Delayed Tree Locality and the Status of Derivation Structures

Joan Chen-Main
Institute for Research in Cognitive Science
University of Pennsylvania
Philadelphia, PA 19104, USA
chenmain@seas.upenn.edu

Abstract

While the derived trees yielded by TAG derivations are uncontroversially taken to correspond to phrase structure, the status of TAG derivation structures as more than a record of TAG operations is less certain. An attractive possibility is to interpret the derivation structure as some representation of semantic meaning, such as a dependency analysis. However, the literature has identified cases where doing so is problematic (Rambow et al., 1995, Candito and Kahane, 1998, Frank and van Genabith, 2001, Gardent and Kallmeyer, 2003, Kallmeyer and Romero 2008), including what has been referred to as the Missing Link Problem: predicates which should have a dependency link are unconnected in the derivation structure. This paper shows that delayed tree-local MC-TAG (Chiang and Scheffler, 2008) provides a solution for certain types of missing links. Further, we observe that the regular form 2-level TAG solutions to the Missing Link Problem given in (Dras et al., 2004) can be reinterpreted using delayed tree-local MC-TAG: the object level derivations of the 2-level TAG derivations can be converted into legal 1-delayed tree-local MCTAG derivations. Thus, delayed tree-locality maintains the possibility that TAG derivation structures can be more meaning-laden than solely a record of the combination of trees.

1 Introduction

In the mainstream generative approaches to grammatical structure, there is typically no distinction between a derivation of a sentence and its phrase structure. For example, in Chomsky's (1995) Minimalist Program (formalized by Stabler (1997) and its precusors, the history of a valid derivation is taken to be the phrase structure of a grammatical construction: composition determines constituency. In contrast, a derivation in a tree-rewriting formalism, such as TAG, allows for an additional level of representation. As a mathematical object, each TAG derivation yields a string, a derived tree, and a derivation structure. When TAG is used for linguistic analysis, the string and derived tree uncontroversially correspond to the grammatical sentence and its phrase structure, respectively, but the status of the derivation structure as more than a record of TAG operations is less certain.

An attractive possibility is to interpret the derivation structure as some representation of semantic meaning, such as a dependency analysis (e.g. Rambow and Joshi, 1997), and, indeed, there is body of work that explores the degree to which it is possible to equate the TAG derivation tree with a dependency analysis. This line of inquiry has identified cases where doing so is problematic (Rambow et al., 1995, Candito and Kahane, 1998, Frank and van Genabith, 2001, Gardent and Kallmeyer, 2003, Kallmeyer and Romero 2008), including what has been referred to as the Missing Link Problem. This particular mismatch stems from the way extraction is handled in TAG-style analyses when part of a lower clause ends up in the higher clause. In such cases, clausal complementation is carried out using adjoining. The resulting "stretching apart" of substructure in the tree for the lower clause eliminates the need for traces, and is a hallmark of TAG accounts of phenomena such as raising and successive-cyclic *wh*-movement in English.

Proceedings of the 11th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+11), pages 99–107,
Paris, September 2012.

The problem arises when more than one instance of this kind of adjoining occurs in the same tree.

The proposed solutions to the Missing Link Problem can be divided into two kinds of approaches. The Missing Link Problem naturally led to the question of how to address the computation of TAG semantics in general, and the first kind of response can be roughly characterized as modification of the object on which semantics is computed, for example, by "enriching" the derivation structure with additional links (Kallmeyer, 2002), computing meaning based on the derived phrase structure tree instead of the derivation tree (Gardent and Kallmeyer, 2003, Frank and van Genabith, 2001), and encoding meaning in both the derivation and derived tree (Kallmeyer and Romero, 2008). The development of TAG semantics is a significant and related contribution, but for the purposes of this paper, it is important to note that the status of the derivation tree is not the primary concern of this area of research.

A second type of solution to the Missing Link Problem can roughly be characterized as modifications to the grammar such that the derivation structures better align with the desired dependency analyses. These include proposals that are more powerful that TAGs, such as set-local MCTAGs (Weir, 1988) and D-tree Substitution Grammar (Rambow et al., 1995), as well as some that are weakly TAG-equivalent., such as regular form 2-level TAG (Dras et al., 2004). These proposals do not include a full semantics for TAG, but they preserve the intuition that the derivation structure is a meaning-carrying level of representation. The derivation structure need not be the object over which semantics is computed to be useful. Note, for example, that the Prague Dependency Treebank (PDT 2.0) is annotated with multiple layers, with the analytical layer encoding what are deemed "dependency relations" and the tectogrammatical layer encoding what is taken to be the "underlying deep structure" of the sentence (Hajič et al, 2006). It is also worth noting that when the derivations of TAGs and TAG variants are converted into the form of dependency structures (in the style of Kulhmann, 2007, Bodirsky et al., 2005), their formal properties as a class inform us with respect to coverage of dependency treebanks (Chen-Main and Joshi, 2012).

