Research on Language and Computation **00:** 1–25, 2007.

Pied-Piping in Relative Clauses: Syntax and Compositional Semantics using Synchronous Tree Adjoining Grammar

CHUNG-HYE HAN (chunghye@sfu.ca)
Department of Linguistics,
Simon Fraser University, Burnaby BC V5A 1S6, Canada

August 14, 2007

Abstract. In relative clauses, the *wh* relative pronoun can be embedded in a larger phrase, as in "a boy [whose brother] Mary hit" and "a boy [whose brother's friend] Mary hit". In such examples, we say that the larger phrase containing the *wh*-word has pied-piped along with the *wh*-word. In this paper, using a similar syntactic analysis for *wh* pied-piping as in Han (2002) and further developed in Kallmeyer and Scheffler (2004), I propose a compositional semantics for relative clauses of the sort illustrated above using Synchronous Tree Adjoining Grammar, a pairing of Tree Adjoining Grammars. It will be shown that (i) the elementary tree representing the logical form of a *wh* relative pronoun provides a generalized quantifier, and (ii) the semantic composition of the pied-piped material and the *wh*-word is achieved through adjoining the elementary tree representing the logical form of the pied-piped material to the elementary tree representing the logical form of the *wh*-word in the semantics.

Key words: relative clauses, pied-piping, compositional semantics, Synchronous Tree Adjoining Grammar

1. Introduction

In relative clauses, the *wh* relative pronoun can be embedded in a larger phrase. For example, in (1) and (2), the *wh* relative pronoun is in a genitive that is embedded in a larger Determiner Phrase (DP). In such examples, we say that the larger phrase containing the *wh*-word has pied-piped along with the *wh*-word.

(1) a boy [[$_{DP}$ whose brother]$_i$ Mary hit t$_i$]

(2) a boy [[$_{DP}$ whose brother's friend]$_i$ Mary hit t$_i$]

Pied-piping in relative clauses poses an interesting challenge for compositional semantics. An appropriate compositional analysis must correctly capture the predicate-argument relation within the relative clause, as well as the relation between the head noun and the *wh* relative pronoun so that the head noun and the relative clause predicate over the same variable. In a phrase-structure based compositional semantics, along the lines of the one presented in Heim and Kratzer (1998), the meaning of a sentence is computed as a function of the meaning of each node in the syntactic tree. A wide-spread analysis of relative clauses in this approach treats the relative pronoun as an operator that turns the relative clause into a function of a predicate type $< e, t >$: it contributes a λ-operator that binds the variable coming from the gap in the relative clause. For example, in (3a), the relative clause is a predicate that can be represented as in (3b). This composes with the semantics of the head noun *boy*, resulting in (3c).

(3) a. a boy [who$_i$ [Mary hit t$_i$]]
b. $\lambda x.\text{hit}(\text{Mary}, x)$
c. $\lambda x.\text{boy}(x) \wedge \text{hit}(\text{Mary}, x)$

Under this analysis, compositional semantics of relative clauses with pied-piping can only be done by putting the pied-piped material back into the position of the gap in the covert component of the grammar (Heim and Kratzer, 1998). That is, the interpretation of (1) and (2) must be done on the 'reconstructed' counterparts, as in (4a) and (5a). In each of the reconstructed structures, the relative pronoun, as a λ-operator, is able to bind the variable contributed by its trace, turning the relative clause into a predicate. The predicates can be represented as λ-expressions in (4b) and (5b), abstracting away from the semantics of *se brother* and *se brother's friend*.

(4) a. a boy [[$_{DP}$ who]$_i$ Mary hit [t$_i$ se brother]]
b. $\lambda x.\text{hit}(\text{Mary}, \text{the-brother-of}(x))$

(5) a. a boy [[$_{DP}$ who]$_i$ Mary hit [t$_i$ se brother's friend]]
b. $\lambda x.\text{hit}(\text{Mary}, \text{the-friend-of-the-brother-of}(x))$

In this paper, using a similar syntactic analysis for *wh* pied-piping as in Han (2002) and further developed in Kallmeyer and Scheffler (2004) for *wh*-questions, I propose a compositional semantics for relative clauses of the sort illustrated in (1) and (2), using Synchronous Tree Adjoining Grammar (STAG), a pairing of Tree Adjoining Grammars (TAGs). In STAG-based compositional semantics, each elementary tree in the syntax is paired up with one or more semantic elementary trees that represent its logical form, and the semantic composition is done by combining the semantic trees using the same combinatory operations as the ones available in the syntax, namely substitution and adjoining. These operations are explained in detail in section 2. Relying on these assumptions, the two main components of my proposal are that (i) the semantic tree representing the logical form of a *wh* relative pronoun provides a generalized quantifier, and (ii) the semantic composition of the pied-piped material and the *wh*-word is achieved through adjoining the semantic tree of the pied-piped material to the semantic tree of the *wh*-word.

I start the paper with an introduction to the basics of Tree Adjoining Grammar (TAG) for natural language in section 2. I restrict the discussion to the properties of TAG that are most relevant for appreciating the analysis proposed in this paper. In the same section, I also discuss the notion of compositionality appropriate for TAG in general and how it is achieved in STAG in particular. In section 3, I introduce the framework of STAG and STAG-based compositional semantics and clarify my assumptions. I illustrate various aspects of the framework with a simple sentence containing a quantifier and an attributive adjective. I then illustrate how quantifier scope ambiguity is handled in STAG. In section 4, I present my analysis of relative clauses and pied-piping. After discussing the examples of genitive pied-piping in (1) and (2), I discuss how the proposed analysis can be extended to other related cases, such as examples in which the *wh*-word is in a Prepositional Phrase (PP) and corresponding examples in which no pied-piping has taken place. Finally, I end with a brief conclusion in section 5.

2. Tree Adjoining Grammar

Tree Adjoining Grammar (TAG) is a tree-rewriting system, first formally defined in Joshi, Levy and Takahashi (1975). In TAG for natural language, the elementary objects are lexicalized trees called elementary trees that represent extended projections of lexical predicates anchoring the trees and encapsulate syntactic/semantic arguments of the lexical anchor. That is, an elementary tree is an extended projection of a single lexical head with all and only its argument slots appearing as frontier nonterminals. The elementary trees in TAG are therefore said to possess an Extended Domain of Locality.

Frank (2002) formulates the extended projection property of elementary trees as a Condition on Elementary Tree Minimality (CETM), and states that "the syntactic heads in an elementary tree and their projections must form an extended projection of a single lexical head" (p. 54). Following Grimshaw (1991), Frank takes extended projections of a lexical head to include the projections of all functional heads that embed it. This means that an elementary tree anchoring a verb can not only project to Verb Phrase (VP) but to Tense Phrase (TP) and Complementizer Phrase (CP), and an elementary tree anchoring a noun can not only project to Noun Phrase (NP) but to DP and PP. Further, the fundamental thesis in TAG for natural language is that "every syntactic dependency is expressed locally within a single elementary tree" (Frank 2002, p. 22). This allows for a syntactic dependency created by movement to occur within an elementary tree, but not across elementary trees.

The trees in Figure 1 are all examples of well-formed elementary trees. (αsolved) is an elementary tree because it is an extended projection of the lexical predicate *solved* and has argument slots for the subject and the object marked by the downward arrow ($\downarrow$).[1] Moreover, the movement of the subject DP from [Spec,VP] to [Spec,TP], following the VP-internal Subject Hypothesis (Koopman and Sportiche, 1991), is an operation internal to the elementary tree, and therefore represents a syntactic dependency localized to the elementary tree. (αa_student) and (αthe_problem) are valid elementary trees because these DP trees each contain a single lexical head

(N), D being a functional head, which can form an extended projection with a DP, in line with the DP Hypothesis (Abney, 1987).[2]

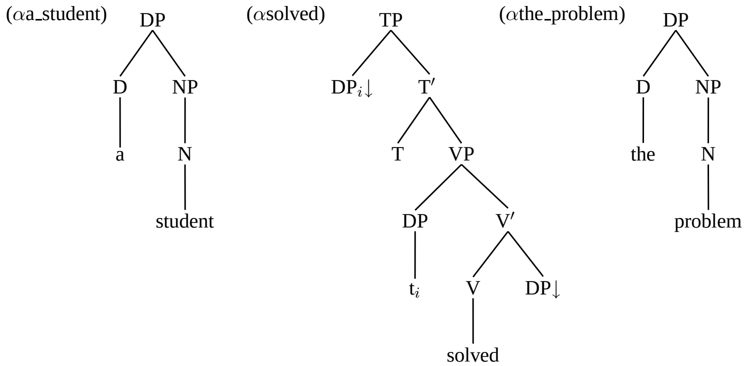


Figure 1. Initial trees in TAG

Elementary trees are of two types: initial trees and auxiliary trees. Initial trees are minimal linguistic structures, for example, trees containing the phrasal structure of simple clauses and DPs. The elementary trees in Figure 1 are examples of initial trees. Auxiliary trees are used to introduce recursive structures, for example, adjuncts or other recursive portions of the grammar. Auxiliary trees have a special non-terminal node called the foot node (marked with an asterisk) among the leaf nodes, which has the same label as the root node of the tree. The auxiliary trees in Figure 2 are well-formed elementary trees, as CETM requires only that syntactic heads and their projections form an extended projection, rendering the presence of the VP root node in (βquickly) and the NP root node in (βsmart) consistent with CETM. Further, following Frank (2002), we can count VP* in (βquickly) and NP* in (βsmart) as arguments of the lexical anchor, as the process of theta-identification (Higginbotham, 1985) obtains between them and the lexical anchor.

