Flexible Composition and Delayed Tree-Locality

David Chiang
USC Information Sciences Institute
4676 Admiralty Way, Suite 1001
Marina del Rey, CA 90292 USA
chiang@isi.edu

Tatjana Scheffler
Department of Linguistics
619 Williams Hall
University of Pennsylvania
Philadelphia, PA 19104-6305 USA
tatjana@ling.upenn.edu

Abstract

Flexible composition is an extension of TAG that has been used in a variety of TAG-analyses. In this paper, we present a dedicated study of the formal and linguistic properties of TAGs with flexible composition (TAG-FC). We start by presenting a survey of existing applications of flexible composition. In the main part of the paper, we discuss a formal definition of TAG-FCs and give a proof of equivalence of TAG-FC to tree-local MCTAG, via a formalism called delayed tree-local MCTAG. We then proceed to argue that delayed tree-locality is more intuitive for the analysis of many cases where flexible composition has been employed.

1 Introduction

Flexible composition (FC) is a way of viewing TAG derivations so that the operation of adjoining of a tree β into a tree γ can be alternatively viewed as attachment of γ to β. That is, γ splits at the adjunction site and wraps around β (see Figure 1b). This "flexible" view of the attachment operation does not have much effect on standard TAG, but has been used in multicomponent TAG (MCTAG) analyses of various linguistic phenomena in order to preserve tree-locality of an otherwise non-local derivation.

First, it has been employed in (Joshi et al., 2003) to derive quantifier-scope restrictions in nested quantifications such as:

(1) Two politicians spy on someone from every city. (Joshi et al., 2003, ex. (6))

Other applications of flexible composition include the modelling of complex noun phrases in pied-piping and stranding of wh-phrases (Kallmeyer and Scheffler, 2004), an analysis of anaphor binding (Ryant and Scheffler, 2006), discourse semantics (Forbes-Riley et al., 2006), and scrambling patterns (Chen-Main and Joshi, 2007).

With the proposal of unification-based semantics for TAG, noun phrase quantifiers have been analysed as multi-component sets, where one component is the lexical quantifier and the other is just an S-node carrying the scopal information for the quantifier. But this kind of analysis can be problematic for tree-local MCTAG, since the two components will in general attach to different elementary trees. For example, see Figure 2a for the sentence

(2) Whom does John like a picture of?
(Kallmeyer and Scheffler, 2004, ex. (2a))

Flexible composition has been used to avoid this problem (Joshi et al., 2003; Kallmeyer and Scheffler, 2004), as shown in Figure 2b. In this derivation, the edge label "rev" (to be defined more precisely in the following section) indicates that the adjunction of $\beta_{\text{a-2}}$ into β_{picture} is reversed. This turns the nonlocal derivation in Figure 2a into a tree-local derivation..

All the other proposals mentioned share this property as well: in each case, flexible composition is used in order to make a potentially nonlocal MCTAG derivation be possible in a tree-local MCTAG. Here, we present a new variant of TAG, called *delayed tree-local* multicomponent TAG, that relaxes the tree-local constraint. We define both formalisms and show that both are weakly equivalent to standard TAG. We then illustrate how

Proceedings of The Ninth International Workshop on Tree Adjoining Grammars and Related Formalisms
Tübingen, Germany. June 6-8, 2008.

