlying subject or appears in subject position as the result of some transformational rule. The subject of a *want*-type predicate, on the other hand, bears a thematic relationship to the matrix predicate and so must appear as subject in underlying structure as well as on the surface. By representing the underlying structure of the two types of predicate as in (3) and (4) and by having transformations like PASSIVE and *there*-INSERTION apply cyclically, the standard theory guaranteed that these differences were captured.

The essential features of the standard theory analysis of the raising/control distinction are preserved throughout the development that culminates in Chomsky’s current Government and Binding framework. Substantial changes in notation and theoretical orientation may obscure this equivalence but it remains the case that the surface subjects of *seem*-type predicates originate as the subjects of the complement infinitive while the subjects of *want*-type predicates undergo no movement. The transformation of EQUI-NP-DELETION has been replaced by conditions on the appearance of the empty category PRO and the transformation of RAISING-TO-SUBJECT has been replaced by the more general rule of MOVE-NP; that is, the NP-movement subcase of MOVE- α . More strikingly, the elaboration of a complex typology of empty categories has made it possible to encode derivational information through coindexing empty elements with antecedent NP’s whose grammatical status determines the type of the empty element. This use of empty categories has made it possible in principle to do away with transformations entirely, but since the coindexing pattern can be made to follow from the application of transformations, the choice of whether to consider the theory transformational or not seems secondary. These changes, while important at the level of the theory of grammar, do not change the way the theory accounts for the pattern described above.

Non-derivational theories like GPSG and LFG, on the other hand, treat the raising/control distinction in a somewhat different way. Since the surface structure sentence is not derived from an underlying level but rather is generated directly by the context-free phrase structure base, the matrix subject must originate in its surface position. To capture the fact that it behaves thematically as the subject of the complement in a sentence with a *seem*-type predicate, these grammars must set up another level of representation at which the semantic complement structure is expressed. For GPSG this is accomplished by different semantic translation rules and for LFG it is the level of functional structure. Thus, in Gazdar 1982 the phrase structure rules that generate the paired sentences of (1) and (2) are essentially identical. The difference between the two sentence types is captured by having a different semantic translation rule for each of them; and, as Sag 1982 points out, this approach necessitates treating the facts of I-IV as semantic in nature. Thus, in Sag's extension of Gazdar's analysis the reason why (6a) is grammatical while (6b) is not is that the semantic translation of *seem* sentence treats the NP *tabs* as an argument of *keep* and not as an argument of *seem* while the translation of *want* treats the subject NP as an argument of both the matrix and complement verbs. He claims that the co-occurrence restrictions on idiom chunks should be treated semantically rather than syntactically and thereby is able to rule out the semantic translation of the *want* sentence as meaningless. It may be that a natural "syntactic" account can be given within the GPSG framework and in fact, we understand that such approaches have been considered in the GPSG work. However, the most recent accounts of distinction appear to us to be semantic (see Klein and Sag 1985).

In LFG the difference between a raising and a control predicate is handled by associating different functional structures with each type. Thus, the lexical entry for a *want*-type predicate will indicate that its subject functions thematically both as the subject of the matrix verb and as the subject of the infinitive while the entry for a *seem*-type predicate will indicate that its subject plays the thematic role of subject of the infinitive and no other. If an expletive or idiom chunk NP appears as the subject of a *want*-type predicate, the functional structure will assign that NP to the predicate argument structure as the first argument of the matrix predicate and so will produce an uninterpretable output since the NP assigned to that position will have no semantically interpretable lexical entry. On the other hand, if such an NP appears as the subject of a *seem*-type predicate, the functional structure will assign the NP as the subject of the lower predicate. If that lower predicate mentions the meaningless item in its lexical entry as carrying a grammatical function (but not filling a slot in predicate argument structure), then the sentence will have an interpretation. This procedure is, of course, recursive so that these elements can be passed down the functional structure of a sentence through any number of predicates.

In spite of many differences, the GPSG and LFG theories of grammar share the property that constituent structure is not a sure guide to thematic relations. In the unmarked case, however, grammatical relations are assigned to NP's according to their constituent structure position; that is, the surface subject of a sentence is interpreted as the thematic subject of its predicate. Thus, there is in each of

these theories an apparatus for assigning thematic roles to NP's not in the canonical position to receive these role assignments. This apparatus is then brought into play to account for such constructions as raising. It is clearly the contention of GB theory that such an apparatus is unnecessary because a correctly defined constituent structure augmented by suitable empty categories will allow thematic roles to be read directly from the (surface) syntactic representation, a point elaborated at some length in Williams 1984. It is of interest, therefore, that the TAG notation allows analyses of raising in which thematic roles are read from constituent structure without the invocation of semantic or functional structure mechanisms.

As we have shown in the preceding sections of this paper, TAGs factor recursion and dependencies in a different way than other grammatical notations developed since *Syntactic Structures*. Instead of using recursive symbols in a context free grammar to generate embedded structures, they embed sentences under other sentences by the use of an adjunction operation reminiscent of the generalized embedding transformations of Chomsky's early work. Relying on this device, we can develop the analysis promised above by defining an appropriate set of auxiliary trees to handle infinitival complements. Consider first the analysis of *want*-type predicates. Under Chomsky's analysis *want* or *try* appears in a structure like the following:

$$(9) \quad [S \text{ I } [VP \text{ want } [\bar{S} \text{ COMP } [S \text{ PRO } [VP \text{ to eat}]]]]]$$

In this analysis, the lexical entry for a control verb will state that it subcategorizes for an infinitival \bar{S} complement. That the subject of the complement is empty is guaranteed by the case filter, since a lexical NP would not receive case as the subject of an infinitival \bar{S} . Similarly, the fact that a control verb has a thematic subject need not be mentioned since that is the default case. This analysis of control predicates can be carried over directly into a TAG if we allow them to appear in auxiliary trees like β_2 of section 3. The adjunction operation will produce the proper constituent structure and the parallel semantic composition will yield a correct semantic translation under which the subject of the matrix verb and the \bar{S} complement to the verb are arguments of the matrix predicate. The interpretation of PRO as obligatorily controlled by the matrix subject can be guaranteed in standard way by translating the complement \bar{S} as an open sentence and writing the rule of semantic composition with the matrix auxiliary tree so that it substitutes the translation of the matrix subject (or a variable bound by a quantifier on the matrix subject) for the free variable, in accordance with a version of the theory of control.⁶ The fact that these predicates take infinitival complements rather than tensed ones can be coded straightforwardly as a local constraint either on the adjunction of the auxiliary tree to the initial sentential tree or on the lexical insertion of the predicate, as can the requirement that the subject of the infinitive be PRO. Indeed, as far as we can see, the entire case theory of GB grammar can be incorporated directly into a TAG as a set of local constraints.

It is also possible in a TAG to analyze the infinitive complements of control predicates as lacking syntactic subjects. Under this analysis the structure for control

predicates given in (9) would be replaced by that in (10) and the sentential auxiliary tree needed for the PRO analysis would be replaced by a VP-rooted one:

(10) [S I [VP want [VP to eat]]]

Semantically, it would be interpreted in the standard way; namely, by composing the matrix verb with the complement VP to yield a derived propositional function that takes the matrix subject as its single argument. This alternative analysis, the one adopted in GPSG⁷, allows one to eliminate the empty category PRO from infinitival complements (See May and Koster 1982 and Baltin 1984 for some of the problems with this approach). The basic contrast between control predicates and raising predicates is that the latter assign no thematic role to their surface subjects. This fact can be captured neatly in a TAG by requiring that raising predicates be inserted in an auxiliary tree with no subject; that is, a tree like (11) with VP as its root:⁸

(11)
$$\begin{array}{c} \text{VP} \\ \swarrow \quad \searrow \\ \text{V} \quad \text{VP} \end{array}$$

If the lexical item *seem* is inserted into this tree and it is then adjoined to an infinitival tree without a complementizer, the resultant tree will give a reasonable constituent structure for a raising sentence. Thus, an example like (1a) would be derived by adjoining the auxiliary tree in (11) to the initial tree in (12), yielding (13) as a result.

(12)
$$\begin{array}{c} \text{S} \\ \swarrow \quad \searrow \\ \text{NP} \quad \text{VP} \\ | \quad \swarrow \quad \searrow \\ \text{John} \quad \text{V} \quad \text{AP} \\ \quad \swarrow \quad \searrow \quad | \\ \quad \text{to} \quad \text{be} \quad \text{happy} \end{array}$$

(13)
$$\begin{array}{c} \text{S} \\ \swarrow \quad \searrow \\ \text{NP} \quad \text{VP} \\ | \quad \swarrow \quad \searrow \\ \text{John} \quad \text{V} \quad \text{VP} \\ \quad | \quad \quad \quad \swarrow \quad \searrow \\ \quad \text{seems} \quad \quad \quad \text{to} \quad \text{be} \quad \text{happy} \end{array}$$

The semantic translation rule associated with this adjunction will, of course, simply apply the raising predicate to the complement sentence interpreting it as an operator that takes a single, propositional argument. Under this analysis the properties of raising predicates fall out automatically. The properties of the constituent structure matrix subject are determined entirely by the complement predicate because prior to adjunction it was the subject of the complement. The raising verb places no constraints on the subject because in the lexicon, where thematic roles are specified, it is marked as not assigning a thematic role to the subject. (See section 5 for further discussion of the interaction of syntax with thematic role assignment.) Thus, this analysis accomplishes the proper treatment of raising without the use

of either extended semantic mechanisms or transformational movements and it is stated within a grammatical formalism of provably restricted power.

One objection that might be raised to the analysis of raising that we have presented is that our grammar provides no direct syntactic encoding of the relationship between sentence pairs like the following:

- (14) a. John seems to like those strawberries.
- b. It seems that John likes those strawberries.

In a GB grammar these two sentence types are directly related because they both derive from deep structures in which *seem* has an empty subject and takes a complement sentence. The fact that the complement is untensed in (14a) and tensed in (14b) forces raising to apply in the former case and prevents it from applying in the latter. The expletive *it* in (14b) is inserted by a late rule so that the sentence will not lack a subject. The GB analysis of these sentences is attractive, but there is reason to think that treating the two cases as parallel is a mistake. As Bresnan points out (1982), very few raising predicates allow full sentential complements as well as raised infinitival ones. Thus, consider the following pairs listed by Bresnan:

- (15) a. Fred tends to ignore Mary.
- b. *It tends that Fred ignores Mary.
- (16) a. Louise is apt to lose her temper.
- b. *It is apt for Louise to lose her temper. (on the relevant interpretation)
- (17) a. There is going to be a movie made about us.
- b. *It is going that there will be a movie made about us.

If we accept Bresnan's claim that the relationship between raised infinitive and full sentence complements is not productive in English, then there is no reason to represent the two types similarly in the syntactic component of the grammar. The fact that (14a) and (14b) mean the same thing will be expressed in the lexicon because *seem* will be subcategorized for two complement types that wind up having the same semantic translation because of the semantic character of *seem*; namely, that it is a modal-like operator on sentences and so does not assign a thematic role its subject.

4.2 The problem of nominalizations

The most interesting consequence of a TAG analysis of raising predicates is that it allows us to explain the often noted fact that there are no raising nominalizations. Thus, there is a systematic contrast between sentences of the type of (18) and of (19):

- (18) a. John appears to have left.

- b. John is likely to have left.
- (19) a. * John's appearance to have left surprised us.
- b. * John's likelihood to have left surprised us.

Chomsky (1970) notes this contrast and proposes that it be explained as a consequence of a general principle to the effect that sentences but not noun phrases allowed freely undergo transformations. In more recent accounts, this principle is no longer invoked and it is instead claimed that nouns cannot properly govern traces while verbs and adjectives can. Thus, the syntactic representation in (20) is well-formed because the trace e is properly governed by the verb *appear* while in (21) the trace is not properly governed and so violates the Empty Category Principle:

- (20) [_S John appears [_S e to have left]]
- (21) [_S [_{NP} John's appearance [_S e to have left]] surprised us]]

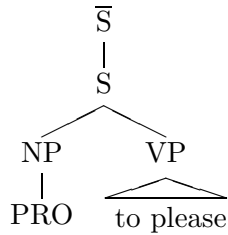
Unfortunately, this account, like the earlier one, amounts to stipulating that nouns cannot function as raising predicates. Since nominalizations of predicates that take infinitival complements with PRO subjects are possible, as is shown in (22) and (23), the stipulation is specific to the raising construction:

- (22) a. John is eager to please.
- b. John attempted to please.
- (23) a. John's eagerness to please surprised us.
- b. John's attempt to please surprised us.

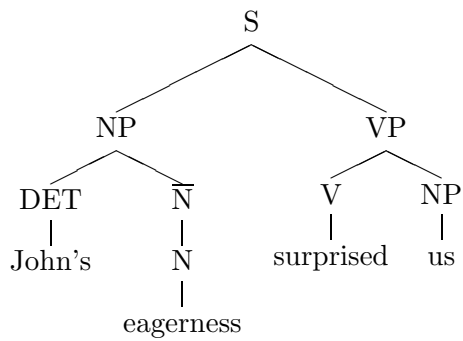
To find a real solution to the problem of the non-existence of raising nominals obviously depends on deriving the different behavior of nouns and verbs/adjectives in this case from general differences between the categories⁹ as well as showing how the difference in the structures assigned to raising and control predicates entails that latter can be nominalized while the former cannot. We shall now proceed to show that TAG provides a solution of the type desired. Indeed, the reader will see that a TAG is inherently incapable of generating raising nominals or of excluding control nominals. Moreover, he/she will here see the potential empirical value of the formalism, as it functions to exclude certain unwanted configurations without invoking any principle beyond its own rules of formation and the subcategorization requirements of lexical items.¹⁰

In a TAG the syntax of derived nominals and sentences will not be parallel. Consider, for example, how a TAG of English would derive sentences (22a) and (23a). Because initial trees always have S or \bar{S} as their root node, the initial trees for these two sentences would have to look like (24) and (25) respectively:

(24)

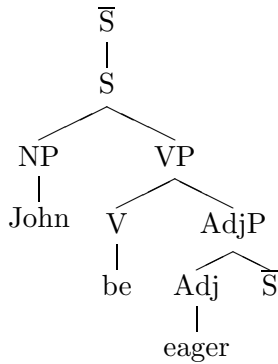


(25)

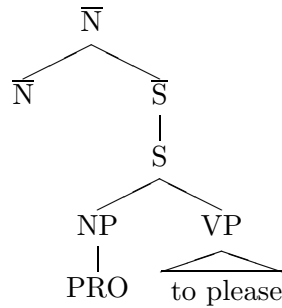


The auxiliary trees for these sentences would then be (26) and (27):

(26)



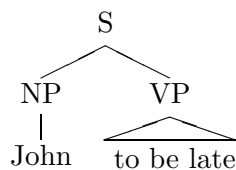
(27)



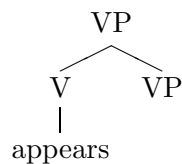
Because the syntactic adjunction in the two cases attaches an infinitive complement to *eager(ness)*, the semantic composition will assign the same interpretation to the noun+complement as it does to the adjective+complement. The fact that the role of initial and auxiliary trees is inverted in the case of the nominalization affects neither syntactic structure nor interpretation.