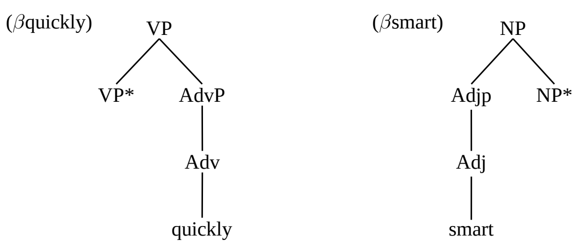


Figure 2. Auxiliary trees in TAG

These elementary trees are combined through two derivational operations: substitution and adjoining. In the substitution operation, the root node on an initial tree is merged into a matching non-terminal leaf node marked for substitution (↓) in another tree. This is illustrated in Figure 3.
In an adjoining operation, an auxiliary tree is grafted onto a non-terminal node in another elementary tree that matches the root and foot nodes of the auxiliary tree.

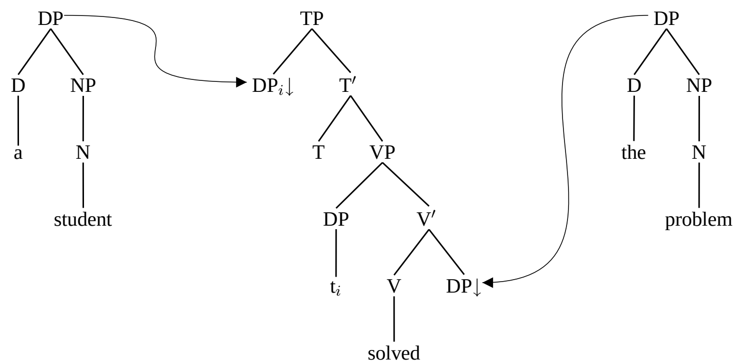


Figure 3. Substitution in TAG

For example, Figure 4 illustrates (βquickly) adjoining to the VP node in (αsolved), and (βsmart) adjoining to the NP node in (αa_student) which in turn substitutes into (αsolved).

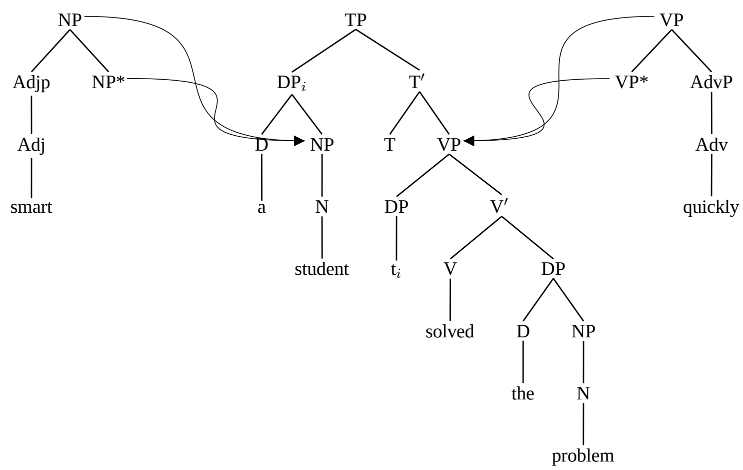


Figure 4. Adjoining in TAG

TAG derivation produces two structures: a derived tree and a derivation tree. The derived tree is the conventional phrase structure tree. For instance, combining the elementary trees in Figures 1 and 2 through substitution and adjoining as in Figures 3 and 4 generates the derived tree in Figure 5 (left). The derivation tree represents the history of composition of the elementary trees. In a derivation tree, each node is an elementary tree, and the children of a node N represent the trees which are adjoined or substituted into the elementary tree represented by N. The link connecting a pair of nodes is annotated with the location in the parent elementary tree where adjoining or substitution has taken place.[3] An example of a derivation tree is given the in Figure 5 (right). Figure 5 (right) records the history of composition of the elementary trees to produce the derived tree in Figure 5 (left): (βsmart) adjoins to (αa_student) at NP, (αa_student) and (αthe_problem) substitute

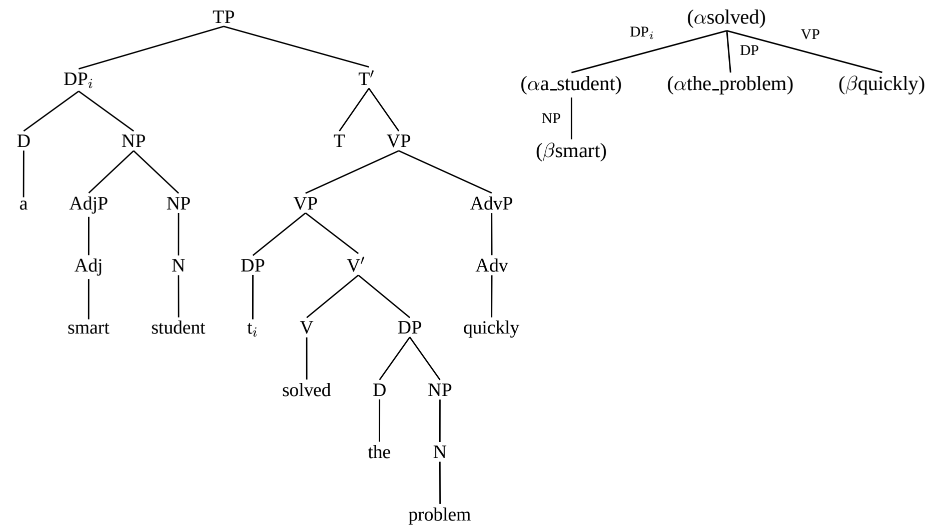


Figure 5. Derived tree and derivation tree in TAG

into (αsolved) at DP$_i$ and DP respectively, and (βquickly) adjoins to (αsolved) at VP.

As explicated in Joshi and Vijay-Shanker (1999) and Joshi and Kallmeyer (2003), the input to compositional semantics in TAG is the derivation tree, not the derived tree. While phrase-structure based compositional semantics computes the meaning of a sentence as a function of the meaning of each node in the syntactic tree, TAG-based compositional semantics computes the meaning of a sentence as a function of the meaning of elementary trees put together to derive the sentence structure. Each syntactic elementary tree is associated with a semantic representation, and following the history of how the elementary trees are put together to derive the sentence structure, the corresponding semantic representation is computed by combining the semantic representations of the elementary trees. Therefore, compositional semantics in TAG is done on the derivation tree, not on the derived tree.

There are two main approaches to doing compositional semantics on the derivation tree: (i) flat semantics (Joshi and Vijay-Shanker, 1999; Kallmeyer and Joshi, 2003; Romero and Kallmeyer, 2005; Kallmeyer and Romero, to appear); and (ii) STAG (Shieber and Schabes, 1990; Shieber, 1994; Abeillé, 1994). Under the flat semantics approach, in the style of Minimal Recursion Semantics (Copestake, Flickinger, Sag and Pollard, 2005), the main operation for semantic composition is the conjunction of the semantic representations associated with each elementary tree along with the unification of variables contributed by these semantic representations. In Romero and Kallmeyer (2005) and Kallmeyer and Romero (to appear), derivation trees are augmented with feature structures to enforce variable unification. The theory of semantic representations developed by Kallmeyer and Romero has been used in a series of empirical work with success: pied-piping of *wh*-phrases (Kallmeyer and Scheffler, 2004), focus (Babko-Malaya, 2004), questions (Romero, Kallmeyer and Babko-Malaya, 2004), VP coordination (Banik, 2004), among others.

Under the STAG approach, on the other hand, the semantic representations are structured trees with nodes on which substitution and adjoining of other semantic representations can take place. Compositionality in STAG obtains with the requirement that the derivation tree in syntax and the corresponding derivation tree in semantics be isomorphic, as specified in Shieber (1994). This isomorphism requirement guarantees that the derivation tree in syntax determines the meaning components needed for semantic composition, and the way in which these meaning components are combined. The semantic objects and the composition of these objects parallel those already utilized in syntax, and so computing semantics only requires the operations of substitution and adjoining used to build the syntactic structures. These properties of STAG allow us to define a simple and elegant syntax-semantics mapping. This has been shown to be the case by Nesson and Shieber (2006), who provide an STAG analysis for various linguistic phenomena, including quantifier scope, long distance *wh*-movement, subject-to-subject raising, and nested quantifiers and inverse linking. In this paper, I extend the empirical coverage of STAG-based compositional semantics to relative clauses and pied-piping. TAG semantic analyses for relative clauses using flat semantics have been proposed before by Han (2002) and Kallmeyer (2003). However, to obtain the right semantics, both Han and Kallmeyer introduce elementary trees in syntax that violate CETM and other well-formedness conditions on elementary trees. The analysis presented here does not have this problem.

3. STAG and STAG-based Compositional Semantics

In this section, I illustrate the framework of STAG and STAG-based compositional semantics and clarify my assumptions, using a simple sentence that contains an existential quantifier and an attributive adjective as in (6), and a sentence with two quantified DPs that show scope ambiguity as in (7).

(6) John kicked a tall boy.

(7) Some student likes every course. ($\exists > \forall, \forall > \exists$)

I use STAG as defined in Shieber (1994). In STAG, each syntactic elementary tree is paired with one or more semantic trees that represent its logical form with links between matching nodes. A synchronous derivation proceeds by mapping a derivation tree from the syntax side to an isomorphic derivation tree on the semantics side, and is synchronized by the links specified in the elementary tree pairs. In the tree pairs given in Figure 6, the trees on the left side are syntactic elementary trees and the ones on the right side are semantic trees. In the semantic trees, F stands for formulas, R for predicates and T for terms. I assume that these nodes are typed (e.g., the F node in (α'kicked) has type t), and I represent predicates as unreduced λ-expressions. The linked nodes are shown with boxed numbers. For the sake of simplicity, in the elementary tree pairs, I only include links that are relevant for the derivation of given examples.