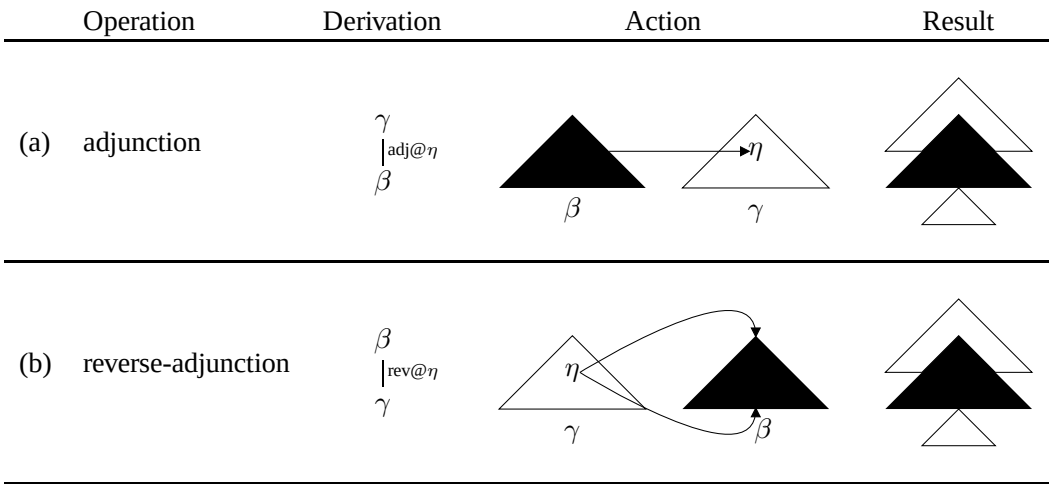

Figure 1: TAG-FC composition operations. (a) Adjunction. (b) Reverse-adjunction.

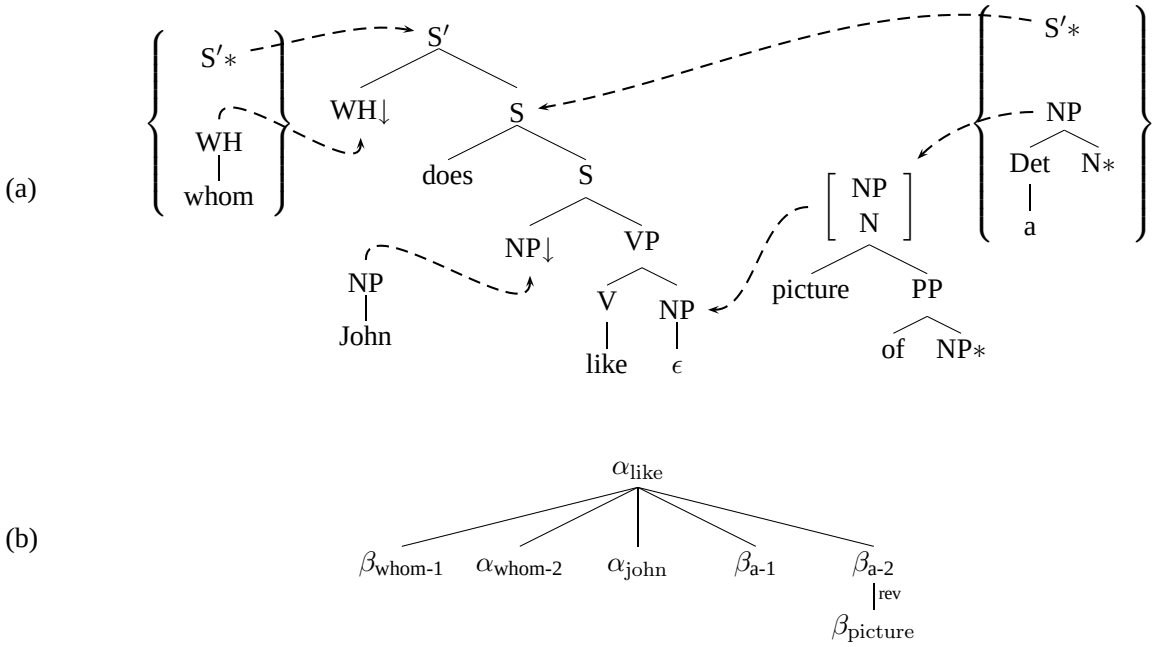

Figure 2: Derivation of "Whom does John like a picture of?" using flexible composition. (a) Syntactic analysis given in (Kallmeyer and Scheffler, 2004, Fig. 4). (b) Derivation tree, according to the notation used in this paper. The derivation is tree-local with flexible composition: The tree for "picture of" β_{picture} wraps around (reverse-adjoins into) the tree for "a" $\beta_{\text{a}-2}$, which then adjoins into the complement NP node of α_{like}.

linguistic analyses using flexible composition can be instantiated in our new formalism and argue that in many cases this new formulation is better.

2 Flexible composition

We present here a formal definition of TAG-FC, to our knowledge the first such definition.

Definition 1. A *TAG with flexible composition (TAG-FC)* is a TAG with two composition operations: adjunction and reverse-adjunction. A derivation of a TAG-FC is represented by a tree with labeled edges: each edge is labeled with an operation (adj for adjunction or rev for reverse-adjunction) and an adjunction site η. An edge labeled adj@η with γ above and β below, where η is a node of γ (see Figure 1a), represents adjunction at η. An edge labeled rev@η with β above and γ below, where η is again a node of γ (see Figure 1b), represents reverse-adjunction at η, in which γ is split at η and wraps around β.

