Binding Variables in English: An Analysis Using Delayed Tree Locality*

Dennis Ryan Storoshenko
Department of Linguistics
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6, Canada
dstorosh@sfu.ca

Chung-hye Han
Department of Linguistics
Simon Fraser University
8888 University Drive
Burnaby, BC V5A 1S6, Canada
chunghye@sfu.ca

Abstract

This paper presents an analysis of bound variable pronouns in English using Synchronous Tree Adjoining Grammar. Bound variables are represented as multi-component sets, composing in delayed tree-local derivations. We propose that the observed anti-locality restriction on English bound variables can be formalised in terms of a constraint on the delay in the composition of the bound variable multi-component set. While most cases are captured in a derivation making use of two simultaneous delays, maintaining weak equivalence with flexible composition, our analysis is open to derivations with an unlimited number of simultaneous delays.

1 Introduction

The English pronouns in (1a) and (1b) do not have the same function as referential pronouns. Instead, they function as bound variables, their references determined by the c-commanding antecedent. The relationship between the antecedent (binder) and the bound variable is difficult to capture in standard TAG, as the dependency between them is necessarily non-local. The predicate in (1a) intervenes between the variable and its binder, and this dependency is even further stretched in (1b) where two predicates intervene.

(1) a. Every girl$_i$ loves **her**$_i$ father.
b. Every girl$_i$ knows that **she**$_i$ is smart.

To capture these cases, a TAG variant is needed which will allow for this type of non-local derivation without excessively increasing generative capacity. In this paper, we show that Delayed Tree-Local Multi-Component (MC) TAG, demonstrated by Chiang and Scheffler (2008) to be weakly equivalent to standard TAG, permits exactly this kind of non-local derivation. We show that 2-delayed tree-local derivation is sufficient to handle core cases such as in (1), though a generalization to k-delayed tree-local derivation is needed to handle complicated cases where a bound variable is embedded in a DP that has another bound variable. Our analysis of bound variable anaphora in English also makes use of Synchronous Tree Adjoining Grammar (STAG) as formulated by Shieber (1994), augmented with syntactic feature agreement (Vijay-Shanker and Joshi, 1988). In Section 2, we show our analysis of the core cases such as (1a) and (1b). We then show, in Sections 3 and 4, how semantic and syntactic well-formedness constraints work together to rule out certain ungrammatical cases, and argue for the necessity of an anti-locality constraint based on the size of delays. In Section 5, we briefly discuss the cases that require generalization to k-delayed tree-local derivation.

2 The Analysis of Core Cases

Elementary trees for (1a) are presented in Figure 1. In the semantic trees, nodes are labelled as (T)erms, (R)elations, and (F)ormulae. Indices are included on substitution sites not only as a mark of syntactic movement, but also to identify substitution sites in derivation trees.

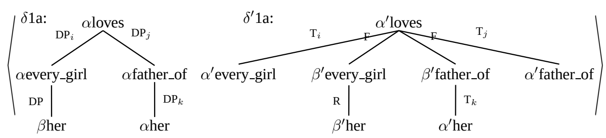

Figure 2: Derivation trees for *Every girl$_i$ loves her$_i$ father.*

Derivation trees for (1a) are shown in Figure 2. The syntactic tree (αevery_girl) treats the quantifier as a single DP, but crucially, the semantic side

*We thank the anonymous reviewers of TAG+10 for their insightful comments. All remaining errors our ours. This work was partially supported by NSERC RGPIN/341442 to Han.