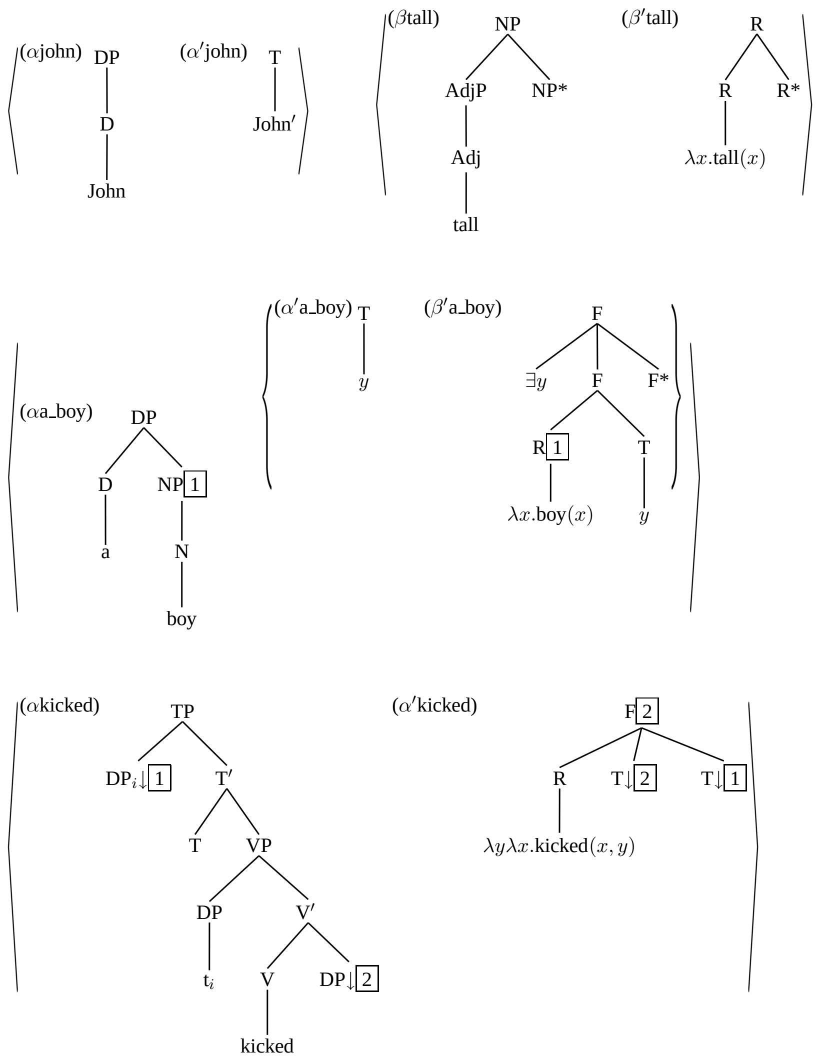


Figure 6. Elementary trees for *John kicked a tall boy*

Figure 6 contains elementary trees required to generate the syntactic structure and the logical form of (6). All the syntactic elementary trees satisfy Frank's (2002) CETM. The proper name tree in (αjohn) is paired with a tree representing a term on the semantics side, and the attributive adjective tree in (βtall) is paired with an auxiliary tree on the semantics side that represents a one-place predicate to be adjoined to another one-place predicate. For quantified DPs, I follow Shieber and Schabes (1990) and Nesson and Shieber (2006), and use tree-local Multi-Component Tree Adjoining Grammar (MC-TAG), an extension of TAG, on the semantics side. In tree-local MC-TAG, the basic objects of derivation are not only individual elementary trees, but also (possibly a singleton) set of such trees, called a multi-component set. At each step in a derivation, all of the trees in a multi-component set must adjoin or substitute simultaneously into a single elementary tree. Restricted in this way, MC-TAG is shown to be identical to basic TAG in both weak and strong generative power (Weir, 1988). Thus, the DP in (αa_boy) is paired with a multi-component set $\{$(α'a_boy), (β'a_boy)$\}$ on the semantics side: (α'a_boy) provides an argument variable, and (β'a_boy) provides an existential quantifier with the restriction and scope. The transitive tree in (αkicked) is paired with a semantic tree representing a formula that consists of a two-place predicate and two term nodes. The links, notated with boxed numbers, guarantee that whatever substitutes into DP$_i$, its corresponding semantic tree will substitute into the term node marked with $\boxed{1}$, and whatever substitutes into DP is paired up with a multi-component set on the semantics side where one of the components will substitute into the term node marked with $\boxed{2}$ and the other will adjoin to the F node marked with $\boxed{2}$. The syntactic and semantic derivation trees are given in Figure 7, and the derived trees are given in Figure 8. Technically, there is only one derivation tree because the syntactic and semantic derivations are isomorphic. In this paper, I provide two derivation trees (one for syntax and the other for semantics) throughout to make the tree-local derivation explicit.[4]

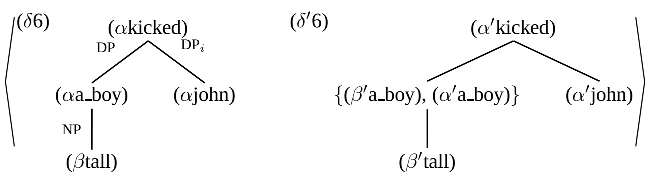


Figure 7. Derivation trees for *John kicked a tall boy*

The semantic derived trees can be reduced by applying λ-conversion, as the nodes dominate typed λ-expressions and terms. When reducing the semantic derived trees, in addition to λ-conversion, I propose to use Predicate Modification, as defined in Heim and Kratzer (1998) in (8).

(8) Predicate Modification

If α has the form $[_{\alpha}\ \beta\ \gamma]$, and $[\![\beta]\!]^s$ and $[\![\gamma]\!]^s$ are both in $D_{<e,t>}$,

then $[\![\alpha]\!]^s = \lambda x_e [\![\beta]\!]^s(x) \wedge [\![\gamma]\!]^s(x)$.

The application of Predicate Modification and λ-conversion reduces (γ'6) to the formula in (9).

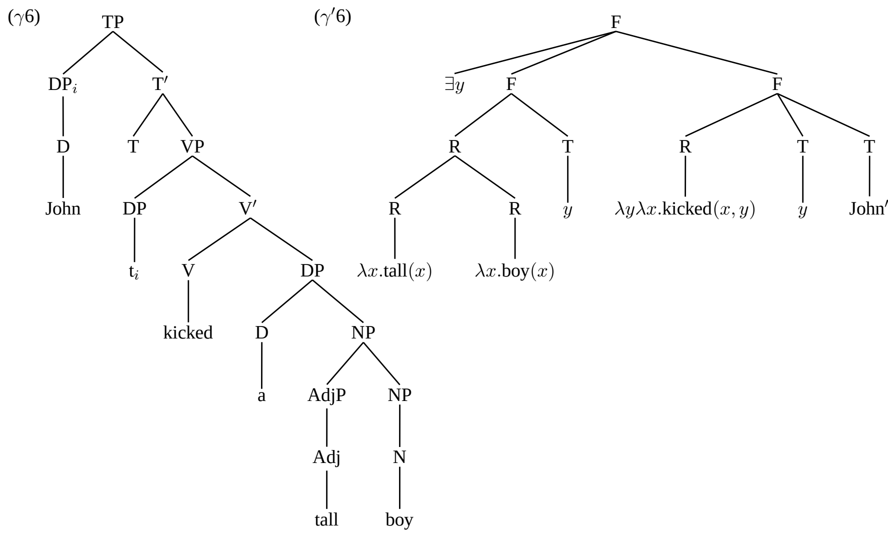


Figure 8. Derived trees for *John kicked a tall boy*

(9) $\exists y[\text{tall}(y) \wedge \text{boy}(y)]\ [\text{kicked}(\text{John}', y)]$

Figure 9 contains elementary trees required to generate the syntactic structure and the logical form of (7). Following Nesson and Shieber (2006), scope ambiguity is accounted for by incorporating multiple adjoining in the derivation on the semantics side. Multiple adjoining allows multiple auxiliary trees to be adjoined at the same node in an elementary tree, as defined in Schabes and Shieber (1994). So, on the semantics side, both (β'some_student) and (β'every_course) adjoin to the root F node in (α'likes). Links $\boxed{1}$ and $\boxed{2}$ marked on the root F node in (α'likes) guarantee this multiple adjoining. Note that the order in which the two trees adjoin to the F node in (α'likes) is unspecified, thereby providing an underspecified representation for scope ambiguity in the semantic derivation tree, as in (δ'7) in Figure 10. This therefore produces multiple semantic derived trees, each representing a different scopal interpretation: (γ'7a) in Figure 11 represents the $\forall > \exists$ reading and (γ'7b) in Figure 12 represents the $\exists > \forall$ reading. On the syntax side, (αsome_student) substitutes into the subject DP$_i$ in (αlikes), and (αevery_course) substitutes into the object DP in (αlikes), as represented in the syntactic derivation tree (δ7). This produces a single syntactic derived tree in (γ7) in Figure 11 (for $\forall > \exists$) and repeated in Figure 12 (for $\exists > \forall$).

The application of Predicate Modification and λ-conversion reduces (γ'7a) to the formula in (10a), and (γ'7b) to the formula in (10b).