The observations reported here fall under the second kind of response. The big picture goal is to understand the role of the derivation structure. With the intuition that the derivation structure's role reaches beyond a record of operations as a backdrop, this paper begins to pursue the kind of answers afforded us by the recently introduced delayed tree-local MC-TAG formalism (Chiang and Scheffler, 2008). Delayed tree-local MCTAG is weakly-equivalent to standard TAG,[1] but it permits linguistic dependencies to be retained that are not necessarily retained in alternative TAG variants. This has already proven useful in linguistic analyses of anaphor binding (Chiang and Scheffler, 2008), non-local right node raising (Han et al., 2010), binding variables (Storoshenko and Han, 2010), and clitic climbing (Chen-Main et al. 2012). Here, we explore how the formalism deals with constructions whose standard TAG (or MCTAG) derivations result in missing links. We also observes that the solutions to the Missing Link Problem given in (Dras et al., 2004) can be reinterpreted as 1-delayed tree-local MCTAG derivations. We see that the increased flexibility of delayed tree-locality is advantageous not only for syntactic analyses, but also for maintaining the possibility that TAG derivation structures can be more meaning-laden than solely a record of the combination of trees.

This paper is structured as follows. Section 2 reviews two situations in which The Missing Link Problem arises. Section 3 first reviews delayed tree-local MCTAG before turning to a solution to one of the types of the Missing Link Problem. Section 4 addresses the second type of the Missing Link Problem. Following a brief review the regular form 2-level TAG solution to given in (Dras et al., 2004), we see how the solution can be recast as a delayed tree-local MCTAG derivation. Section 5 includes further discussion, raising some open questions, and concludes the paper.

2 The Missing Link Problem Revisited

Consider the construction in (1) (from Dras et al., 2004), in which raising and cyclic *wh*-movement co-occur.

[1] Delayed tree-local MCTAG is related to tree-local MCTAG with flexible composition (Joshi et al., 2003). Chiang and Scheffler (2008) show their weak equivalence by showing how any derivations in MCTAG with flexible composition can be converted into a 2-delayed tree-local MCTAG derivation. However, delayed tree-local MCTAG is not a formalization of MCTAG with flexible composition.

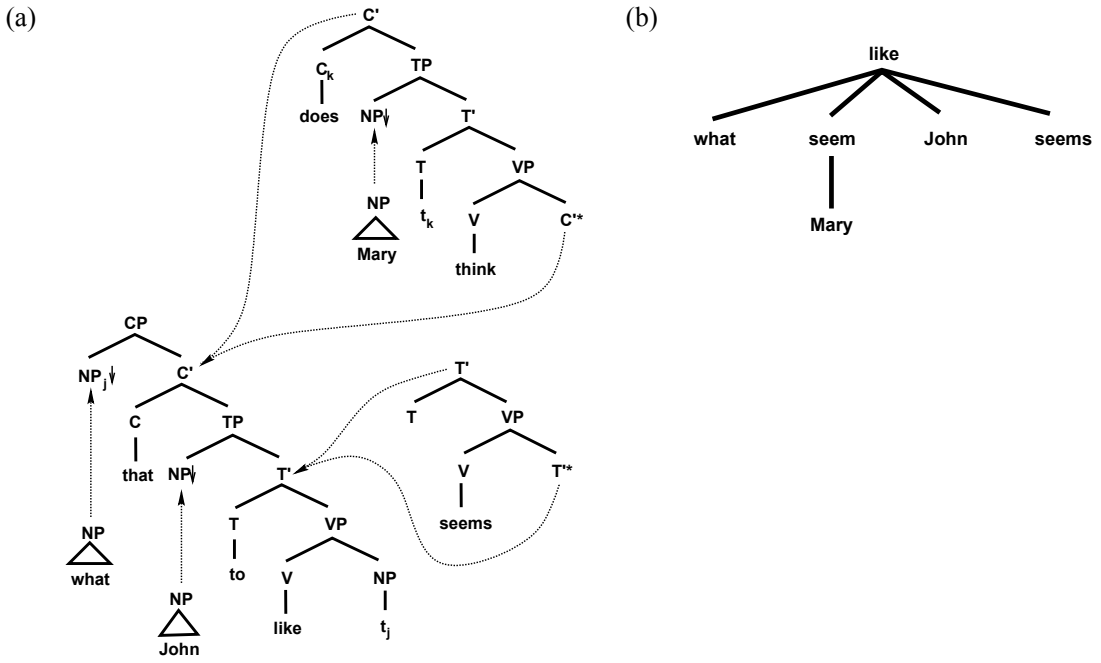

Figure 1. a) TAG derivation for *What does Mary think that John seems to like?*
b) Derivation structure for (a)