Ambiguity arises in TAG-FC derivations whenever two elementary trees reverse-adjoin around the same elementary tree, or when an elementary tree both adjoins and is reverse-adjoined around (see Figure 3). In these cases a different derived tree will result depending on the order of operations. Thus, we simply rule out the former case,[1] and in the latter case, we stipulate that the reverse-adjunction occurs first.

Flexible composition generalizes to tree-local multicomponent TAG (Weir, 1988) in the obvious way. Note that there are two ways of defining tree-local MCTAG derivation trees: one in which the derivation nodes are elementary *tree sets* (as in Weir's definition), and the other in which the derivation nodes are elementary *trees*. We use the latter notion.

Definition 2. A *multicomponent TAG (with flexible composition)* is a TAG (with flexible composition) whose elementary trees are partitioned into *elementary tree sets*. In a derivation of a multicomponent TAG, the nodes of the derivation are also partitioned into sets such that each partition is an instance of a complete elementary tree set.

Definition 3. A *tree-local multicomponent TAG (with flexible composition)* is a multicomponent TAG (with flexible composition) whose derivations have the following property: for each elementary tree set instance, all the member derivation nodes are sisters.

[1] We are not aware of any examples of this case in the literature. If this case should prove to be useful, the definitions and results in this paper would need to be modified. We leave this possibility for future work.

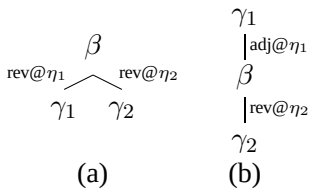

Figure 3: Ambiguity in TAG-FC derivations. (a) Multiple reverse-adjunction is disallowed. (b) The reverse-adjunction of γ_2 takes place before the adjunction of β.

In other words, all the members of an elementary tree set must adjoin at the same time, and must adjoin into the same elementary tree.

3 Delayed tree-locality

Next, we present another variant of MCTAG that relaxes the tree-locality constraint without losing weak equivalence with standard TAG, but uses only standard adjunction, not reverse adjunction.

Definition 4. A *k-delayed tree-local multicomponent TAG* is a multicomponent TAG whose derivations have the following property. Let the *destination* of an elementary tree set instance S be the lowest derivation node that dominates all the members of S. Let the *delay* of S be the union of the paths from the destination down to each member of S, minus the destination itself. Then no derivation node can be a member of more than k delays.

See Figure 4. Intuitively, this means that the members of an elementary tree set can adjoin into different trees, arriving at the same elementary tree (the destination) after some delay; and there can be at most k delays at any point in the derivation. (Note that this definition also allows one member of an elementary tree set to adjoin into another.) For a more practical example, observe that the derivation in Figure 2a is a 1-delayed tree-local MCTAG derivation.

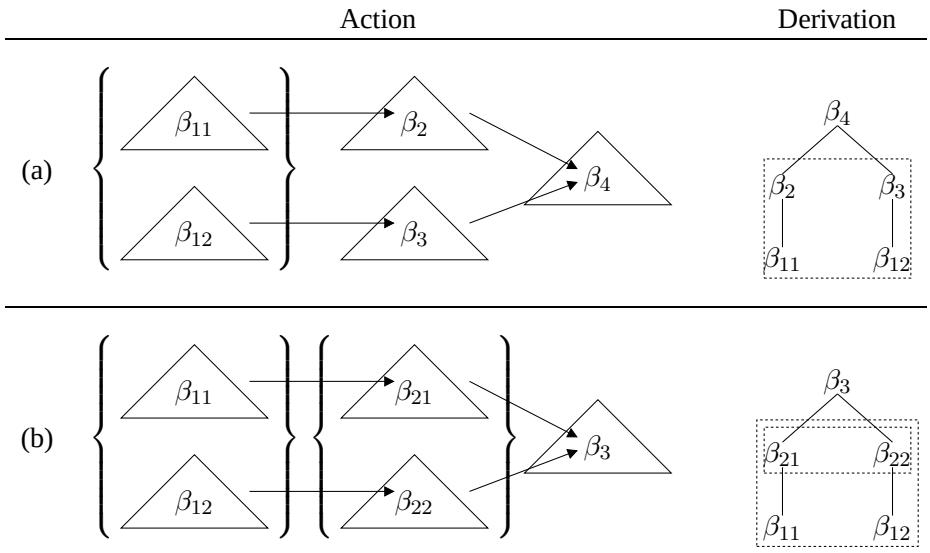

Figure 4: Delayed tree-locality. Nonlocal adjunction of an elementary tree set is allowed as long as the members eventually compose into the same elementary tree. The dashed boxes mark the delays. (a) One simultaneous delay. (b) Two simultaneous delays are allowed in 2-delayed tree-local MCTAG but not 1-delayed tree-local MCTAG.