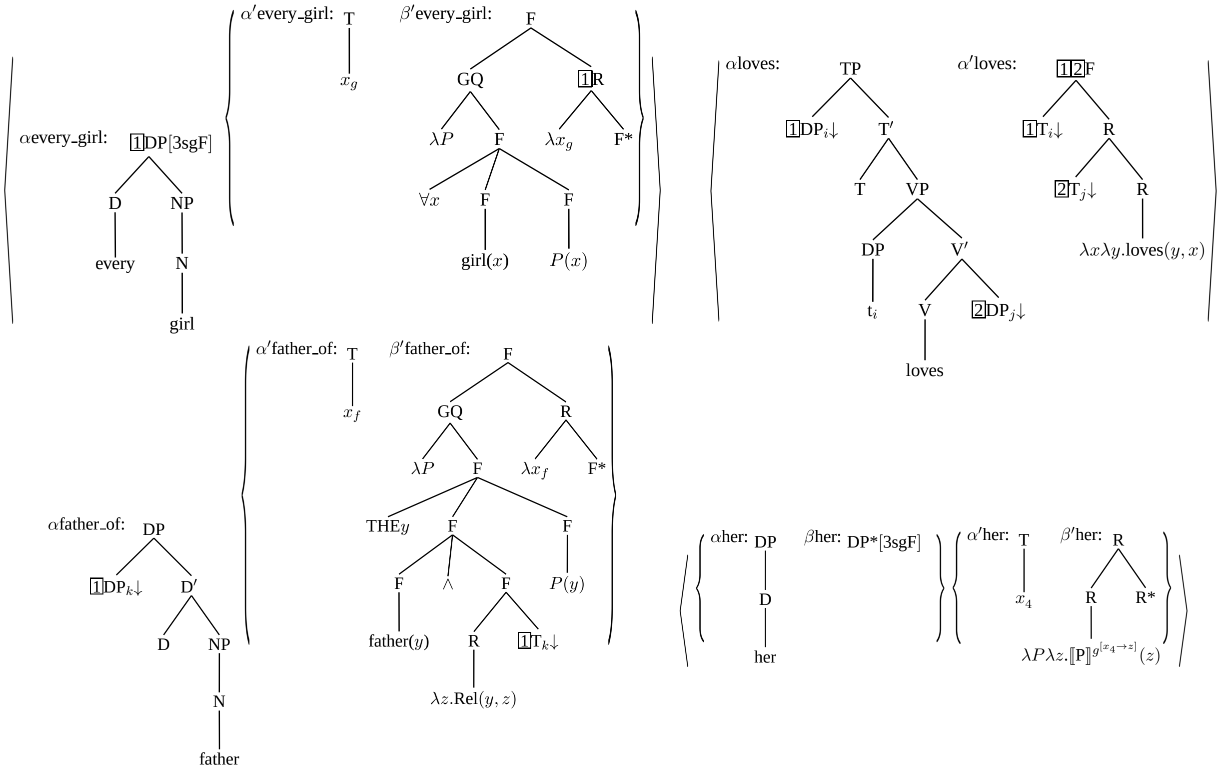

Figure 1: Elementary trees for *Every girl$_i$ loves her$_i$ father.*

is an MC set. (α'every_girl) is a variable which substitutes into an argument position in (α'loves). (β'every_girl) is an auxiliary tree which adjoins at the root of (α'loves), taking advantage of the multiple links (indicated by boxed numerals) between the syntax and semantics trees. A syntactic argument position links to two positions in the semantics: one for the argument variable, and another at the predicate's root where scope is calculated. In this way, isomorphism of the derivations is maintained despite one syntactic tree corresponding to an MC set in the semantics. (β'every_girl) presents a generalised quantifier (GQ) analysis (Barwise and Cooper, 1981), as implemented for STAG in Han et al. (2008). The trees for *father_of* are similar, implementing a GQ analysis for possession. Following Shieber and Schabes (1990) and Kallmeyer and Joshi (2003), we leave unspecified the order of adjoining for the scope portions of the GQs at the root of (α'loves). The possessor is the bound variable *her*, an MC tree set in both syntax and semantics. (αher) is a DP, which substitutes into (αfather_of). There is a defective auxiliary tree (βher) which adjoins at the root of (αevery_girl); syntactic agreement is captured in the union of ϕ features at this adjoining site. The semantic side follows the same derivation: (α'her) substitutes into the linked argument position in (β'father_of), and (β'her) adjoins into (β'every_girl), between the GQ and the binder, λx_g. (β'her) contains a condensed representation of the binder index evaluation rule presented in Büring (2005), using one function to show both steps of altering the assignment function on the relation created by the binder portion of (β'every_girl), and then re-binding the remaining variable inside. This derivation is licit under the definition of 2-Delayed Tree-Local MC-TAG, in that there are no more than two simultaneous delays. Delays are defined as sets of derivation tree nodes along the shortest path between members of an MC set, excluding the lowest node dominating both members of the MC set. As shown in (2), there are three delays in the semantic derivation, but no one node in the derivation tree participates in more than two delays.

