Syntax and Semantics of *It*-Clefts: a Tree Adjoining Grammar Analysis*

Chung-hye Han and Nancy Hedberg
Department of Linguistics
Simon Fraser University
{chunghye, hedberg}@sfu.ca

Abstract

In this paper, we examine two main approaches to the syntax and semantics of *it*-clefts as in *It was Ohno who won*: an expletive approach where the cleft pronoun is an expletive and the cleft clause bears a direct syntactic or semantic relation to the clefted constituent, and a discontinuous constituent approach where the cleft pronoun has a semantic content and the cleft clause bears a direct syntactic or semantic relation to the cleft pronoun. We argue for an analysis using Tree Adjoining Grammar (TAG) that captures the best of both approaches. We use Tree-Local Multi-Component Tree Adjoining Grammar (MC-TAG) to propose a syntax of *it*-clefts and Synchronous Tree Adjoining Grammar (STAG) to define a compositional semantics on the proposed syntax. It will be shown that the distinction TAG makes between the derivation tree and the derived tree, the extended domain of locality characterizing TAG, and the direct syntax-semantics mapping characterizing STAG allow for a simple and straightforward account of the syntax and semantics of *it*-clefts, capturing the insights and arguments of both the expletive and the discontinuous constituent approaches. Our analysis reduces the syntax and semantics of *it*-clefts to copular sentences containing definite description subjects, such as *The person that won is Ohno*. We show that this is a welcome result, as evidenced by the syntactic and semantic similarities between *it*-clefts and the corresponding copular sentences.

*We thank the audience at TAG+8 in Sydney, 2006, for comments and questions on the previous version of this paper. We are also extremely indebted to the two anonymous reviewers for their insightful comments that were crucial in improving this paper. All remaining errors are ours. This work was supported by SSHRC 410-2003-0544 and NSERC RGPIN341442 to Han, and SSHRC 410-2007-0345 to Hedberg.

1 Introduction

The extant literature on the syntax of *it*-clefts, as in (1), can be classified into two main approaches. First, the cleft pronoun *it* is an expletive, and the cleft clause bears a direct syntactic or semantic relation to the clefted constituent, such as one of predication (Jespersen, 1937; Chomsky, 1977; Williams, 1980; Delahunty, 1982; Rochemont, 1986; Heggie, 1988; Delin, 1989; É. Kiss, 1998). Second, the cleft clause bears a direct syntactic or semantic relation to the cleft pronoun and is spelled out after the clefted constituent through extraposition or by forming a discontinuous constituent with the cleft pronoun (Jespersen, 1927; Akmajian, 1970b; Emonds, 1976; Gundel, 1977; Wirth, 1978; Hedberg, 1990; Percus, 1997; Hedberg, 2000). Under this second approach, the cleft pronoun is not necessarily expletive but rather has a semantic function such as that of a definite article.

(1) It was OHNO [who won].
cleft pronoun + copula + clefted constituent + cleft clause

In this paper, we argue for an an analysis using Tree Adjoining Grammar (TAG) that captures the best of both traditional analyses by making use of the distinction in TAG between the derivation tree on which syntactic dependencies between elementary objects and compositional semantics are defined, and the derived tree on which aspects of surface constituency are defined. An illustration of the derivation tree and derived tree in TAG is given in section 3.1. In our analysis, as in the expletive approach, at the level of surface syntax (the derived tree), the clefted constituent and cleft clause form a syntactic constituent. As in the discontinuous constituent approach, however, at the level of syntactic dependencies (the derivation tree), the cleft pronoun and the cleft clause form a syntactic unit, and a semantic unit as a definite description. This aspect of our analysis reduces the syntax and semantics of *it*-clefts to copular sentences containing definite description subjects. We show that this reduction is supported by the fact that *it*-clefts and the corresponding copular

sentences pattern alike both syntactically and semantically. In particular, we use Tree-Local Multi-Component Tree Adjoining Grammar (MC-TAG) to propose a syntax of *it*-clefts and Synchronous Tree Adjoining Grammar (STAG) to define a compositional semantics on the proposed syntax. It will be shown that the distinction TAG makes between the derivation tree and the derived tree, the extended domain of locality characterizing TAG and the direct syntax-semantics mapping characterizing STAG allow for a simple and straightforward account of the syntax and semantics of *it*-clefts, capturing the insights and arguments of both the expletive and the discontinuous constituent approaches.

The paper is organized as follows. In section 2, we present arguments supporting the discontinuous constituent analysis as well as some arguments supporting the expletive analysis. We also discuss connectivity effects in *it*-clefts and parallel effects in copular sentences instantiated by binding and agreement. In section 3, we introduce the basics of TAG for doing natural language syntax and present our TAG analysis of the syntax of *it*-clefts. In section 4, we introduce STAG and show how compositional semantics is done using STAG, and present our analysis of the semantics of *it*-clefts. In section 5, we show how our TAG analysis can account for the connectivity effects in *it*-clefts instantiated by binding and agreement.

2 The tension between the expletive and the discontinuous constituent analyses

In this section, we review five main syntactic and semantic properties of *it*-clefts: semantic content of the cleft pronoun, internal structure of the cleft clause, presence of existential and exhaustive presuppositions, presence of equative and predicational readings, and connectivity. For each property, we discuss how the expletive analysis and the discontinuous constituent analysis fare. The arguments presented in this section are taken from the existing literature on *it*-clefts.

First, it has been shown in Hedberg (1990; 2000) that the cleft pronoun can be replaced with *this* or *that*, as in (2), depending on the discourse contextual interpretation of the cleft clause. The fact that the choice of the cleft pronoun is subject to pragmatic constraints indicates that the cleft pronoun is not an expletive element.

(2) a. This is not Iowa we're talking about. (Hedberg 2000, ex. 17)

b. That's the French flag you see flying over there. (Hedberg 2000, ex. 20)

In (2), the proximal demonstrative pronoun is selected when the content of the cleft clause indicates that the referent of the clefted constituent is close to the speaker, and the distal demonstrative is selected when the content of the cleft clause indicates that the referent is far from the speaker. Reversing the cleft pronouns would lead to infelicity. The discontinuous constituent analysis allows the cleft pronoun to be treated as having the semantic content of a determiner. Thus, we can view the cleft pronoun and cleft clause in (2) as working together to function as a demonstrative description as in (3).

(3) a. This [place] we're talking about is not Iowa.

b. That [thing] you see flying over there is the French flag.

Second, the cleft clause has the internal structure of a restrictive relative clause. This is supported by the fact that the initial element in the cleft clause may be realized either as a *wh*-word (1) or as *that* (4a), or it may be absent altogether when the gap is not in the subject position (2, 4b). It may even be in the form of a genitive *wh*-word as in (4c).

(4) a. It was Ohno that won.

b. It was Ohno Ahn beat.

c. It was Ohno whose Dad cheered.

The cleft clause, however, does not relate to the clefted constituent in the way that a restrictive relative clause relates to its head noun, as first noted in Jespersen (1927). This is

because the clefted constituent can be a proper noun, unlike a head noun modified by a restrictive relative clause, as illustrated in (5). Many expletive analyses (e.g., Delahunty, 1982; Rochemont, 1986; Heggie, 1988) thus do not consider the cleft clause to have the internal structure of a restrictive relative clause. The discontinuous constituent analysis, on the other hand, allows the cleft clause to be treated as such, as argued for in Hedberg (1990), because it assumes that the relative clause forms a constituent with the cleft pronoun.

(5) * Ohno that won is an American.

Even so, as pointed out first in Delahunty (1982), there is some syntactic evidence that the clefted constituent and the cleft clause do form a surface syntactic constituent. The examples in (6), from Hedberg (2000), show that the two together can be deleted as a unit, as in (6a), and coordinated as a unit, as in (6b).

(6) a. I said it should have been [Bill who negotiated the new contract], and it should have been.

b. It must have been [Fred that kissed Mary] but [Bill that left with her].

It will be shown in section 3.2 that our analysis resolves this tension between the discontinuous constituent analysis and the expletive analysis by making use of TAG's distinction between the derivation tree, on which compositional semantics and syntactic dependencies between elementary objects are defined, and the derived tree, on which surface syntactic relations are defined. On our analysis, the clefted constituent and the cleft clause form a constituent in the derived tree, and the cleft pronoun and the cleft clause form a syntactic unit in the derivation tree.

Third, *it*-clefts pattern with copular sentences containing definite description subjects syntactically and semantically. Semantically, *it*-clefts have existential and exhaustive presuppositions, just as definite descriptions do, as pointed out in Percus (1997) and Hedberg (2000). The inference in (7c) associated with (7a) survives in the negative counterpart in

(7b). This is exactly the way the presupposition associated with the definite description *the king of France* behaves: the presupposition spelled out in (8c) survives in both the affirmative (8a) and the negative counterpart in (8b).

(7) a. It was Ohno who won.

b. It was not Ohno who won.

c. Someone won, and only one person won.

(8) a. The king of France is bald.

b. The king of France is not bald.

c. There is one and only one king of France.

Both Percus and Hedberg argue that this parallelism between definite descriptions and *it*-clefts can be accounted for if the cleft pronoun and the cleft clause form a semantic unit, with *it* playing the role of the definite article and the cleft clause the descriptive component. What this translates to syntactically is that the cleft clause is a restrictive relative clause which is situated at the end of the sentence, forming a discontinuous constituent with the cleft pronoun. On this view, the syntax and semantics of *it*-clefts reduce to that of copular sentences with definite description subjects.

Fourth, it has been observed that *it*-clefts can have equative and predicational interpretations (Ball, 1977; DeClerck, 1988; Hedberg, 1990, 2000), both of which are readings attested in simple copular sentences, as shown in (9):

(9) a. The teacher is Sue Johnson.

b. The teacher is a woman.

This observation follows under the discontinuous constituent analysis, as *it*-clefts there reduce to ordinary copular sentences, unlike some expletive analyses where the copula is

treated as a focus marker (É. Kiss, 1998). For instance, (7a) (repeated as (10a)) can be paraphrased as (10b), and corresponds to a typical equative sentence. And (11a) can be paraphrased as (11b), and corresponds to a typical predicational sentence. According to the analysis we will present in section 4, (10a) will be assigned the semantic representation in (10c), and (11a) will be assigned the semantic representation in (11c).

(10) a. It was Ohno who won.

b. The one who won was Ohno.

c. THEz [won(z)] [z = Ohno′]

(11) a. It was a kid who beat John.

b. The one who beat John was a kid.

c. THEz [beat(z, John′)] [kid(z)]

Fifth, Percus (1997) points out that *it*-clefts pattern with copular sentences containing definite description subjects with regard to SELF-anaphor binding and negative polarity item (NPI) licensing. In the absence of c-command, a SELF-anaphor in the clefted constituent position can be bound by an antecedent inside the cleft clause, as shown in (12a). Also a pronoun in the clefted constituent position cannot be bound by an antecedent inside the cleft clause, as shown in (13a). Copular sentences with definite description subjects exhibit the same pattern, as in (12b) and (13b). An NPI can occur in the clefted constituent position, licensed by a matrix negative element, as shown in (14a), but it is not licensed by a negation in the cleft clause, as in (15a). This pattern of NPI licensing is attested in copular sentences, as shown in (14b) and (15b).

(12) a. It was himself$_i$ who John$_i$ nominated.

b. The one that John$_i$ nominated was himself$_i$.

(13) a. * It was him$_i$ who John$_i$ nominated.

b. * The one that John$_i$ nominated was him$_i$.

(14) a. It isn't anyone I know that John saw.

b. The one that John saw isn't anyone I know.

(15) a. * It is anyone I know that John didn't see.

b. * The one that John didn't see is anyone I know.

Since *it*-clefts and copular sentences with definite description subjects exhibit the same pattern of binding and NPI licensing, a uniform explanation for the two cases can be sought if the cleft pronoun and the cleft clause together form a definite description.[1]

The NPI facts are not difficult to explain, as the NPI in (14) is c-commanded by the negative element, and the NPI in (15) is not c-commanded by the negative element. However, the SELF-anaphor in (12) and the pronoun in (13) are at first sight mysterious under the discontinuous constituent analysis. This is an example of connectivity, whereby the clefted constituent appears to behave as it would if it were generated inside the cleft clause, thus lending support for the expletive analysis. In section 5, we present a solution to this problem by incorporating Reinhart and Reuland's (1993) Binding Conditions to our TAG analysis, and also arguing that the SELF-anaphor in (12) is a discourse anaphor of focus.