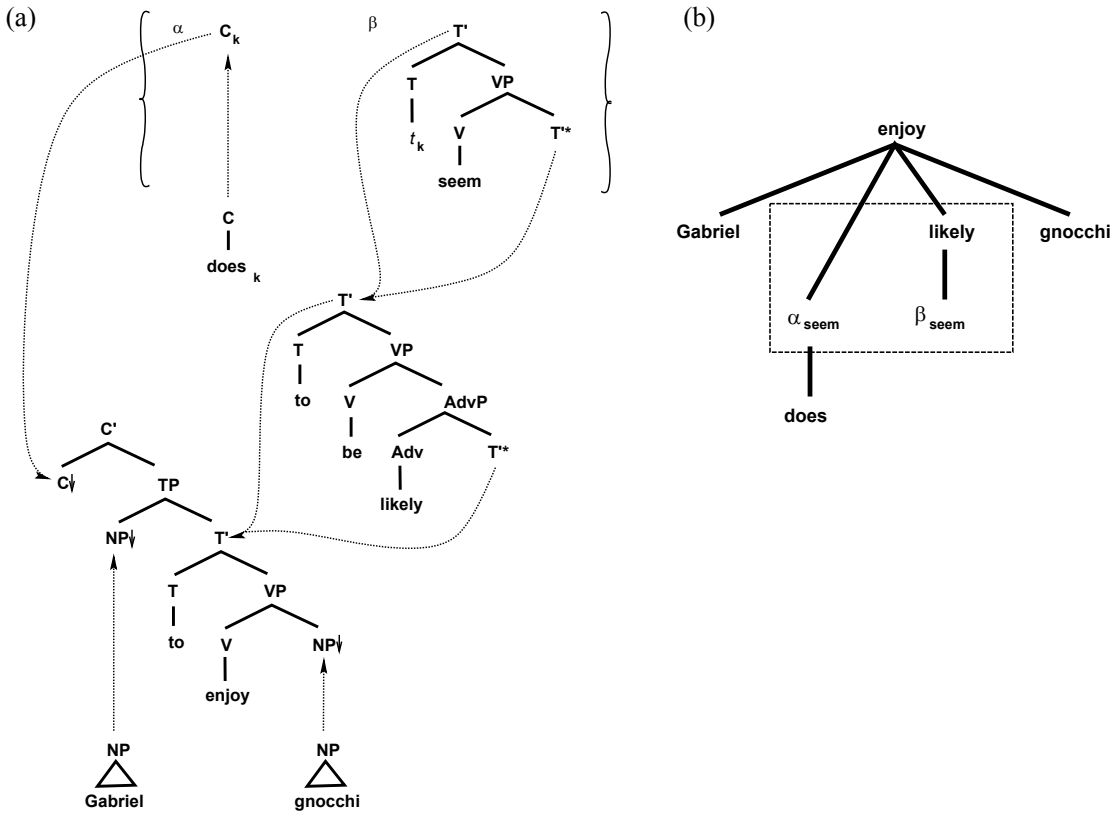

Figure 2. a) Derivation for *Does Gabriel seem to be likely to enjoy gnocchi?*: Legal in 1-delayed tree-local MCTAG, illegal in tree-local and set-local MCTAG
(b) Derivation structure for (a) with delay marked with a dashed box

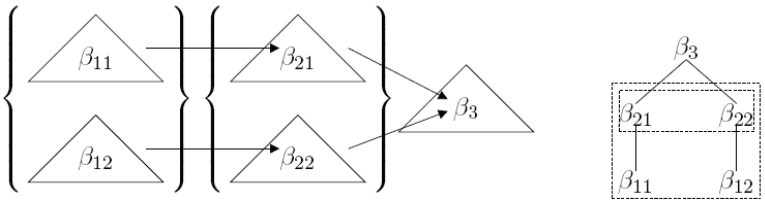

Figure 3: A 2-delayed tree-local MCTAG derivation. Delays are marked with dashed boxes. *(Figure taken from Chiang and Scheffler (2008).)*

(1) What does Mary think that John seems to like?

The derivation for (1), shown in Figure 1a, combines the standard TAG treatment of both these phenomena: Both the *seems*-tree for the raising construction and the *thinks*-tree for the bridge construction adjoin into the *like*-tree. The resulting two-fold problem can be seen in the corresponding derivation structure shown in Figure 1b: 1) the bridge verb *think* and the most embedded verb *like,* which have no semantic dependency, are connected with an edge while 2) the bridge verb *think* and the raising verb *seems*, which should have a semantic dependency, are unconnected.