(2) Delay for *every_girl*:
$\{\alpha'$every_girl, β'every_girl$\}$
Delay for *father_of*:

{α'father_of, β'father_of}
Delay for *her*:
{α'her, β'her, β'father_of, β'every_girl }

In the syntactic derivation, only one delay is present:

(3) Delay for *her*:
{αher, βher, αfather_of, αevery_girl }

While this delay is not identical to the semantic one, it is set-isomorphic in that both delays for *her* contain members of the *father_of* and *every_girl* sets. The difference is that on the syntax side, composition of (βher) is with (αevery_girl) while (β'her) is composed with (β'every_girl), which has no equivalent in the syntax.

The final derived trees are shown in Figure 3. Recalling the ambiguous ordering of adjoining at the root of (α'loves), we only show the derived semantic tree for the ordering where (β'father_of) adjoins before (β'every_girl); though the alternate order is available, it results in the x_4 variable remaining unbound, and we assume this is blocked by a constraint against unbound variables. Semantic composition on the tree in (γ1a) yields the formula in (4), showing the binding relationship between *every girl* and *her*.

(4) $\forall x$[girl(x)][THEy[father(y) $\wedge$ Rel(y,x)][loves(x,y)]]

A similar derivation is possible for the example in (1b), with additional trees shown in Figure 4. Following the derivation in Figure 5, we arrive at the derived trees in Figure 6. Again, the derivation has no more than two simultaneous delays. The final semantic form is shown in (5), and the expected variable binding comes through the derivation.

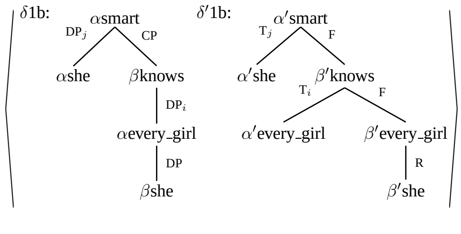

Figure 5: Derivation trees for *Every girl$_i$ knows that she$_i$ is smart.*

(5) $\forall x$[girl(x)][knows(x,smart(x))]

3 Blocking Spurious Derivations

There are some derivations which our analysis must block, shown in (6). For the case of (6a), the standard explanation is that the variable is not c-commanded by its quantifier. Making use of previously-presented elementary trees, the derivation of (6a) is shown in Figure 7.

(6) a. * **She**$_i$ thinks that every girl$_i$ is smart.
b. * Every girl$_i$ loves **her**$_i$
c. Every girl$_i$ loves **herself**$_i$

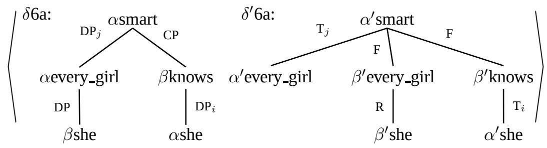

Figure 7: Derivation trees for **She$_i$ knows that every girl$_i$ is smart.*

Note that there is nothing about the derivation itself which blocks (6a): the same delays are observed as in (1b). However, performing semantic composition on the derived semantic tree in Figure 8 yields (7), which leaves the x_2 variable unbound, similar to the blocked derivation for (1a).

(7) thinks($x_2, \forall x$[girl(x)][smart(x)])

The situation in (6b) is more complex. This example can be derived using familiar elementary trees, with derivation trees shown in Figure 9. The derived trees in Figure 10 result in the semantic form given in (8); all variables are bound, and the intended reading comes out, yet the example is ungrammatical.

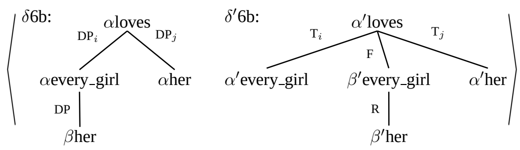

Figure 9: Derivation trees for **Every girl$_i$ loves her$_i$*.

(8) $\forall x$[girl(x)][loves(x, x)]

For this, we propose a constraint on the derivation itself, based on the delays. Nesson and Shieber (2009) propose that locality on MC sets can be measured in terms of the size of a delay. For all the previous examples, the cardinality of a delay for a