If we now consider the sentences of (18) and (19), however, we find a different pattern. In accord with our earlier exposition, the initial and auxiliary trees of sentence like (18a) would be those of (28) and (29):

(28)



(29)



The ungrammatical sentence (19a), on the other hand, is underivable. For instance we cannot use the initial and auxiliary tree structures that underlie the *eagerness*

sentence to generate (19a), for the subcategorization requirements on *appearance* will block its co-occurrence with an infinitival \bar{S} complement. Alternatively, if we attempt to construct a raising derivation for the sentence, we find that move blocked precisely by the fact that *appearance* is of the grammatical category ‘noun’ rather than ‘verb’. The raising analysis entails that *John(’s)* be the subject of *to be late* at the start of the derivation, and this requirement in turn entails that *appearance* be part of an auxiliary tree which has a VP node at the frontier. That is to say it must appear in a tree with the form of (30):



This structure is, however, impossible because a noun cannot function as the head of a verb phrase; and if we require that members of the set of elementary (initial and auxiliary) trees conform to the \bar{X} phrase structure schema, we will insure that no such structure will be available to a TAG derivation. The reader may check for him/herself that no other derivation will yield the ungrammatical (9a).

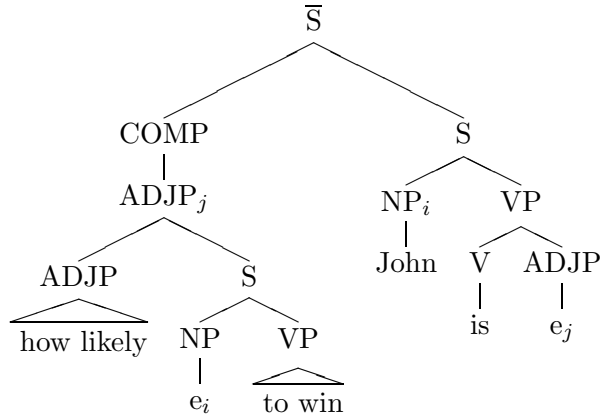
4.3 Further considerations

Let us now discuss certain difficulties raised by the analysis proposed above. For instance, consider the raising sentence in (31), which we give in an embedded context to avoid the complications introduced by subject-aux inversion:¹¹

(31) (I know) how likely to win John is.

While we will not discuss the analysis of *wh*- constructions in detail until section 6, this particular sentence poses difficulties for our analysis of raising and so we will discuss it here on the basis of the brief presentation in sections 2 and 3. As we indicated there, *wh*- constructions will be handled in a TAG with the linking device of section 2.2. These links are introduced into a TAG in a strictly local way; that is, they are only introduced into elementary trees. Since in our analysis of raising constructions the raising predicate is introduced through adjunction of a tree with a VP root (see (11) above) and since the *wh*- phrase in (31) is the raising predicate itself, there is no way for us to derive (31) as a raising construction under the assumptions we have been making. This sentence poses a similar problem for the GB account of raising, as can be seen if we look at the structure of the sentence under GB assumptions, here given in (32):

(32)



The problem for GB is simply that the empty category in the subject position of the raising complement is not *c*-commanded by its antecedent, the NP *John*. Since GB binding theory requires that traces be bound at s-structure¹² by a *c*-commanding antecedent, this structure is ill-formed and the sentence that it corresponds to is falsely predicted to be ungrammatical.¹³

Although there are ways around the problem for the TAG formalism that is posed by sentence (31), the solutions have an ad hoc character. Instead of pursuing these, we prefer to take the predictions of our formalism seriously and to ask whether raising is in fact, involved in the derivation of (31). Consider the following sentences:

(33) Hope is slow to die.¹⁴

(34) John's likelihood/probability of winning is slight.

(35) John became sick.¹⁵

The sentence in (33) and the nominalization in (34) share with raising constructions the property that their subjects bear no apparent thematic relation to the predicate with which they are construed syntactically. However, in these cases there can be no question of proposing a raising analysis since the key syntactic evidence for raising, the appearance of expletive and idiomatic NPs as raising subjects, is absent, as the ungrammaticality of the sentences in (36) and (37) shows:

(36) a. *There are slow to be solutions found for difficult problems.

b. *There's likelihood/probability of being a depression is difficult to estimate.

(37) a. *Headway was slow to be made on the problem.

b. *Headway's likelihood/probability of being made is uncertain.

c. *Headway became made on the problem.

From these facts we conclude that the semantic diagnostic for raising, that is, the failure of a predicate to assign a thematic role to its subject, is an inadequate one and that there can be cases which meet the semantic requirement for raising without

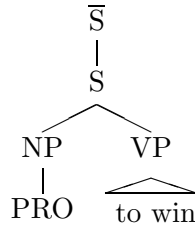
in fact exhibiting syntactic raising. This is a problem for linguistic theory and we do not pretend to have a ready solution for it, except to say that it may be a reflection of the relative autonomy of syntax and semantics. The point we wish to make is that by the syntactic criterion of the behavior of expletive and idiomatic subjects, the structure of (31) is not a raising structure, as the ungrammaticality of the following sentences shows:

- (38) a. *I know just how likely to be a problem there is.
 b. *I know just how likely to be made on the problem real headway is.

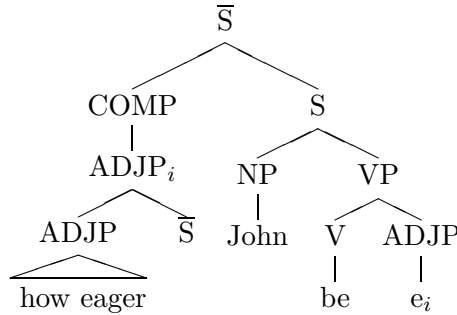
This evidence leads to the conclusion that (31) is a control structure and that it is derived just as (39) below is, namely from the adjunction of an initial tree with the structure of (40) and an auxiliary tree with the structure of (41):

- (39) (I know) how eager to win John is.

(40)



(41)



The problem of how and when raising constructions become reanalyzed as control structures we must leave to further research.

5 The Passive in a TAG

5.1 The link between raising and the passive

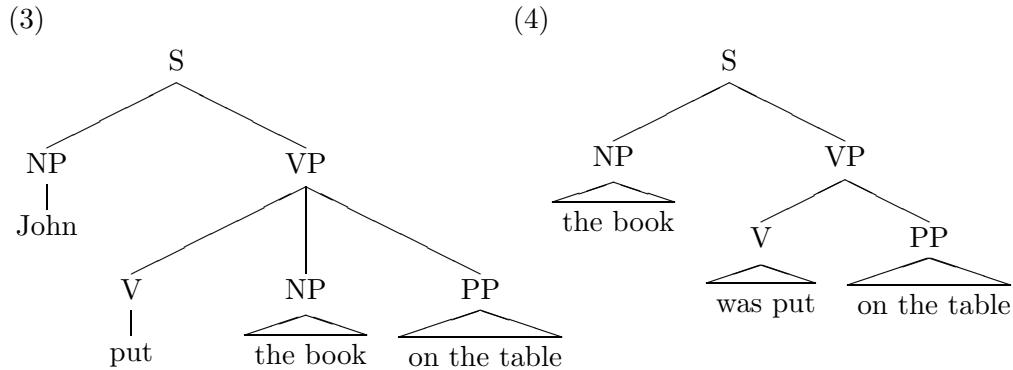
As we have shown, a TAG instantiates the difference between raising and control verbs in a natural way, but we have captured this difference without the use of links or empty categories. In consequence, we have not reconstructed in any direct way the notion ‘NP trace’, which plays such a crucial role in a GB grammar. We must ask, therefore, how the range of possible syntactic analyses implementable in a TAG is affected by the absence of NP trace and whether the available analyses are adequate. Although we cannot hope to investigate the entire range of such cases in this paper, we shall look in some detail at what is perhaps the most central one, the passive. The passive can be handled in a TAG in at least two ways, which correspond roughly to a transformational and lexical analysis in a transformational

grammar. To emulate a transformation analysis of the passive it is necessary simply to define an operation on elementary trees that converts the trees for active sentences into ones for passive sentences suppressing the direct object NP in the active sentence. This operation is one of family of rules that can be defined as mappings from elementary trees to elementary trees. These mappings may be thought of as producing derived simplex sentence trees from canonical trees and, therefore, as similar to cyclic transformations in standard transformational theory.¹⁶ The rule for the passive might look as follows:

- (1) $[_S NP_1 [_{VP} V NP_2 X]] \longrightarrow [_S NP_2 [_{VP} V X]]$
 for X any maximal projection

The labeled bracketings here should be treated as functions which, for any value of the variable X, specify a tree. Thus, a sentence like (2a) below will have the tree (3) and the passive sentence in (2b) will have the tree in (4):

- (2) a. John put the book on the table.
 b. The book was put on the table.



The proper interpretation of the passive is accomplished by coding the fact that the sentence is a passive in the index numbers of the derived trees produced by this rule. Then the interpretive component of the grammar can be directed to interpret a passive sentence by reference to the active tree from which it is derived.

This analysis in terms of pseudo-transformations works on ordinary passives like (3a) and can be extended in a natural way to dative, prepositional, and idiomatic passives, like those in (5) below:

- (5) a. She was sent a letter.
 b. The bed hasn't been slept in.
 c. He was taken advantage of

If the dative alternation is captured by a transformational rule in the TAG, dative passives like (5a) will be produced by applying the passive rule to the derived initial

tree produced by the dative rule. If the dative is not treated transformationally then the passive rule will, of course, apply to the double object dative in the same way that it applies to ordinary transitive sentences. Prepositional passives and idiomatic passives can be handled by allowing initial trees to undergo reanalysis by which sequences like *sleep in* or *take advantage of* are bracketed together as complex verbs. Bresnan (1982) and Baltin (1978) give good arguments for treating such passives as cases of reanalysis.

Unfortunately, the pseudo-transformational analysis fares less well with raising passives like those in (2c) above. The difficulty is, of course, that the subject of the passive is not the thematic direct object of the active sentence. Under the standard transformational theory of the 1960's, this problem was solved by a transformation of raising to object which altered the constituent structure of a sentence like (6) from to (7b):

- (6) Mary believes Sarah to have left.
 (7) a. [S Mary believes [S Sarah to have left]]
 b. [S Mary believes [NP Sarah][VP to have left]]

Because the interpretation of grammatical relations takes place at deep structure in a standard theory grammar, this transformation allows the shifted NP to behave thematically as an infinitive subject and syntactically as the object of the matrix verb. Recent non-transformational analyses adopt the constituent structure in (7b) and rely on a level of functional structure or semantic interpretation to capture the fact that the true thematic role of the direct object NP is as subject of the infinitive clause. Under all of these analyses, raising passives are assimilated to simple sentence passives and pose no special problems. This option is available in the TAG formalism as well, since the constituent structure in (7b) is subject to rule (1).¹⁷

There are, however, substantial reasons for arguing that the constituent structure in (7b) is incorrect and that the structure in (7a) is to be preferred. As Baltin (1984) has shown, there is a contrast between *believe*-type and *persuade*-type verbs in the possibilities for extraposition of the infinitive VP. Compare, for example, the following sets of sentences:

- (8) a. I myself believe Mary to be the best choice.
 b. I believe Mary to be the best choice myself.
 c. ??I believe Mary myself to be the best choice.
 (9) a. I myself persuaded Mary to accept the nomination.
 b. I persuaded Mary to accept the nomination myself.
 c. I persuaded Mary myself to accept the nomination.

Arguing in a transformational framework, Baltin claims that an emphatic reflexive on subject position, as in (8a) and (9a), can shift to the end of the verb phrase, as in

(8b) and (9b), but not to a position interior to the VP. Comparison of the sentences in (10) below establishes this point:

- (10) a. I myself put the book on the table.
 b. I put the book on the table myself.
 c. ??I put the book myself on the table.

In sentences like (9c), which appear to violate the rule for placement of emphatic reflexives, Baltin argues that the transformation of \bar{S} extraposition has applied. This rule is needed to account for such alternations as those in (11) and (12):

- (11) a. Mary stated that the plane was safe to the committee.
 b. Mary stated to the committee that the plane was safe.
 (12) a. Mary called the people who were involved up.
 b. Mary called the people up who were involved.

Extraposition will apply to derive a sentence like (9c) from the structure underlying (9b) because the latter has the structure (13):

- (13) [I [VP persuaded Mary [\bar{S} PRO to accept the nomination] myself]]

Applying the transformation will produce the structure in (14):

- (14) [I [VP persuaded Mary e myself] [\bar{S} PRO to accept the nomination]]]

In a *believe*-type sentence like (8b) the extraposition rule cannot apply in a parallel way since the infinitive does not have a PRO subject.¹⁸

If we assign the same constituent structure to the complements of *believe* and *persuade*, however, this explanation of the difference between them is no longer available. In an LFG or a GPSG grammar the constituent structure of a *persuade* sentence, like that of a *believe* sentence, contains a subjectless infinitive. In consequence, there is no natural way to prevent the insertion of an emphatic reflexive between the accusative NP and the infinitive in the *believe* sentence without also blocking it in the *persuade* sentence.¹⁹

5.2 A TAG analysis of the passive

The above line of argument seems to us persuasive enough to justify trying to develop a TAG analysis of raising passives which does not require the infinitive subject to be a matrix direct object in constituent structure. Such an analysis, however, cannot be of the pseudo-transformational kind sketched above. Since TAG pseudo-transformations apply only to elementary trees, a structure with an embedded infinitival sentence is not a possible input to such a rule. It turns out, interestingly,

that we can achieve our goal by implementing an analysis of the passive which is lexical in character.

The distribution of raising passives suggests that such a lexical analysis is, moreover, to be preferred on empirical grounds. The reason for this is that there are a number of cases where the raising passive exists but its active counterpart does not (Bach 1980). Thus, none of the following examples with infinitive complements have a direct active counterpart:

- (15) a. John is said to be a crook.
b. John is reputed to have absconded with the payroll from his last job.
c. John was heard to call for “liberty or death.”
d. John was made to rescind the proclamation.

Similarly, the following examples where the matrix passive has an adjectival complement sound very odd in the active:

- (16) a. How many Americans are known dead in the embassy bombing?
b. *How many Americans do they know dead in the embassy bombing?
a. The Americans are believed involved in the coup.
b. *They believe the Americans involved in the coup.