Agreement facts constitute another example of connectivity, in that when the cleft clause has a subject gap, the verb in the cleft clause agrees in number and person with the clefted constituent. Note also that in equative clefts the copula agrees with the singular cleft pronoun and not with a plural clefted constituent. These facts are shown in (16).

[1]Percus shows that *wh*-clefts differ from both *it*-clefts and copular sentences with definite description subjects in that only in the former can post-copular NPIs be licensed by embedded negation. See the examples in (15) and (i). The grammaticality of (i), as opposed to the ungrammaticality of (15), shows that *it*-clefts should not be treated as deriving from *wh*-clefts, as was argued, for example, in Akmajian (1970b).

(i) What John didn't see was anything I might recognize.

(16) a. It is John and Mary that like Pete.

b. * It is John and Mary that likes Pete.

c. * It are John and Mary that like Pete.

The agreement connectivity between the clefted constituent and the cleft clause favors expletive analyses that analyze the clefted constituent as adjoined to or extracted from the cleft clause.

Interestingly, as first pointed out in Ball (1977), in predicational clefts, a plural clefted constituent triggers a plural cleft pronoun and the copula agrees with this plural cleft pronoun, while the verb in the cleft clause again agrees with the clefted constituent, as shown in (17).

(17) a. They're just fanatics who are holding him.

b. These are students who are rioting.

c. Those are kids who beat John.

This difference in cleft pronoun choice between equative and predicational clefts with plural clefted constituents shows that the distinction is a real one and emphasizes the parallelism between *it*-clefts and ordinary copular sentences, which also exhibit the distinction, as shown above in (9).[2] It would be difficult for an expletive analysis that assumes that the copula as well as the cleft pronoun is semantically inert, to account for the distinction between the predicational and equative *it*-clefts. In section 5, we use agreement features and

[2]An anonymous reviewer suggests that the indefinite plural clefted constituent examples in (17) could also be produced with a singular pronoun and copula. While we agree that this might be possible, we have the strong intuition that such examples are equative in nature. Thus, in (i), it is no longer the case that the property of being fanatics is being predicated of a set of people independently identified as those who are holding him. Instead, the question of who is holding him is being answered by identifying these people as a group of fanatics.

(i) It's just fanatics who are holding him.

feature unification in TAG to account for the connectivity in agreement and the difference in agreement behavior between equative and predicational *it*-clefts, again showing that our TAG analysis can capture the best of both the discontinuous constituent analysis and the expletive analysis.

3 Syntax of *it*-clefts

3.1 Introduction to TAG syntax

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 a lexical anchor. These trees are minimal in that all and only the syntactic/semantic arguments of the lexical anchor are encapsulated and all recursion is factored away. 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 project to Verb Phrase (VP) but also to Tense Phrase (TP) and Complementizer Phrase (CP), and an elementary tree anchoring a noun can project to Noun Phrase (NP) but also to Determiner Phrase (DP) and Prepositional Phrase (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. (αsaw) is an elementary tree because it is an extended projection of the lexical predicate *saw* and has argument slots for the subject and the object marked by the downward arrow (↓). 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. (αJohn) and (αa_movie) are valid elementary trees because these DP trees each contain a single lexical head, *John* for (αJohn) and *movie* for (αa_movie), that can form an extended projection with a DP, in line with the DP Hypothesis (Abney, 1987).[3]

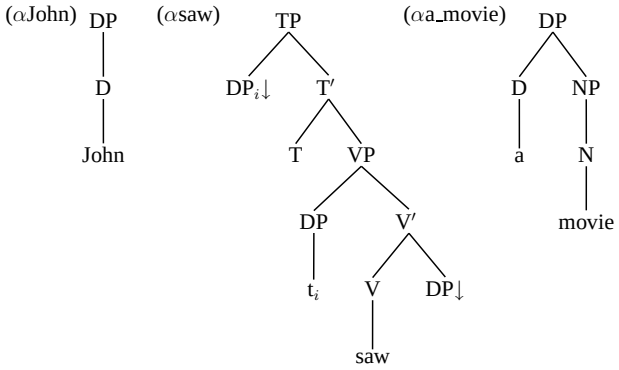
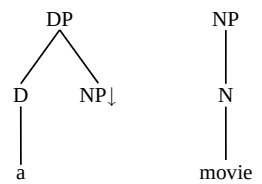


Figure 1: Initial trees in TAG

Elementary trees are of two types: initial trees and auxiliary trees. A derivation in TAG starts with initial trees such as trees for simple clauses and nominal phrases. The elementary trees in Figure 1 are examples of initial trees. Auxiliary trees are used to introduce recursive

[3]In principle, trees such as (αa_movie) could be broken down into trees for determiners and trees for NPs, as in (i). Under this approach, an NP tree anchoring a noun would substitute into a DP tree anchoring a determiner. But strictly speaking, this violates Frank's (2002) formulation of CETM, as the DP tree in (i) is a projection of a functional head (D), not a lexical head.

(i)



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 (βreluctantly) and the NP root node in (βscary) consistent with CETM. Further, following Frank (2002), we can count VP* in (βreluctantly) and NP* in (βscary) as arguments of the lexical anchor, as the process of theta-identification (Higginbotham, 1985) obtains between them and the lexical anchor.[4]

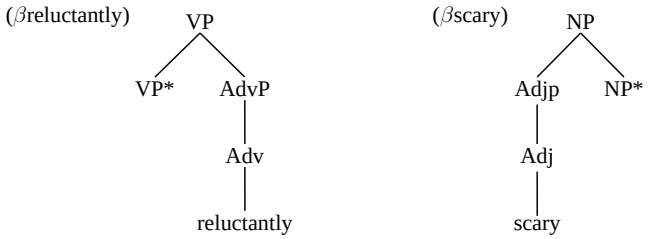


Figure 2: Auxiliary trees in TAG

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 ($\downarrow$) 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. For example, Figure 4 illustrates (βreluctantly) adjoining to the VP node in (αsaw), and (βscary) adjoining to the NP node in (αa_movie) which in turn substitutes into (αsaw).

TAG derivation produces two structures: a derived tree and a derivation tree. The derived tree is the conventional phrase structure tree and represents surface constistuency.

[4]By convention, names of initial trees are prefixed with α, and names of auxiliary trees are prefixed with β.

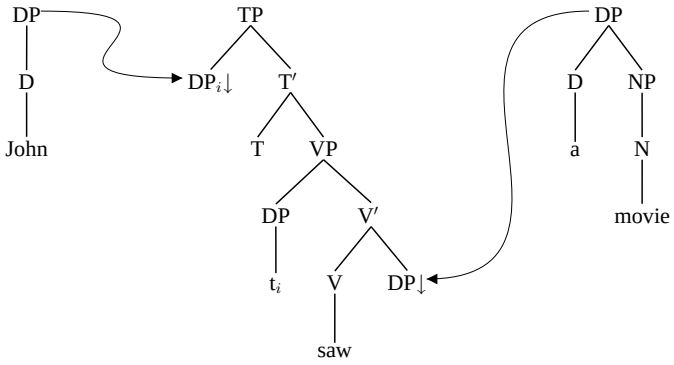


Figure 3: Substitution in TAG

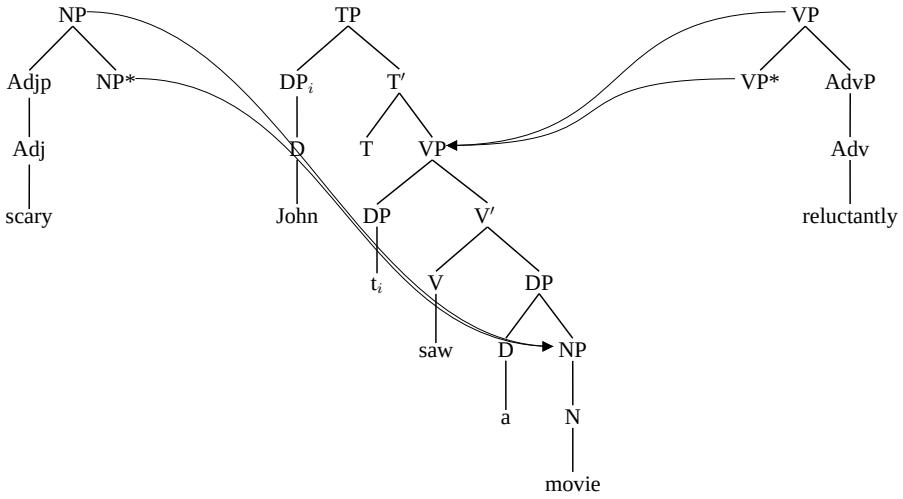


Figure 4: Adjoining in TAG

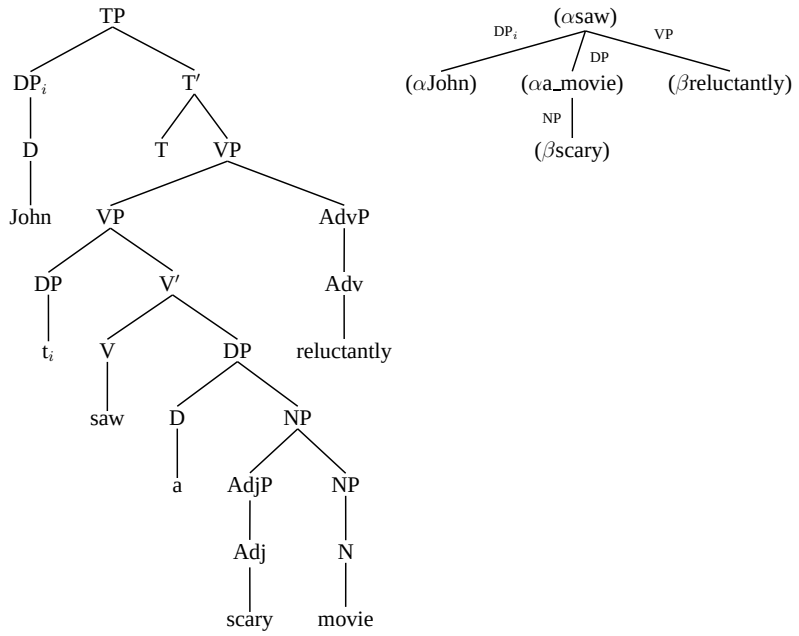


Figure 5: Derived tree and derivation tree in TAG

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 and the dependencies between 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.[5] An example of a derivation tree is given 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): (βscary) adjoins to (αa_movie) at NP, (αJohn) and (αa_movie) substitute into (αsaw) at DP$_i$ and DP respectively, and (βreluctantly) adjoins to (αsaw) at VP.

As first shown by Joshi (1985) and Kroch and Joshi (1985), and explored further in Frank (2002), the properties of TAG permit us to provide computationally feasible accounts

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

for various phenomena in natural language syntax. For example, TAG's extended domain of locality and its factoring of recursion from elementary trees lead, among other things, to a localization of unbounded dependencies. TAG is a mildly context-sensitive grammar (Joshi et al., 1991), formally sitting between context-free and context-sensitive grammar, and is able to generate unbounded cross-serial dependencies such as those that occur between the arguments and verbs in Dutch and Swiss German in a natural way. In section 3.2, we show that TAG's extended domain of locality allows us to provide an elegant syntactic account of the discontinuous constituency of the cleft pronoun and the cleft clause without adopting a movement-based account of the extraposition of the cleft clause. At the same time, TAG's distinction between the derivation and derived trees allows us to account for the surface syntactic constituency of the clefted constituent and the cleft clause.