(10) a. $\forall y[\text{course}(y)]\ [\exists x[\text{student}(x)]\ [\text{likes}(x, y)]]$

b. $\exists x[\text{student}(x)]\ [\forall y[\text{course}(y)]\ [\text{likes}(x, y)]]$

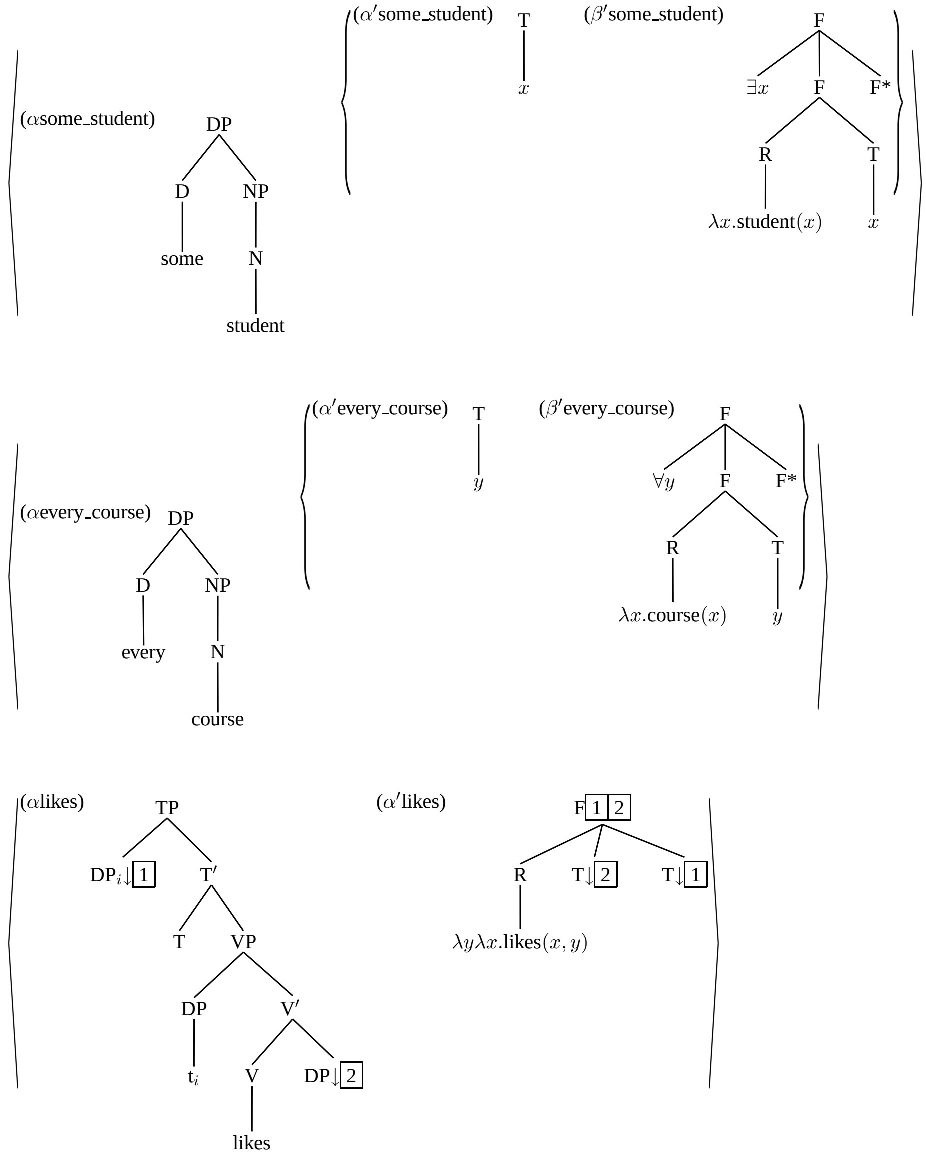


Figure 9. Elementary trees for *Some student likes every course*

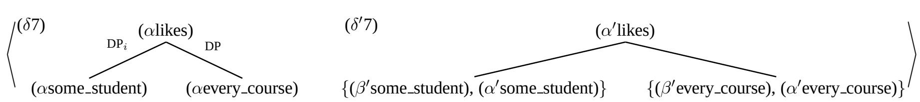


Figure 10. Derivation trees for *Some student likes every course*

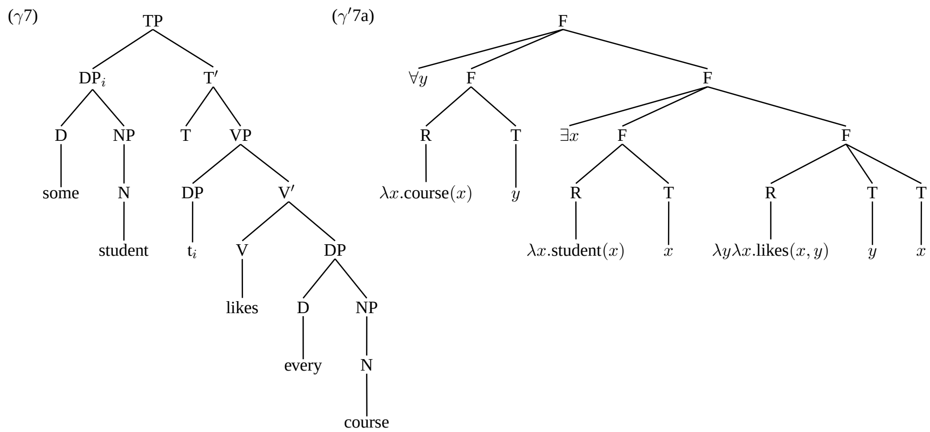


Figure 11. Derived trees for *Some student likes every course*: $\forall > \exists$

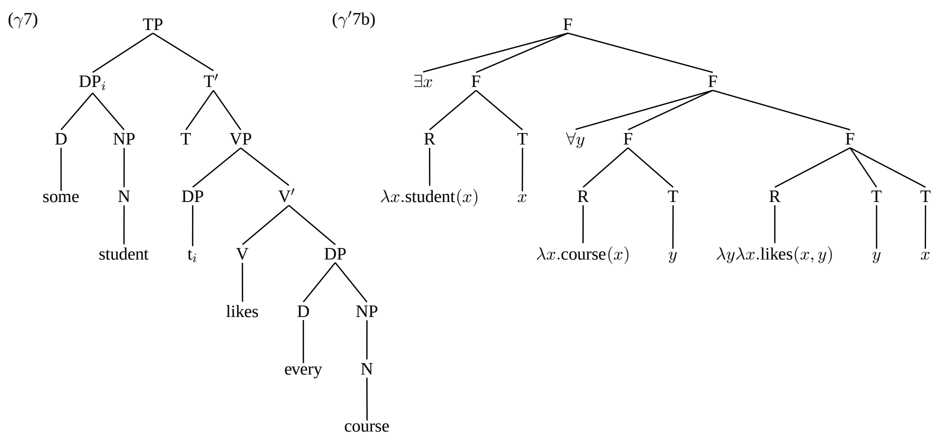


Figure 12. Derived trees for *Some student likes every course*: $\exists > \forall$

4. An STAG analysis of pied-piping in relative clauses

In this section, I present my analysis of relative clauses and pied-piping. I first discuss, in section 4.1, how the proposed analysis accounts for genitive pied-piping exemplified in (1) and (2). I then extend the analysis to cases in which the *wh*-word is embedded in a PP in section 4.2 and corresponding examples in which no pied-piping has taken place in section 4.3.

4.1. *wh*-WORD EMBEDDED IN A GENITIVE DP

I propose the elementary tree pairs in Figure 13 for the syntactic derivation and semantic composition of the relative clause in (1), repeated here as (11).

(11) a boy [[$_{DP}$ whose brother]$_i$ Mary hit t$_i$]

On the syntax side, (βhit) represents an object relative clause. It is anchored by the lexical head of the clause *hit* with two argument slots, DP_i for the subject and DP_j for the object. This tree satisfies CETM, as CP and TP are extended projections of the verb *hit*, and the presence of the NP root node does not violate CETM because CETM requires only that syntactic heads and their projections form an extended projection. Within this elementary tree, DP_j has been moved from the base object position to [Spec,CP] to represent object relativization. This is an auxiliary tree to be adjoined to an NP anchoring the head noun *boy*, which is the noun that it modifies. (αwho) represents the relative pronoun. It substitutes into DP_j in (βhit), and the pied-piping of the rest of the DP is achieved by adjoining (β's_brother) to (αwho). The tree in (β's_brother) is a widely-accepted genitive structure according to the DP hypothesis, where the genitive *'s* heads the DP tree. This satisfies CETM, as a DP is an extended projection of a noun. Substituting (αmary) into DP_i in (βhit) completes the derivation of the relative clause. The derivation tree for the relative clause is given in (δ1) in Figure 14 and the derived tree is given in (γ1) in Figure 15.

Semantically, we must make sure that the variable coming from the *wh*-word is also the one being predicated of the head noun (*boy* in (11)), and yet the same variable does not serve as an argument of the predicate (*hit* in (11)) in the relative clause. I argue that the introduction of a generalized quantifier (GQ) node in the semantic tree in (β'who) and an adjoining of (β''s_brother) to the GQ node give us the desired result. I define the logical form of a *wh* relative pronoun as an auxiliary tree given in (β'who). In (β'who), λx binds x in the generalized quantifier, $\lambda P.P(x)$. The logical form of the object relative clatuse (β'hit) is defined as an auxiliary tree anchoring a two place predicate, to which the logical form of the subject DP (α'mary) will substitute and the logical form of the relative pronoun (β'who) will adjoin. Adjoining (β'who) to (β'hit) essentially has the effect of abstracting over the variable coming from the *wh*-word in the relative clause, turning the relative clause into a one-place predicate. This therefore ensures that the relative clause and the head noun are predicating over the same variable, deriving the interpretation of the relative clause as a modifier of the head noun. The meaning of the pied-piped material *'s brother* is added onto the meaning of *who* by adjoining the auxiliary tree defined in (β''s_brother) to the GQ node in (β'who). In (β''s_brother), λy ensures that the variable coming from the DP* (*who*) is in some relation with the variable coming from the lexical head of the pied-piped DP (*brother* in *whose brother*), and λQ, by turning *whose brother* into a GQ, ensures that the variable coming from the lexical head of the pied-piped DP (*brother* in *whose brother*) is the argument of the predicate (*hit*) that the DP (*whose brother*) combines with. In (β''s_brother), I use the predicate Rel to refer to the relation the genitive expresses. This is usually a possession relation, but not always. As the exact nature of the relation is determined by the context, I leave this relation underspecified. The derivation tree and the derived tree on the semantics side are given in (δ'1) in Figure 14 and (γ'1) in