This case is an example of the more general problem that can arise when several trees are adjoined into distinct nodes of the same tree. Thus, we also see the Missing Link Problem arise when other long distance dependencies co-occur with raising, e.g. Rambow et al.'s (1995) example, given in (2), in which topicalization out of an embedded clause interacts with raising.

(2) Small spicy hotdogs, he claims Mary seems to adore.

Dras et al. (2004) discuss a similar case involving subject-auxiliary inversion in conjunction with raising. A yes-no question like (3) is typically handled used multi-component TAG, with the structure for *does* and structure of *seem* as members of the same elementary tree set. The difficulty arises when another level of embedding is added, as in (4).

(3) Does Gabriel seem to enjoy gnocchi?
(4) Does Gabriel seem to be likely to enjoy gnocchi?

A dependency ought to link *seem* to *to be likely*, with another link between *to be likely* and *enjoy*. To derive (4) with the desired dependencies, the tree set containing *does* and *seem* must combine with *to be likely* before combing into the *enjoy*-tree. This derivation is shown in Figure 2. However, this derivation is neither tree-local nor set-local. (The *to be likely* tree cannot adjoin into the *seem* tree without adjoining into a foot node, and would also not yield the desired dependencies.) An alternative would be to permit multiple adjoining (Schabes and Shieber, 1994), but with predicative trees. A derivation structure for such a derivation, however, would link both *seem* and *to be likely* directly to *enjoy*, without the desired link between *seem* and *to be likely*.

3 Delayed Tree-Local MCTAG and Desired Links

3.1 *k*-Delayed Tree-Local MCTAG

The delayed tree-local variant of MCTAG specifies a way to relax the restriction that all components of a multi-component set must combine into the same tree during the same derivational step. Each use of a multicomponent set introduces a *delay* into the derivation. A delay is the union of the paths in the derivation structure from each component of an MC-set *S* to the lowest node that dominates all members of *S*. A *k*-delayed tree-local MCTAG permits each node in the derivation structure to be a member of at most *k* delays. Figure 3 replicates the example of a 2-delayed tree-local derivation given in Chiang and Scheffler (2008). The dashed boxes mark the delays. Thus, a valid *k*-delayed tree local MCTAG derivation permits members of the same MC set to compose into different trees, so long as all members of the MC set eventually compose into the same tree without requiring any node to belong to more than *k* delays.

3.2 A Solution for Raising and Subj-Aux Inversion

In contrast to the traditional tree-local and set-local variants of MCTAG, 1 delayed tree-local MCTAG does permit the derivation given in Figure 2 for our problematic raising and subj-aux inversion example. The derivation structure is given in Figure 2b, with the delay indicated by a dashed box. The two components of the *seem* tree set eventually both combine into the *enjoy* tree. However, the β component first adjoins into the *to be likely* tree and the combined phrase structure adjoins into the *enjoy* tree, while the α component (into which *does* substitutes) combines directly into the *enjoy* tree. This yields a link *seem* between *to be likely* and a link between *to be likely* and *enjoy*.[2]

This example illustrates how delayed tree-locality provides a straightforward solution in cases where two predicative trees are ultimately contiguous in the derived tree, but tree-locality and set-locality do not permit a derivation. The added flexibility of delayed tree-locality allows for a derivation in which the two predicative trees are combined, yielding the desired link in the derivation structure.

4 Borrowing from Regular Form Two-level TAG

We turn now to how to deal with the more typical cases of the Missing Link Problem, where several trees are adjoined into distinct nodes of the same tree. To allow the predicative trees to combine in the desired order, we will need to modify the shapes of the tree. The strategy is to conform the derivational shape of this case to the derivational shape of the case above where the two predicative trees are ultimately contiguous in the derived tree. We appeal to the solution given in Dras et al. (2004) and show how it can be recast as a 1-delayed tree-local MCTAG derivation.

4.1 Regular Form Two-level TAG Solution for Raising and Wh-movement

Dras et al. (2004) propose a regular form 2-level TAG, with a meta-level grammar that generates possible derivation structures and an object level grammar that yields derived phrase structures. Consider the object level derivation given in Figure 5 for example (1). At the object level, the derivation looks similar to a standard TAG derivation in that the combinatory operations combine pieces of phrase structure. The derivation structure for the object level (shown in Figure 5b), however, is the end product of a derivation at the meta-level. At the meta-level, the trees are pieces of object-level derivations. Figure 4a shows how the meta-level grammar generates the object-level derivation structure in Figure 5b. Figure 4b shows the history of this meta-level derivation. It is the derivation structure at this level which Dras et al. (2004) take to encode dependencies. Their goal is to match the meta-level derivation structure with a reasonable dependency analysis.