3.2 Our TAG analysis of the syntax of *it*-clefts

Inspired by work of Kroch and Joshi (1987) and Abeillé (1994) on discontinuous constituents resulting from extraposition, we propose an analysis for the syntax of *it*-clefts using tree-local MC-TAG, an extension of TAG. 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. All of the trees in a multi-component set are restricted to adjoin or substitute simultaneously into a single elementary tree, at each step in a derivation. With this restriction, MC-TAG is shown to be identical to basic TAG in terms of strings and structural descriptions it generates: that is, MC-TAG has the same weak and strong generative capacity as the basic TAG (Weir, 1988). In addition to extraposition, MC-TAG has been used in the analyses of West Germanic verb raising (Kroch and Santorini, 1991), Romance clitic climbing (Bleam, 2000), and extraction of an object *wh*-phrase from a *wh*-island (Kroch, 1989; Frank, 2002). The trees in a multi-component set can be thought of as a single elementary tree decomposed into two or more trees. As these trees substitute

or adjoin into different positions in another elementary tree, the effect of discontinuous constituency can be produced. Further, the locality of the syntactic dependencies that exist between these trees is maintained, as they are restricted to compose simultaneously with a single elementary tree, contributing to the restricted generative capacity of MC-TAG.

We propose that the elementary trees for the cleft pronoun and the cleft clause in the derivation of *it*-clefts such as (10a) (repeated beows as (18)) and (11a) (repeated below as (19)) form a multi-component set, as in {(αit), (βwho_won)} and {(αit), (βwho_beat)} in Figure 6.

(18) It was Ohno who won.

(19) It was a kid who beat John.

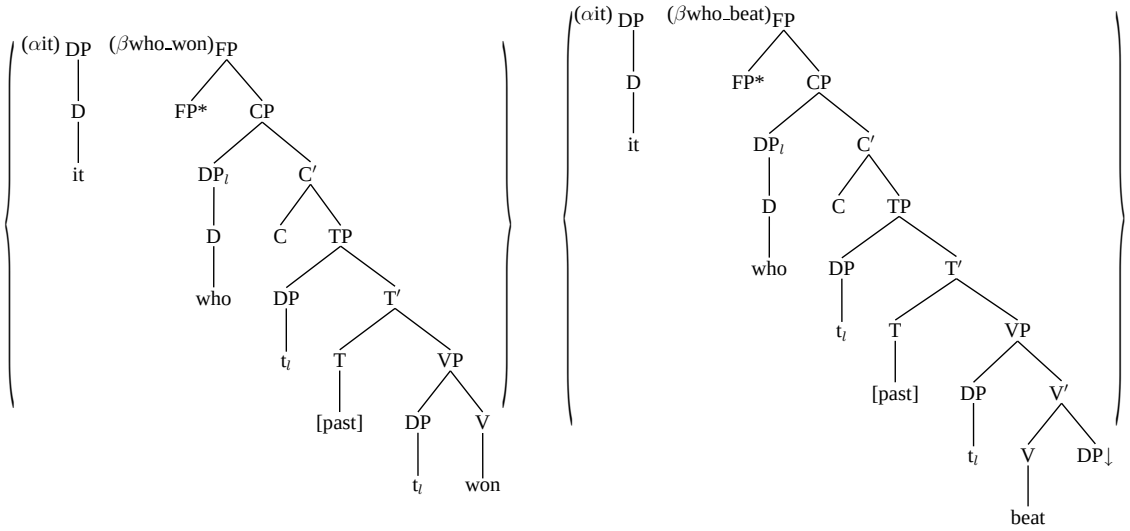


Figure 6: Multi-component sets of cleft pronoun and cleft clause

We capture the intuition that the cleft pronoun and the cleft clause form a syntactic unit by placing the elementary trees for them in a single a multi-componenet set. And as these are two separate trees, they are able to substitute and adjoin onto two different places in a single elementary tree, producing the effect of discontinuity. The first component of each set introduces a determiner, and the second component of each set introduces a rel-

ative clause anchoring the lexical predicate.[6] The multi-component set can be thought as a DP tree decomposed into two parts: a functional projection of a determiner and a lexical domain on which the determiner operates. That is, the two parts are comparable to a projection of D and a projection of N in a simple DP tree such as (αa_movie) in Figure 1: like *a* in (αa_movie), *it* in (αit) is a determiner that heads a DP, and like the NP (*movie*) in (αa_movie), (βwho_won) and (βwho_beat) include the lexical domains on which the determiner operates. Moreover, just like simple DP trees like (αa_movie), the two components in the sets {(αit), (βwho_won)} and {(αit), (βwho_beat)} together comply to CETM: each set has a single lexical head, the verb, and all other syntactic heads and their projections, TP, CP and DP form extended projections of the verb. The presence of FP does not violate CETM, as CETM requires only that syntactic heads and their projections in an elementary tree form an extended projection of the anchor.

For the derivation of equative *it*-clefts as in (18), we adopt the equative copular tree in ($\alpha$was) in Figure 7, a tree similar to the one proposed in Frank (2002) for copular sentences. In this tree, FP is a small clause of the copula from which the two DPs being equated originate.

(18) is derived by substituting (αit) into DP0 in ($\alpha$was), adjoining (βwho_won) into FP in ($\alpha$was), and substituting (αOhno) into DP1 in ($\alpha$was), as illustrated in Figure 8. The syntactic derivation tree and the derived tree for (18) are given in (δ18) and (γ18) respectively in Figure 9.[7] In (δ18), the elementary trees for the cleft pronoun and the cleft clause form a unit, represented as a single node, and in (γ18), the clefted constituent and the cleft clause form a constituent.

[6] Strictly speaking, the elementary trees representing the cleft clause in the two multi-component sets in Figure 6 should have a substitution site in [Spec,CP] to be substituted in by a separate DP elementary tree anchoring a relative pronoun. Here, to simplify the derivation, we have already substituted in the relative pronoun DP tree.

[7] By convention, names of derivation trees are prefixed with δ, and names of derived trees are prefixed with γ.

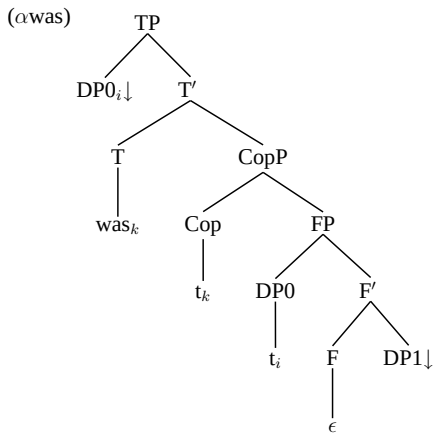


Figure 7: Equative copula elementary tree

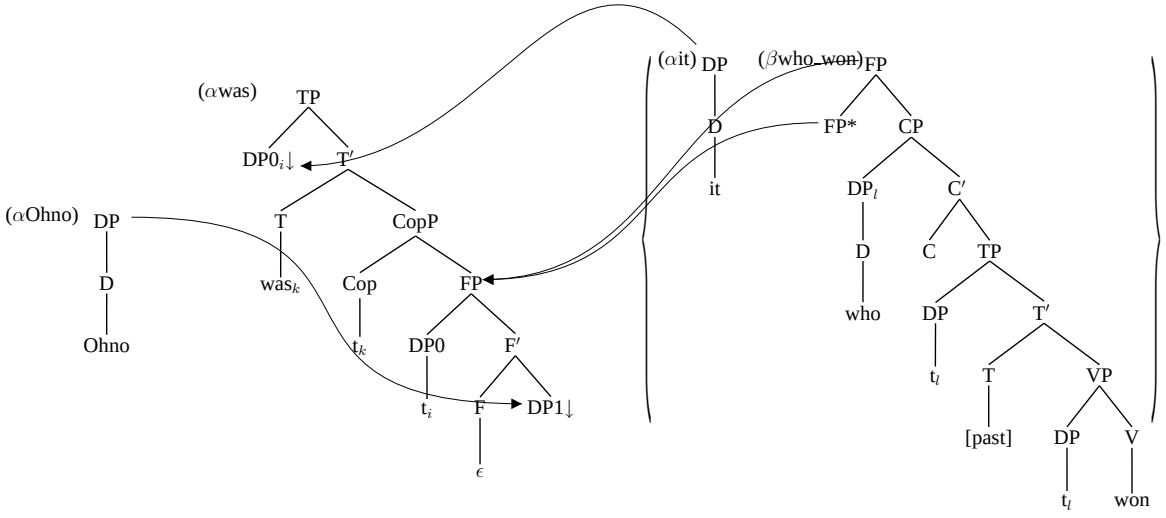


Figure 8: Elementary trees for *It was Ohno who won*

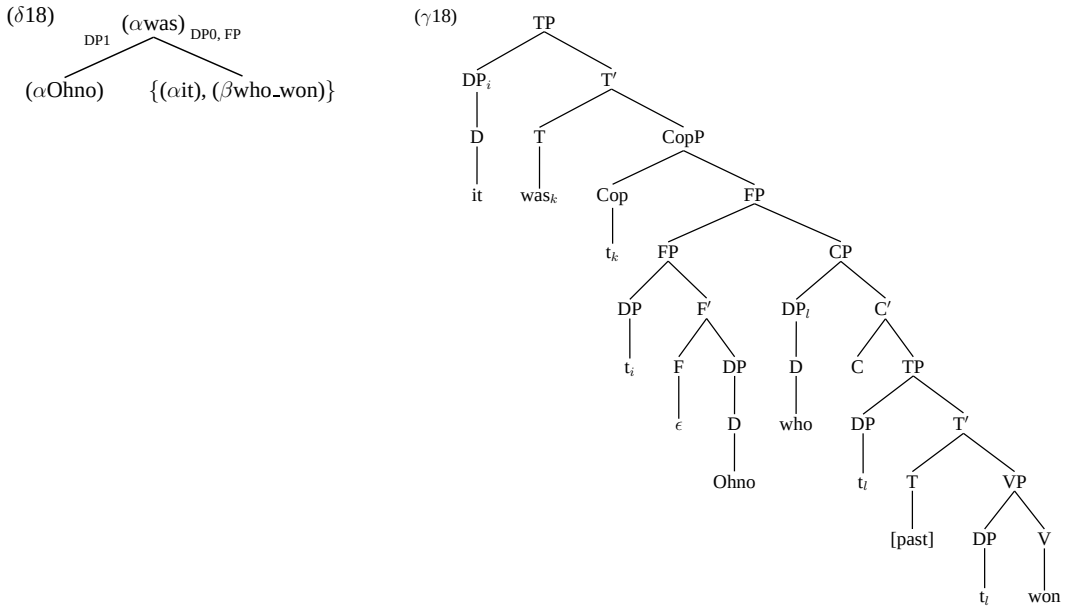


Figure 9: Derivation and derived trees for *It was Ohno who won*

Postulating separate projections for the copula (CopP) and the small clause (FP) in ($\alpha$was) can account for the fact that the clefted constituent and the cleft clause form a constituent, as illustrated in (6ab) (repeated below as (20ab)), and yet they can be separated by an adverbial phrase, as in (20c). In our analysis, (20ab) are possible because the bracketed parts are the higher layers of the FPs in the derived tree. (20c) is possible because an adverbial phrase can adjoin onto FP or F′ in the equative copula tree, in which case, the clefted constituent and the cleft clause would be separated by the adverbial phrase in the derived tree.[8]

(20) a. I said it should have been [Bill who negotiated the new contract], and it should have been.

b. It must have been [Fred that kissed Mary] but [Bill that left with her].

c. It was Kim, in my opinion, who won the race.

[8]See Han and Hedberg (2006) for a TAG analysis of coordination in *it*-clefts, as exemplified in (20b).

For the derivation of predicational *it*-clefts as in (19), we adopt a predicational copula tree (αwas_kid) in Figure 10. The predicational copula tree in (αwas_kid) is similar to the equative copula tree in ($\alpha$was) in that in both trees, the copula combines with a small clause FP. But the two trees have different anchors and different number of argument substitution sites. In (αwas_kid), the noun (*kid*) is the predicate requiring a single argument, and thus the noun (*kid*) is the lexical anchor of the tree and the subject DP is an argument substitution site. But in ($\alpha$was), both the subject and the non-subject DPs are argument substitution sites as they are arguments of an equative predicate.