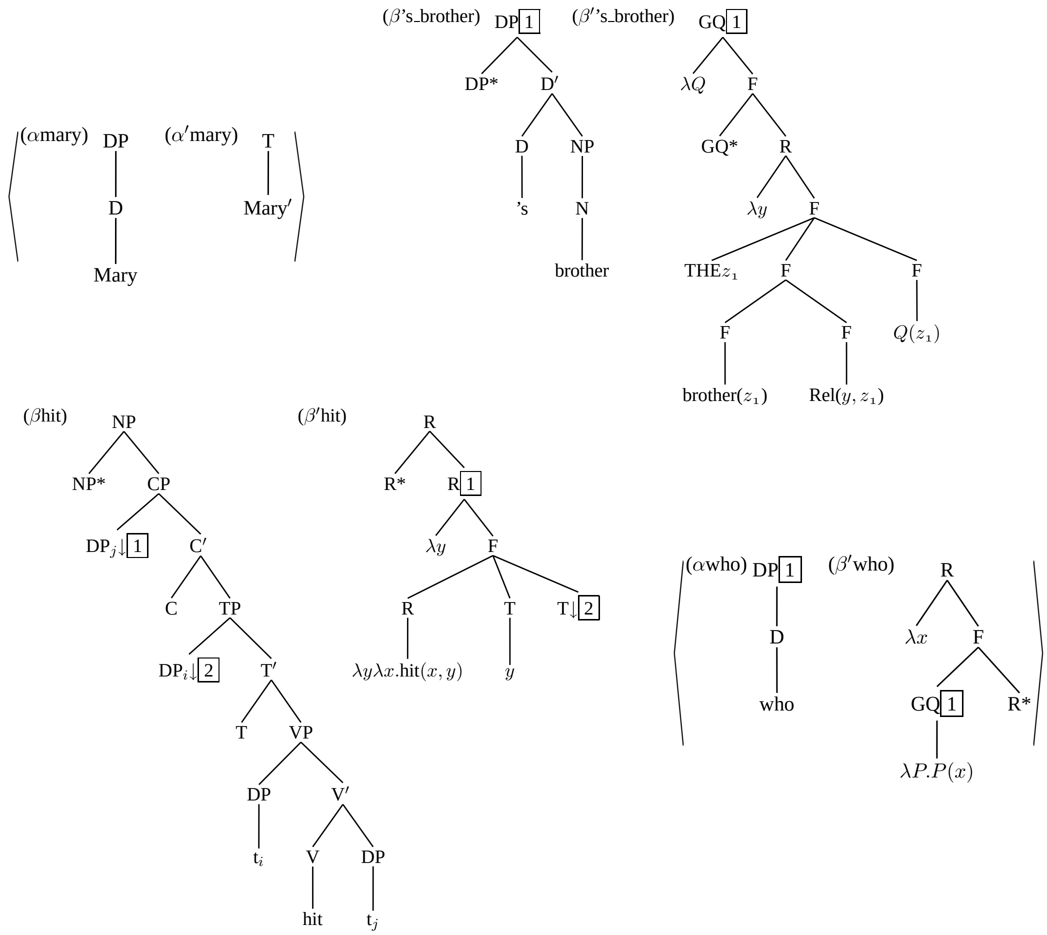


Figure 13. Elementary trees for *whose brother Mary hit*

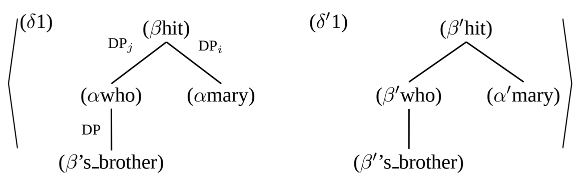


Figure 14. Derivation trees for *whose brother Mary hit*

Figure 15. After all the λ-conversions have applied, $(\gamma'1)$ can be reduced to the expression in (12).

(12) $\lambda x.\text{THE}z_1[[\text{brother}(z_1) \wedge \text{Rel}(x, z_1)]\ [\text{hit}(\text{Mary}', z_1)]]$

The expression in (12) is a one-place predicate which can be paraphrased as a set of all x's such that there is a unique brother z_1 and x is in some relation with z_1 and Mary hit z_1. As the semantics of relative clauses is defined to be a one-place

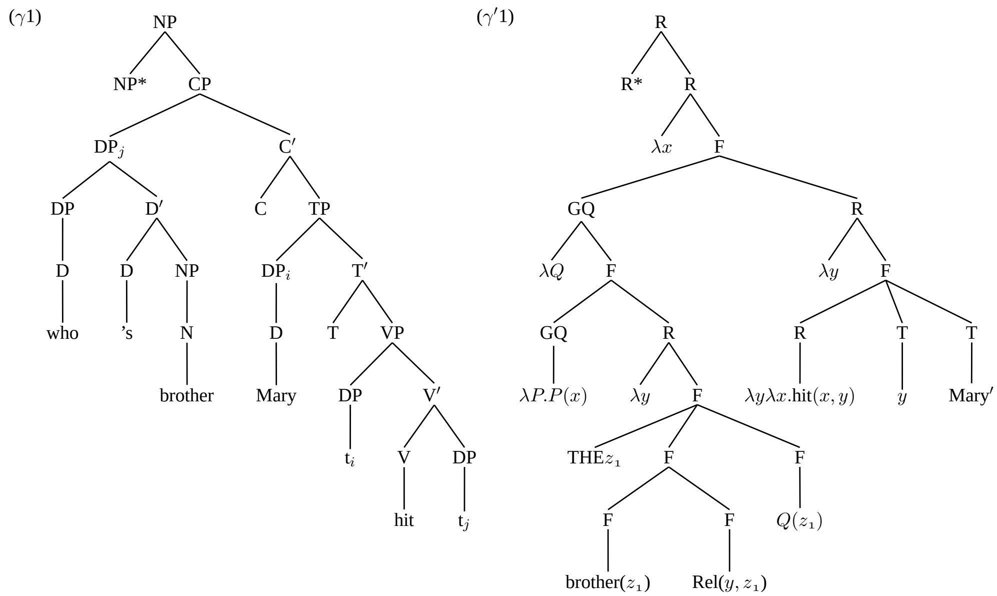


Figure 15. Derived trees for *whose brother Mary hit*

predicate, it is just like the semantics of attributive adjectives. This means that the semantic tree resulting from adjoining ($\gamma'1$) to the logical form of the head noun *boy* can be reduced to the expression in (13), through Predicate Modification.

(13) $\lambda x.\text{boy}(x) \wedge \text{THE}z_1[[\text{brother}(z_1) \wedge \text{Rel}(x, z_1)]\ [\text{hit}(\text{Mary}', z_1)]]$

The proposed semantics for the object relative clause and the relative pronoun also handles simple cases where the relative pronoun is the relativized object, as in (14). The derivation trees and the derived trees are given in Figures 16 and 17. Adjoining the semantic tree ($\gamma'14$) to the logical form of the head noun *boy*, and reducing the output through λ-conversion and Predication Modification will give us the expression in (15).

(14) a boy [who$_i$ Mary hit t$_i$]

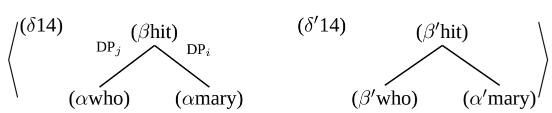


Figure 16. Derivation trees for *who Mary hit*

(15) $\lambda x.\text{boy}(x) \wedge \text{hit}(\text{Mary}', x)$

For the syntactic derivation and the compositional semantics of the relative clause in (2), repeated below as (16), all we need to do is add the tree pair in Figure 18 to the set of elementary tree pairs in Figure 13.

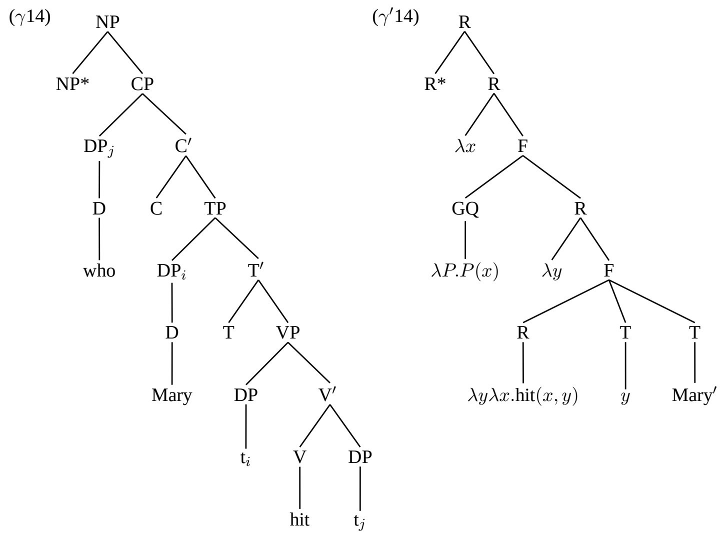


Figure 17. Derived trees for *who Mary hit*

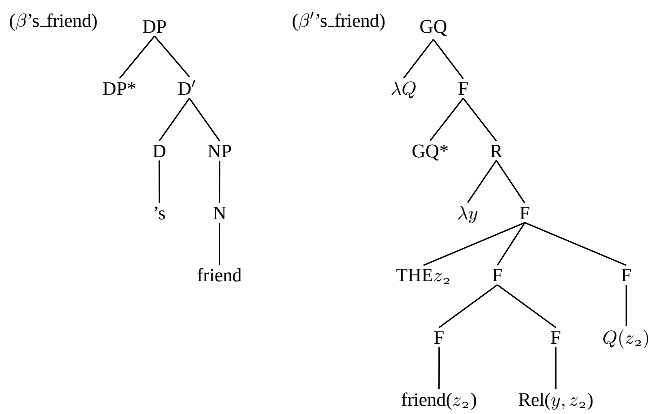


Figure 18. Elementary trees for *'s friend*

(16) a boy [[$_{DP}$ whose brother's friend]$_i$ Mary hit t$_i$]

On the syntax side, (β's_friend) adjoins to (β's_brother) and on the semantics side, (β''s_friend) adjoins to (β''s_brother), as shown in the derivation trees in Figure 19. The derived trees are given in Figure 20.
The semantic derived tree (γ'2) can be reduced to the expression in (17) through λ-conversions. This can be paraphrased as a set of all x's such that there is a unique brother z_1 and x is in some relation with z_1 such that there is a unique friend z_2 and z_1 is in some relation with z_2 and Mary hit z_2.