A key aspect to their analysis is that the tree anchored by a verb can be split into two parts. Consider the $\mathcal{A}$[*like*] meta-level tree in Figure 4a. The α_S[like] node and $\beta_{S/VP}$[like] node correspond to separate pieces of structure at the object level. As can be seem in Figure 5a, the α_S[like] tree contains the verb itself while the $\beta_{S/VP}$[like] tree contains the position for the subject. It is in the meta-grammar that the two parts are elementary tree local. As the authors themselves note, this is strikingly similar to a multi-component TAG approach, but their proposed derivation would not be tree-local in the original sense.

Note also that Dras et al. (2004) modifies the shapes of the trees by using a feature unification TAG where all non-terminals have the label X, but have top and bottom features that must be identical at the end of the derivation. A non-terminal node's part-of-speech or phrasal category is no longer its label, but rather, one of its features. However, in the figures that follow, the part-of-speech or phrasal category feature is graphically represented as a node label. It is crucial to adopt the modified shape of the trees to allow the predicative trees to combine as desired. The material that previously intervened between the *think* tree and the *seem* tree, forcing the two trees to be adjoined into different nodes in the *like* tree in Figure 1, is moved in two ways: 1) "that" is moved into a different tree, the *seems* tree, and 2) the position for the subject of *like* is extracted from the *like* tree as a separate piece of structure. Now, *think* may adjoin into *seems*, which later adjoins into *like*. The subject position

[2] When using multi-component sets, the question arises as to how to interpret multiple links from the same set. I assume that the link between the component containing the lexical anchor of the set and its target is the primary link for the set. Something more may need to be said about links to the other components, but my chief concern here is to ensure that the previously missing links are now present.

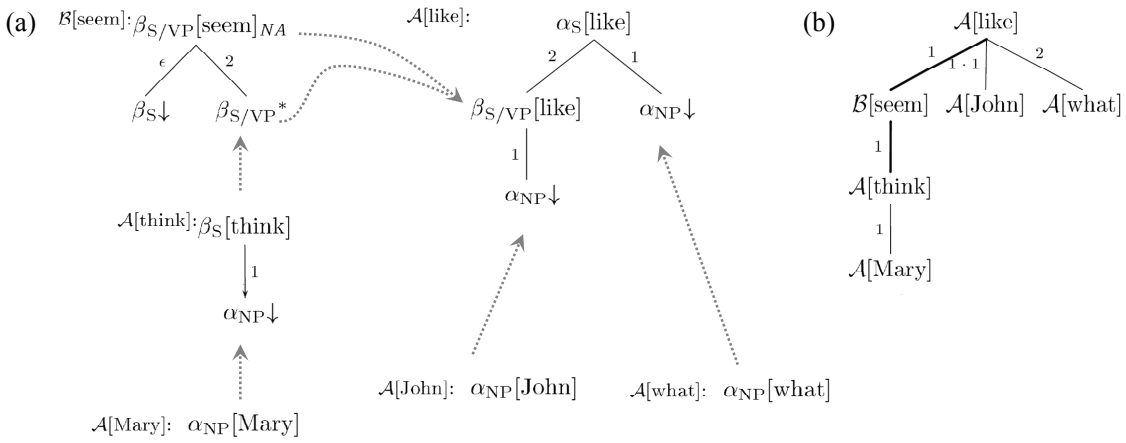

Figure 4: a) 2LTAG meta-level derivation for *What does Mary think that John seems to like?*
b) derivation structure for (a)
(Adapted from Dras et al. (2004))