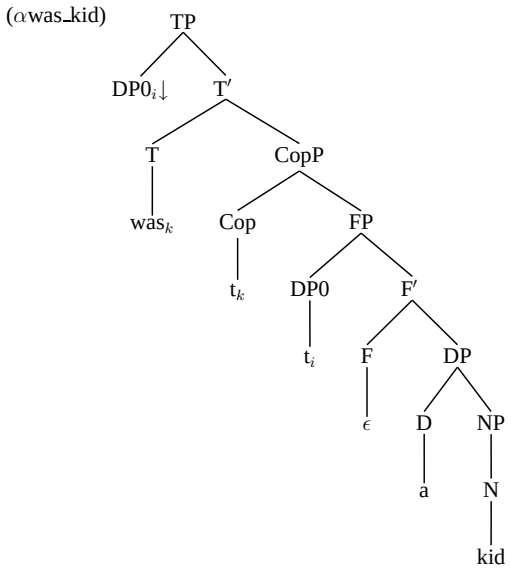


Figure 10: Predicational copula tree

As illustrated in Figure 11, (19) is derived by substituting (αit) into DP0 and adjoining (βwho_beat) onto FP in (αwas_kid), and substituting (αJohn) into DP in (αwho_beat). The syntactic derivation tree and the derived tree for (19) are given in (δ19) and (γ19) respectively in Figure 12. Just as in the derivation tree and the derived tree for the equative *it*-cleft in Figure 12, in (δ19), the elementary trees for the cleft pronoun and the cleft clause form a unit, represented as a single node, and in (γ19), the clefted constituent and the cleft clause form a constituent.

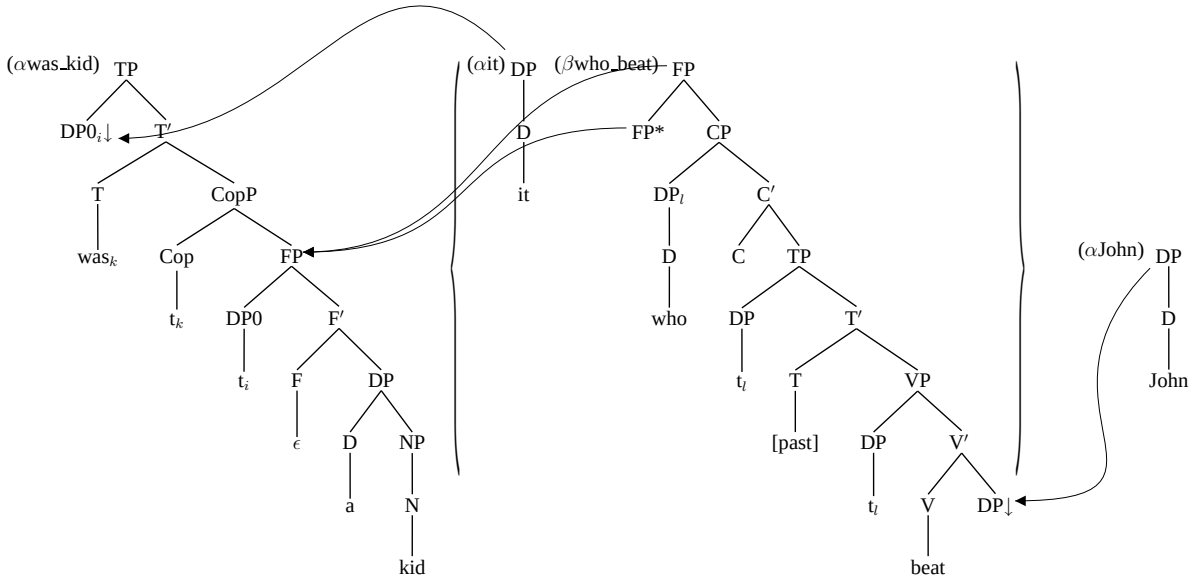


Figure 11: Elementary trees for *It was a kid who beat John*

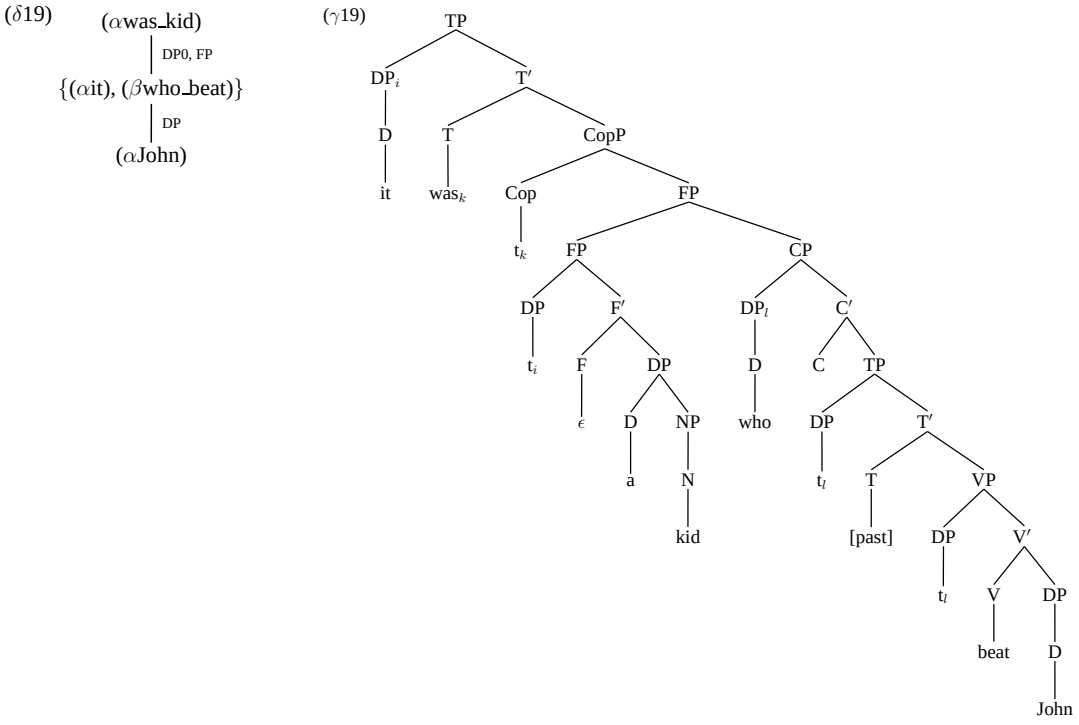


Figure 12: Syntactic derivation and derived trees for *It was a kid who beat John*

4 Semantics of *it*-clefts

In TAG, the derivation tree, not the derived tree, serves as the input to compositional semantics (Joshi and Vijay-Shanker, 1999; Kallmeyer and Joshi, 2003). 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.

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, 2008); 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 (2008), 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: pied-piping of *wh*-phrases (Kallmeyer and Scheffler, 2004), focus (Babko-Malaya, 2004), questions (Romero et al., 2004), VP coordination (Banik, 2004), among others.

In this paper, however, we use STAG, a pairing of a TAG for the syntax and a TAG for the semantics, to propose a compositional semantic analysis for *it*-clefts. In STAG-based

compositional semantics, the semantic representations are structured trees with nodes on which substitution and adjoining of other semantic representations can take place. Compositionality 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 these meaning components are combined. Since the semantic representations are structured trees, 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, as 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, and Han (2007), who provide an STAG analysis for relative clauses and pied-piping. In section 4.1, we introduce the basics of STAG and STAG-based compositional semantics, and in section 4.2, we present our proposed analysis for the semantic composition of *it*-clefts.

4.1 Introduction to STAG and compositional semantics

We illustrate the framework of STAG and STAG-based compositional semantics and clarify our assumptions, using (21), a simple sentence that contains an existential quantifier and an attributive adjective. A similar example was used in section 3 to illustrate the syntactic derivation in TAG.

(21) John saw a scary movie.

We use STAG as defined in Shieber (1994). In STAG, each syntactic elementary tree is paired with one or more semantic trees that represent its meaning 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 13, 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. We assume that these nodes are typed (e.g., the F node in (α'saw) has type t and the lowest R node in (α'saw) has type $<e,<e,t>>$), and we represent predicates as unreduced λ-expressions, following the notation in Han (2007). Making use of unreduced λ-expressions in semantic trees allows the reduction of semantic derived trees to logical forms through the application of λ-conversion and other operations defined on λ-expressions. The linked nodes are shown with boxed numbers. For the sake of simplicity, in the elementary tree pairs, we only include links that are relevant for the derivation of given examples.[9]

Figure 13 contains elementary trees required to generate the syntactic structure and the logical form of (21). The proper name tree in (αJohn) is paired with a tree representing a term on the semantics side, and the attributive adjective tree in (βscary) 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, we follow Shieber and Schabes (1990) and Nesson and Shieber (2006), and use tree-local MC-TAG on the semantics side. Thus, the DP in (αa_movie) is paired with a multi-component set {(α'a_movie), (β'a_movie)} on the semantics side: (α'a_movie) provides an argument variable, and (β'a_movie) provides an existential quantifier with the restriction and scope. The transitive tree in (αsaw) is paired with a semantic tree representing a formula that consists of a two-place predicate and two

[9]By convention, names of semantic elementary trees are prefixed with α' or β', names of semantic derivation trees are prefixed with δ', and names of semantic derived trees are prefixed with γ'.

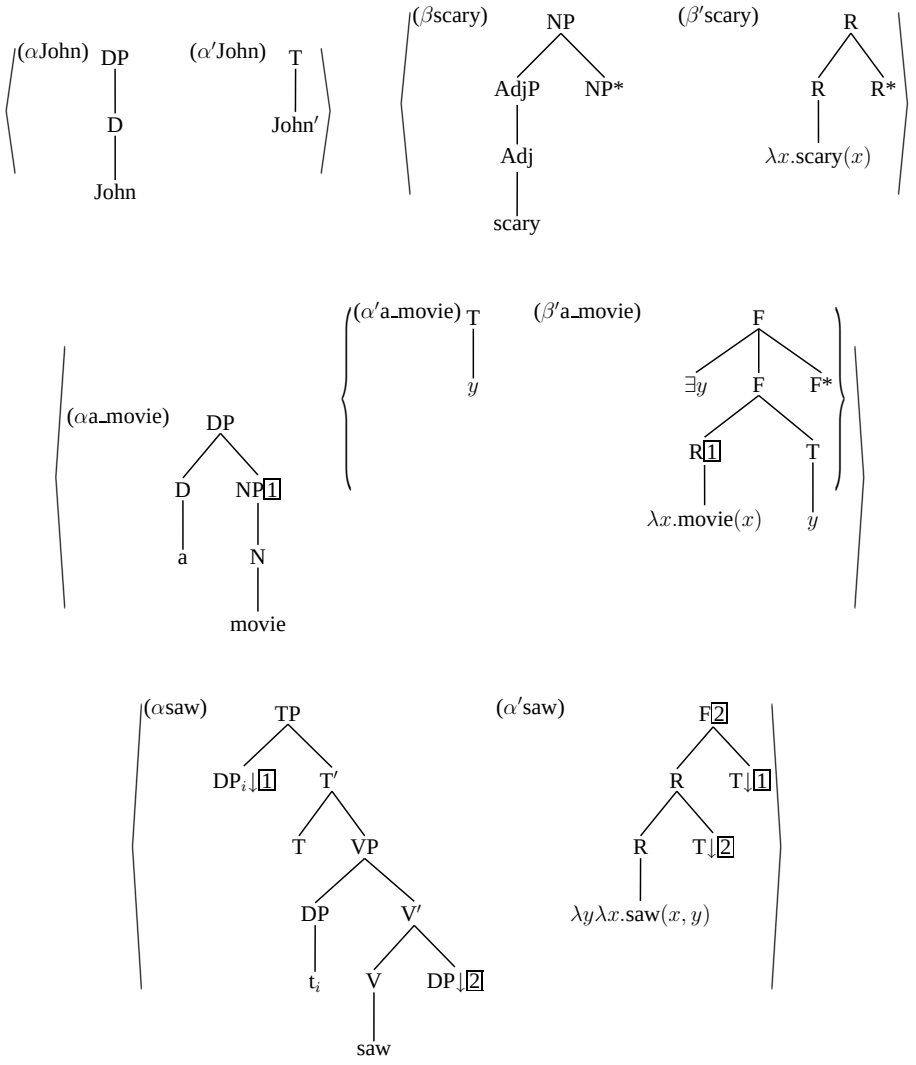


Figure 13: Syntactic and semantic elementary trees for *John saw a scary movie*

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 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 2 and the other will adjoin to the F node marked with 2. The syntactic and semantic derivation trees are given in Figure 14, and the derived trees are given in Figure 15. Technically, there is only one derivation tree because the syntactic and semantic derivations are isomorphic.