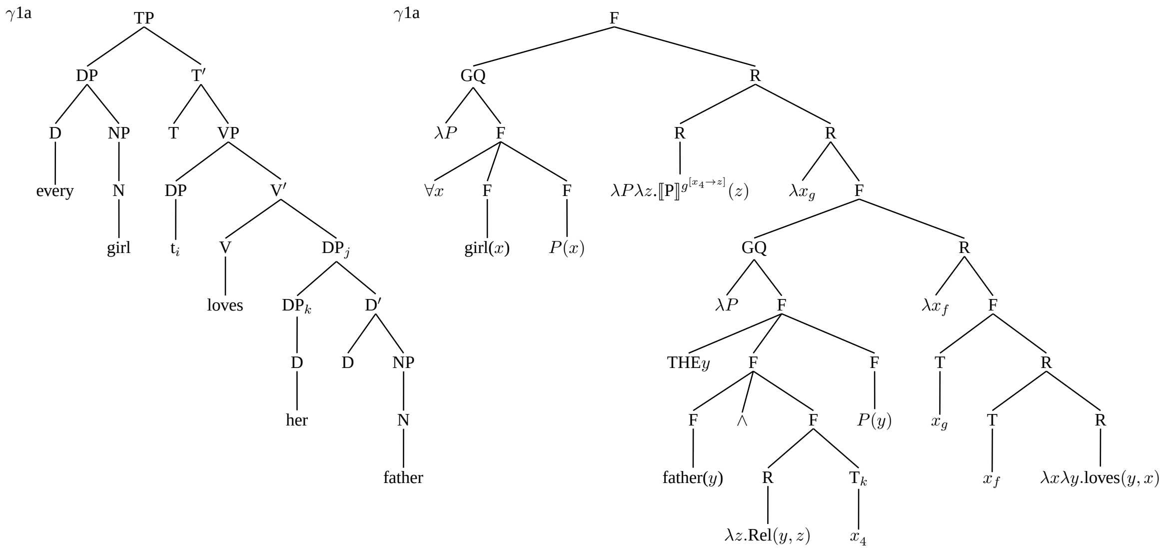

Figure 3: Derived trees for *Every girl$_i$ loves her$_i$ father.*

bound variable was at least four. For (6b), the delay is smaller, with a cardinality of only three. We thus propose a constraint on derivations containing bound variable trees in English: the cardinality of the delay of an MC set for a bound variable must be at least four, imposing a minimum distance between the variable and its antecedent. The grammatical equivalent of (6b), using a reflexive in (6c), can be captured with the analyses of either Frank (2008) or Storoshenko et al. (2008).

4 Capturing Crossover

In the literature on bound variable anaphora, a widely-known constraint is that against crossover, coming in two flavours, weak and strong. For both cases, the analysis is that an antecedent in a derived position binds a variable it did not originally c-command. Looking at the examples in (9), crossover will result after quantifier raising. In strong crossover, the variable c-commands the quantifier's base position, shown in (9a), but in weak crossover, the (9b) case, this is not so.

(9) a. * **She**$_i$ loves every girl$_i$

b. * **Her**$_i$ father loves every girl$_i$

(9a), derived according to Figure 11, is semantically identical to (6b) after all composition has been completed on the derived trees in Figure 12. The same constraint on the delay will rule out this example, as the cardinality of the delay of the MC set for the bound variable is again just three. Furthermore, Condition C, implemented for STAG, would rule out such an example.

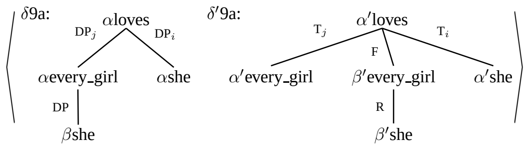

Figure 11: Derivation trees for **She$_i$ loves every girl$_i$*.

However, the same constraints will not account for (9b). Recalling the discussion of (1a), there are two possible derivations where there are two GQs, one of which leaves the variable contributed by (α'her) unbound. However, a perfectly legitimate derivation is possible, shown in Figure 13. This example cannot be blocked on the basis of the delay size constraint, as the delay of the MC set for the bound variable has a cardinality of four. Semantic composition from the derived trees in Figure 14 results in the semantic form in (10) with the variable bound, and the intended meaning intact.

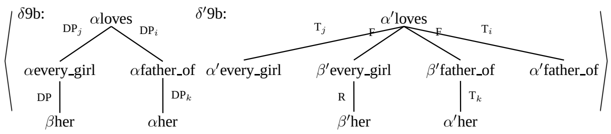

Figure 13: Derivation trees for **Her$_i$ father loves every girl$_i$*.