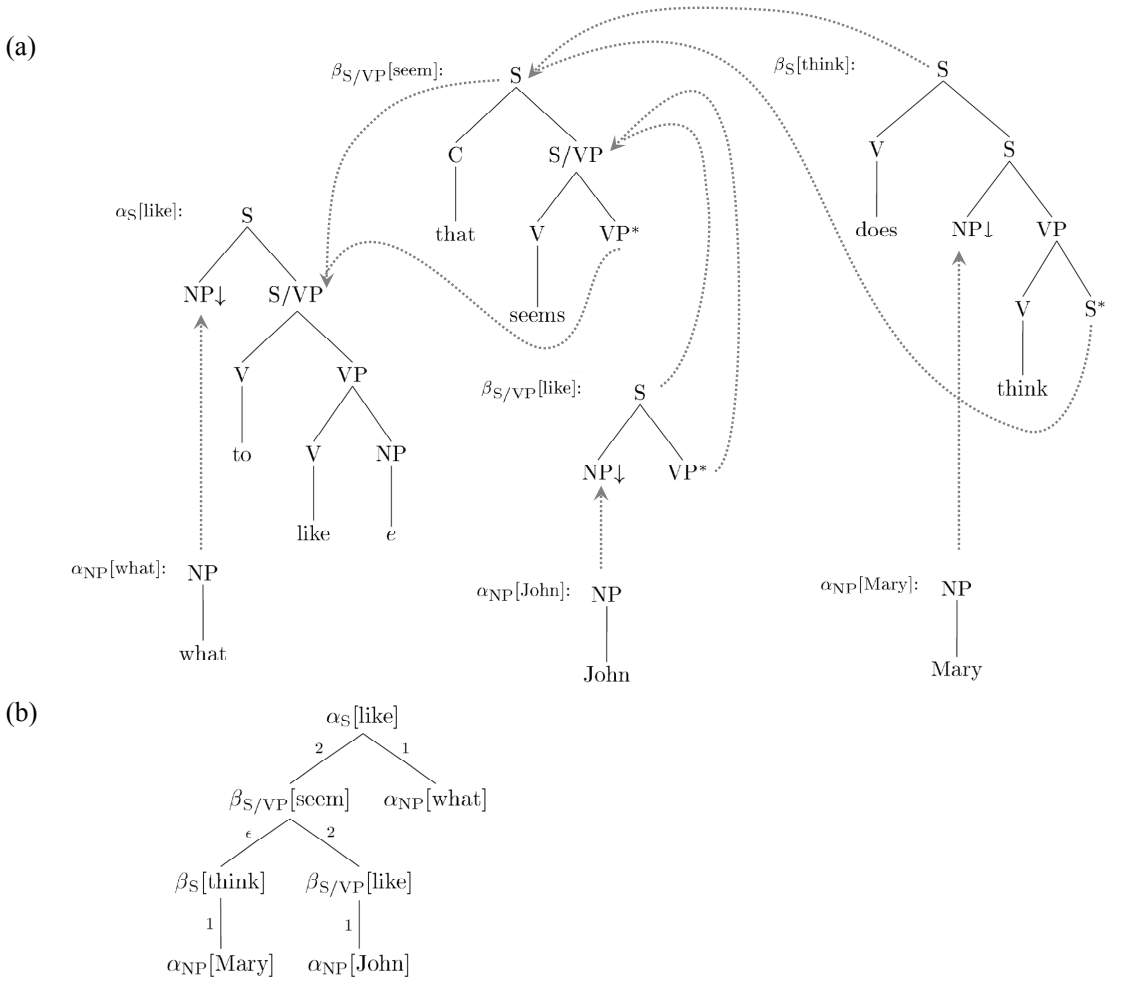

Figure 5: a) 2LTAG object-level derivation for *What does Mary think that John seems to like?*
b) derivation structure for (a)
(Adapted from Dras et al. (2004))