4 Formal results

In this section, we show the equivalence of both tree-local MCTAG-FC and delayed tree-local MCTAG to standard TAG.

Proposition 1. *Any tree-local MCTAG with flexible composition G can be converted into a 2-delayed tree-local MCTAG G' that is weakly equivalent to G and has exactly the same elementary structures as G.*

The fact that G' has the same elementary structures as G means that if we convert an analysis from tree-local MCTAG-FC to delayed tree-local MCTAG, its domains of locality will be preserved. However, the dependencies between them will in general be different.

Proof. The conversion is trivial: G' has exactly the same elementary structures as G. In order to demonstrate weak equivalence, we show how to convert any TL-MCTAG-FC derivation into a nonlocal MCTAG derivation, and then show that this derivation is a 2-delayed TL-MCTAG derivation.

Given a TL-MCTAG-FC derivation, consider the subgraph formed by erasing all adjunction edges and keeping only the reverse-adjunction edges. Call the components of this subgraph the *reverse chains* (see Figure 5a).

It is easy to see from the definition of TAG-FC that reverse chains are all subpaths; thus, to convert the derivation to a nonlocal MCTAG derivation, we simply invert all the reverse chains. We continue to refer to the inverted reverse chains in the new derivation as reverse chains, even though they are only definable with reference to the original derivation (see Figure 5b).

Now we must show that this derivation is a 2-delayed TL-MCTAG derivation. Actually, we prove a stronger claim, by induction on the height of the derivation tree: (i) no node belongs to more than two delays, and moreover (ii) the nodes in the root's reverse chain belong to no more than one delay. (See Figure 5c for an example.)

Let R be the root's reverse chain, and let C be those nodes which are children of nodes in R but are not themselves in R. Apply the transformation to the subderivations rooted by nodes in C. By the induction hypothesis, the transformation creates (i) no more than two delays for the nodes in those subderivations, and (ii) no more than one delay for the reverse chains of the nodes in C.

Next, reverse R itself. For a node η in R that belongs to an elementary tree set, a new delay is created that comprises η and the reverse chains of all the other members of the elementary tree set.

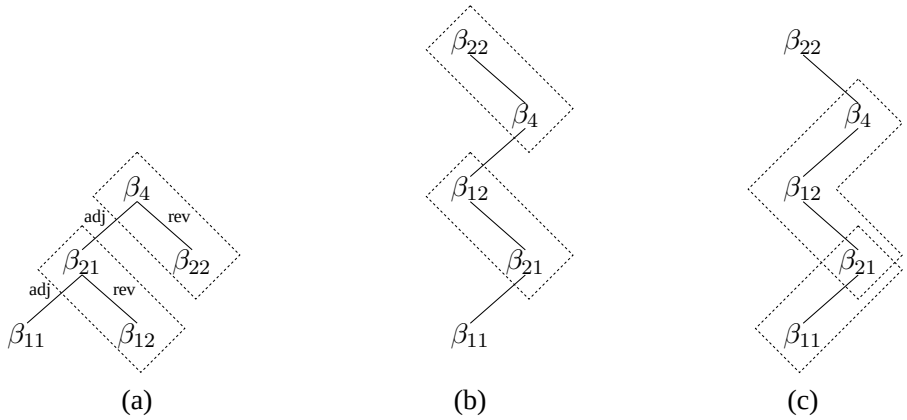

Figure 5: (a) Example tree-local MCTAG-FC derivation tree with reverse chains marked. (b) Result of conversion to delayed tree-local MCTAG derivation tree, again with reverse chains marked. (c) Same derivation tree but with delays marked.

But by (ii), the nodes in those reverse chains belonged to no more than one delay already, so even after creating this new delay, they still belong to no more than two delays.