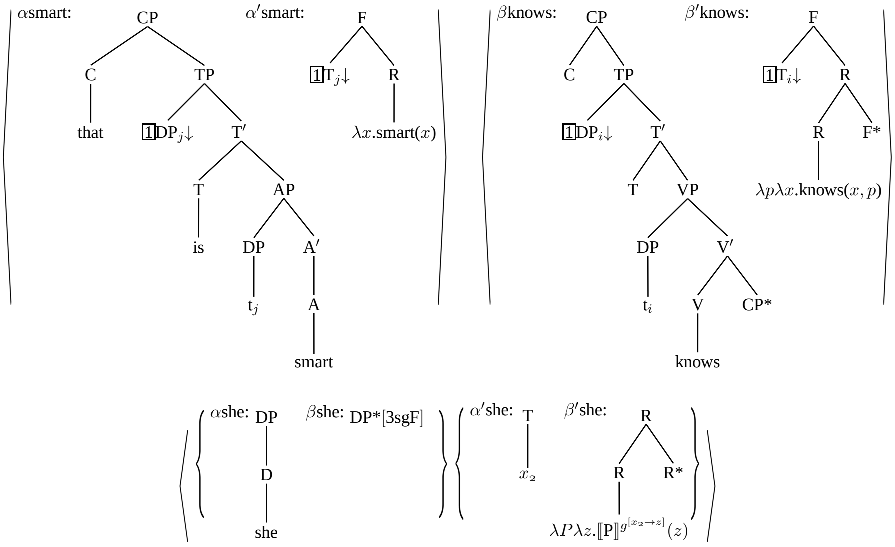

Figure 4: Additional elementary trees for *Every girl$_i$ knows that she$_i$ is smart.*

(10) $\forall x$[girl(x)][THEy[father(y) $\wedge$ Rel(y, x)][loves(y, x)]]

To block this, we impose one final constraint on the syntax of the bound variable, a c-command constraint between the elementary trees of the bound variable MC set: in the derived syntactic tree, the defective DP* elementary tree must c-command the argument DP tree. In (9b), (βher) is adjoined at the root of (αevery_girl), while (αher) substitutes at a higher position in (αloves); the necessary c-command relation does not hold, ruling out this sentence. The same constraint will also rule out (9a), and it will likewise rule out (6a), both of which violated other constraints as well.

5 Complicated Cases

The examples presented in this paper so far have all been restricted to 2-delayed tree-local derivations. There are however examples which, if treated under our present analysis, will require more than 2 simultaneous delays in the derivation. These are cases where more than one bound variable is embedded in a DP, as in (11).[1]

(11) a. Every girl$_i$ showed a boy$_j$ some picture of him$_j$ by her$_i$.

b. Every girl$_i$ told a boy$_j$ that some professor$_k$ liked a picture of him$_j$ that she$_i$ gave him$_k$.

For instance, as can be seen from the semantic derivation tree of (11a) in Figure 15, (α'some_picture_of) occurs in 3 delays, those of *some_picture_of*, *him* and *by_her*. And in (11b), it occurs in 4 delays, those of *some_picture_of*, *him$_j$*, *she*, and *him$_k$*. So, as the number of bound variables embedded in a DP increases, so does the number of simultaneous delays in the derivation. As embedding is in principle unbounded, we cannot put a formal bound on the number of simultaneous delays required to handle bound variables, though Tatjana Scheffler (p.c.) points out that the number of elementary trees will ultimately limit the number of delays in a given derivation—it's not the case that any one derivation will have an unbounded number of delays. Still, we speculate that as the number of simultaneous delays increases, so does the processing load in deriving the sentence. Speakers encountering a 4-delay example such as (11b) may have difficulty in reaching the desired interpretation.[2]

[1]Thanks to a TAG+10 reviewer for pointing this out to us and providing us with these examples.

[2]Chiang and Scheffler (2008) has shown that 2-delayed tree-local MC-TAG is weakly equivalent to MC-TAG with flexible composition. The existence of such examples as in (11) which require even further simultaneous delays can be argued to show that delayed tree local derivations are preferable to derivations using flexible composition in that they permit such sentences to be formed.