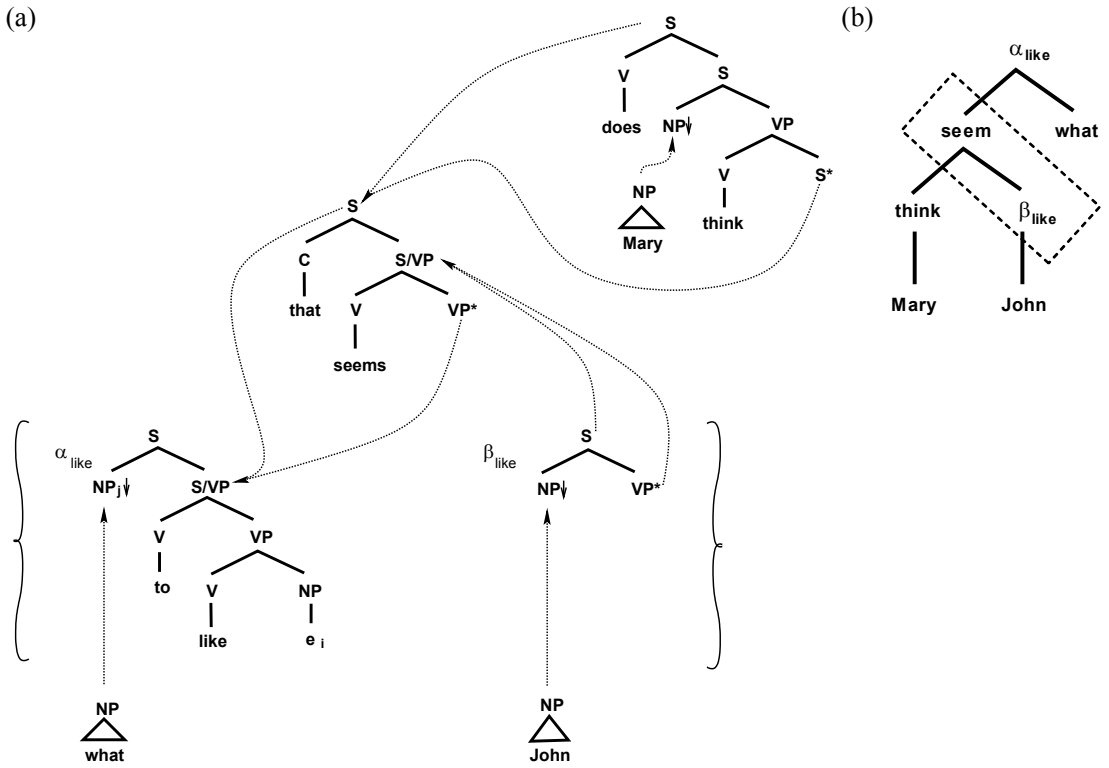

Figure 6. (a) 1-delayed tree-local MC-TAG Derivation for *What does Mary think that John seems to like?* (b) Derivation structure for (a) with delay marked with a dashed box

of *like* adjoins into *seems*, achieving the correct word order.

4.2 Translating into Delayed Tree-local MCTAG

It is straightforward to convert Dras et al.'s object level derivation into an MC-TAG derivation that abides by 1-delayed tree-locality. Instead of treating the tree for *like* and the tree for its subject as part of the same object at the meta-level, the delayed tree-local approach treats the elementary object for *like* as a 2-component set. Figure 6 is almost identical to Figure 5. The only differences are the braces denoting the 2-component set and the dashed box indicating the delay. [3] Note that delayed tree-locality as formalized in (Chiang and Scheffler, 2008) does permit a component to be combine into another component from the same MC-set. This means that the derivational steps that were prohibited under tree-locality and set-locality, i.e. the β component of the *like* set adjoins into the *seems* tree, which in turn adjoins into the α component of the *like* set, are now legal. As in the regular form 2-level TAG solution, the desired link between *think* and *seems* is no longer missing and the undesired link between *think* and *like* is no longer present

We have already presented a solution for constructions involving both raising and subject-Aux inversion, but we note that the alternative regular form 2-level TAG solution given in Dras et al. (2004) can also be converted into a 1-delayed tree-local derivation.

5 Conclusion and Discussion

The observation that delayed tree-locality can provide a solution for at least two types of missing links bears on questions specific to the formalism as well as more general issues. With respect to *k*-delayed tree-local MCTAG, we can frame this work as complementing work exploiting the formalism for linguistic analysis. Whereas the analyses in Chiang and Scheffler (2008), Han et al., (2010), Storoshenko and Han (2010), and Chen-Main et al. (2012) can tell us something about the coverage of *k*-delayed tree-local MCTAG (for a specific *k*) at the phrase

[3] See footnote 2 for comments on interpreting links from the same MC-set.

structure level, this paper demonstrates the kind of "coverage" that the formalism provides at the derivation structure level.

We also see how one of the derivational sequences that is, thus far, unique to delayed tree-local MCTAGs can be utilized. As shown in Figure 6, one component of an MC set ultimately combines into another component belonging to the same set. This allows what other TAG variants previously treated as contiguous pieces of structure to be treated instead as a multicomponent set.

Additionally, the observation that the regular form 2-level TAG derivations given by Dras et al. (2004) can be straightforwardly viewed as legal 1-delayed tree-local MCTAG derivations adds a linguistic dimension to Chiang and Scheffler's (2008) assertion that it is possible to give a formulation of TAG with flexible composition as a special case of regular-form 2-level TAG. As suggested by a reviewer, a sensible future avenue would be to see if other analyses that use multiple levels (or dimensions) can be restated in delayed tree-local MCTAG, particularly for phenomena that have been challenging for standard TAG (e.g. Rogers' (2004) analysis of scrambling).

Turning to broader issues, this paper revisits the question of what linguistic information, if any, is encoded in a derivation structure. It also raises the related question of what exactly a dependency analysis is and what linguistic information it carries. Chen-Main and Joshi (2012) show how TAG derivation structures (interpreted in the form of dependency structures) can be the basis for measuring complexity and a means for assessing coverage of large scale corpora, but they steer away from claims about the meaning that might be encoded. As noted in the introduction, work on TAG semantics appears to have reached a consensus that the derivation structure is not the appropriate representation for computing semantics. The introduction of delayed tree-local MCTAG, however, renews the viability of interpreting the derivation tree as a dependency analysis. The degree to which this is possible can lead to two additional research avenues. One is that we retain the current non-derivation structure based approach to TAG semantics and wrestle with distinguishing between a dependency analysis and semantic analysis. The other is to reevaluate the coverage that is possible when TAG semantics uses a delayed tree-local MCTAG derivation structure as the object on which semantics is computed. Either avenue should lead to a greater understanding of the role of the derivation structure.