(17) $\lambda x.\text{THE}z_1[[\text{brother}(z_1) \wedge \text{Rel}(x, z_1)]\ [\text{THE}z_2[[\text{friend}(z_2) \wedge \text{Rel}(z_1, z_2)]$
$[\text{hit}(\text{Mary}', z_2)]]]]$

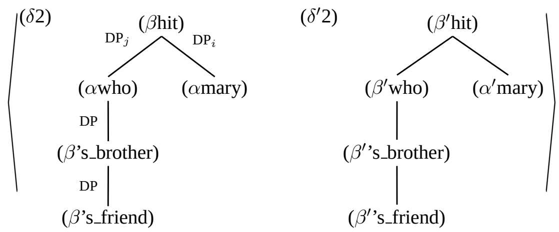


Figure 19. Derivation trees for *whose brother's friend Mary hit*

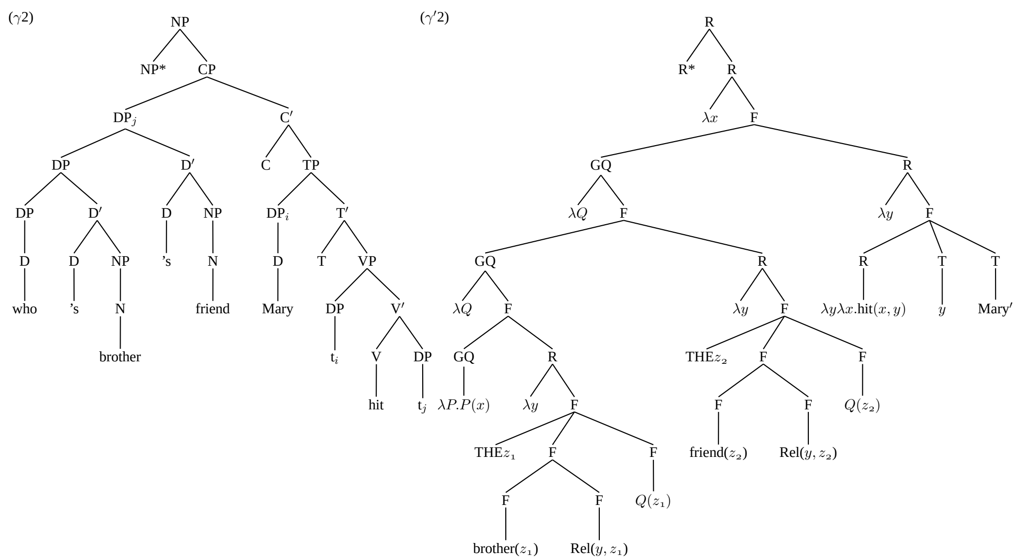


Figure 20. Derived trees for *whose brother's friend Mary hit*

4.2. *wh*-WORD EMBEDDED IN A PP

The proposed syntax and semantics of pied-piping can straightforwardly be extended to cases in which the *wh*-word is embedded in a PP, as in (18).

(18) a boy [[$_{DP}$ the brother [$_{PP}$ of whom]]$_i$ Mary hit t$_i$]

To derive the syntax of (18), we need to change two of the elementary tree pairs in Figure 13 slightly. The elementary tree pairs $<$ $(\alpha$who$), (\beta'$who$)$ $>$ and $<$ $(\beta$'s_brother$), \beta'$'s_brother$)$ $>$ need to be replaced with the pairs in Figure 21. Since the relative pronoun in (18) is *whom*, we use a DP tree anchoring *whom* in (αwhom). The corresponding semantic tree (β'whom) remains exactly the same as before. (βthe_brother_of) represents the pied-piped material in DP. It is a well-formed elementary tree according to CETM: it has a single lexical head *brother* and DP is an extended projection of this head, and PP is not subject to CETM because P is a functional (not a lexical) head. Moreover, DP* is licensed as it is an argument of the lexical head *brother*, as argued in Kroch (1989). The semantics of *the brother of whom* is equivalent to *whose brother*, and therefore, we pair up (βthe_brother_of) with the exact same semantic tree as (β''s_brother).

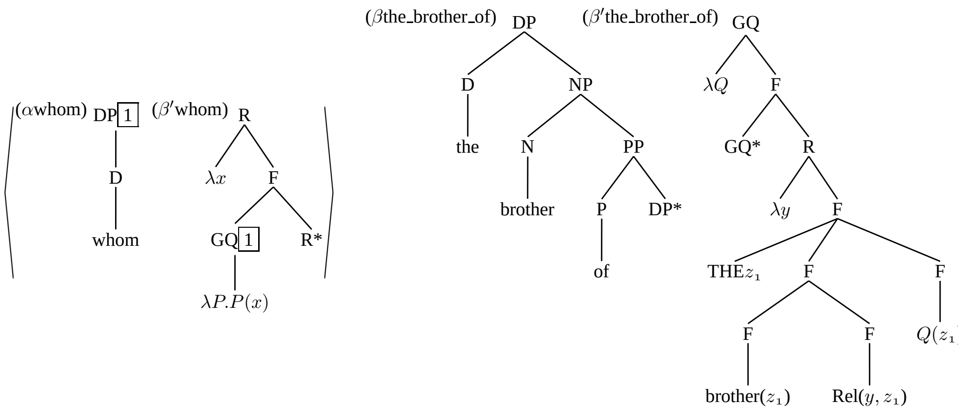


Figure 21. Elementary trees for *whom* and *the brother of*

The derivation trees for the relative clause in (18) are given in Figure 22. They look exactly the same as the ones for the relative clause in (11), except for the names of the elementary trees in a few nodes. The derived trees are given in Figure 23. While the syntactic derived tree (γ18) is different from (γ1) in Figure 15 in that the pied-piped DP contains a PP, the semantic derived tree (γ'18) looks exactly the same as (γ'1) in Figure 15. This is as it should be given that the meanings of (11) and (18) are equivalent.

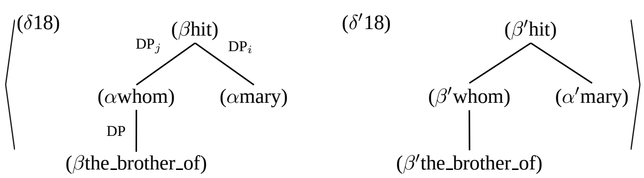


Figure 22. Derivation trees for *the brother of whom Mary hit*

4.3. STRANDING THE LARGER DP

When the larger DP containing the relative pronoun is indefinite or non-specific, the DP can be stranded, as in (19). This effectively gives us a configuration in which a *wh*-word has moved out of a DP.

(19) a boy [whom$_i$ Mary hit [$_{DP}$ a friend of t$_i$]]

Since we now have a DP with an indefinite article, we need the tree pair in Figure 24, for the syntactic derivation and the semantic composition of the relative clause in (19).

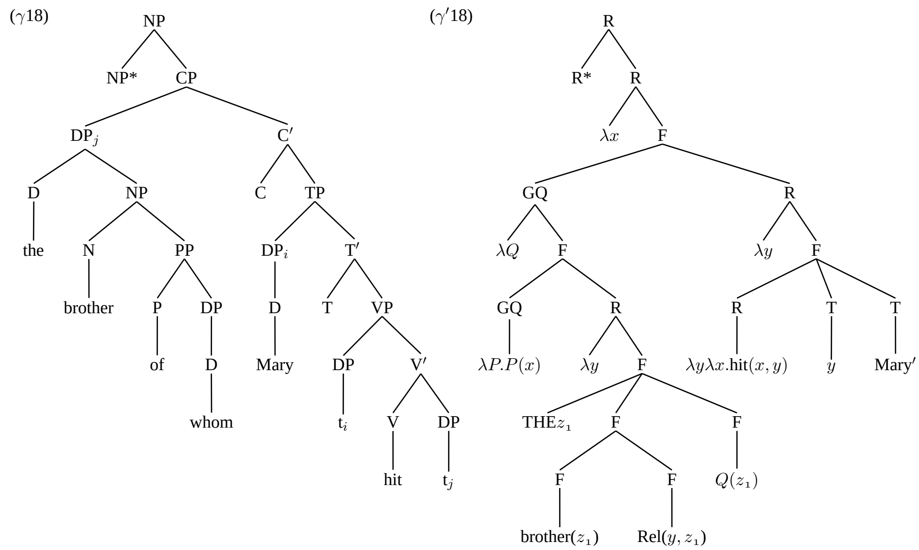


Figure 23. Derived trees for *the brother of whom Mary hit*

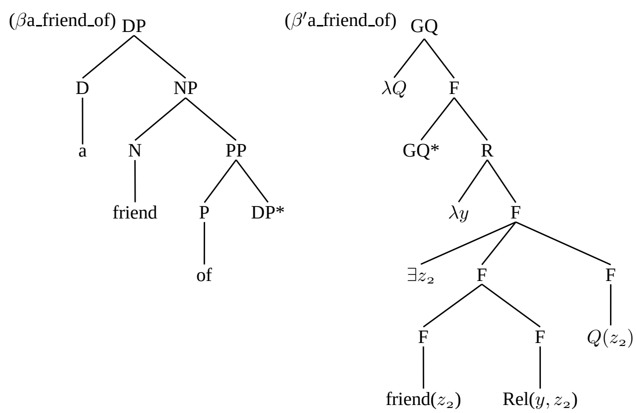


Figure 24. Elementary trees for *a friend of*

Using the semantic tree (β'a_friend_of), the semantic composition of the relative clause in (19) can proceed as before: the semantic tree (β'a_friend_of) adjoins to the semantic tree (β'whom) in Figure 21, which then adjoins to (β'hit) in Figure 13. In the syntax, however, we must make sure that (βa_friend_of) does not adjoin to the relative pronoun *whom*, because if it did, we would end up with the string *a friend of whom*. Instead, what we need is for (βa_friend_of) to adjoin to the DP dominating the trace of the extracted object in (βhit). This however is not a valid derivation in STAG, as elementary trees from two different syntax-semantics pairs, $<(\alpha\text{who}), (\beta'\text{who})>$ and $<(\beta\text{hit}), (\beta'\text{hit})>$, are composing with elementary trees in a single syntax-semantics pair, $<(\beta\text{a\_friend\_of}), (\beta'\text{a\_friend\_of})>$. A slight modification in the syntactic elementary tree for the relative pronoun (αwhom) in