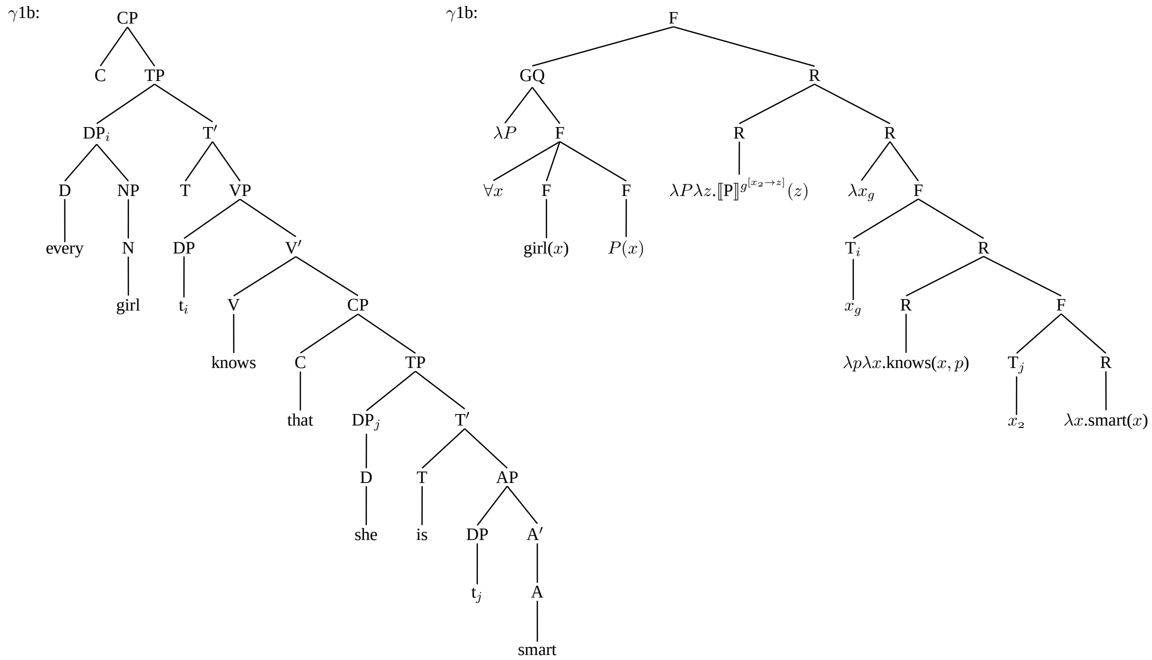

Figure 6: Derived trees for *Every girl$_i$ knows that she$_i$ is smart.*

6 Conclusion and Implications

In this paper, we have presented an analysis of bound variable anaphora for English in STAG. This analysis presents the bound variable as an MC set in the syntax and the semantics, and crucially makes use of delayed tree-locality in the derivation. We have proposed three different constraints on the derivations: a syntactic constraint which was necessary to rule out weak crossover, a semantic constraint against derivations with unbound variables, and a derivation constraint which enforces a degree of anti-locality, to account for the case where a reflexive must be used. While some derivations violate multiple constraints, each constraint is vital in ruling out at least one ungrammatical example. The syntactic and semantic constraints are quite standard in the literature on bound variables, and are relatively uncontroversial. In future work, we hope to explore possible parametric variation in the delay constraint, accounting for languages where bound variables are either more strictly local, or more flexible in their use than in English. Our analysis has not touched on co-referential, rather than bound, uses of English pronouns. These we assume to be captured under an STAG implementation of Condition B, possibly along the lines of the LTAG binding theory proposed in Champollion (2008). Finally, acknowledging that our present analysis requires a c-command constraint between the variable and its antecedent, we leave for future work English cases such as *Someone from every city$_i$ is proud of its$_i$ history*, in which a pronoun with a bound variable interpretation is not c-commanded by its antecedent.

References

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.

Büring, Daniel. 2005. *Binding theory*. Cambridge University Press.

Champollion, Lucas. 2008. Binding theory in LTAG. In *Proceedings of TAG+9*, 1–8.

Chiang, David, and Tatjana Scheffler. 2008. Flexible composition and delayed tree-locality. In *Proceedings of TAG+9*, 17–24.

Frank, Robert. 2008. Reflexives and TAG semantics. In *Proceedings of TAG+9*, 97–104.

Han, Chung-hye, David Potter, and Dennis Ryan Storoshenko. 2008. Compositional semantics of coordination using Synchronous Tree Adjoining Grammar. In *Proceedings of TAG+9*, 33–41.