Thus, (i) holds for all nodes in the derivation. The nodes in R that belong to an elementary tree set belong to only one delay, satisfying (ii), and the other nodes in R do not belong to any delays, also satisfying (ii). □

Next we show that k-delayed tree-local MCTAG is, in turn, weakly equivalent to standard TAG.

Proposition 2. *Any k-delayed tree-local MCTAG can be converted into a weakly equivalent TAG.*

Proof. The construction is a generalization of the conversion of tree-local MCTAG to TAG. We consider 1-delayed tree-local MCTAG first. First, we normalize the grammar so that all adjunction is obligatory and no adjunction is allowed at root/foot nodes, following Lang (1994): for each auxiliary tree, create new null-adjunction root and foot nodes; and for each nonterminal X, create a trivial auxiliary tree with a single null-adjunction X that is both root and foot. Next, create a new feature $tree$ whose values are of the form $S^\bullet$ or $S_\bullet$, where S is a multiset of elementary trees. We replace each elementary tree γ with copies of γ that have the $tree$ feature set in all possible ways that satisfy the following properties:

- The top of each interior node has $tree = S^\bullet$ and the bottom of each interior node has $tree = S_\bullet$, where S is a nonempty proper subset (without duplicates) of an elementary tree set.
- If γ is an auxiliary tree, the top/bottom of the root node of γ has $tree = S^\bullet$ and the top/bottom of the foot node has $tree = S_\bullet$, where S is as above, and is equal to:
  - $\{\gamma\}$,
  - plus the union of the values of the $tree$ features of all the interior nodes,
  - minus any complete elementary tree sets.
- If γ is an initial tree, we define S as for auxiliary trees, but require that S be empty.

The effect of the $tree$ feature is to keep track of any incomplete elementary tree sets that have been used in a subderivation. Each elementary tree combines the $tree$ features of the elementary trees adjoining into it, and discharges any complete elementary tree sets that are formed. If the resulting S contains elementary trees from more than one set, there would be more than one simultaneous delay, so the construction rules out this case. In an initial tree, S is required to be empty because there can be no outstanding delays at the top of the derivation.

To move from 1-delayed tree-locality to k-delayed tree-locality, we simply allow S to be the multiset union of k nonempty proper subsets of elementary tree sets. □

5 Discussion

As noted above, flexible composition has been used in TAG analyses of linguistic phenomena when the description necessitated by the linguistic facts would lead to a non-local (or set-local) derivation. As we have shown, this move is useful because adding flexible composition increases the descriptive power of TL-MCTAG, but not the weak generative power.

In a linguistic analysis, flexible composition can be used to reverse a non-local attachment edge (or path) and thus make the derivation tree-local. However, this process also makes the derivation hard to read and linguistically unintuitive if it creates attachment edges between non-dependent lexical items in the derivation tree. As we have shown above, any derivation that uses flexible composition can alternatively be expressed in a 2-delayed tree-local MCTAG. The advantage of using this alternative formalism directly is that the linguistic dependencies can be retained. In effect, we have shown that non-local MCTAG derivations are benign in many cases that are needed for linguistic analyses of certain phenomena, such as complex noun phrases, binding, and scrambling. This kind of non-locality is handled by a delayed tree-local MCTAG.[2]

It might be objected that 2-delayed tree-local MCTAG imposes an somewhat arbitrary limit on the number of simultaneous delays. We would agree that 1-delayed tree-locality is a more natural constraint, and believe that it is probably sufficient in practice, and that the example of Figure 5, which requires two simultaneous delays, is unusual.

On the other hand, there may be some cases where there is a 1-delayed tree-local analysis, but no analysis using TL-MCTAG with flexible composition. For example, consider the following sentence (3):

(3) John believes himself to be a decent guy.
(Ryant and Scheffler, 2006, ex. (10))

In the TAG-FC derivation previously proposed (see Figure 6a), α_{dg} is attached to $\alpha_{\mathrm{himself}}$ by reverse-substitution, and the result of this is attached to β_{believe} by reverse-adjunction. However, the reverse-adjunction site (S) does not come from $\alpha_{\mathrm{himself}}$, and therefore the reverse-adjunction of $\alpha_{\mathrm{himself}}$ into β_{believe} is not allowed according to our definition of flexible composition (Definition 1), since reverse-adjunction of γ into β at node η requires γ to be split at η, which must be a node in γ.