Thus, the relationship between the active and passive exhibited by verbs like *believe* and *expect* is not entirely general. If anything, the number of verbs taking infinitive complements for which both active and passive forms exist is smaller than the number for which only one or the other exists. This circumstance suggests that raising passives should be entered in the lexicon as separate entries from their active counterparts.²⁰

In order to develop a lexical analysis in the TAG framework, we must first consider the question of how the lexicon of our grammar should be structured.²¹ Suppose that, from the perspective of syntax, we consider each lexical item to itself be a tree, with the lexical category of the item as the root and the word (or words in the case of complex lexical items like idioms) as the terminal node. Lexical insertion would then be a substitution operation under which a terminal node of the syntactic tree, which in TAG is a lexical category, has a lexical tree substituted for it.²² The resulting tree will then just be a syntactic tree with actual words as its terminal nodes. Subcategorization constraints and/or thematic role assignments (in the sense of Gruber 1985 as extended by Jackendoff 1972) can be considered as conditions on the lexical insertion operation. Furthermore, we can incorporate the case-marking properties of predicates into their lexical entries by labelling the subcategorized NP's with the case marking that the predicate assigns or with the particular preposition that the verb requires. In this way, we insure that predicates are inserted only into environments where their complement NP's are properly case marked and governed by the proper prepositions.

Under this treatment of lexical insertion, a lexical passive rule can easily be defined that will take the place of the pseudo-transformational rule we introduced initially. This lexical rule might be formulated as in (18), stating that for any verb which assigns accusative case to an adjacent object NP in its basic subcategorization, there exists a morphologically related passive verb in whose subcategorization the object NP is suppressed:

$$(18) \quad V: \langle _ _ NP[acc] X \rangle \longrightarrow V[+passive]: \langle _ _ X \rangle$$

Since subcategorization is a constraint on lexical insertion, this rule insures that the passive form of any verb will appear in the proper syntactic environment. It is not, however, sufficient to guarantee that the interpretation of the passive form will be correct. We must still capture the apparently universal fact that the suppression of the direct object entails the suppression of the subject thematic role in the passive sentence and the assignment of the direct object's thematic role to the passive subject (Burzio's generalization (Burzio 1981)). In the pseudo-transformational analysis presented above this result is achieved by the indexing of the subject and object NP's and the interpretation of the passive tree in terms of the active one from which it is derived. This machinery is rather creaky and ill-motivated, and there is a much more attractive solution available. Let us suppose that the lexical passive rule, instead of simply suppressing the direct object, assigns its thematic role to the subject position. Since each argument position in a sentence bears one and only one thematic role (the functional uniqueness principle of Freidin 1978 or the theta criterion of Chomsky 1981a), the suppression of both direct object and the thematic role assigned to the subject by the active verb follows directly, as does the correct interpretation of the passive sentence. We may assume, furthermore, that the loss of accusative case-marking ability by the passive participle is a direct consequence of its loss of thematic role assignment to a direct object. A modified lexical rule, which would now mention thematic roles as well as constituent structure might be formulated as in (19):

$$(19) \quad \begin{array}{ccc} [S \ NP[\theta_1] \ [VP \ NP[\theta_2] \ X \]] & \longrightarrow & [S \ NP[\theta_2] \ [VP \ X \]] \\ [+nom] \quad \quad \quad [+acc] & & \quad \quad \quad [+nom] \end{array}$$

θ_1 and θ_2 are variables ranging over the thematic roles assignable to the subjects and objects respectively of active sentences.

Once again the case marking in the rule will serve as a constraint on lexical insertion but the thematic role marking has to be treated differently. The most straightforward interpretation of it is as a feature marking imposed on the phrase structure tree at the time of lexical insertion, which constrains the semantic interpretation of the tree. Other interpretations are also possible, depending on how the interface between syntax and semantics is conceived; but the considerations involved in choosing among the alternatives are beyond the scope of this discussion.

5.3 Impersonal and raising passives

Before we extend this analysis to the raising passive case, let us briefly discuss how our analysis might be extended to the impersonal passive. Although this construction does not exist in English, it is common enough in other languages and extending our analysis to accommodate it will help to simplify our treatment of the English raising passive. Consider the following examples from French and German:

- (20) a. Es wurde getanzt. (“There was dancing”)
b. Man tanzte.
- (21) a. Il a été tiré sur le bateau. (“The boat was shot at.”)
b. On a tiré sur le bateau.

In both of these cases the subject of the passive is an expletive and the active sentence (here the b sentence) to which the passive corresponds has an intransitive verb. The first fact clearly depends on the second since the absence of a direct object in the active form implies that no promotion of the direct object’s role to the subject position can occur, but it does not directly follow from it. Indeed, it is not obvious why the subject of an intransitive verb must disappear in the passive. There is no direct object theta role to displace the θ_1 role on the subject and no direct object to displace the subject syntactically. One might, therefore, have found that the subject of the active could remain as the subject of the passive. In other words, the suppression of the thematic subject appears in these cases to be a more fundamental feature of the passive than the promotion of the direct object to subject. But this apparently fundamental characteristic of the passive can hardly be fundamental in fact, for the passive must basically be a device for the promotion of direct objects. After all, every language with passive construction allows the promotion of direct objects but only some permit an impersonal passive. On these grounds, therefore, we propose the following treatment of impersonal passives, similar in spirit to the relational grammar analysis of Perlmutter and Postal 1984. Suppose that the suppression of the thematic subject in the impersonal passive is, contrary to appearances, actually due to its replacement by another thematic role; namely, the null role. Under this hypothesis the impersonal passive would necessarily be subjectless since the appearance of a lexical subject would violate the theta criterion.²³ The presence of such a null theta role can be derived if we assume that the proper analysis of passivizable intransitive verbs in languages with impersonal passives is as transitives which assign the null theta role to the direct object position (compare Roeper 1984). Instead of lacking a direct object at every level of representation, they have a direct object slot in their lexical representation; but since the null theta role is assigned to that slot, the direct object can never surface because the null theta role does not fulfill the theta criterion.

As Perlmutter and Postal point out, an analysis of the impersonal passive as involving a ‘dummy’ direct object is attractive on a number of grounds. Some of their arguments are entirely internal to relational grammar, but the gist of them –

that this analysis accounts for the many parallels between impersonal and personal passives – is applicable to any theory of thematic role assignment and grammatical relations. From our perspective, the most important feature of the analysis is that it can be extended directly to the raising passive cases of English. Let us suppose that the thematic and case marking features of a verb like *believe* are given in the following partial lexical entry:

$$(22) \text{ believe: } \langle \text{ -- } [_s \text{ NP}[\text{null } \theta] \text{ to VP }] \rangle \\ \text{ [+acc]}$$

This subcategorization frame has an exceptional feature; namely, it mentions nodes that are not sisters to the verb it subcategorizes. Classically, (see Chomsky 1965) it has been assumed that only arguments of a verb, in a phrase structure grammar its sisters, may subcategorize it. In the case of *believe*-type verbs, however, the verb exceptionally assigns case to an NP which is neither its argument nor its phrase structure sister. This exceptional behavior must be accommodated in any grammar of English; and given our incorporation of case marking into subcategorization frames, we can most easily accomplish this by allowing the subcategorization frame of a *believe*-type verb to mention the internal structure of its sentential complement at the level of the NP to which the verb assigns case. We can now ask ourselves what thematic role the verb *believe* assigns to the accusative NP. Obviously, it can assign no role since that NP is thematically marked as the subject of the infinitive verb. But given our analysis of the impersonal passive, we now have the choice of saying that *believe* fails to assign a role and saying that it assigns the null role. Since the null role does not count as fulfilling the theta criterion, it will also not cause the accusative NP to be excluded for bearing two roles. To block the appearance impersonal passives in English, we will want to say that intransitives do not assign an object position theta role at all; but in the case of raising passives we will say that superficially transitive verbs of the *believe*-type assign the null theta role to their objects, perhaps because case assignment can only occur when a predicate mentions a NP position in its thematic assignments.

Once we allow the subcategorization given in (22), the passive rule will automatically apply to it. If we assume a pruning convention that eliminates non-branching phrasal categories, the application of the passive rule yields the following result:

$$(23) \text{ believe[+passive]: } \langle [_s \text{ NP}[\text{null } \theta] [_{VP} \text{ -- to VP }]] \rangle$$

The VP constituent of this lexical entry is the same as the subcategorization frame given earlier to the raising verbs (*seem* etc.) If we augment that frame to include thematic role assignment information, we need only specify that raising verbs assign the null role to their subjects to guarantee absolute identity between *believe*-type passives and raising verbs. The only augmentation that we need make to our earlier analysis of raising to make it work out technically is to specify how thematic role marking and lexical insertion interact with adjunction, since the elementary tree for

the raising construction does not contain a subject. This interaction can be handled in more than one way, but the simplest is to make lexical insertion “free,” in the sense that it will be possible at any stage of a syntactic derivation. The principle governing the lexical insertion of predicates could be just that such insertion occurs as soon as all of the syntactic positions mentioned in the lexical entry for the predicate are available for checking in the structure into which the predicate is to be inserted.²⁴ This interleaving of lexical insertion with syntactic derivation is natural to a TAG and represents one of ways in which TAG’s differ from standard context-free phrase structure grammars, in which lexical insertion follows the generation of the syntactic tree. Since, in fact, the constraints on lexical insertion are always local, it is a point in favor of the TAG formalism that it treats insertion locally.²⁵

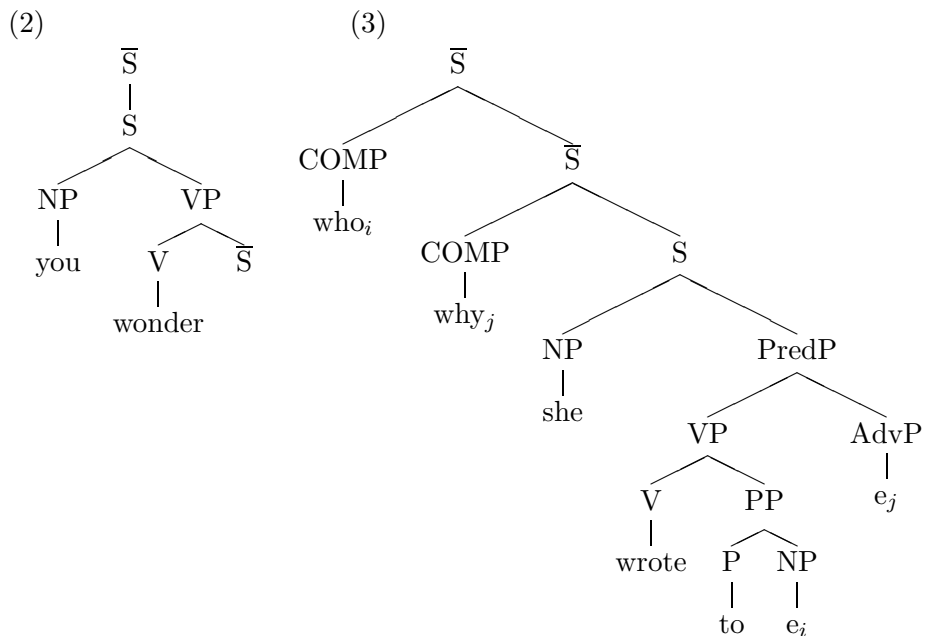
6 *Wh*- movement in a TAG

6.1 Subjacency in a TAG

6.1.1 As we illustrated in section 3, *wh*- movement can be accommodated in a TAG by the addition of linked trees to the inventory of initial trees. The unbounded character of *wh*- movement constructions will then follow from the nature of the adjunction operation, by which matrix predicates can be adjoined “between” the *wh*- in COMP and the embedded S, as in Example 3.3. One interesting consequence of using links with adjunction to express long-distance *wh*- dependencies is that some island constraints on extraction described in Ross 1967, especially those for which Chomsky proposed the principle of subjacency, are storable in an extremely natural way. Consider, for example, a *wh*- island violation like the following:

- (1) * Who_{*i*} did you wonder why she wrote to e_{*i*}?

This sentence cannot be generated by the TAG formalism as it stands. Since long-distance *wh*- movement is always the result of adjunction of a matrix predicate above an embedded clause, (1) could only be produced by the adjunction of the auxiliary tree (2) to the initial tree (3), as below:

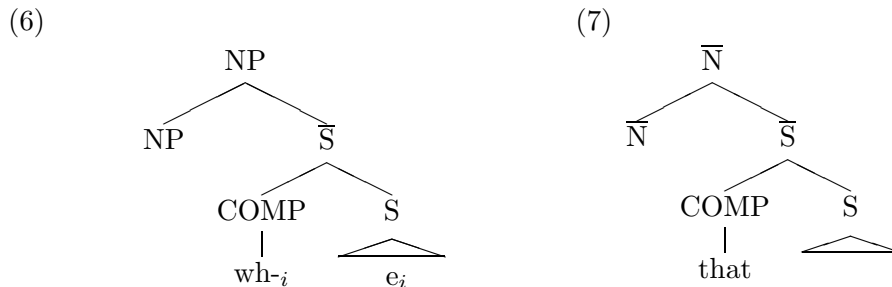


Obviously, however, such a derivation would be illicit since the initial tree here has two *wh*- phrases in COMP, which English sentences never allow. Since the constraints on the “derivation” of linked initial trees will have to rule out such configurations to correctly generate simplex sentence questions, the ungrammatical of *wh*- island violations follows directly.

Similar reasoning applies in the case of other subadjacency violations like the complex NP constraint in (4) and (5):

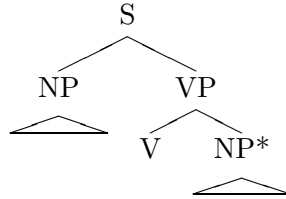
- (4) * Who_i did he reject the idea that the company might hire e_i?
 (5) * Who_i did she hit the dog that chased e_i?

We indicated in section 3 that relative clauses are treated in a TAG as auxiliary trees with the form of (6) and a similar structure, perhaps that in (7), will generate nominal complements:



The initial trees that these auxiliary trees would adjoin to would both have the form of (8):

(8)



NP* here indicates the noun phrase into which the adjunction of the relative clause or nominal complement auxiliary tree will occur, according to the structure of each tree. If the sentences to be generated did not contain questioned elements this adjunction would be possible and allow the generation of sentences like the following:

(9) He rejected the idea that the company might hire Bill.

(10) She hit the dog that chased Bill.