Figure 21 can resolve this issue. I propose to do this by turning (αwhom) into a multi-component set of tree-local MC-TAG, $\{$(αwhom), (βwhom)$\}$ as in Figure 25. An auxiliary tree like (βwhom), which does not dominate any other nodes, is called a degenerate tree. Degenerate trees have been used in Frank (2002) to handle extraction from a *wh*-island such as *[Which car]$_i$ does Sally wonder how to fix t$_i$?* Here, in syntax, to derive the relative clause in (19), (αwhom) substitutes into DP$_j$ in (βhit) as before, and (βwhom) adjoins to the DP dominating the trace of the extracted object in (βhit), as shown in the derivation tree (δ19) in Figure 26. And in semantics, (β'whom) adjoins to (β'hit) as before, as shown in (δ'19) in Figure 26. Subsequently, in syntax (βa_friend_of) adjoins to (βwhom) giving us the DP *a friend of t$_j$*, and in semantics (β'a_friend_of) adjoins to (β'whom). Thus, by using the multi-component set $\{$(αwhom), (βwhom)$\}$, we now have a valid derivation in STAG where elementary trees in a single syntax-semantics pair, $< \{(\alpha\text{whom}), (\beta\text{whom})\}, (\beta'\text{whom}) >$, are composing with elementary trees belonging to another syntax-semantics pair, $< (\beta\text{a\_friend\_of}), (\beta'\text{a\_friend\_of}) >$. The syntactic and the semantic derived trees are given in Figure 27. After λ-conversions, (γ'19) can be reduced to the expression in (20).

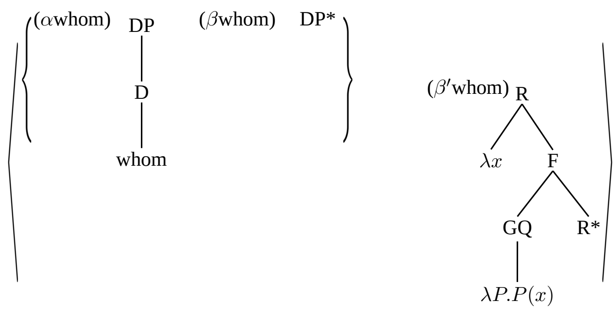


Figure 25. Multi-component set of elementary trees for *whom*

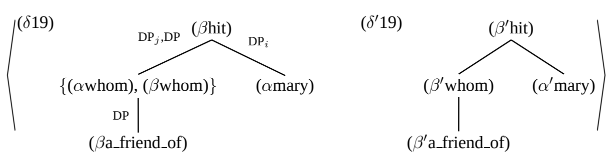
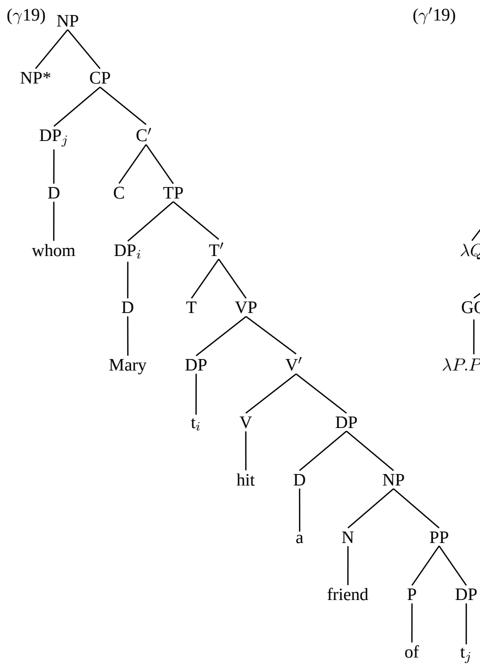


Figure 26. Derivation trees for *whom Mary hit a friend of*

(20) $\lambda x.\exists z_2[[\text{friend}(z_2) \wedge \text{Rel}(x, z_2)]\ [\text{hit}(\text{Mary}', z_2)]]$

It is also possible to mix pied-piping and stranding, resulting in partial stranding, as in (21). Here, out of the DP *a friend of the brother of whom*, *the brother of whom* has been extracted, stranding *a friend of*.

(21) a boy [[the brother of whom]$_i$ Mary hit [$_{DP}$ a friend of t$_i$]]



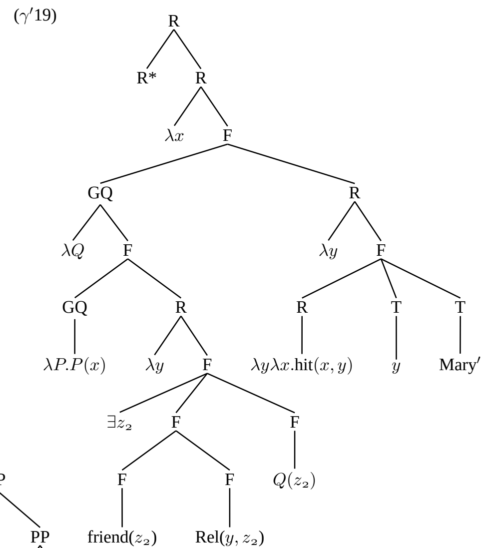


Figure 27. Derived trees for *whom Mary hit a friend of*

Such partial stranding can be handled with set-local MC-TAG, an extension of tree-local MC-TAG. In set-local MC-TAG, at each step in a derivation, all of the elementary trees in a multi-component set must adjoin or substitute into elementary trees belonging to another multi-component set. In the case at hand, we need to turn the elementary tree for *the brother of* in Figure 21 into a multi-component set with a degenerate DP, as in Figure 28. The partial stranding example in (21) can be derived by composing this multi-component set with the multi-component set for *whom* in Figure 25. The isomorphic syntactic and semantic derivation trees are given in Figure 29. In syntax, (αwhom) substitutes into DP$_j$ in (βhit) and (βwhom) adjoins to the DP dominating the trace of the extracted object in (βhit). In semantics, (β'whom) adjoins to (β'hit). Subsequently, in syntax, (βthe_brother_of1) adjoins to the DP of (αwhom) and (βthe_brother_of2) adjoins to the DP of (βwhom). In semantics, (β'the_brother_of) adjoins to (β'whom). Lastly, in syntax, (βa_friend_of) adjoins to the DP of (βthe_brother_of2) and in semantics (β'a_friend_of) adjoins to (β'the_brother_of). This derivation gives us the DP *the brother of whom* in [Spec,CP], and the DP *a friend of* t_j in the object position in the relative clause. The syntactic and semantic derived trees are given in Figure 30, and the reduced λ-expression is given in (22).[5]

(22) $\lambda x.\text{THE}z_1[[\text{brother}(z_1) \wedge \text{Rel}(x, z_1)]\ [\exists z_2[[\text{friend}(z_2) \wedge \text{Rel}(z_1, z_2)]$
$[\text{hit}(\text{Mary}', z_2)]]]]$