In this paper, we provide two derivation trees (one for syntax and the other for semantics) throughout to make the tree-local derivation explicit.[10]

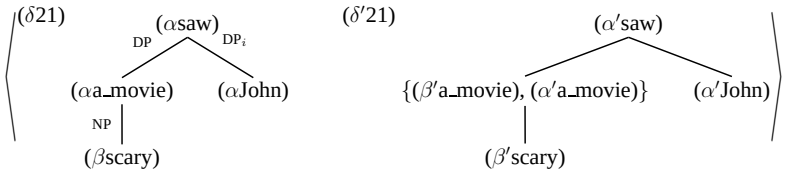


Figure 14: Syntactic and semantic derivation trees for *John saw a scary movie*

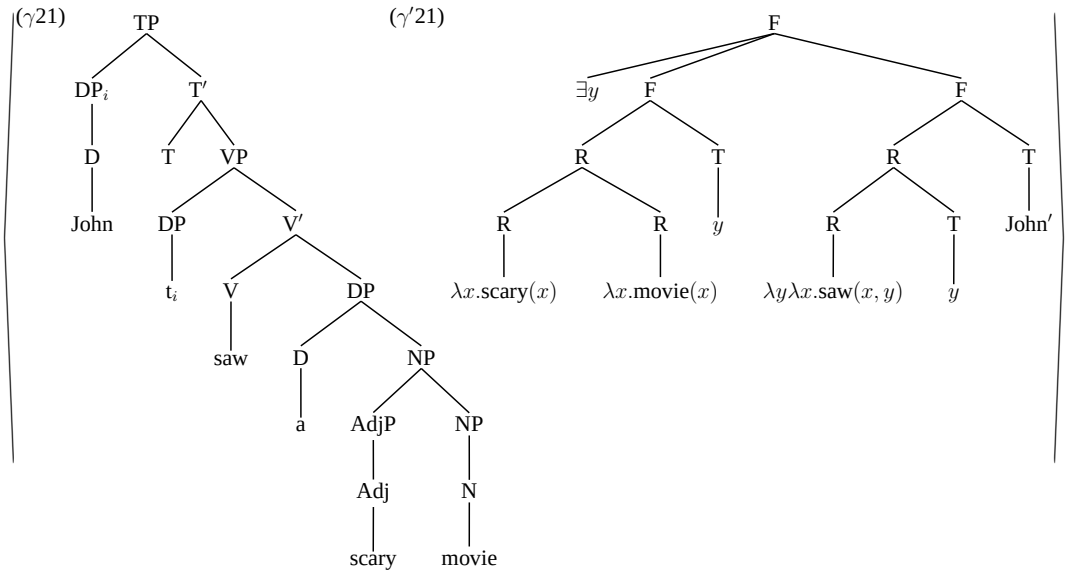


Figure 15: Syntactic and semantic derived trees for *John saw a scary movie*

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, we propose to use Predicate Modification, as defined in Heim and Kratzer (1998) in (22).

(22) Predicate Modification

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

[10] In semantic derivation trees, we 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.

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

The application of Predicate Modification and λ-conversion reduces ($\gamma'21$) to the formula in (23).

(23) $\exists y[\text{scary}(y) \wedge \text{movie}(y)]$ [saw(John$'$, y)]

4.2 Our TAG analysis of the semantics of *it*-clefts

The elementary tree pairs required for the syntax-semantics mapping of the equative *it*-cleft in (18) are given in Figure 16. (α'it) and (β'who_won) in the multi-component set in Figure 16 together define the semantics of definite quantification, where the former contributes the argument variable and the latter the definite quantifier, the restriction and scope, and ($\alpha'$was) represents the semantics of equative sentences.[11] The derivation tree for the semantics of (18) is given in ($\delta'18$) in Figure 17, and the semantic derived tree is given in ($\gamma'18$) in Figure 18. Note that the semantic derivation tree in ($\delta'18$) is isomorphic to the syntactic one in ($\delta 18$). The semantic derived tree in ($\gamma'18$) can be reduced to the formula in (24) after the application of λ-conversion.

(24) THEz [won(z)] [z = Ohno$'$]

The elementary tree pairs required for the syntax-semantics mapping of the predicational *it*-cleft in (19) are given in Figure 19. The difference between the semantics of equative sentences and predicational sentences is represented by the two different semantic trees, ($\alpha'$was) in Figure 16 and (α'was_kid) in Figure 19. While ($\alpha'$was) in Figure 16 represents the semantics of equative sentences and has two term nodes with a two-place equative

[11] In (β'who_won), the R node represents the semantics of the relative clause *who won*. This is a product of composing the semantics of the relative pronoun *who* and the semantics of the rest of the relative clause. Here, to simplify the derivation and to streamline the discussion, we skipped a step in the derivation with separate semantic trees for the relative pronoun and the rest of the relative clause. For a detailed analysis of the compositional semantics of relative clauses using STAG, see Han (2007).

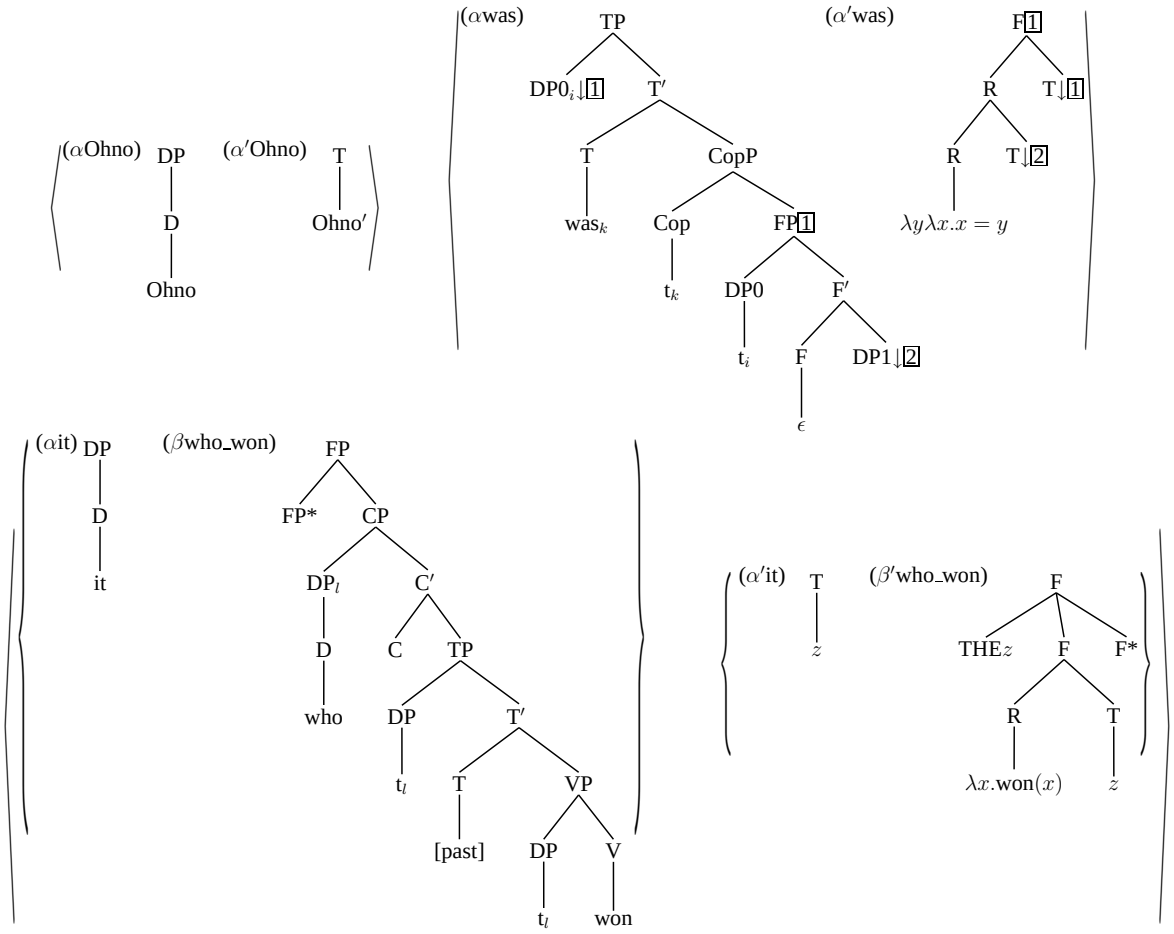


Figure 16: Syntactic and semantic elementary trees for *It was Ohno who won*

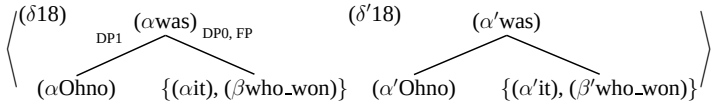


Figure 17: Syntactic and semantic derivation trees for *It was Ohno who won*

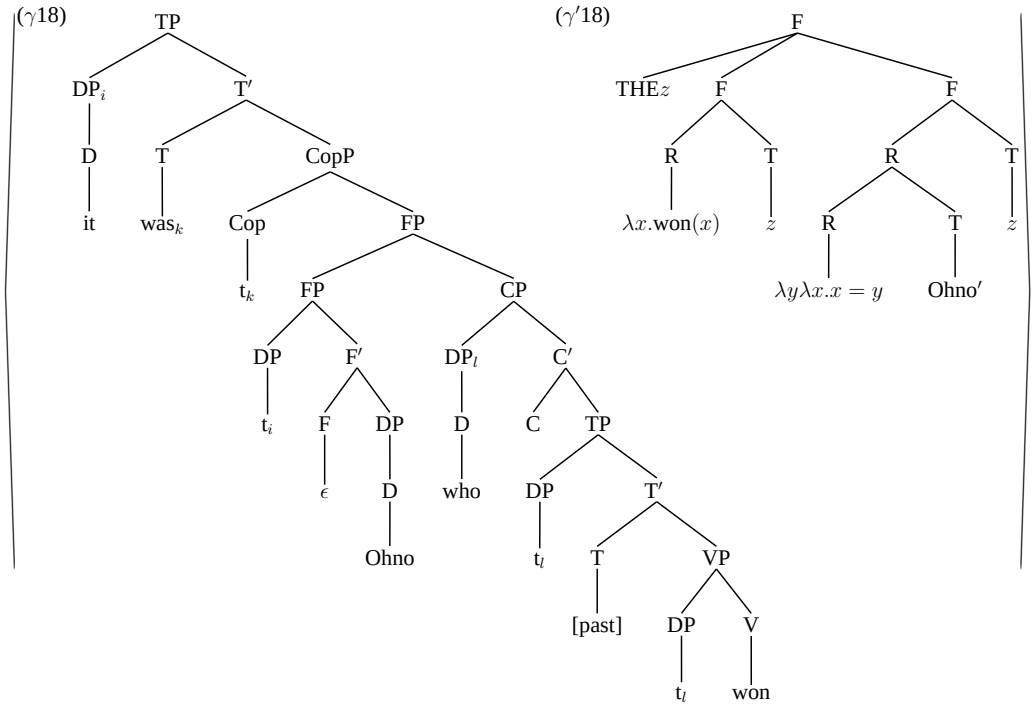


Figure 18: Syntactic and semantic derived trees for *It was Ohno who won*

predicate anchoring the tree, (α'was_kid) in Figure 19 represents the semantics of predicational sentences and has one term node with a one-place predicate, $\lambda x.\text{kid}(x)$, anchoring the tree. The syntactic and semantic derivation trees for (19), which are isomorphic, are given in $<$(δ19), (δ'19)$>$ in Figure 20, and the corresponding derived trees are given in $<$(γ19), (γ'19)$>$ in Figure 21. The semantic derived tree in (γ'19) can be reduced to the formula in (25) after the application of λ-conversion.