Since, however, the ungrammatical sentences under discussion do contain *wh*- question phrases and the gap positions linked to them are found inside the auxiliary trees, a TAG requires that the *wh*- phrases originate in the auxiliary trees. In the tree (6), there is no place to put the *wh*- phrase in question because the relative clause *wh*- phrase already fills the COMP position and so sentence (5) cannot be generated. In the tree (7), the COMP position of the nominal complement clause is available for a *wh*- phrase, but the adjunction of (7) to (8) will not produce a tree with the *wh*- phrase in the matrix COMP. Therefore, sentence (4) also cannot be generated.²⁶

6.1.2 There is an apparent problem with the treatment of subjacency violations given above; namely, it works too well. Consider, for example, the following sentences:

(11) a. Who did you get a notion to visit e?

b. Who did you get a notion that you should visit e?

(12) a. Which problem do you know how to solve?

b. Which problem do you know how you're going to solve e?

These sentences violate either the complex NP or the *wh*- island constraint, yet they are perfectly, or almost perfectly, acceptable. Because the TAG we have presented incorporates the constraints accurately, it wrongly predicts that the sentences should be out. Ross already noted the existence of complex NP constraint violations like (11) and suggested that they involved the reanalysis of the head of the nominal complement as part of the verb, thereby transforming the nominal complement clause into an object complement. That these cases involve such a reanalysis has been widely accepted since Ross proposed it and we merely wish to point out that if such reanalysis is allowed the resultant structure will admit of extraction out the complement clause in a TAG just as it would in any other. Accounting for the acceptability of the *wh*- island violations in (12) is slightly more difficult. In order

to do so we will have to make a small but interesting changes in the grammar we have developed. Up to now we have been assuming that the initial trees of our TAG are all simplex sentences; but all that the definition of a TAG requires is that the set of initial trees be finite. Suppose, therefore, that we allow sentences with one level of embedding as initial trees, so that a sentence (13) is treated as basic rather than as derived:

(13) You know how to solve that problem.

We could then apply the *wh*-linking rule to the direct object of the embedded verb in the structure of (13) and directly produce a structure for (11) and (12). Of course, if we allow an additional level of embedding in all sentences, we will also incorrectly generate true *wh*-island violations like (1), an unwelcome result about which the TAG formalism, like all others that we know of, has nothing to say.²⁷

There is more than one reason for thinking that the approach of adding a level of embedding to the initial trees may be on the right track, even though it leaves unanswered the important question of when such an added level is possible. First, there are languages in which extraction from *wh*-islands is in general acceptable and our analysis seems to account well for them. Thus, let us consider Italian, as described by Rizzi (1982). Rizzi points out that in Italian, it is possible to extract out of a single embedded question but not across two embedded questions. Thus, the relative clause in (14) [Rizzi's (6b)] is acceptable but the one in (15) [Rizzi's (15b)] not:

(14) Tuo fratello, [a cui]_i mi domando [che storie]_j
 Your brother, [to whom]_i I myself ask [which stories]_j (they)
 abbiano raccontato e_j e_i, era molto preoccupato.
 have told, e_j e_i, was very worried
 'Your brother, to whom I wondered which stories they told, was very worried.'

(15) *Questo argomento, [di cui]_i mi sto domandando [a chi]_j
 This topic, [on which]_i I myself am asking [to whom]_j I
 potrei chiedere e_j [quando]_k dovrò parlare e_i e_k, me sembra
 could ask e_j when_k I will have to speak e_i e_k, me seems
 sempre più complicato.
 always more complicated
 'This topic, which I am wondering who I could ask when I have to speak about it, seems to me ever more complicated.'

Rizzi points out that these facts can be accounted for by fixing \bar{S} as the bounding node for subjacency in Italian as opposed to English, where S is the relevant bounding node. For us it is of interest that exactly the same result is predicted by TAG which allows one level of embedding to be coded in the initial trees. The status of the following sentences, also discussed by Rizzi, is equally predicted by such a TAG:

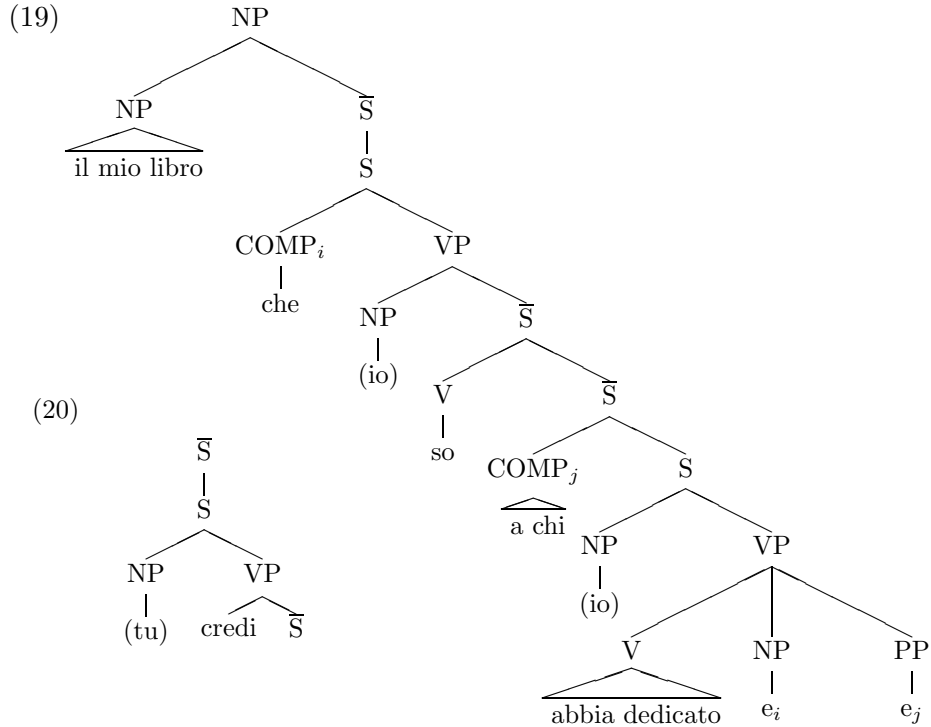
- (16) [=Rizzi's (18a)] Il mio primo libro, che_i credo che tu sappia [a
My first book, which_i I believe that you know [to
chi]_j ho dedicato e_i e_j, mi è sempre stato molto caro.
whom]_j I dedicated e_i e_j, me is always been very dear
- (17) * [=Rizzi's (18b)] Il mio primo libro, che_i so [a chi]_j
my first book, which_i I know [to whom]_j you
credi che abbia dedicato e_i e_j, mi è sempre stato molto caro.
believe that I have dedicated e_i e_j, me is always been very dear.

Again, this pattern follows from subadjacency if we assume \bar{S} to be the bounding node. In a TAG analysis, (16) will be grammatical so long as (14) is, because its structure is derived legitimately from the structure of (14) by the adjunction of an auxiliary tree above “tu sappia \bar{S} ” (i.e., “you know \bar{S} ”), just as in derivations like that in Example 3.3. Sentence (17), on the other hand, is not derivable, at least as long as we assume that syntactic adjunction is done in such a way as to preserve compositional semantics.

The case of (17) is exactly parallel to Example 3.4, in which sentence (18) was derived:

- (18) [_{S₃} John persuaded Bill [_{S₂} PRO to try [_{S₁} PRO to invite Mary]]]

In discussing this example, we pointed out that from a purely syntactic point of view it could be derived in two ways. The first, and intuitively correct, way would be to compose the initial tree S_1 with the auxiliary tree containing *try* to produce S_2 and then to compose S_2 with the auxiliary tree containing *persuade* to produce S_3 . The other way would be to compose S_1 with the *persuade* tree and then to compose the result with the *try* tree. We noted that this second derivation would be undesirable because it would complicate the statement of a compositional semantics for TAG's and we blocked it by imposing a null adjoining local constraint on the foot nodes of the auxiliary trees. The problem with the second derivation is that it would prevent our interpreting the sentence we were deriving in a step by step way as it was built up because at one stage of the derivation the proposition that Bill would invite Mary would be a complement to *persuade* and at a later stage it would be a complement to *try*. If we insist that the semantic interpretation of sentences be built up in tandem with syntactic adjunction and that partial semantic structures be indelible as to their predicate argument relations, surely an attractive requirement, then the second derivation will be ruled out without the imposition of an ad hoc mechanism.²⁸ Returning now to the Italian sentence (17), we see that the relative clause in it can only be derived by adjoining into a tree like (19) one like (20), given that (19), by our account of *wh*-island violations, must be an elementary tree:

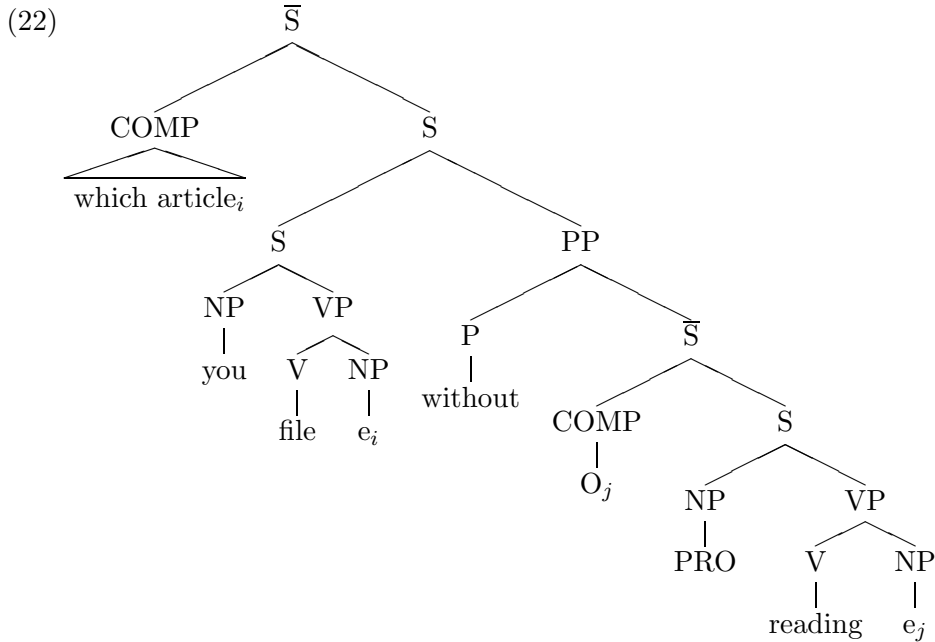


This derivation is excluded, however, because it violates the condition on semantic compositionality we have imposed on derivations.

6.1.3 Aside from the account of subadjacency violations allowed by our relaxation of the constraint that initial trees correspond to simplex sentences, there is a second reason for allowing this change. If we allow more complex initial trees, we can give an interesting account of parasitic gaps. Consider the following by now classic examples:

- (21) a. Which article did you file e without reading e ?
 b. He's a man who everyone who knows e feels sorry for e .

Assuming that the gerund *reading* heads a clause with a PRO subject, the parasitic gaps in both sentences are located in subordinate clauses. There is, therefore, no way for our grammar to relate them to the fronted *wh*- elements by a link unless the subordinate and main clauses are jointly present in initial trees. We could, of course, follow Contreras 1984 and propose that parasitic gaps are produced by the movement of a phonetically null empty operator to an A-bar position, as in the following tree we have drawn for example (21a):



Then some rule of interpretation will identify the indices i and j in the course of construction of a semantic representation for the sentence. While this may, in fact, be the correct analysis, it does have drawbacks, including the fact that it cannot account for such cases as (23), often assimilated in the literature to the cases in (21):

(23) Who did you give a picture of e to e ?

Let us, therefore, sketch out our alternative that follows the line we have been taking on subadjacency violations and that will explain a troublesome fact about parasitic gap sentences that is not always sufficiently appreciated. Theoretical discussion of parasitic gaps has emphasized the fact that they seem to occur in islands; that is, in environments where ordinary extraction is blocked. In fact, however, at least postverbal parasitic gaps seem most natural in those islands from which extraction marginally possible (Engdahl 1983). Thus, for example, (24) below is no less acceptable than (21a):

(24) Which article did you go to sleep without reading e ?

Notice that if we add another level of embedding in either of the sentences of (21) the sentence becomes much less acceptable:

- (25) a. *Which article did you file e without making sure that you had read e ?
 b. *He's a man who everyone who knows that the boss fired e feels sorry for e .

Allowing a single level of embedding in our initial trees and extending our use of linking to generate cases in which link daughters may have more than one link

parent,²⁹ produces a grammar in which the sentences in (25) will not be generated while those in (21) will be. Of course, the grammar will also have to contain the basic constraint governing the occurrence of parasitic gaps, namely that the parasitic gap and the primary gap cannot c-command one another (Engdahl 1983, Taraldsen 1979). This principle is easily stated as a condition on the well-formedness of initial trees but unlike subjacency, it does not follow from the TAG formalism itself.

6.2 The *that*-trace effect in a TAG

6.2.1 Let us now consider another extraction constraint, the well-known *that*-trace effect (Perlmutter 1971, Chomsky and Lasnik 1977, Pesetsky 1981), which cannot be subsumed under subjacency. It too can be incorporated into a TAG, although the principle governing it, a clause of the Empty Category Principle (ECP), must be stipulated, as it is in other treatments. The sentences below illustrate the effect in question:

- (26) Who_i did you say e_i was coming?
 (27) * Who_i did you say that e_i was coming?

In transformational accounts of this phenomenon, it is made to follow from the cyclic character of *wh*- movement (see, for example Pesetsky 1981). If *wh*- movement is cyclic and movement always leaves a trace, then the proper s-structure representation of sentences (26) and (27) must be as in (28) and (29) respectively:

- (28) [_S Who_i [_S did you say [_S e_i [_S e_i was coming]]]]]
 (29) [_S Who_i [_S did you say [_S e_i [_S that [_S e_i was coming]]]]]]

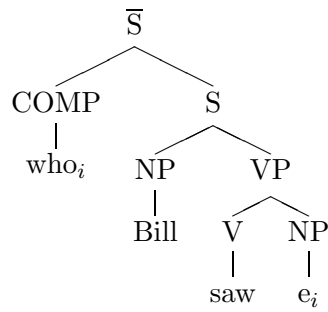
The difference between (28) and (29) is that the intermediate trace in the former but not in the latter properly governs the embedded subject trace. This difference allows the structure in (29) to be ruled out by the ECP, which is formulated to require that non-lexically governed empty categories like the subjects of tensed clauses be locally bound by a c-commanding antecedent. In a TAG grammar, however, this solution is not available. The cyclicity of *wh*- movement appears in the derivation of long-distance extractions through the repeated adjunction of matrix sentences, but no intermediate traces exist. Instead the ECP must have its effect either as a well-formedness condition on initial trees or by imposing local constraints on the auxiliary trees that adjoin to linked trees.