Kallmeyer, Laura, and Aravind K. Joshi. 2003. Fac-

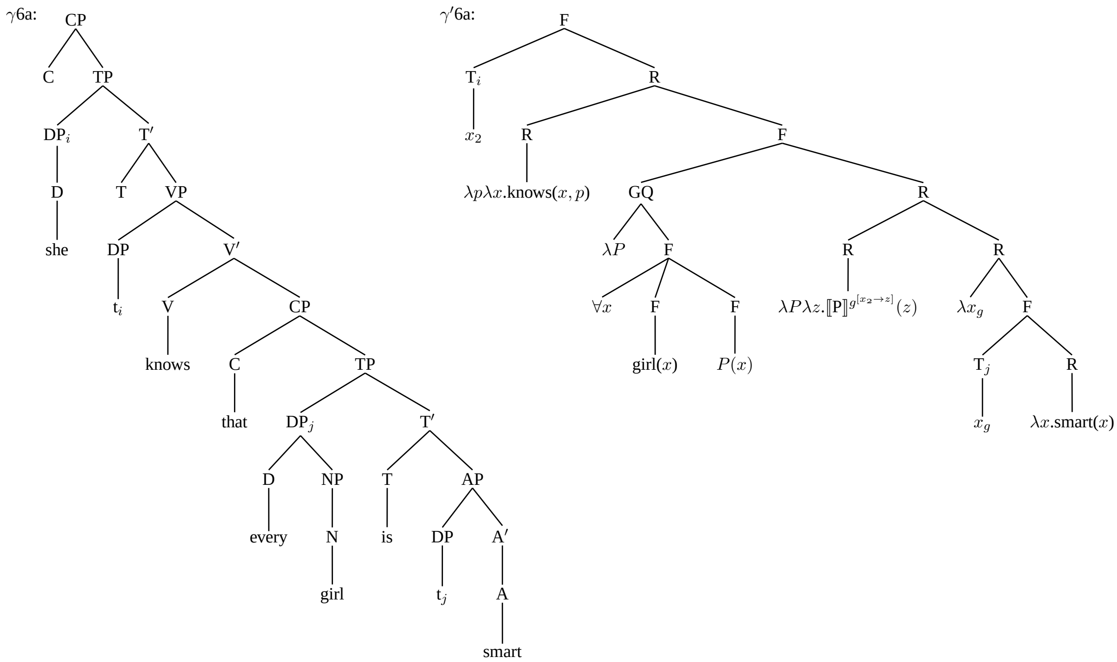

Figure 8: Derived trees for **She$_i$ knows that every girl$_i$ is smart.*

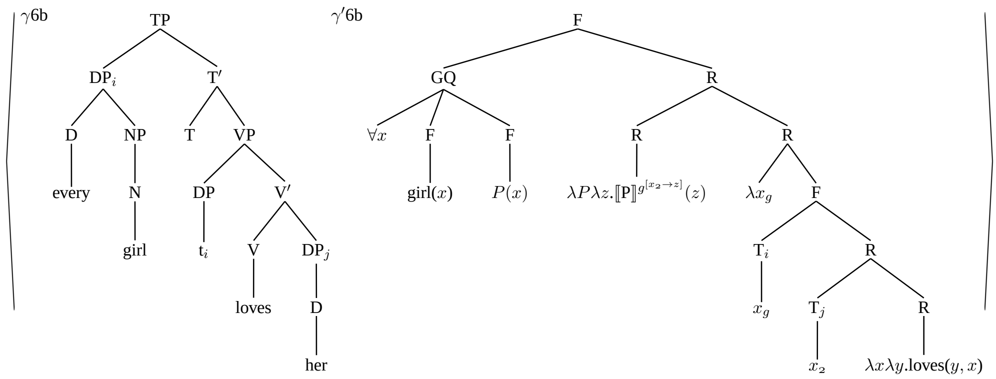

Figure 10: Derived trees for **Every girl$_i$ loves her$_i$.*

toring predicate argument and scope semantics: Underspecified semantics with LTAG. *Research on Language and Computation* 1:3–58.

Nesson, Rebecca, and Stuart M. Shieber. 2009. Efficiently parsable extensions to Tree-Local Multicomponent TAG. In *Proceedings of NAACL 2009*, 92–100.

Shieber, Stuart, and Yves Schabes. 1990. Synchronous Tree Adjoining Grammars. In *Proceedings of COLING'90*, 253–258.

Shieber, Stuart M. 1994. Restricting the weak generative capacity of Synchronous Tree Adjoining Grammars. *Computational Intelligence* 10:371–385.

Storoshenko, Dennis Ryan, Chung-hye Han, and David Potter. 2008. Reflexivity in English: An STAG analysis. In *Proceedings of TAG+9*, 149–157.