(25) THEz [beat(z, John′)] [kid(z)]

5 Connectivity

5.1 Agreement

In equative *it*-clefts, the cleft pronoun is always singular and agrees with the copula, but the clefted constituent can be either singular or plural. Further, when the cleft clause is

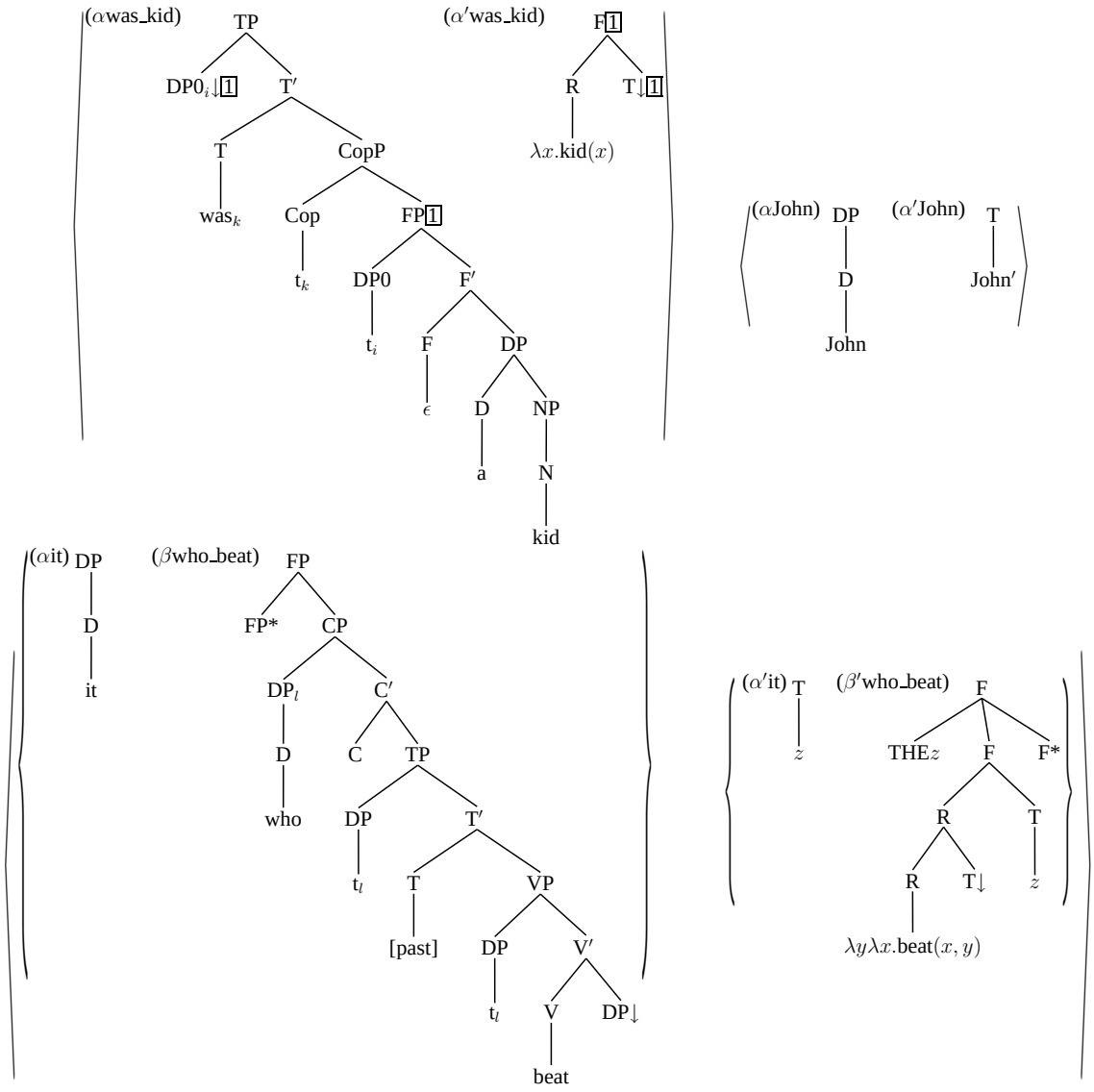


Figure 19: Syntactic and semantic elementary trees for *It was a kid who beat John*

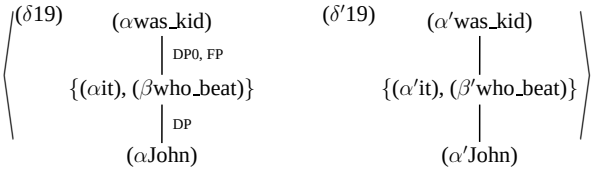


Figure 20: Syntactic and semantic derivation trees for *It was a kid who beat John*

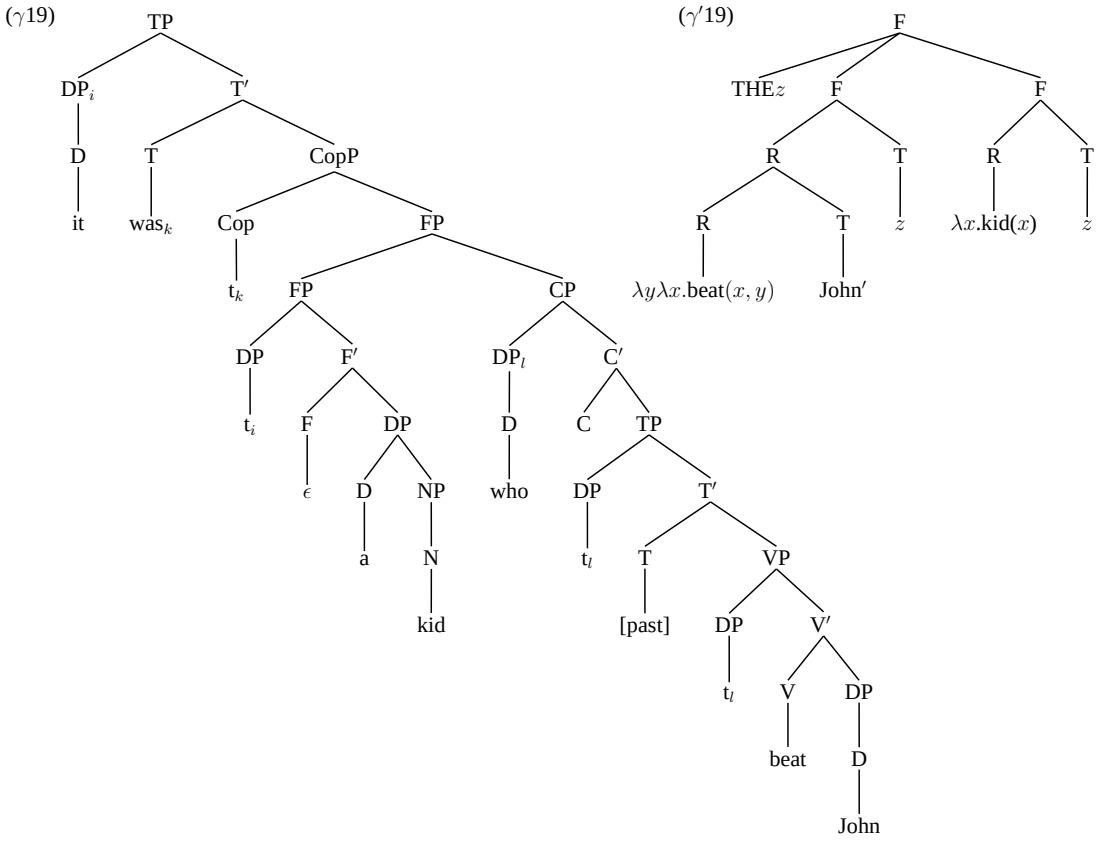


Figure 21: Syntactic and semantic derived trees for *It was a kid who beat John*

a subject relative clause, the clefted constituent agrees with the verb in the cleft clause in person and number. This is illustrated in (16), repeated here as (26). This apparent agreement between the clefted constituent and the verb in the clefts clause, even though they are not in the same clause in our analysis, gives rise to a connectivity effect.

(26) a. It is John and Mary who like Pete.

b. * It is John and Mary who likes Pete.

c. * It are John and Mary who like Pete.

We point out that agreement across clauses is not unique to *it*-clefts. In (27), the subject of the main clause *John and Mary* agrees with the copula of the non-restrictive relative clause. So, there is independent motivation for a mechanism in the grammar that allows

agreement across clauses in appropriate syntactic contexts.

(27) John and Mary, who are students, came to see me.

The agreement phenomena in *it*-clefts can be easily accommodated by our TAG analysis, with the addition of feature unification (Vijay-Shanker and Joshi, 1988). We will postulate an agreement feature attribute, Agr, that can have feature values such as 3rd person singular (3sg) or 3rd person plural (3pl) feature. This Agr feature can also be unspecified in an elementary tree and obtain a value through feature unification as it composes with another elementary tree. An unspecified Agr feature has an arbitrary index as a temporary value, and Agr features with the same indices must have the same value at the end of the derivation.

Figure 22 illustrates how our TAG analysis can capture the agreement between the cleft pronoun *it* and the copula *is*, and the clefted constituent *John and Mary* and the verb of the cleft clause *like* in (26a).[12] To simplify the discussion, we have already derived the DP coordination tree for *John and Mary* and refer to it as ($\alpha$and), and substituted the DP tree anchoring *Pete* into (βwho_like). The substitution of (αit) into DP0 in ($\alpha$is) is licensed because DP in (αit) has [Agr:3sg] feature which unifies with [Agr:3sg] in DP0 in ($\alpha$is). And the agreement between *it* and *is* is guaranteed as both DP0 and T in ($\alpha$is) tree have the same agreement features, as indicated by the co-indexation between the agreement feature on DP0 and the 3rd person singular feature in T. As ($\alpha$and) tree substitutes into DP1 in ($\alpha$is), the [Agr:4] feature on FP is valued as 3pl. As (βwho_like) tree adjoins onto FP in ($\alpha$is), DP_l and T in (βwho_like) are valued as 3pl as well. This will guarantee the agreement between *John and Mary* and *like*. The derived tree with all the Agr features valued and unified is in Figure 23.

[12] An anonymous reviewer asks why the agreement feature on T in (βwho_like) is not valued as plural. We chose to leave it unspecified, as it is compatible with 3rd person plural as well as 2nd and 1st person singular and plural.

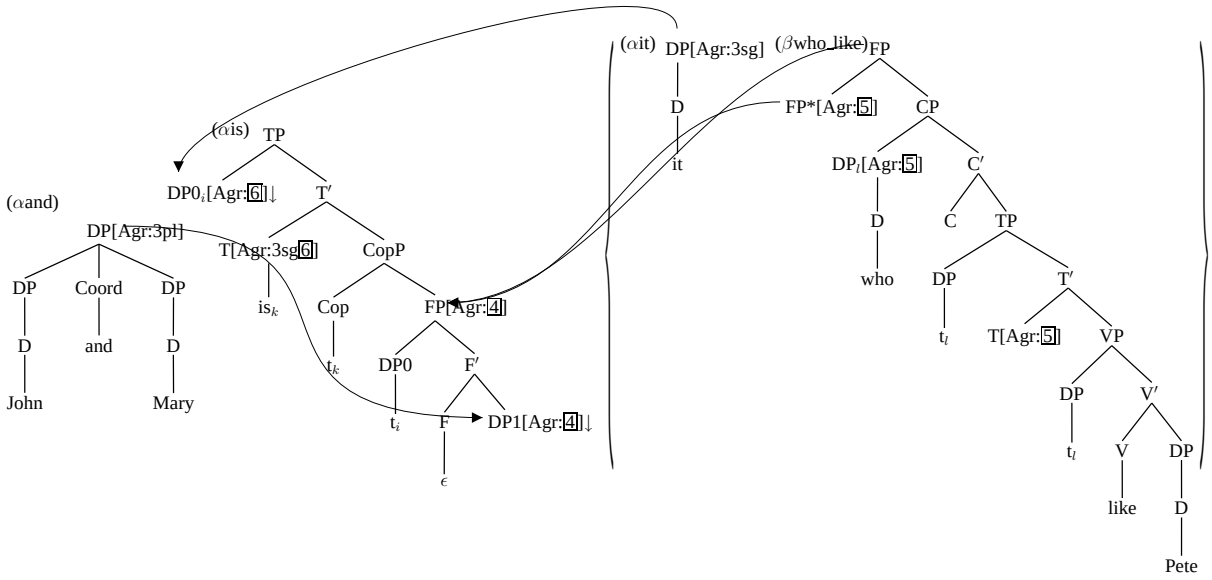


Figure 22: Derivation of *It is John and Mary who like Pete*

In predicational *it*-clefts, the cleft pronoun can be plural, and it must agree with the copula as well as the clefted constituent. Moreover, if the cleft clause is a subject relative clause, then the clefted constituent must agree with the verb of the cleft clause, even though they are not in the same clause in our analysis, giving rise to a connectivity effect. This is illustrated in (17), repeated here as (28).