Of the two possible ways of instantiating the ECP, the more interesting alternative is the first. Consider the following simple example:

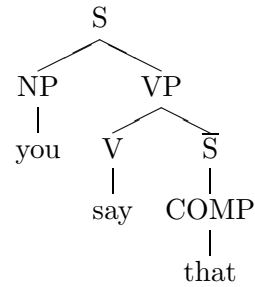
- (30) Who_i did you say that Bill saw e_i?

This sentence could be derived from the initial tree (31) and the auxiliary tree (32):

(31)

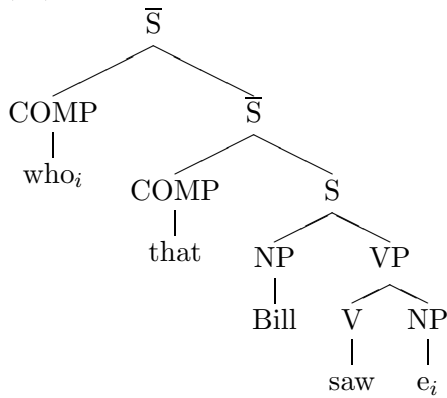


(32)

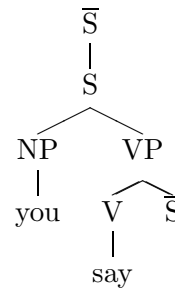


But there is another option for the derivation of the sentence. If we treat the complementizer as part of the initial, rather than of the auxiliary tree, the two trees will appear as follows:³⁰

(33)

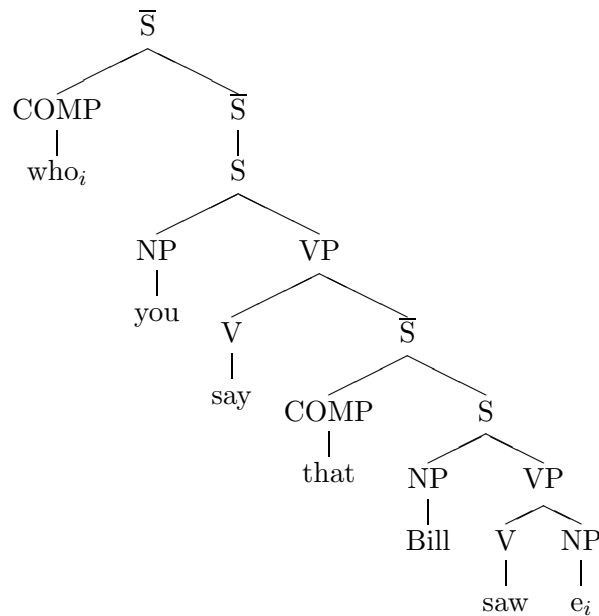


(34)



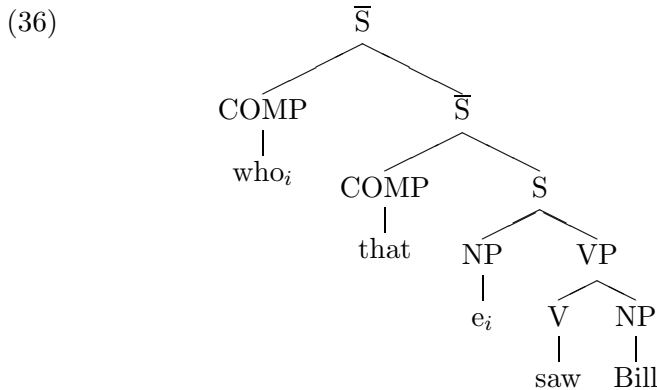
Performing the adjunction operation³¹, we obtain the following result:

(35)



This tree gives a reasonable structural representation of sentence (30), ignoring, as we have throughout, the complications introduced by subject-aux inversion. Its one peculiarity, the non-branching \bar{S} node that appears as a sister to *who*, would be rectified by pruning that node automatically.³²

Having thus redefined sentential embedding, let us examine the TAG derivation available for the ungrammatical (27) to see how it might be ruled out. The auxiliary tree will, of course, be (34), as in the example just presented, but the initial tree would have to look like (36):



Note that in this tree the empty category in subject position is separated from its antecedent *wh*- phrase by a complementizer *that*. If we formulate the clause of the ECP relating to government by a local c-commanding antecedent so as to exclude this configuration (see Aoun and Sportiche 1983 for a formulation which would accomplish this), we can state the ECP as a constraint on the well-formedness initial linked trees, as in (37):

- (37) Empty categories (link parents) must be properly governed in initial trees, where proper government is either government by a lexical head or local binding by the link daughter (cf. Chomsky 1981a).

Since trees like (36) will not be well-formed under (37), sentences exhibiting *that*-trace violations will simply not be generated. Of course, if the complementizer *that* is absent from the initial tree, it will be well-formed and the adjunction of a matrix predicate auxiliary tree will be possible, yielding a grammatical derivation sentences like (38), as we would wish:

- (38) Who_i did you say e_i saw Bill.

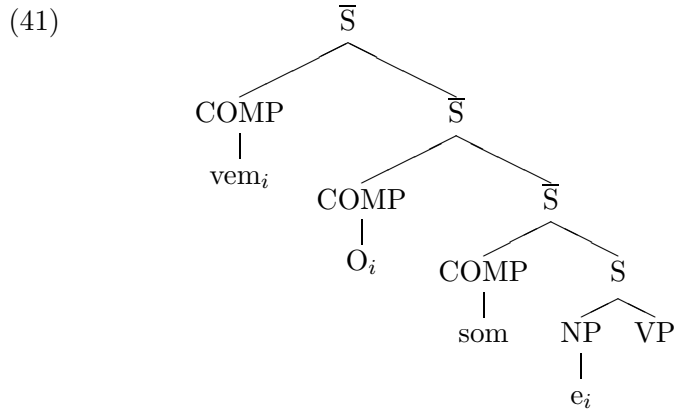
The analysis presented above is analogous to several proposed within the GB framework but here it is especially attractive formally because it allows the ECP to be stated on a single level and independently of the embedding mechanism. Therefore, as we have indicated, the locality of proper government is achieved without the postulation of intermediate traces.

6.2.2 The analysis of the *that*-trace effect we have proposed is made more interesting because of a certain empirical prediction that it makes. There are many languages which do not exhibit the effect. Some of these arguably involve post-position of the subject into a configuration from which it is lexically governed (or alternatively lexical government by the inflection node Chomsky 1981a); but others, for example Dutch, seem not to be amenable to this analysis. For such languages the TAG can generate *that*-trace violating sentences only by allowing initial trees with structures like (36), in which a lexically filled double COMP appears. What the exact structure of such a “doubly-filled” COMP should be is unclear, as is the nature of the variation across languages that putatively do and do not allow doubly-filled COMP’s. Nonetheless, it is of great interest that Koopman, in her recent analysis, is able to show that Dutch allows government of the subject position from a double COMP (Koopman 1983), just as the TAG formalism predicts.

Of course, the relationship between government of subject position from doubly-filled COMP’s and violations of the *that*-trace prohibition is more complex than our brief remarks will have suggested; but the complications should not vitiate our approach. Thus, there are languages which appear to countenance antecedent government from doubly-filled COMP’s but which do not allow violations of the *that*-trace prohibitions. One such case is standard Swedish, which, while prohibiting *that*-trace violations, appears to allow doubly-filled COMP’s in certain interrogative clauses. Direct questions never allow doubly-filled COMP’s; but indirect questions do allow them. In fact, when the subject of the indirect question undergoes *wh*-movement, the double COMP becomes obligatory. The following examples from Anward 1982 and Engdahl 1982 illustrate the point (see also Andersson 1975):

- (39) a. Jag vet inte vem_i *(som) e_i äger den.
 I know not who *(that) owns it.
- b. Han undrade vem_i (som) vi träffade e_i i stan.
 he asked who (that) we met in town
- (40) Vilken elev_i trodde ingen (*att) e_i skulle fuska?
 which student thought no one (*that) would cheat

As Engdahl points out, Swedish is exceptional among the Scandinavian languages in exhibiting the *that*-trace effect. The other languages of the group, including at least one dialect of Swedish itself, do not obey it; and, like Dutch, they give evidence of allowing government of subjects from doubly-filled COMP’s. The exceptional character of standard Swedish suggests to us that a reanalysis of the structure of COMP may have taken place in that language and that the apparently doubly-filled COMP may have been reanalyzed as a different syntactic configuration. There is some evidence, in fact, that Swedish questions with *wh*+*som* are actually cleft-like structures with the *wh*-phrase outside the sentence with which they are construed. In other words, Swedish questions like the embedded question in (39a) may actually have a structure like (41):



If this structure turns out to be correct, then Swedish, like English, must be said not to allow doubly-filled COMP's and, therefore, not to provide a counterexample to our analysis of the *that*-trace phenomenon. Instead a subject position empty category in Swedish indirect questions would have to be governed in the way that the subject position in English *that* relatives is. As is well known, it cannot be the empty operator in COMP (here represented as O_i) which serves as the governor because that operator is in the configuration which blocks government in the GB analysis that we have incorporated into our TAG account. Several solutions to the problem of how the subject position empty category is governed have been proposed in the literature; and we shall not attempt to choose among them for they all amount to stipulations and can all be stated, if not motivated, in the TAG formalism.

Although we are not in a position to demonstrate conclusively that reanalysis of the structure of embedded questions has taken place, we can present certain considerations that favor it.³³ First of all, both the literature (Lie 1982) and informants we have questioned state that the most natural way of forming a question in the Scandinavian languages, Swedish included, is as a cleft sentence, as in (42):

- (42) Vad var det (som) han köpte?
 what was it (that) he bought

Given that we know that cleft structures without the introductory expletive are used for questions in several languages (e.g., the Celtic languages (see McCloskey 1979) and vernacular Montreal French (see Lefebvre 1982)), it seems plausible that standard Swedish might have reanalyzed its embedded questions in this direction under the influence of the spoken vernacular. Matrix questions, which obey the verb second constraint that governs word order in root sentences, would be immune to the reanalysis, which would have demoted the interrogative clause to non-root status. In the second place, we find an interesting difference between questions and relative clauses in matter of the permissibility of *wh*-+*som*. While questions allow this configuration quite freely, relative clauses do not allow it at all (Anward 1982). If the configuration were the sign of a cleft structure, this difference would be expected since the cleft structure could not function as a relative clause. Finally, the reanalysis hypothesis explains why *som* is optional in non-subject questions

like (39b) but obligatory in subject questions like (39a). It seems that both the relativizer *som* and the *som* which appears in overt cleft sentences are obligatory when the subject of the embedded clause has a gap in subject position, just as a complementizer is obligatory in corresponding (standard) English cases like those of (43):

- (43) a. The box *(that) fell on me was heavy.
 b. It was the box *(that) fell on me.

Obviously, if Swedish questions are analyzed as cleft structures, the fact that *som* obligatory when the subject is questioned follows directly.

6.2.3 The non-existence of intermediate traces is a consequence of the TAG formalism so it is encouraging that we are able to give reasonable alternative accounts of the *that*-trace effect without appealing to them. In addition, however, the elimination of intermediate traces is attractive because it avoids certain well-known problems that their existence provokes so that the fact that they cannot appear in a TAG becomes an argument for the formalism. The problems we refer to are several. One is that, at least under some conceptions of bounding, one cannot simultaneously maintain subjacency and an S-bar deletion analysis of exceptional case-marking verbs like *believe*. The problem (see Chomsky 1981a) is that if S-bar deletion precedes *wh*-movement or if these verbs are allowed to exceptionally subcategorize for an S complement in the base, a sentence like (44) will falsely be predicted to be a subjacency violation, given that S is the bounding node for subjacency in English:

- (44) Who_i [_S does John believe [_S Bill to have seen e_i]]

On the other hand, if *wh*- movement precedes S-bar deletion, it will produce a configuration like (45), in which the presence of an *wh*- trace would presumably block S-bar deletion since \bar{S} would now be a branching node:

- (45) Who_i [_S does John believe [\bar{S} e_i [_S Bill to have seen e_i]]]

There are, of course, a number of ways of avoiding this problem, including (1) the analysis in Kayne 1981b, under which *believe*-type verbs subcategorize for a null complementizer and (2) redefining what counts as a bounding node for subjacency (Chomsky 1981b). In a TAG, however, the problem disappears entirely and the otherwise attractive option of stating the exceptional character of *believe* as feature of its subcategorization becomes available.

A second, and more important, difficulty raised by the existence of intermediate traces is how they are to be governed. Since they are empty categories, they must be properly governed; but their structural position is incompatible with their being either lexically governed like a complement to a verb or governed by a local antecedent like the subject of a tensed sentence. That they cannot in general be assumed to be antecedent-governed is shown by sentences like (46), in which the

intermediate trace e^3_i , which would be the antecedent governor of the intermediate trace e^2_i , is in a doubly-filled COMP and so in a configuration which blocks antecedent government:

- (46) Who_{*i*} did John say [_S e^3_i that Bill thought [_S e^2_i that Sally would invite e^1_i]]

On the other hand, intermediate traces seem not to be lexically governed since the only available lexical governor is the verb that c-commands them and they are separated from this verb by a maximal projection boundary. Although government across the S-bar boundary has been proposed to handle this case (Kayne 1981a), it seems a clear advantage not to have to weaken the definition of government in this way.³⁴

A final problem introduced by intermediate traces is the well-known case of *wanna* contraction (Postal and Pullum 1978, Jaeggli 1980). The basic fact is that *want to* can contract to *wanna* in casual speech when the subject of the complement sentence is PRO but not when it is an empty category bound by *wh*-. The contrast given in (47):

- (47) a. I want/wanna PRO to see the circus.
 b. Who_{*i*} do you want/*wanna e_i to visit the circus?

Given current assumptions about rule ordering, these facts argue for a trace-theoretic account of *wh*- movement. As Postal and Pullum point out, however, the trace putatively left by raising (NP-movement) does not block contraction, as the sentences in (48) show:

- (48) a. John seemsta like pickles.
 b. I hafta leave now.

Postal and Pullum use this fact to argue against a trace-theoretic account of raising; and since in our analysis there are no NP-movement-type traces, the contraction pattern poses no problem for us. Those who wish to maintain the existence of NP traces have pointed out that NP traces are not the only ones that are transparent to contraction (Jaeggli 1980). In particular, intermediate *wh*- traces in COMP do not block contraction, as the following example shows:

- (49) a. Who_{*i*} do you want [_S e_i [_S PRO to visit e_i]]
 b. Who do you wanna visit?

On this basis Jaeggli argues that intermediate traces, like NP traces, are not case marked. This solution seems to us unattractive. Not only are there arguments for intermediate traces that depend on them being case marked in at least one language (Hungarian, see footnote 33), but intermediate traces are part of a chain along which case and other syntactic features are said to be propagated. In the TAG analysis the apparent transparency of intermediate *wh*- traces and raising traces to phonological processes follows from their non-existence.