This operation was not explicitly excluded under previous definitions of flexible composition.[3] But if we tried to modify our definition of TAG-FC to allow such an operation, it is not clear how one would write the derivation trees, or whether the results obtained above would still hold.

In contrast, there is a straightforward 1-delayed TL-MCTAG derivation for the example. This derivation is shown in Figure 6b. In addition to readability, all the intuitive dependencies are retained explicitly in this derivation, for example the dependency between β_{believe} and α_{dg}.

6 Conclusion

This paper takes a closer look at the mechanism of flexible composition, which has been employed in TAGs for linguistic analysis for some time. Based on a survey of existing applications of flexible composition, we provide a formal definition of TAG-FC. We then prove the weak equivalence of tree-local MCTAG-FC to standard TAG via a variant called delayed tree-local MCTAG introduced here. Finally, we argue that delayed tree-local MCTAG is more intuitive than flexible composition for linguistic analyses that need slightly more descriptive power than tree-locality.

It remains for future work to reformulate existing analyses that use TAG-FC to use delayed tree-locality instead, and to compare the resulting analyses against the originals. On the formal side, it is also possible to give a formulation of TAG-FC as a special case of regular-form two-level TAG (Dras, 1999; Dras et al., 2003; Rogers, 2004; Rogers, 2006), a connection that deserves to be explored further.

[2] It needs to be tested more thoroughly how well the additional descriptive power of delayed tree-local MCTAG fares for other linguistic analyses, in particular those cases that have been claimed to necessitate non-local analyses in regular MCTAG (Bleam, 2000, for clitic climbing, for example).

[3] The definition in (Joshi et al., 2003) merely requires that the goal of reverse-adjoining is an elementary tree, but the reverse-adjoining tree may be a derived tree resulting from previous attachments.