(28) a. They're just fanatics who are holding him.

b. Those are students who are rioting.

c. Those are kids who beat John.

How our TAG analysis can capture the agreement phenomena in predicational *it*-clefts is illustrated in Figure 24.[13] To simplify the discussion, we have already substituted the DP tree anchoring *John* into (βwho_beat). In our TAG analysis, the lexical anchor of a predicational copula elementary tree is the predicative noun, as in (αare_kids). In this tree, the agreement between the cleft pronoun, the copula and the predicative noun is guaranteed:

[13]We left the agreement feature on T in (βwho_beat) unspecified for the same reason we left it unspecified in (βwho_beat): it is compatible with 3rd person plural, and 2nd and 1st person singular and plural.

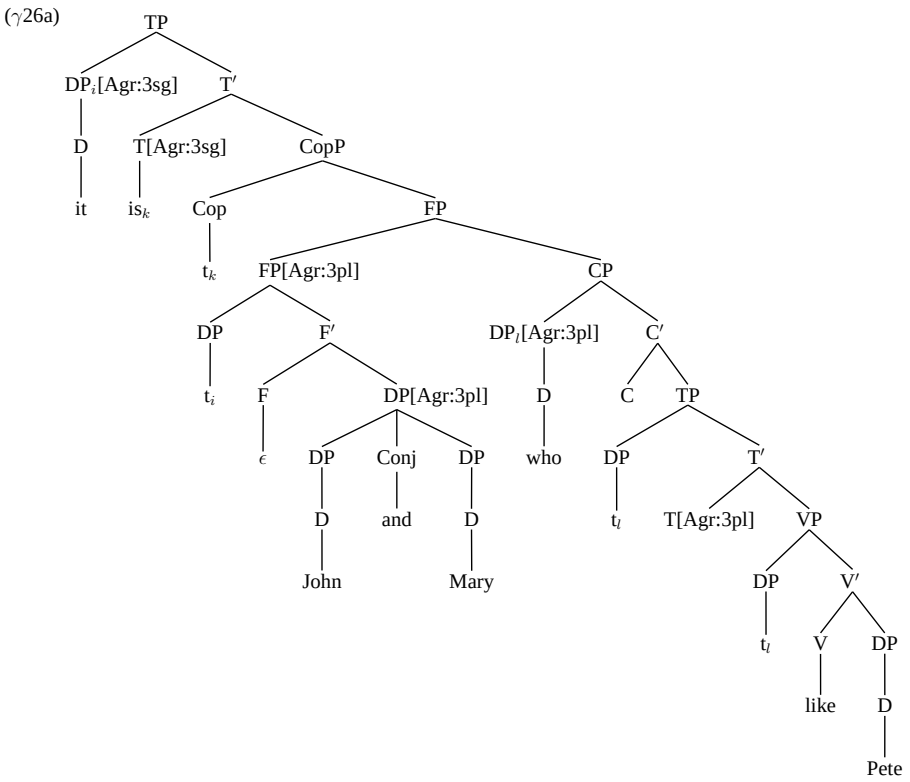


Figure 23: Syntactic derived tree for *It is John and Mary who like Pete*

DP0, T and DP all have the same agreement features as they all have the same indices. Here, they all have 3rd person plural features as the DP containing the predicative noun is specified with the 3rd person plural feature. The substitution of (αthose) tree into DP0 in (αare_kids) is licensed because DP in (αthose) has [Agr:3pl] feature which unifies with the 3rd person plural feature in DP0 in (αare_kids). As (βwho_beat) tree adjoins onto FP in (αare_kids), DP$_l$ and T in (βwho_beat) will obtain 3pl value as well. This will guarantee the agreement between *kids* and *beat*. The derived tree with all the Agr features valued and unified is given in Figure 25.[14]

[14]Why equative clefts require singular cleft pronouns when they contain a plural clefted constituent doesn't follow from our theory and remains a puzzle. However, the fact that different agreement patterns occur shows that there are clearly two types of *it*-cleft.

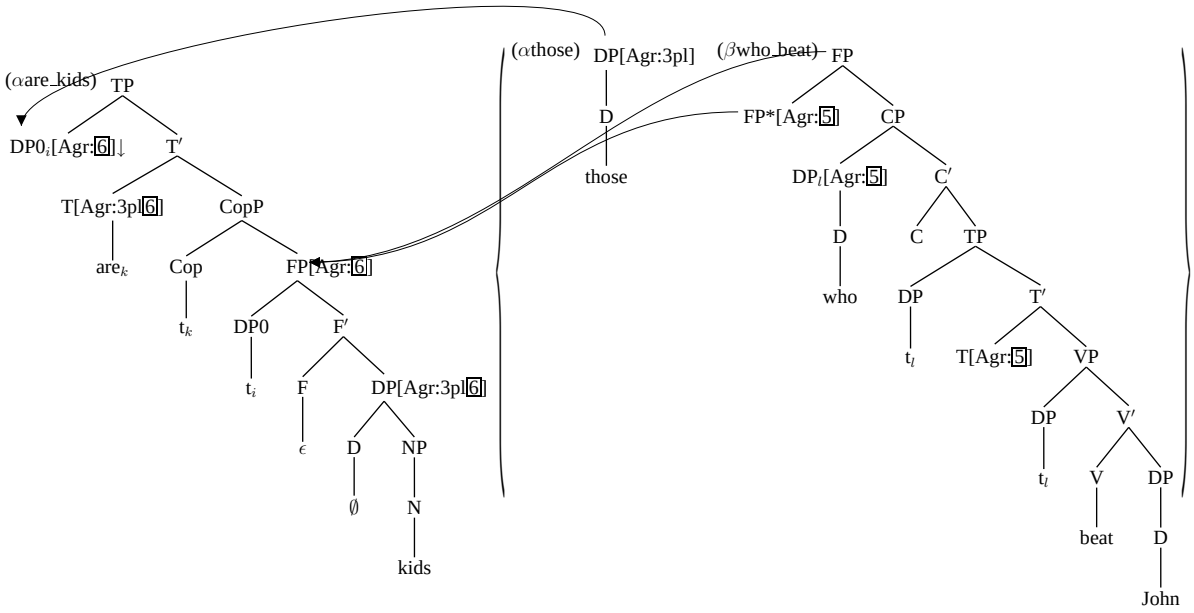


Figure 24: Derivation of *Those are kids who beat John*

5.2 Binding

In *it*-clefts, even though the clefted constituent is not c-commanded by the subject of the cleft clause, a SELF-anaphor in the clefted constituent can be co-indexed with the subject in the cleft clause as in (12a), repeated here as (29a), and a pronoun in the clefted constituent cannot be co-indexed with the subject in the cleft clause as in (13a), repeated here as (29b). In other words, the SELF-anaphor and the pronoun behave as if they are inside the cleft clause as in (30a) and (30b), giving rise to a connectivity effect.

(29) a. It was himself$_i$ who John$_i$ nominated.

b. * It was him$_i$ who John$_i$ nominated.

(30) a. John$_i$ nominated himself$_i$.

b. * John$_i$ nominated him$_i$.

We will use the Binding Conditions defined in Reinhart and Reuland (1993) to account for this phenomenon. Reinhart and Reuland's formulation of Binding Conditions and the

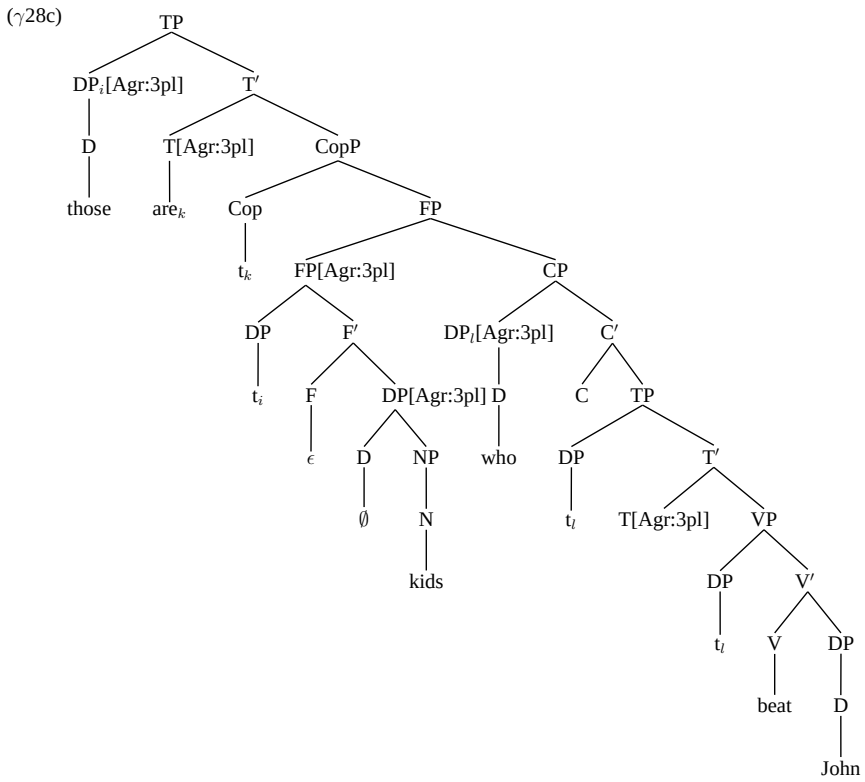


Figure 25: Syntactic derived tree for *Those are kids who beat John*

definitions needed to understand it are given in (31) and (32). Condition A constrains the distribution of SELF-anaphors, and Condition B constrains the distribution of pronouns.

(31) Binding Conditions (Reinhart and Reuland, 1993)

a. A: If a syntactic predicate is reflexive-marked, it is reflexive.

b. B: If a semantic predicate is reflexive, it is reflexive-marked.

(32) Definitions (Reinhart and Reuland, 1993)

a. The syntactic predicate formed of a head P is P, all its syntactic arguments (the projections assigned theta-roles/case by P), and an external argument of P.

b. The semantic predicate of P is P and all its arguments at the relevant semantic level.

c. P is reflexive iff two of its arguments are coindexed.

d. P is reflexive-marked iff either P is lexically reflexive, or one of P's arguments is a SELF-anaphor.

According to Reinhart and Reuland, Condition A successfully applies to (30a) because the syntactic predicate *John nominated himself* is reflexive-marked, as one of the arguments, *himself*, is a SELF-anaphor, and it is also reflexive, as two of its arguments, *John* and *himself*, are co-indexed. However, (30b) is ruled out by Condition B. In (30b), the semantic predicate *nominated(John, John)* is reflexive, as two of its arguments are coindexed, but it is not reflexive-marked, as *nominated* is not lexically reflexive and none of *nominated*'s arguments is a SELF-anaphor.