6.3 A further example

Subjacency and the *that*-trace effect are not the only *wh*- movement phenomena that the TAG formalism accounts for in an interesting way. As a further illustration, we shall here consider the case of stylistic inversion in French as discussed in Kayne and Pollock 1978. K&P point out that aside from occurring in main clause questions, stylistic inversion occurs in embedded *wh*- clauses, both questions and relatives, but not in embedded declaratives or in questions with a *wh*- in situ. Thus the examples in (50) are grammatical but the ones in (51) are not:³⁵

- (50) a. Je me demande quand partira ton ami.
I me ask when will-leave your friend
- b. La maison où habite cet homme est tres jolie.
the house where lives this man is very pretty
- (51) a. *Marie pense qu'a crié Pierre.
Marie thinks that has shouted Pierre
- b. *Partira ton ami quand?
will-leave your friend when

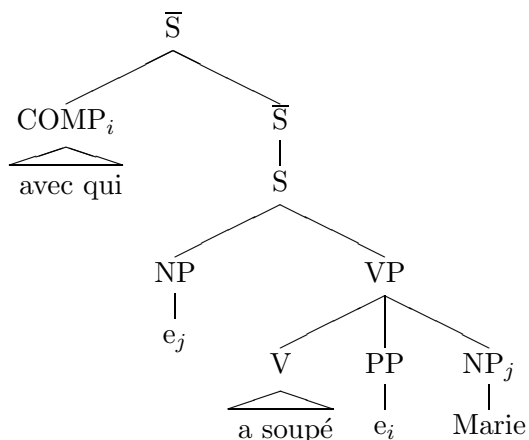
These facts and others cited by K&P show that the grammar of French must contain a constraint that makes the appearance of an inverted subject dependent on appearance of a *wh*- in the COMP of the same simplex sentence.

The phenomenon of stylistic inversion becomes more interesting theoretically when one notices the fact that inversion can take place not only within simplex sentences introduced by *wh*- COMP's but also in more complex cases. Consider, for instance, the following example:

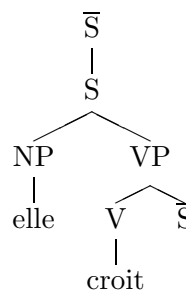
- (52) Avec qui croit-elle qu'a soupé Marie?
with whom believes she that has dined Marie

As K&P point out, this sentence is most easily accounted for by assuming that at some stage of the derivation the *wh*- phrase is in the COMP of the lower sentence. In a transformational grammar this is accomplished because *wh*- movement applies cyclically;³⁶ but the same analysis is naturally statable in a TAG because the initial trees from which (52) derives must be (53) and (54), abstracting away from interrogative inversion of verb and pronominal subject, a different syntactic process:

(53)



(54)



Clearly, the constraint on stylistic inversion, however formulated, will apply to (53) as to any simplex sentence; and the adjunction of (54) to it just has no effect. Thus, we see again that a TAG directly reproduces the effect of cyclic movement *wh*-without transformations.

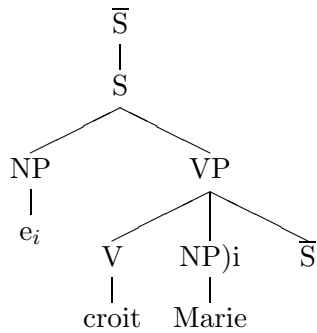
Now let us consider the following sentences, the most interesting cases discussed by K&P:

- (55) a. ? Comment sait Marie que Luc est mort?
 how knows Marie that Luc is dead
- b. ? Avec qui a prétendu Marie que sortirait Jean?
 with whom has claimed M that would-leave J

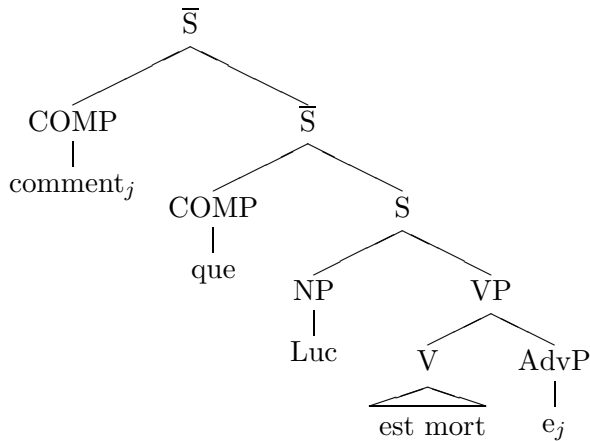
(56) * Sait Marie comment Luc est mort?³⁷

The sentences in (55) are of strained acceptability; but we shall assume, along with others who have considered this phenomenon, that the reasons for their unnaturalness are independent of the points at issue here. In these sentences inversion can occur not only in the clause in which the *wh*- appears but also in the clause in which it ends up. If, in a transformational account, the application of inversion is interspersed with *wh*- movement, the pattern of facts described by K&P is predicted.³⁸ In our TAG account, the acceptability of (55) and the unacceptability of (56) require that a local constraint be imposed on the adjunction of an inverted auxiliary tree to an initial tree. If the initial tree contains a fronted *wh*-, the inverted auxiliary tree will be adjoinable to the right of the *wh*- COMP. Otherwise, it will not be. Thus, sentence (55a) will be derived from the auxiliary and initial trees below; but sentence (56) be undervivable because tree (57) is blocked by a local constraint from adjoining to a position where it will not be in the domain of a *wh*- element:

(57)



(58)



At this point the reader will have noticed that we have stated the constraint on the appearance of stylistic inversion in two places in our grammar. It appears first as a well-formedness constraint on inverted initial trees and then again as a local constraint on the adjunction of inverted auxiliary trees. Clearly, this duplication is undesirable, and it can be eliminated. To do so, however, requires that we provide a general way of integrating local constraints on adjunction with well-formedness constraints on initial trees. Although further research will be necessary before we can give a definitive formulation of this integration, the general lines of a solution are clear. From a purely formal point of view, there is no need in a TAG to specify any relationship between auxiliary and initial trees. They simply belong to two different finite sets, the sources of which are beyond the scope of a TAG's formal definitions. From a linguistic point of view, on the other hand, the failure to specify how these sets are constituted and how they relate to one another would be unacceptable, since many linguistic generalizations can and should be stated as constraints on all possible elementary trees. Indeed, well-formedness constraints will generally apply equally to both types of elementary tree, initial and auxiliary, since the two types must equally exhibit the properties of well-formed phrase structure. Ordinarily, the well-formedness constraints can be stated in such a way as to apply jointly to initial and auxiliary trees without difficulty; but sometimes it will happen that the configurational information available in an initial tree will not be present in the corresponding auxiliary tree. This circumstance is illustrated by the case of French stylistic inversion. The inverted auxiliary tree (57) does not contain the *wh*-element necessary to license its structure and instead the tree is associated with a local constraint limiting its appearance to contexts which supply the absent *wh*.

The correspondence between an absence of context within an elementary tree and the need for a local constraint that we find in this case can, we believe, be stated as general convention on well-formedness conditions which we can express in the following terms. Suppose that certain elementary trees are thought of as being derived from others by some sort of structure altering, pseudo-transformational,

operation, as we suggested in section 2.3. In the case of trees containing elements this operation will be a Chomsky-adjunction to COMP that generates a link, or equivalently, a coindexed empty category. In the case of stylistic inversion the operation can be seen as moving a subject NP into VP and leaving a coindexed empty category in subject position. These pseudo-transformations can be defined so that they occur freely whenever the minimal conditions necessary for their application are met; that is to say, without regard to the presence of contextual material ‘unaffected’ by the rule (in the sense of Bresnan 1976). The effects of context on pseudo-transformations can then be introduced as well-formedness constraints on the trees output by them. If the context necessary to license a certain operation is present in the elementary tree derived by applying that operation, then that tree immediately meets the well-formedness constraint associated with it by virtue of having undergone the operation. If, on the other hand, the tree does not contain the necessary context, the well-formedness constraint is not satisfied; and for this case we propose that the unsatisfied well-formedness constraint be interpreted, by general convention,³⁹ as a local constraint on adjoining that tree to any other. This derived local constraint will simply require that the original well-formedness constraint be met by the structure that results from the adjunction. If there is no tree to which the derived tree can be adjoined so as to meet this constraint, it will simply never appear in any derivation. In the case of stylistic inversion it is immediately obvious that this convention will produce the desired results. In fact, within the transformational paradigm, K&P present substantial arguments that inversion should apply without reference to context and that the well-formedness of the output of the operation should be controlled by a filter. Our proposal is clearly very similar, if not equivalent, to this.

7 Conclusions

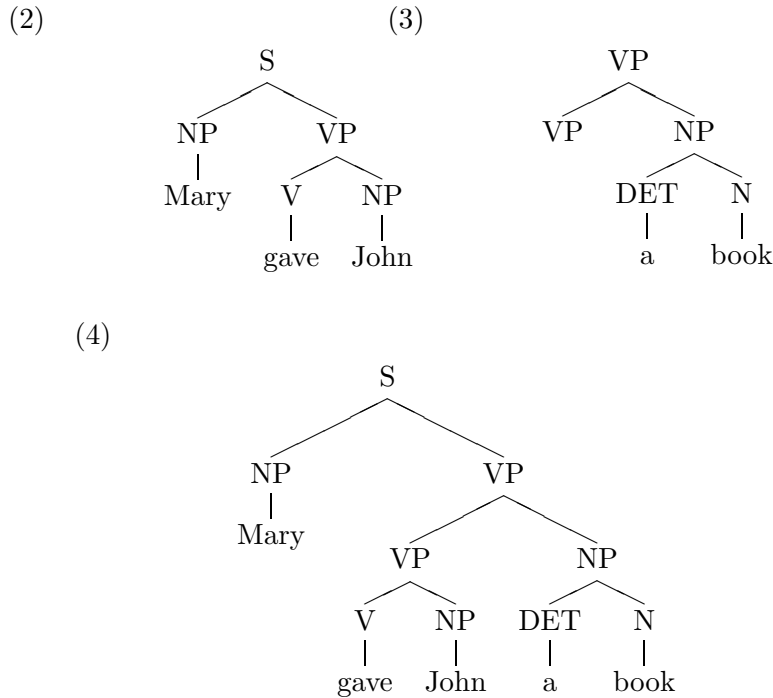
In the course of the preceding discussion we hope to have demonstrated two main points. The first of these is that the TAG formalism’s partitioning of the device responsible for recursion from the mechanism for generating the elementary clausal structures of language is linguistically attractive. All of the analyses presented in this paper rely crucially on this feature of TAG’s so that its significance should be clear to the reader. Nonetheless, it is perhaps worth re-emphasizing two points. First, this partitioning of recursion from co-occurrence restrictions is directly responsible for increased generative capacity of TAG over context-free grammar. It is therefore, noteworthy that the same device that enables us to implement transformational-style analyses in a TAG also provides the power needed to accommodate those linguistic constructions which require more than context-free generative capacity. Secondly, this partitioning makes possible an extremely constrained theory of phrase structure, which would only have to specify the form of elementary trees. We have not developed such a theory in this paper because it can be formulated to be essentially independent of the TAG formalism itself. Moreover, since it will not have to incorporate recursion, it will be a finite device stateable in many formally equivalent

ways.

One interesting question that will have to be resolved in the future is what the proper boundaries are between the domain of the rules that generate elementary structures and the adjunctions that generate complex ones. To give one example, let us consider double-object sentences like (1):

- (1) Mary gave John a book.

It would be possible in a TAG to generate this sentence either as an initial tree or by composing the trees in (2) and (3) below to yield (4):



This analysis, however, is not linguistically desirable. As the analysis in Bach 1979 suggests, the auxiliary tree in (3) assigns the wrong constituency to the VP because it does not allow the direct object to be introduced into the derivation as the sister of its verb. To explain how both complement NP's in (1) can be direct objects at some stage of the derivation of the sentence, Bach, within a Montague framework, proposes a syntactic composition rule called "right wrap," which in essence composes the VP *gave a book* with the NP *John* to produce another VP. If Bach's analysis is correct, as we believe it to be, then we would want to implement it in our grammar, which we could do by constructing our elementary trees using a modified categorial grammar (see Jacobson 1985 for one approach). The point of most importance is that in such an implementation the right wrap operation would be limited to entirely local contexts and would not apply to unbounded contexts like raising. The still unanswered question, however, is precisely what considerations determine which

linguistic phenomena should be treated as tree adjunctions and which by the theory governing the well-formedness of elementary trees.

Our second concluding point is that the Tree Adjoining Grammar formalism permits the implementation of analyses that are “transformational” in their linguistic spirit within a mathematically explicit and constrained formalism. Although the linguistic significance of constraining generative capacity is controversial, there can be no doubt that it is interesting to discover the linguistic analyses favored by contemporary transformational grammar do not require a formalism of extravagant generative capacity. We have in this paper succeeded in providing analyses of the NP-movement and *wh*-movement phenomena central to contemporary syntax which preserve the central features of transformational, including current Government Binding theory, analyses. They give a configurational account of grammatical relations rather than treating them as primitive or as semantic, and they treat the boundedness of long-distance movement through constraints on derivations rather than through constraints on representation. Of course, there are significant differences between the analyses that we have proposed and those that are at present current in GB. To the extent that these differences reflect the constrained formal power of TAG, they should become the focus of future attempts to evaluate the success of our enterprise. Indeed, if such theoretical entities as NP trace and intermediate *wh*-trace, which are difficult or impossible to implement in a TAG, turn out not to be needed for linguistic analysis, as more than one linguist has claimed or suspected, then the claims of the TAG formalism to linguistic relevance will be strong indeed.

Let us now end with a remark on the future of our investigations into the linguistic relevance of TAG. If the TAG formalism is to be used as the formal notation of linguistic theory, it will, of course, have to show its capacity to allow the expression of linguistically motivated analyses of a much wider range of constructions in a much wider range of languages than we have been able to consider here. What these constructions are is fairly clear, and we have begun to work out TAG analyses for some of them which we hope to present in future work. Especially important topics for future work include rightward movement phenomena (e.g., extraposition), Dutch and German verb raising, and the extension of the formalism to so-called “non-configurational” languages. We are hopeful in each of these cases that the TAG formalism will turn out to constrain the set of possible analyses in interesting ways. Thus, in the case of extraposition, we expect to be able to provide an analysis from which the strict boundedness of rightward movement follows naturally. In the case of verb raising, we know that the TAG formalism will generate the cross-serial dependencies required by Dutch, and we hope to show that the structural descriptions generated by the TAG, which involve true verb raising, can be motivated linguistically.⁴⁰ Finally, the problem of non-configurationality can be addressed in an interesting way with the TAG formalism because its additional power beyond context-free grammar allows the generation of more word order variations than are possible within CFG. In this domain we hope to pursue the effects of implementing in the TAG formalism the partitioning of immediate domination from linear precedence

relations proposed in Pullum 1982.