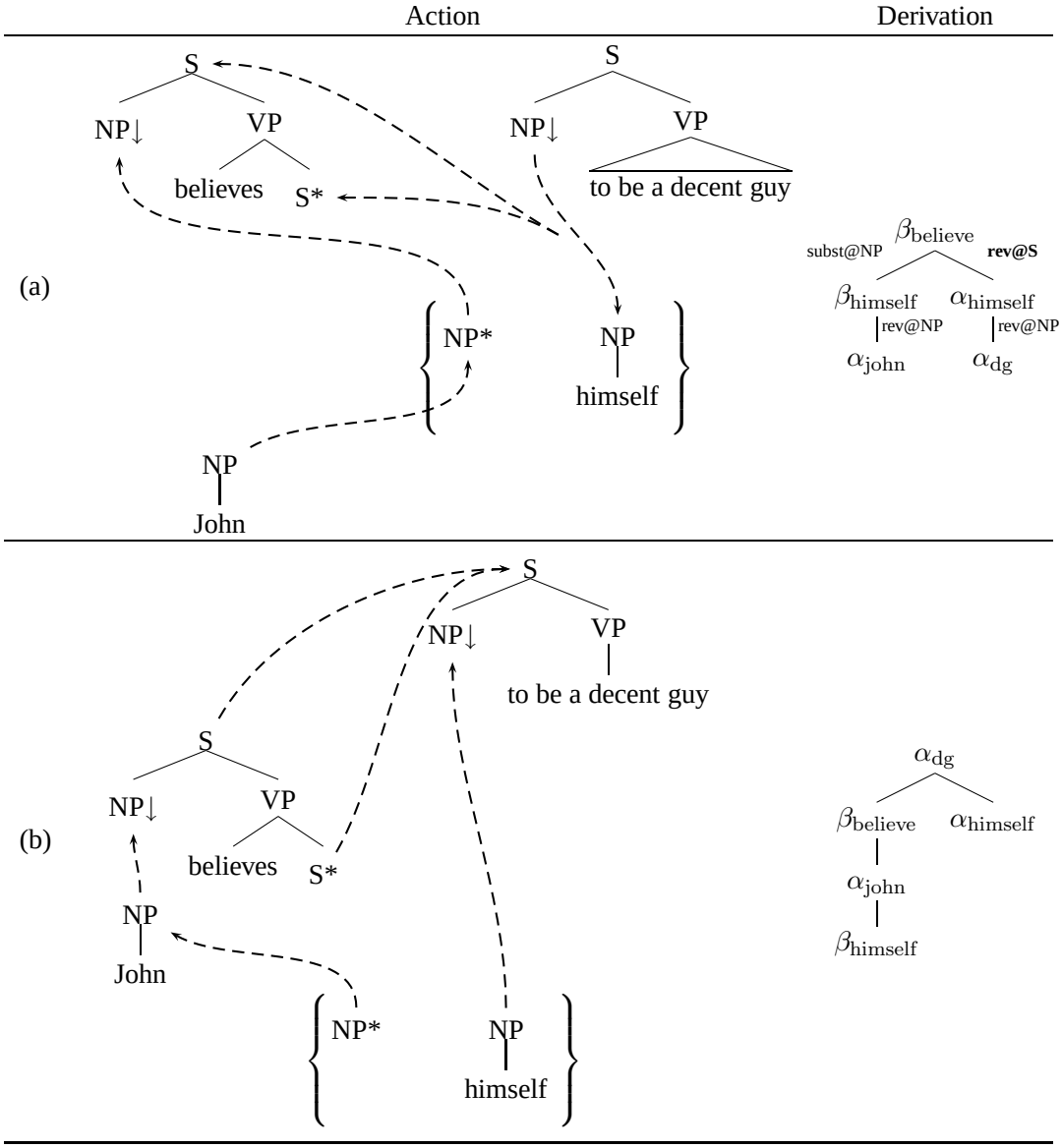

Figure 6: Derivation of "John believes himself to be a decent guy." (a) Illegal use of flexible composition, proposed in (Ryant and Scheffler, 2006): α_{himself} is claimed to reverse-adjoin at the S-node, but there is no S-node in α_{himself} (it originates from α_{dg}). (b) Straightforward analysis using 1-delayed TL-MCTAG.

Acknowledgements

The first author would like to acknowledge the person who first suggested delayed tree-locality. He thought it was Seth Kulick, but Seth thinks it must be somebody else.

The second author would like to thank the members of the XTAG weekly meeting at Penn, in particular Aravind Joshi, Joan Chen-Main, Lucas Champollion, and Joshua Tauberer, for comments and discussion about flexible composition.

References

Bleam, Tonia. 2000. Clitic climbing and the power of tree-adjoining grammar. In Abeillé, A. and O. Rambow, editors, *Tree Adjoining Grammars: Formalisms, Linguistic Analysis and Processing*. CSLI Publications, Stanford.

Chen-Main, Joan and Aravind K. Joshi. 2007. Some observations on a "graphical" model-theoretical approach and generative models. In *Proceedings of the Workshop on Model-Theoretic Syntax*, pages 53–62, Trinity College, Dublin. European Summer School in Logic, Linguistics, and Information.

Dras, Mark, David Chiang, and William Schuler. 2003. On relations of constituency and dependency grammars. *Research on Language and Computation*, 2:281–305.

Dras, Mark. 1999. A meta-level grammar: redefining synchronous TAG for translation and paraphrase. In *Proceedings of ACL 1999*, pages 80–87.

Forbes-Riley, Katherine, Bonnie Webber, and Aravind Joshi. 2006. Computing discourse semantics: The predicate-argument semantics of discourse connectives in D-LTAG. *Journal of Semantics*, 23(1):55–106.

Joshi, Aravind K., Laura Kallmeyer, and Maribel Romero. 2003. Flexible composition in LTAG: Quantifier scope and inverse linking. In Bunt, Harry, Ielka van der Sluis, and Roser Morante, editors, *Proceedings of the Fifth International Workshop on Computational Semantics (IWCS-5)*, pages 179–194, Tilburg, Netherlands.

Kallmeyer, Laura and Tatjana Scheffler. 2004. LTAG analysis for pied-piping and stranding of wh-phrases. In *Proceedings of the Seventh International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+7)*, pages 32–39, Vancouver, BC, Canada.

Lang, Bernard. 1994. Recognition can be harder than parsing. *Computational Intelligence*, 10(4):484–494. Special Issue on Tree Adjoining Grammars.

Rogers, James. 2004. One more perspective on semantic relations in TAG. In *Proceedings of TAG+6*.

Rogers, James. 2006. On scrambling, another perspective. In *Proceedings of TAG+7*.

Ryant, Neville and Tatjana Scheffler. 2006. Binding of anaphors in LTAG. In *Proceedings of the Eighth International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+8)*, pages 65–72, Sydney, Australia.

Weir, David J. 1988. *Characterizing Mildly Context-Sensitive Grammar Formalisms*. Ph.D. thesis, University of Pennsylvania.