Vijay-Shanker, K., and Aravind Joshi. 1988. Feature structure based Tree Adjoining Grammars. In *Proceedings of COLING'88*, 714–719.

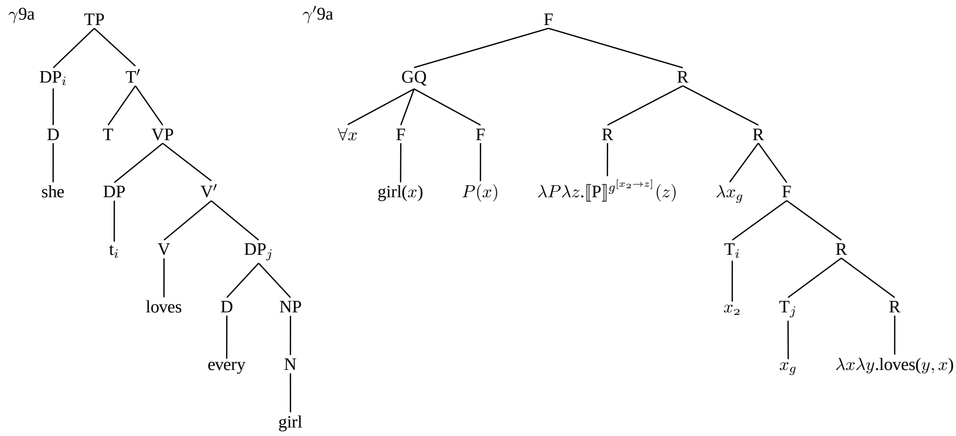

Figure 12: Derived trees for *She_i *loves every* $girl_i$.

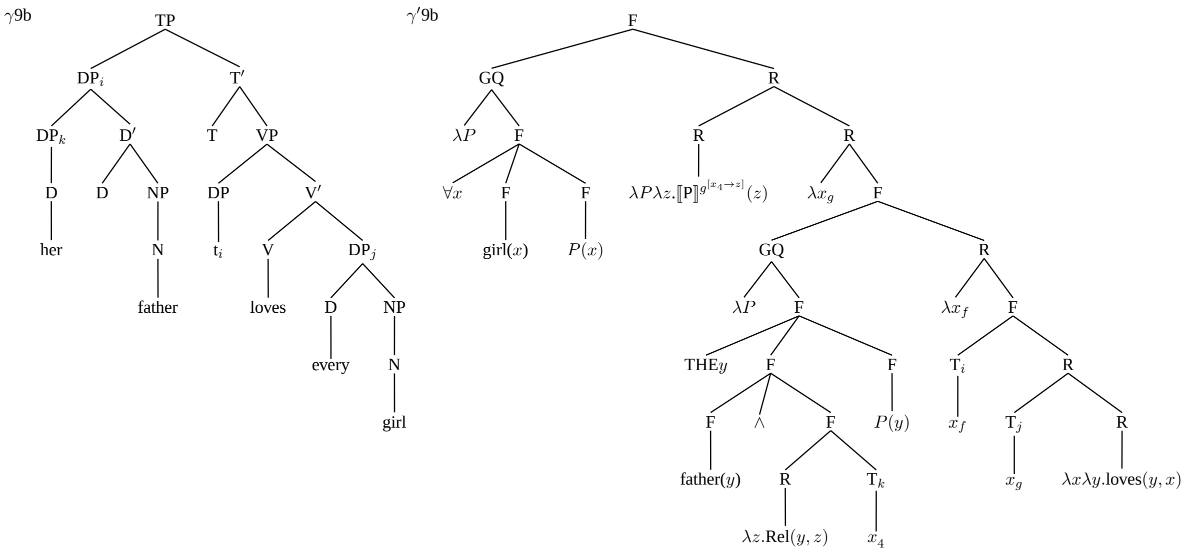

Figure 14: Derived trees for *Her_i *father loves every* $girl_i$.

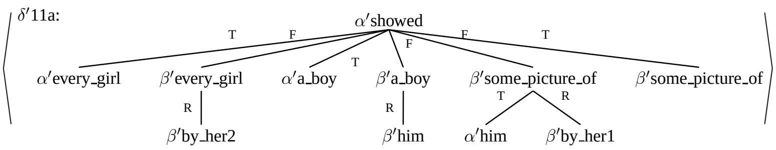

Figure 15: Semantic derivation tree for *Every* $girl_i$ *showed a* boy_j *some picture of* him_j *by* her_i