8 References

Anderson, S. and P. Kiparsky, eds. 1973. *FESTSCHRIFT FOR MORRIS HALLE*. New York: Holt, Rinehart and Winston.

Andersson, L.-G. 1975. 1975; *FORM AND FUNCTIONING OF SUBORDINATE CLAUSES IN SWEDISH*. University of Goteborg dissertation.

Anward, J. 1982. Basic Swedish. In E. Engdahl and E. Ejerhed, eds. *READINGS ON UNBOUNDED DEPENDENCES IN SCANDINAVIAN LANGUAGES*. Stockholm: Almqvist and Wiksell International.

Aoun, J. and D. Sportiche. 1983. On the formal theory of government. *THE LINGUISTIC REVIEW* 2:3 pp. 211-137.

Bach, E. 1979. Control in Montague grammar. *LINGUISTIC INQUIRY* 10:4 pp. 515-531.

Bach, E. 1980. In defense of passive. *LINGUISTICS AND PHILOSOPHY* 3:3 pp. 297-341.

Baltin, M. 1978. *TOWARD A THEORY OF MOVEMENT RULES*. unpublished MIT dissertation.

Baltin, M. 1982. A landing site theory of movement rules. *LINGUISTIC INQUIRY* 13:1 pp. 1-38.

Baltin, M. 1983. The proper notion of proper government. unpublished paper NYU.

Baltin, M. 1984. Review article on J. Bresnan, ed. *THE MENTAL REPRESENTATION OF GRAMMATICAL RELATIONS*. to appear in *LANGUAGE*.

Belletti, A., L. Brandi, and L. Rizzi eds. 1979. *THE THEORY OF MARKEDNESS IN GENERATIVE GRAMMAR*. Proceedings of the 1979 Glow conference. Pisa Scuola Normale Superiore.

Berwick, R. and A. Weinberg. 1984. *THE GRAMMATICAL BASIS OF LINGUISTIC PERFORMANCE*. Cambridge: MIT Press.

Bresnan, J. 1978. On the form and functioning of transformations. *LINGUISTIC INQUIRY* 7:1 pp. 3-40.

Bresnan, J. ed. 1982. *THE MENTAL REPRESENTATION OF GRAMMATICAL RELATIONS*. Cambridge: MIT Press.

Bresnan, J. 1982. The passive in lexical theory. In J. Bresnan ed. *THE MENTAL REPRESENTATION OF GRAMMATICAL RELATIONS*. Cambridge: MIT Press.

Bresnan, J. and J. Grimshaw. 1978. The syntax of free relatives in English. *LINGUISTIC INQUIRY* 9:3 pp. 331-301.

- Bresnan, J., R. Kaplan, S. Peters, and A. Zaenen. 1982. Cross-serial dependencies in Dutch. *LINGUISTIC INQUIRY* 13:4 pp. 613-636.
- Burzio, L. 1981. *INTRANSITIVE VERBS AND ITALIAN AUXILIARIES*. unpublished MIT dissertation.
- Chomsky, N. 1965. *ASPECTS OF THE THEORY OF SYNTAX*. Cambridge, MA: MIT Press.
- Chomsky, N. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum. *READINGS IN ENGLISH TRANSFORMATIONAL GRAMMAR*. Waltham, MA: Blaisdell.
- Chomsky, N. 1973. Conditions on transformations. In S. Anderson and P. Kiparsky, eds. *FESTSCHRIFT FOR MORRIS HALLE*. New York: Holt, Rinehart and Winston.
- Chomsky, N. 1977. On *wh*-movement. In P. Cullicover, T. Wasow, and A. Akmajian. *FORMAL SYNTAX*. New York: Academic Press.
- Chomsky, N. 1980. *RULES AND REPRESENTATIONS*. New York: Columbia University Press.
- Chomsky, N. 1981a. *LECTURES ON GOVERNMENT AND BINDING*. Cinnaminson, NJ: Foris.
- Chomsky, N. 1981b. Principles and parameters in linguistic theory. In N. Hornstein and D. Lightfoot, eds. *EXPLANATION IN LINGUISTICS*. London: Longmans.
- Chomsky, N. 1984. *KNOWLEDGE OF LANGUAGE: ITS NATURE, ORIGINS AND USE*. unpublished manuscript MIT.
- Chomsky, N. and H. Lasnik 1977. Filters and control. *LINGUISTIC INQUIRY* 8:3 pp.425-504.
- Contreras, H. 1984. A note on parasitic gaps. *LINGUISTIC INQUIRY* 15:4 pp. 698-701.
- Cullicover, P., T. Wasow, and A. Akmajian. 1977. *FORMAL SYNTAX*. New York: Academic Press.
- Dowty, D., L. Karttunen, and A. Zwicky, eds. to appear. *NATURAL LANGUAGE PROCESSING: PSYCHOLINGUISTIC, COMPUTATIONAL AND THEORETICAL PERSPECTIVES*. New York: Cambridge University Press.
- Engdahl, E. 1983. Parasitic gaps. *LINGUISTICS AND PHILOSOPHY*. 6:1 pp. 5-34.
- Engdahl, E. 1982. Restrictions on unbounded dependencies in Swedish. In E. Engdahl and E. Ejerhed, eds. *READINGS ON UNBOUNDED DEPENDENCES IN SCANDINAVIAN LANGUAGES*. Stockholm: Almqvist and Wiksell International.
- Engdahl, E. and E. Ejerhed, eds. 1982. *READINGS ON UNBOUNDED DEPENDENCES IN SCANDINAVIAN LANGUAGES*. Stockholm: Almqvist and Wiksell International.

- Freidin, R. 1978. On cyclicity and the theory of grammar. *LINGUISTIC INQUIRY* 9:4 pp. 519-550.
- Freidin, R. 1983. X-bar theory and the analysis of English infinitivals. *LINGUISTIC INQUIRY* 14:4 pp. 713-722
- Gazdar, G. 1982. Phrase structure grammar. In P. Jacobson and G. Pullum, eds. *THE NATURE OF SYNTACTIC REPRESENTATION*. Boston: D. Reidel.
- Gazdar, G., Klein, J.M.E., Pullum, G.K., and Sag, I.A. 1985. *GENERALIZED PHRASE STRUCTURE GRAMMAR*. Oxford: Blackwell.
- Groos, A. and H. van Riemsdijk. 1979. Matching effects in free relatives. In A. Belletti, L. Brandi, and L. Rizzi eds. *THE THEORY OF MARKEDNESS IN GENERATIVE GRAMMAR*. Proceedings of the 1979 Glow conference. Pisa: Scuola Normale Superiore.
- Gruber, J. 1965. *STUDIES IN LEXICAL RELATIONS*. unpublished MIT dissertation.
- Heny, F., ed. 1981. *BINDING AND FILTERING*. Cambridge: MIT Press.
- Higginbotham, J. 1984. English is not a context-free language. *LINGUISTIC INQUIRY* 15:2 pp. 225-235.
- Hoekstra, T., H. van der Hulst, and M. Moortgat, eds. 1980. *LEXICAL GRAMMAR*. Cinnaminson NJ: Foris.
- Hornstein, N. and D. Lightfoot, eds. 1981. *EXPLANATION IN LINGUISTICS*. London: Longmans.
- Jackendoff 1972. *SEMANTIC INTERPRETATION IN GENERATIVE GRAMMAR*. Cambridge: MIT Press.
- Jacobs, R. and P. Rosenbaum. 1970. *READINGS IN ENGLISH TRANSFORMATIONAL GRAMMAR*. Waltham, MA Blaisdell.
- Jacobson, P. 1985. Grammatical relations and phrase structure. unpublished paper Brown University.
- Jacobson, P. and G. Pullum, eds. 1982. *THE NATURE OF SYNTACTIC REPRESENTATION*. Boston: D. Reidel.
- Jaeggli, O. 1980. Remarks on *to* contraction. *LINGUISTIC INQUIRY* 11:1 pp. 239-247.
- Joshi, A. 1983. 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*. New York: Cambridge University Press.
- Joshi, A., L. Levy, and M. Takahashi. 1975. Tree adjunct grammars. *JOURNAL OF THE COMPUTER AND SYSTEM SCIENCES* 10:1 pp. 136-163.

- Joshi, A. and L. Levy. 1982. Phrase structure trees bear more fruit than you would have thought. *AMERICAN JOURNAL OF COMPUTATIONAL LINGUISTICS* 6:2 pp.272-284.
- Kayne, R. 1980. Extensions of binding and case marking. *LINGUISTIC INQUIRY* 11:1 pp. 75-9B6.
- Kayne, R. 1981a. ECP extensions. *LINGUISTIC INQUIRY* 12:1 pp. 93-133.
- Kayne, R. 1981b. On certain differences between French and English. *LINGUISTIC INQUIRY*. 12:3 pp.349-371.
- Kayne, R. and J.-Y. Pollock. 1978. Stylistic inversion, successive cyclicity, and move NP in French. *LI* 9:4 pp.595-621.
- Keyser, S. ed. 1978. *RECENT TRANSFORMATIONAL STUDIES IN EUROPEAN LANGUAGES*. Linguistic Inquiry Monograph No. 3.
- Klein, J.M.E. and Sag, I.A. 1985. Type-driven translation. *LINGUISTICS AND PHILOSOPHY* 8:2 pp 163-202.
- Koopman, H. 1983. *THE SYNTAX OF VERBS*. Cinnaminson, NJ: Foris.
- Koster, J. 1978. Why subject sentences don't exist. In S. Keyser, ed. *RECENT TRANSFORMATIONAL STUDIES IN EUROPEAN LANGUAGES*. Linguistic Inquiry Monograph No. 3.
- Koster, J. and R. May. 1982. On the constituency of infinitives. *LANGUAGE* 58:1 pp.117-143.
- Lasnik, H. and Fiengo, R. 1974 Complement object deletion. *LINGUISTIC INQUIRY* 5:4 pp.535-571.
- Lefebvre, C. 1982. *LA SYNTAXE COMPAREE DU FRANCAIS STANDARD ET POPULAIRE: APPROCHES FORMELLE ET FONCTIONELLE*. Gouvernement du Québec: Office de la langue française.
- Lie, S. 1982. Discontinuous questions and subjacency in Norwegian. In E. Engdahl and E. Ejerhed, eds. *READINGS ON UNBOUNDED DEPENDENCIES IN SCANDINAVIAN LANGUAGES*. Stockholm: Almqvist and Wiksell International.
- Lightfoot, D. 1976. Trace theory and twice moved NPs. *LINGUISTIC INQUIRY* 7:4 pp. 55-582.
- Lightfoot, D. 1980. The History of NP Movement. In T. Hoekstra, H. van der Hulst, and M. Moortgat, eds. *LEXICAL GRAMMAR*. Cinnaminson NJ: Foris.
- Lightfoot, D. and N. Hornstein. 1984. Movement in NP's: some consequences. Paper presented at the XVth annual meeting of the North East Linguistics Society. Brown University.
- Maling, J. and A. Zaenen. 1981. Germanic word order and the format of surface filters. In F. Heny, ed. *BINDING AND FILTERING*. Cambridge: MIT Press.
- McCloskey, J. 1979. *TRANSFORMATIONAL SYNTAX AND MODEL THEORETIC SEMANTICS: A CASE STUDY IN MODERN IRISH*. Boston: D. Reidel.

- Osherson, D., M. Stob, and S. Weinstein. 1983. Learning theory and natural language. *COGNITION* 13:2 pp. 1-25.
- Perlmutter, D. 1971. *DEEP AND SURFACE STRUCTURE CONSTRAINTS IN SYNTAX*. New York: Holt, Rinehart and Winston.
- Perlmutter, D. and P. Postal. 1984. Impersonal passives and some relational laws. In D. Perlmutter and C. Rosen, eds. *STUDIES IN RELATIONAL GRAMMAR 2*. Chicago: University of Chicago Press.
- Perlmutter, D. and C. Rosen, eds. 1984. *STUDIES IN RELATIONAL GRAMMAR 2*. Chicago: University of Chicago Press.
- Pesetsky, D. 1981. Complementizer-trace phenomena and the Nominative Island Condition. *THE LINGUISTIC REVIEW* 1:3 pp. 297-344.
- Peters, S. and R. Ritchie. 1982. Phrase linking grammars. unpublished paper. U. of Texas.
- Pollard, C. 1984. *HEAD GRAMMARS*. unpublished Stanford University dissertation.
- Postal, P. and G. Pullum. 1978. Traces and the description of English complementizer contraction. *LINGUISTIC INQUIRY* 9:1 1-29.
- Postal, P. and Langendoen, D. 1984. English and the class of context free languages. *COMPUTATIONAL LINGUISTICS* 10:4 pp. 177-181.
- Pullum, G. 1982. Free word order and phrase structure rules. In J. Pustejovsky and P. Sells, eds. *PROCEEDINGS OF THE 12TH ANNUAL MEETING OF THE NORTH EAST LINGUISTICS SOCIETY*. Amherst: Graduate Linguistics Students Association of the University of Massachusetts.
- Pullum, G. 1984. On two recent attempts to show that English is not a CFL. *COMPUTATIONAL LINGUISTICS* 10:4 pp. 182-186.
- Reinhart, T. 1981. Definite NP anaphora and c-command domains. *LINGUISTIC INQUIRY* 12:4 pp. 605-636.
- Rizzi, L. 1982. *ISSUES IN ITALIAN SYNTAX*. Cinnaminson NJ: Foris.
- Roach, K. 1984. Formal properties of Head Grammars. unpublished paper, Stanford University.
- Roeper, T. 1984. Implicit thematic roles in the lexicon and syntax. unpublished paper, University of Massachusetts.
- Rosenbaum, P. 1967. *THE GRAMMAR OF ENGLISH PREDICATE COMPLEMENT CONSTRUCTIONS*. Cambridge: MIT Press.
- Ross, J. 1967. *CONSTRAINTS ON VARIABLES IN SYNTAX*. unpublished MIT dissertation.
- Sag, I. 1982. A semantic theory of NP-movement dependencies. In P. Jacobson and G. Pullum, eds. *THE NATURE OF SYNTACTIC REPRESENTATION*. Boston: D. Reidel.