We first apply Reinhart and Reuland's Condition B to rule out (29b), repeated below as (33a). According to our TAG analysis, (33a) would map onto an equative semantic representation as in (33b). Since the clefted constituent *him* is co-indexed with *John*, they co-refer, and so the variable from the cleft pronoun, z, would be equated with *John*$'$. We will represent this as z=*him*$_{John'}$, just to be explicit about the fact that the form of the clefted constituent here is *him*. This in turn means that the semantic predicate *nominated(John*$'$*,* z*)* is reflexive. But it is not reflexive-marked, as *nominated* is not lexically reflexive and none of its arguments is a SELF-anaphor.

(33) a. * It was him$_i$ who John$_i$ nominated.

b. * THEz [nominated(John$'$, z)] [z=him$_{John'}$]

We now turn to (29a). According to our TAG analysis, (29a) is also an equative sentence. We thus have a syntactic predicate whose head is the equative copula and with two syntactic arguments, *it* and *himself*. But then Condition A should rule out this sentence because even though the syntactic predicate is reflexive-marked, it is not reflexive, as *it* and *himself* are not co-indexed.

Reinhart and Reuland point out that focus anaphors can occur in an argument position without a binder, appearing to be exempt from Condition A. Such anaphors are also known as discourse anaphors of focus or emphatic anaphors (Kuno, 1987; Zribi-Hertz, 1989). Some examples are given in (34).

(34) a. This letter was addressed only to myself. (Reinhart and Reuland 1993, ex. 27a)

b. "Bismarck's impulsiveness has, as so often, rebounded against himself." (Reinhart and Reuland 1993, ex. 27c, originally quoted in Zribi-Hertz 1989)

We note that the clefted constituent is a focused position (Akmajian, 1970a; Prince, 1978). This means that a SELF-anaphor in a clefted constituent position is always focused, and so it can be exempt from Condition A. A further support for this view comes from examples as in (35). These examples are acceptable even though *myself* and *yourself* do not have possible binders in the sentences in which they occur.

(35) a. It was myself who John nominated.

b. It was yourself who John nominated.

A question remains though as to why the clefted constituent cannot be occupied by just any SELF-anaphor. For instance, (36) is degraded where *herself* in the clefted constituent position does not have a binder.

(36) * It was herself who John nominated.

This implies that even though a focus anaphor in the clefted constituent position is not subject to Condition A, its distribution is constrained by discourse factors. The exact nature of the discourse constraints on the distribution of focus anaphors in *it*-clefts remains to be investigated.

6 Conclusion

We have proposed a syntax and semantics of *it*-clefts, using tree-local MC-TAG and STAG. We accounted for the equative and predicational interpretations available to *it*-clefts, the two readings available to simple copula sentences as well, by postulating two types of copula sentences in English, an equative one and a predicational one (Heycock and Kroch, 1999). The two types of copula sentences are represented by two different pairs of syntactic and semantic elementary trees. Our analysis thus contrasts with the inverse analysis of Williams (1983), Partee (1986), Moro (1997), and Mikkelsen (2005), according to which specificational clauses (our equatives) are inverted predicational clauses. On some versions of this analysis, both orders derive from an underlying embedded small clause, with either the subject or the predicate raising to matrix subject position.

In our TAG analysis, the derivation of *it*-clefts start either with an equative copula elementary tree or a predicational copula elementary tree. The copula tree then composes with the elementary tree for the cleft pronoun and the elementary tree for the cleft clause. In our analysis, the cleft pronoun and the cleft clause bear a direct syntactic relation because the elementary trees for the two parts belong to a single multi-component set. They do not actually form a syntactic constituent in the derived tree, but as the elementary trees for the two belong to the same multi-component set, the intuition that they form a syntactic unit is captured, represented in the derivation tree as a single node. At the same time, the surface syntactic constituency is represented in the derived tree where the clefted constituent and the cleft clause form a constituent. Further, the semantics of the two trees in the multi-component set is defined as a definite quantified phrase, capturing the intuition that they form a semantic unit as a definite description. We have also shown that our TAG analysis can account for connectivity effects instantiated by binding and agreement: for binding, we applied Reinhart and Reuland's (1993) Binding Conditions and exploited the fact that the

clefted constituent is a focused position, and for agreement, we added feature unification to our TAG analysis.

The distinction in TAG between the derivation tree and the derived tree enabled us to resolve the tension between the surface constituency and the syntactic and semantic dependency in *it*-clefts: in the derived tree, the cleft clause forms a constituent with the clefted constituent, not with the cleft pronoun, capturing the insight from the expletive approach, but in the derivation tree, the cleft clause and the cleft pronoun form a syntactic/semantic unit, capturing the insight from the discontinuous constituent approach. The extended domain of locality of TAG and the ability to decompose an elementary tree to a set of trees in MC-TAG enabled us to provide a straightforward syntactic account of the discontinuous constituent property of the cleft pronoun and the cleft clause without having to adopt movement to produce the effect of extraposition of the cleft clause. Moreover, the derivation-tree-based compositional semantics and the direct syntax-semantics mapping in STAG enabled us to provide a simple compositional semantics for *it*-clefts without using an ad-hoc interpretive operation to associate the meaning coming from the cleft pronoun and the meaning coming from the cleft clause. It remains as future work to extend our analysis to *it*-clefts that have non-DP clefted constituents, such as *It was to the library that John went* and *It was happily that John quit his job*.

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. Doctoral Dissertation, MIT.

Akmajian, Adrian. 1970a. Aspects of the grammar of focus in English. Doctoral Disserta-

tion, MIT.

Akmajian, Adrian. 1970b. On deriving cleft sentences from pseudo-cleft sentences. *Linguistic Inquiry* 1:149–168.

Babko-Malaya, Olga. 2004. LTAG semantics of focus. In *Proceedings of TAG+7*, 1–8. Vancouver, Canada.

Ball, Catherine N. 1977. Th-clefts. *Pennsylvania Review of Linguistics* 2:57–69.

Banik, Eva. 2004. Semantics of VP coordination in LTAG. In *Proceedings of TAG+7*, 118–125. Vancouver, Canada.

Bleam, Tonia. 2000. Clitic climbing and the power of Tree Adjoining Grammar. In *Tree Adjoining Grammars: Formalisms, linguistic analysis, and processing*, ed. Ann Abeillé and Owen Rambow, 193–220. Stanford, Calif.: CSLI.

Chomsky, Noam. 1977. On wh-movement. In *Formal syntax*, ed. P. W. Culicover, T. Wasow, and A. Akmajian, 71–132. New York: Academic Press.

Copestake, Ann, Dan Flickinger, Ivan A. Sag, and Carl Pollard. 2005. Minimal recursion semantics: An introduction. *Journal of Research on Language and Computation* 3:281–332.

DeClerck, Renaat. 1988. *Studies on copular sentences, clefts and pseudo-clefts*. Dordrecht: Foris.

Delahunty, Gerald P. 1982. *Topics in the syntax and semantics of English cleft sentences*. Bloomington: Indiana University Linguistics Club.

Delin, Judy L. 1989. Cleft constructions in discourse. Doctoral Dissertation, University of Edinburgh.

É. Kiss, Katalin. 1998. Identificational focus versus information focus. *Language* 74:245–273.

Emonds, Joseph E. 1976. *A transformational approach to English syntax*. New York: Academic Press.

Frank, Robert. 2002. *Phrase structure composition and syntactic dependencies*. Cambridge, MA: MIT Press.

Grimshaw, Jane. 1991. Extended projection. Manuscript, Brandeis University.

Gundel, Jeanette K. 1977. Where do cleft sentences come from? *Language* 53:53–59.

Han, Chung-hye. 2007. Pied-piping in relative clauses: syntax and compositional semantics using Synchronous Tree Adjoining Grammar. *Research on Language and Computation* 5:457–479.

Han, Chung-hye, and Nancy Hedberg. 2006. A Tree Adjoining Grammar Analysis of the Syntax and Semantics of *It*-clefts. In *Proceedings of the 8th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+ 8)*, 33–40. COLING-ACL workshop, Sydney, Australia.

Hedberg, Nancy. 1990. Discourse pragmatics and cleft sentences in English. Doctoral Dissertation, University of Minnesota.

Hedberg, Nancy. 2000. The referential status of clefts. *Language* 76:891–920.

Heggie, Lorie A. 1988. The syntax of copular structures. Doctoral Dissertation, University of Southern California, Los Angeles.

Heycock, Caroline, and Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface. *Linguistic Inquiry* 30:365–397.

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

Jespersen, Otto. 1927. *A modern English grammar*, volume 3. London: Allen and Unwin.

Jespersen, Otto. 1937. *Analytic syntax*. London: Allen and Unwin.

Joshi, Aravind K. 1985. Tree Adjoining Grammars: How much context sensitivity is required to provide a reasonable structural description. In *Natural Language Parsing*, ed. D. Dowty, L. Karttunen, and A. Zwicky, 206–250. Cambridge, U.K.: Cambridge University Press.

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 *Proceedings of the Third International Workshop on Computational Semantics (IWCS-3)*, ed. H. C. Blunt and E. G. C. Thijsse, 131–145. Tilberg.

Joshi, Aravind K., K Vijay-Shanker, and David Weir. 1991. The convergence of mildly context-sensitive grammatical formalisms. In *Foundational issues in natural language processing*, ed. Peter Sells, Stuart Shieber, and Tom Wasow, 31–82. Cambridge, MA: MIT Press.

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. 2008. Scope and situation binding in LTAG using semantic unification. *Research on Language and Computation* 6:3–52.

Kallmeyer, Laura, and Tatjana Scheffler. 2004. LTAG analysis for pied-piping and stranding of wh-phrases. In *Proceedings of TAG+7*, 32–39. Vancouver, Canada.

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 *Alternative conceptions of phrase structure*, ed. Mark Baltin and Anthony Kroch, 66–98. Chicago: University of Chicago Press.

Kroch, Anthony, and Aravind Joshi. 1985. Linguistic relevance of Tree Adjoining Grammar. Technical report, Department of Computer and Information Sciences, University of Pennsylvania. MS-CS-85-16.

Kroch, Anthony, and Beatrice Santorini. 1991. The derived constituent structure of the West Germanic verb raising construction. In *Principles and parameters in comparative*

grammar, ed. Robert Freidin, 269–338. Cambridge, MA: MIT Press.

Kroch, Anthony S., and Aravind K. Joshi. 1987. Analyzing extraposition in a Tree Adjoining Grammar. In *Discontinuous constituents*, ed. G. Huck and A. Ojeda, volume 20 of *Syntax and Semantics*. Academic Press.

Kuno, Susumu. 1987. *Functional syntax: Anaphora, discourse and emphathy*. Chicago: University of Chicago Press.

Mikkelsen, Lina. 2005. *Copular clauses: Specification, predication and equation*. John Benjamins.

Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.

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

Partee, Barbara. 1986. Ambiguous pseudoclefts with unambiguous be. In *Proceedings of NELS*, ed. S. Berman, J. Choe, and J. McDonough, volume 16, 354–366. University of Massachusetts, Amherst: GLSA.

Percus, Orin. 1997. Prying open the cleft. In *Proceedings of the 27th Annual Meeting of the North East Linguistics Society*, ed. K. Kusumoto, 337–351. GLSA.

Prince, Ellen. 1978. A comparison of wh-clefts and it-clefts in discourse. *Language* 54:883–906.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657–720.

Rochemont, Michael. 1986. *Focus in generative grammar*. Amsterdam: John Benjamins.

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*, 186–193. Vancouver, Canada.

Shieber, Stuart. 1994. Restricting the weak-generative capacity of Synchronous Tree-Adjoining Grammars. *Computational Intelligence* 10.

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

Vijay-Shanker, K, and Aravind K. Joshi. 1988. Feature structure based Tree Adjoining Grammars. In *Proceedings of the 12th International Conference on Computational Linguistics*, 714–719. Budapest.

Weir, David. 1988. Characterizing mildly context-sensitive grammar formalisms. Doctoral Dissertation, University of Pennsylvania.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11:203–238.

Williams, Edwin. 1983. Semantic vs. syntactic categories. *Linguistics and Philosophy* 6:423–446.

Wirth, Jessica R. 1978. The derivation of cleft sentences in English. *Glossa* 12.

Zribi-Hertz, Anne. 1989. A-type binding and narrative point of view. *Language* 65:695–727.