Acknowledgments

The author would like to thank Aravind K. Joshi and the TAG+ reviewers for their comments, despite not yet having found the time to properly address them.

References

Manuel Bodirsky, Marco Kuhlmann, and Mathias Möhl. 2005. Well-nested drawings as models of syntactic structure. In *10th Conference of Formal Grammar and 9th Meeting on Mathematics of Language*, Edinburgh, UK.

Marie-Hélène Candito and Sylvain Kahane. 1998. Can the TAG derivation tree represent a semantic graph? An answer in the light of Meaning-Text Theory. In *Proceedings of TAG+4,* Philadelphia, USA.

Joan Chen-Main, Tonia Bleam, and Aravind K. Joshi. 2012. Delayed Tree-Locality, Set-locality, and Clitic Climbing. In *Proceedings of TAG+11*, Paris, France.

Joan Chen-Main and Aravind.K. Joshi. 2012. A Dependency Perspective on the Adequacy of Tree Local Multi-component Tree Adjoining Grammar. *Journal of Logic and Computation*. doi:10.1093/logcom/exs012

David Chiang and Tatjana Scheffler. 2008. Flexible Composition and Delayed Tree-Locality. In *Proceedings of TAG+9*, Tübingen, Germany.

Noam Chomsky. 1995. *The Minimalist Program*. MIT Press, Boston, Massachusetts.

Mark Dras, David Chiang and William Schuler. 2004. On Relations of Constituency and Dependency Grammars. *Research on Language and Computation* 2(2), 281-305. Hermes Science Publishers, Paris, France.

Anette Frank and Josef van Genabith. 2001 GlueTag: Linear Logic based Semantics for LTAG - and what it teaches us about LFG and LTAG. In M. Butt and T. Holloway King, editors, *Proceedings of the LFG01 Conference,* Hong Kong. p 104-126.

Claire Gardent and Laura Kallmeyer. 2003. Semantic Construction in FTAG. In *Proceedings of EACL 2003*, Budapest, Hungary. 123-130.

Jan Hajič, Jarmila Panevová, Eva Hajičová, Jarmila Panevová, Petr Sgall, Petr Pajas, Jan ?těpánek, Jiří Havelka, and Marie Mikulová. 2006. *Prague Dependency Treebank 2.0* LDC2006T01, Linguistic Data Consortium, Philadelphia

Chung-Hye Han, David Potter, and Dennis Ryan Storoshenko. 2010. Non-local Right Node Raising: an Analysis Using Delayed Tree-Local MC-TAG. In *Proceedings of TAG+10*, New Haven, USA.

Aravind K. Joshi, Laura Kallmeyer, and Maribel Romero. 2003. Flexible composition in LTAG: quantifier scope and inverse linking. In H. Bunt and R. Muskens (eds.), *Computing Meaning* 3. Kluwer.

Laura Kallmeyer. 2002. Using an Enriched TAG Derivation Structure as Basis for Semantics In *Proceedings of TAG+6,* Universitá di Venezia.

Laura Kallmeyer and Maribel Romero. 2008. Scope and Situation Binding in LTAG using Semantic Unification. *Research on Language and Computation* 6(1), 3-52. Kluwer Academic Publishers, the Netherlands.

Marco Kuhlmann. 2007. Dependency Structures and Lexicalized Grammars. PhD thesis, Saarland University, Saarbrücken, Germany.

Owen Rambow and Aravind K. Joshi. 1997 A formal look at dependency grammars and phrase structure grammars, with special consideration of word-order phenomena. In L. Wanner, ed. *Recent Trends in Meaning-Text Theory,* 167-190. John Benjamins, Amsterdam and Philadelphia.

Owen Rambow, K, Vijay-Shanker, and David Weir. 1995. D-Tree grammars. In *Proceeding of the 33rd Annual Meeting of the Association for Computational Linguistics (ACL-95)*, Cambridge, USA. 151-158.

James Rogers. 2004. On scrambling, another perspective. In *Proceedings of TAG+7*, Vancouver, Canada.

Yves Schabes and Stuart Shieber. 1994. An alternative conception of tree-adjoining derivation. *Computational Linguistics*, 20(1), 91-121.

Edward P. Stabler. 1997. Derivational minimalism. In *Proceedings of Logical Aspects of Computational Linguistics*, 68-95.

Dennis Ryan Storoshenko and Chung-Hye Han. 2010. Binding Variables in English: An Analysis using Delayed Tree Locality. In *Proceedings of TAG+10*, New Haven, USA.

David Weir. 1988. *Characterizing mildly context-sensitive grammar formalisms*. PhD dissertation, University of Pennsylvania, Philadelphia, USA.