- Shieber, S. 1984. Evidence against the context-freeness of natural language. SRI International technical note no. 330. to appear in LINGUISTICS AND PHILOSOPHY.
- Taraldsen, T. 1979. The theoretical interpretation of a class of marked extractions In A. Belletti, L. Brandi, and L. Rizzi eds. THE THEORY OF MARKEDNESS IN GENERATIVE GRAMMAR. Proceedings of the 1979 Glow conference. Pisa Scuola Normale Superiore.
- Vijay Shankar, K. and Joshi, A.K. 1985. Some computationally significant properties of tree adjoining grammars. to appear in PROCEEDINGS OF THE 25TH ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS.
- Wasow, T. 1977. Transformations and the lexicon. In P. Cullicover, T. Wasow, and A. Akmajian. FORMAL SYNTAX. New York: Academic Press.
- Williams, E. 1980. Predication. LINGUISTIC INQUIRY 11:1 pp. 203-238.
- Williams, E. 1984. Grammatical relations. LINGUISTIC INQUIRY 15:4 pp. 639-674.
- Zaenen, A. 1983. On syntactic binding. LINGUISTIC INQUIRY 14:3 pp. 469-504.

9 Notes

1. This paper has benefited enormously from discussions the authors have had with many colleagues and students, and written comments received from others. We especially want to thank Mark Baltin, Bob Berwick, Dominique Estival, Gerald Gazdar, Richard Larson, Ellen Prince, Geoff Pullum, Vijay Shankar, Tom Wasow, and David Weir for their helpful comments. Of course, they do not necessarily agree with us in all of our claims.

2. Grammars G_1 and G_2 are weakly equivalent if the string language of G_1 , $L(G_1)$ = the string language of G_2 , $L(G_2)$. G_1 and G_2 are strongly equivalent if they are weakly equivalent and for each w in $L(G_1) = L(G_2)$, both G_1 and G_2 assign the same structural description to w . A grammar G is weakly adequate for a (string) language L , if $L(G) = L$. G is strongly adequate for L if $L(G) = L$ and for each w in L , G assigns an “appropriate” structural description to w . The notion of strong adequacy is undoubtedly not precise because it depends on the notion of appropriate structural descriptions.

3. Adjoining reduces to substitution only in the special case where an auxiliary tree adjoins to the root node of another tree so that it “sits on top of” the tree to which it is adjoined. In this special case, the adjoining operation has the same effect as would the substitution of a tree at its root node for the foot node of the auxiliary tree.

4. Note that under this treatment PRO occurs in a given position if and only if that position does not receive case. This is not quite true in GB because that

theory allows for governed positions that are not case marked. As the reader will see in later sections, our TAG analyses do not include any counterpart to NP trace and so they do not countenance governed but non-case-marked positions.

5. This statement is not quite true. In the colloquial language “weather” *it* does appear as the subject of control predicates:

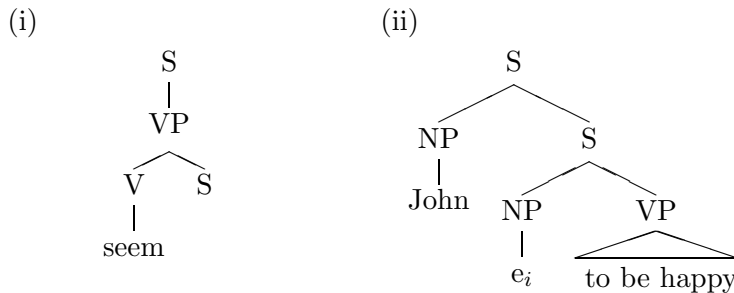
- (i) It tried to rain yesterday, but it couldn’t.

This suggests that “weather” *it* may, be a generalized indefinite subject rather than a true expletive, at least in some contexts.

6. The theory of obligatory control will become a systematic account of these translation rules, which will have to provide both for the general form of these rules and for lexically governed deviations. The single most crucial fact is, of course, that infinitival complements with PRO subjects generally exhibit object control when embedded under matrix predicates with direct objects and subject control elsewhere. The lexical exceptions to this generalization, e.g., *promise*, show that the theory of control is probably lexical in character.

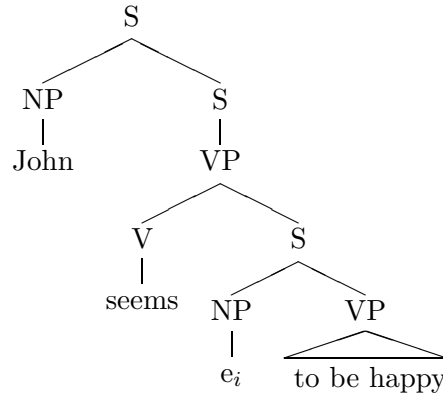
7. The analysis of control predicates in LFG assigns them the constituent structure of (10), but in functional structure the embedded verb is said to have a subject controlled by the matrix subject. Thus, the LFG analysis appears intermediate between the GB and GPSG analyses.

8. As we mentioned in our introduction, it is also possible to give a structure for raising predicates under which they take sentential subjects with empty subjects (NP traces). Under such an analysis sentence (1a) would be derived through the adjunction of an auxiliary tree like (i) to a linked initial tree like (ii):



The resultant tree would appear as in (iii) and a pruning convention would remove the non-branching S node:

(iii)



We shall not pursue this analysis here because it seems to us implausible to derive subject raising sentences from topicalized initial structures like (ii). The analysis may be worth further attention but it is at present unclear what the empirical consequences of choosing this analysis over the one adopted in the text would be. For instance, the argument to be given in section (4.2), which accounts for the non-existence of raising nominals, carries over directly to the analysis sketched in this note.

9. Kayne 1981 tries to link this difference to the fact that nouns do not allow exceptional case marking. But so long as GB theory requires proper government and government for case marking to be distinct, this solution will not work. Since intransitive verbs like *seem* are said to properly govern positions that they do not mark for case, there is nothing to keep nouns from doing the same.

10. The difficulties facing the current standard GB account of the non-existence of raising nominals may go beyond its ad hoc character. Thus, in at least one recent discussion (Lightfoot and Hornstein 1984), it is argued that nouns should be allowed to govern empty categories. If this proposal were adopted, it would undermine the standard account entirely. The substitute explanation offered by Lightfoot and Hornstein is to say that the trace in a structure like (21) has no antecedent because the genitive NP does not really c-command it. This line appears to us similar in spirit to the one we offer.

11. The issues raised by (31) also arise in the analysis of other sentences exhibiting preposing of infinitive complements, such as (i) below:

- (i) Likely to win though John may be, he's not counting his chickens.

The considerations we bring to bear on (31) in this text apply equally to (i).

12. Or, perhaps at logical form, a detail that does not affect the point we are making here.

13. We thank Ken Safir for pointing this difficulty out to us. There are various ways to avoid it but we know of none that is completely satisfactory. For instance, one might follow Reinhart 1981 and claim that the definition of c-command should

be modified so that the matrix subject in (32) will c-command the COMP. Then the trace will have a c-commanding antecedent. the difficulty here is that in a more complex sentence the *why* phrase may be even higher in the tree and so clearly outside the c-command domain of the subject. Sentence (i) below illustrates the point:

(i) How likely to win do you think that John could be?

14. We thank Rita Manzini for pointing out this sentence to us.

15. We thank Mark Baltin for pointing out to us the existence of this sentence type and implications of its existence for the status of raising.

16. ‘Pseudo-transformations’, storable as mappings from elementary trees to elementary trees, can be introduced into a TAG without affecting its mathematical characterization because they amount to redundancy rules. Mappings for which the output is not an elementary tree would, on the contrary, not be compatible with the formalism.

17. We should note here that the transformation of raising to object cannot be instantiated in a TAG in a natural way, because it is not a mapping from an elementary tree to an elementary tree.

18. Of course, we also cannot get a sentence like (i) because the infinitive subject would no longer be adjacent to the matrix verb and so, in recent theory, could not receive case:

(i) *I believe myself Mary to be the best choice.

19. One might propose that *believe* sentences have the structure in (i) and *persuade* sentences the structure in (ii):

(i) [I believe NP Mary] [VP to be the best choice]]

(ii) [I persuaded [NP Mary] [_S PRO to leave early]]

Although emphatic reflexive placement is easily accounted for if these structures are assumed, this solution has not been seriously considered in the literature. The reason that the solution must be ruled out is that once PRO is introduced as a possible subject for an apparently subjectless infinitive VP, the case for allowing infinitives without structural subjects becomes so tenuous that the structure in (i) is no longer well-motivated.

20. It would, of course, be unreasonable to have a grammar of English in which the relationship between raising passives and their active counterparts went entirely unexpressed. for one thing, the passive form is clearly derived from the active in a historical sense. All current passives that do not have active counterparts did have them in the period when the raising passive came into being in English. The difficulty, given the fact that the active and passive forms have somewhat drifted

apart, is deciding how and to what extent this historical relationship should be preserved in a synchronic grammar.

21. There is one obvious move in trying to construct a lexical analysis which will handle raising passives that will not work. One might propose that raising passives were adjectival like the passives in (i):

- (i) a. The door is closed (= “shut”).
- b. George is (very) interested in music.

If this proposal were tenable, then one could argue that raising passives, being adjectives, were related to active verbs by the same sort of category-changing lexical rules that relate the verbs *to close* and *to interest* to the adjectives *closed* and *interested*. Unfortunately for this approach, it is clear that the raising passive forms are verbs and not adjectives. For example, the bare participle cannot be embedded under *seem*, which Wasow (1977) shows to be a diagnostic of adjectival status:

- (ii) a. * John seems seen to have left.
- b. John seems to have been seen to have left.
- (iii) a. * John seems said to have left.
- b. John seems to have been said to have left.
- (iv) a. * John seems made to leave.
- b. John seems to have been made to leave.

22. This substitution can be stated formally as a specially constrained form of the adjunction operation basic to the TAG formalism. If the substitution operation is formalized in this way the constraints on lexical insertion can be stated as ‘local constraints’ in the sense of section 2. Since, however, the form in which the insertion operation is expressed seems not to have consequences for the theory, we shall not pursue the issue here.

23. The expletive subjects which appear in this position do so for syntactic reasons which we shall not discuss here. Suffice it to say that an expletive subject is precisely non-lexical in the relevant sense; that is, it does not receive a thematic role.

24. It will also be necessary to constrain adjunctions that occur after lexical insertion so that the resultant structures do not wind up violating the subcategorization requirements of phrasal heads. This can probably be accomplished best by having phrasal nodes inherit the subcategorization requirements of their heads as local constraints on adjoining. This or some equivalent mechanism will be required if the TAG solution to the non-existence of raising nominals is to work out technically. We thank Dominique Estival for pointing this out to us.

25. Since our TAG unifies the syntax of raising passives and subject-to-subject raising predicates, it implicitly predicts, like other analyses that achieve this unification, that a language with one will also have the other. This prediction is interestingly consistent with the history of English, which seems to have acquired these two constructions at the same period (see Lightfoot 1980)

26. Whether the analysis given here extends to sentential subjects, the third major case of subjacency, as it applies to *wh*-movement depends on the phrase structure assumed for sentences with sentential subjects. If we implement in our TAG either the classical *it*-S analysis of Rosenbaum 1967 or the topicalization analysis of Koster 1978, sentential subject extractions will, of necessity, be impossible. If, on the other hand, we assume the simplest phrase structure for the construction, additional constraints will be required to block the unwanted extraction. We thank Ellen Prince for pointing out the difficulties involved in accommodating this construction.

27. There are a number of structural factors that influence the acceptability of *wh*-island violations in English and perhaps they do so by affecting the plausibility of the analysis of indirect questions as initial trees. Among the contextual factors that favor acceptability are an embedded question that is infinitival rather than tensed, a complement subject that is coreferential with the matrix subject, and having the gap linked to the *wh*-phrase of the embedded question located inside the VP. In addition, it is sometimes but not always the case that extractions with crossing dependencies are less acceptable than those with nested ones.

28. The use of a null adjoining constraint on the foot nodes of auxiliary trees will in any case not work as a general solution. Consider, for example, the sentence in (i) below:

- (i) John believes Bill to be likely to win.

Step by step semantic interpretation requires that the *be likely* tree be adjoined infinitival initial tree before the *believes* tree is. Since *be likely* is a raising predicate, however, putting a null adjoining constraint on the foot node of the *believes* tree will not prevent the *be likely* tree from being adjoined at the VP node of the original infinitival structure at the wrong stage of the derivation.

29. Configurations in which a single link daughter has more than one link parent are allowed by the definition of linking we give in section 2.2.

30. We are assuming here that *wh*-movement is a Chomsky-adjunction to \bar{S} as in Baltin 1982.

31. The reader will have noted that the structure in (33) cannot be allowed to appear as is at the end of a derivation. This result is easily enough guaranteed if we place a local constraint on the lower \bar{S} of the tree which requires that an adjunction take place at this site. This constraint would be similar to the one required to keep an infinitival clause from surfacing as a root sentence. One yet to be investigated aspect of TAG is what the linguistic principles are that should define the set of

possible local constraints; but it is clear that a much more restricted set than those allowed by the mathematical definition of local constraints are actually appropriate for linguistic description. They seem to be needed mainly to accomplish some of the work done by filters in current transformational grammar.

32. Adopting this convention does not affect the (weak) generative capacity of TAGs.

33. We have preferred to assume that the reanalyzed structure of the Swedish embedded question is cleft-like because this structure is attested in other languages. The facts of Swedish, however, are equally consistent with assuming the structure to be that of a free relative. In particular, free relatives show the same pattern of occurrence of *wh+som* that embedded questions do. The following examples illustrate the point:

- (i) Jag åt vad (som) han gav mig.
I ate what that he gave me
- (ii) Jag åt vad *(som) stod på bordet.
I ate what that stood on the table.

The reader will note that, if sound, the discussion in the text argues for a “head” analysis (Bresnan and Grimshaw 1978) rather than a “COMP” analysis (Groos and van Riemsdijk: 1979) of free relatives in Swedish.

34. Kayne attempts to motivate government of intermediate traces across maximal projections on the basis of certain facts concerning the distribution of *combien* phrases in French; and Chomsky 1981a presents evidence from Hungarian case marking to the same end. It is also possible that comparative subdeletion in English should be analyzed as involving such exceptional government (Baltin 1983). There are, not surprisingly, plausible alternatives to government across maximal projection boundaries for all of these cases and we hope to explore them in the future.

35. The examples in this section are taken from Kayne and Pollock’s article.

36. For an alternative account, stated in the notation of LFG, see Zaenen 1983.

37. No sentence of this form appears in K&P’s discussion; but its ungrammaticality is directly implied by their analysis.

38. This not the analysis that K&P wind up proposing, but the differences are not relevant to our discussion.

39. This convention can be conceived as a generalization of the class OA of local constraints defined in section 2.3; that is, of the class of local constraints which specify that a given elementary tree must receive an adjunction if it is to appear in a derivation.

40. See Pollard 1984 for an elegant treatment of this problem in his Head Grammar formalism.