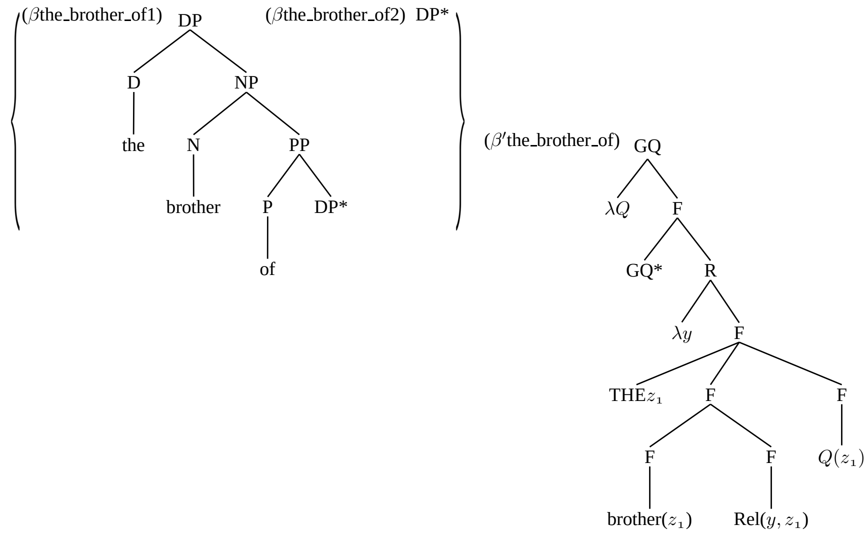


Figure 28. Multi-component set of elementary trees for *the brother of*

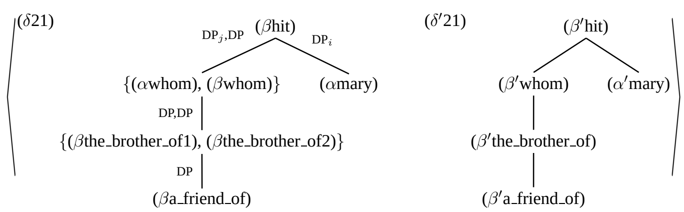


Figure 29. Derivation trees for *the brother of whom Mary hit a friend of*

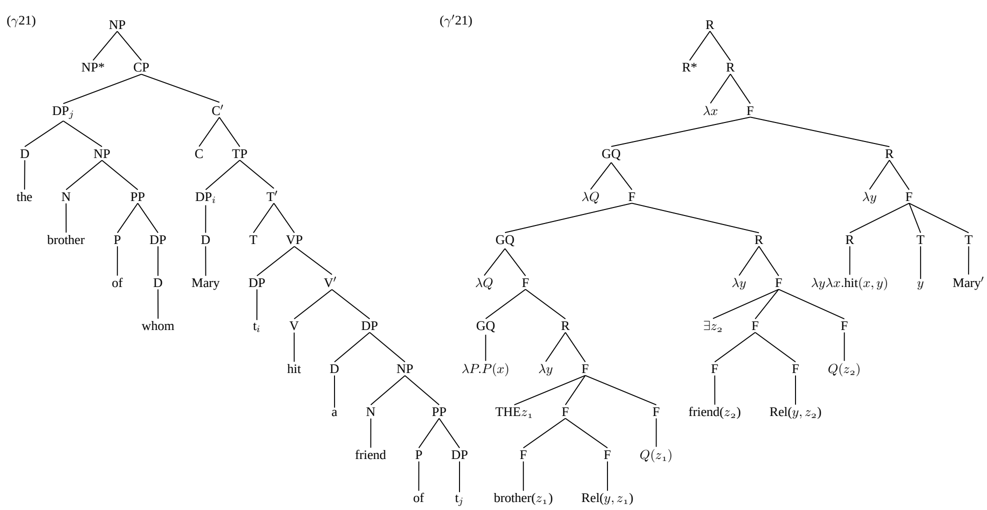


Figure 30. Derived trees for *the brother of whom Mary hit a friend of*

5. Conclusion

I have shown that STAG-based compositional semantics for relative clauses with pied-piping is possible using examples in which the *wh*-word is embedded in a genitive DP, and that the proposed analysis can straightforwardly be extended to cases in which the *wh*-word is embedded in a PP. The main ingredients of the proposed analysis are: in syntax, the pied-piped material adjoins to the *wh*-word, and in semantics, the *wh*-word provides a GQ to which the meaning of the pied-piped material adjoins. I have also shown that a similar analysis can handle cases in which the *wh*-word alone has moved to [Spec,CP], stranding the rest of the DP in situ, if we use tree-local MC-TAG with a multi-component set containing a degenerate DP for the syntax of the relative pronoun. Further, partial stranding can be handled if we use set-local MC-TAG and postulate a multi-component set containing a degenerate DP for the syntax of the pied-pied DP, as well as the multi-component set for the syntax of the relative pronoun. The proposed analysis utilizes composition operations in semantics that are already available in syntax. This makes the syntax-semantics mapping simple and straightforward. It remains as future work to expand further the empirical coverage of STAG-based compositional semantics and compare its coverage to that of a flat semantics based approach.

Acknowledgements

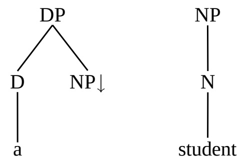
I thank Anoop Sarkar for many helpful discussions on this topic. I also thank Laura Kallmeyer, Dennis Storoshenko and the audience at TAG+8 for comments and questions on the previous version of this paper. I am also extremely indebted to the two anonymous reviewers for their insightful comments that were crucial in improving this paper. All remaining errors are mine.

Notes

[1]Throughout the paper, elementary trees whose names are prefixed with α are initial trees, and those whose names are prefixed with β are auxiliary trees. Names of derivation trees are prefixed with δ, and names of derived trees are prefixed with γ. Semantic trees are named with a prime.

[2]One could ask if elementary trees as in (αa_student) and (αthe_problem) shouldn't be broken down into trees for determiners and trees for NPs, as in (23). Under this approach, an NP tree anchoring a noun would substitute into a DP tree anchoring a determiner. In principle, this is certainly possible. But strictly speaking, this violates Frank's (2002) formulation of CETM, as the DP tree in (23) is a projection of a functional head (D), not a lexical head.

(23)



[3]The location in the parent elementary tree is usually denoted by the Gorn tree address. Here, I use node labels such as DPs or VPs for the sake of simplicity.

[4]In semantic derivation trees, I do not annotate the connections between a mother and a daughter node with the location of adjoining or substitution that has taken place in the mother elementary tree, as this is determined by the links between syntactic and semantic elementary trees.

[5]As pointed out by Laura Kallmeyer and Maribel Romero (personal communication), in examples as in (24), even though the preferred reading is the one where *every boy* takes scope over *a picture*, the analysis presented here as it is only generates the reading in which *a picture* takes scope over *every boy*. This is because in semantics, the semantic tree contributing the universal quantifier, restriction and scope of *every boy* adjoins to the F node of *took*, which ends up in the scope of the existential quantifier contributed by *a picture of*.

(24) a. a soccer star whom every boy took a picture of
b. a soccer star a picture of whom every boy took

A possible analysis may involve set-local MC-TAG in semantics. The semantics of *whom* can be defined as a multi-component set containing a degenerate F tree as well as the tree (β'whom) in Figure 25, and the semantics of *a picture of* can be decomposed into the tree providing a GQ and the tree providing an existential quantifier with its restriction and scope, again forming a multi-component set. Then, the degenerate F tree of *whom* can adjoin to the F node of *took*, to which the existential quantifier tree of *a picture of* and the universal quantifier tree of *every boy* can adjoin. As the order in which the two quantifier trees adjoin to the F node of *took* is not specified, both ‘every>a’ and ‘a>every’ scope can be generated in principle. I leave for future research the exact details of this set-local MC-TAG analysis.

References

Abeillé, Ann: 1994, ‘Syntax or semantics? Handling nonlocal dependencies with MCTAGs or Synchronous TAGs’. *Computational Intelligence* **10**, 471–485.

Abney, Steven: 1987, ‘The English Noun Phrase in its Sentential Aspect’. Ph.D. thesis, MIT.

Babko-Malaya, Olga: 2004, ‘LTAG Semantics of Focus’. In: *Proceedings of TAG+7*. Vancouver, Canada, pp. 1–8.

Banik, Eva: 2004, ‘Semantics of VP Coordination in LTAG’. In: *Proceedings of TAG+7*. Vancouver, Canada, pp. 118–125.

Copestake, Ann, Dan Flickinger, Ivan A. Sag, and Carl Pollard: 2005, 'Minimal Recursion Semantics: An introduction'. *Journal of Research on Language and Computation* **3**(2-3), 281–332.

Frank, Robert: 2002, *Phrase Structure Composition and Syntactic Dependencies*. Cambridge, MA: MIT Press.

Grimshaw, Jane: 1991, 'Extended Projection'. Manuscript, Brandeis University.

Han, Chung-hye: 2002, 'Compositional Semantics for Relative Clauses in Lexicalized Tree Adjoining Grammar'. A talk presented at TAG+6, Venice, Italy, www.sfu.ca/∼chunghye/papers/tag6-rc-slides.pdf.

Heim, Irene and Angelika Kratzer: 1998, *Semantics in Generative Grammar*. Oxford: Blackwell.

Higginbotham, James: 1985, 'On semantics'. *Linguistic Inquiry* **16**, 547–594.

Joshi, Aravind K., L. Levy, and M. Takahashi: 1975, 'Tree Adjunct Grammars'. *Journal of Computer and System Sciences*.

Joshi, Aravind K. and K. Vijay-Shanker: 1999, 'Compositional Semantics with Lexicalized Tree-Adjoining Grammar (LTAG): How much underspecification is necessary?'. In: H. C. Blunt and E. G. C. Thijsse (eds.): *Proceedings of the Third International Workshop on Computational Semantics (IWCS-3)*. Tilberg, pp. 131–145.

Kallmeyer, Laura: 2003, 'LTAG Semantics for Relative Clauses'. In: *Proceedings of the Fifth International Workshop on Computational Semantics (IWCS-5)*. Tilburg.

Kallmeyer, Laura and Aravind K. Joshi: 2003, 'Factoring Predicate Argument and Scope Semantics: Underspecified Semantics with LTAG'. *Research on Language and Computation* **1**, 3–58.

Kallmeyer, Laura and Maribel Romero: to appear, 'Scope and Situation Binding in LTAG using Semantic Unification'. *Research on Language and Computation*.

Kallmeyer, Laura and Tatjana Scheffler: 2004, 'LTAG Analysis for Pied-Piping and Stranding of wh-Phrases'. In: *Proceedings of TAG+7*. Vancouver, Canada, pp. 32–39.

Koopman, Hilda and Dominique Sportiche: 1991, 'The position of subjects'. *Lingua* **85**, 211–258.

Kroch, Anthony: 1989, 'Asymmetries in long-distance extraction in a Tree Adjoining Grammar'. In: Mark Baltin and Anthony Kroch (eds.): *Alternative Conceptions of Phrase Structure*. Chicago: University of Chicago Press, pp. 66–98.

Nesson, Rebecca and Stuart M. Shieber: 2006, 'Simpler TAG Semantics through Synchronization'. In: *Proceedings of the 11th Conference on Formal Grammar*. Malaga, Spain, CSLI.

Romero, Maribel and Laura Kallmeyer: 2005, 'Scope and Situation Binding in LTAG using Semantic Unification'. In: *Proceedings of the Sixth International Workshop on Computational Semantics (IWCS-6)*. Tilburg.

Romero, Maribel, Laura Kallmeyer, and Olga Babko-Malaya: 2004, 'LTAG Semantics for Questions'. In: *Proceedings of TAG+7*. Vancouver, Canada, pp. 186–193.

Schabes, Yves and Stuart M. Shieber: 1994, 'An Alternative Conception of Tree-Adjoining Derivation'. *Computational Linguistics* pp. 167–176.

Shieber, Stuart: 1994, 'Restricting the Weak-Generative Capacity of Synchronous Tree-Adjoining Grammars'. *Computational Intelligence* **10**(4).

Shieber, Stuart and Yves Schabes: 1990, 'Synchronous Tree Adjoining Grammars'. In: *Proceedings of COLING'90*. Helsinki, Finland.

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