

## Reflexivity in English: an STAG analysis

**Dennis R. Storoshenko**

Department of Linguistics  
Simon Fraser University  
dstorosh@sfu.ca

**Chung-hye Han**

Department of Linguistics  
Simon Fraser University  
chunghye@sfu.ca

**David Potter**

Department of Linguistics  
Simon Fraser University  
dkpl@sfu.ca

### Abstract

In this paper, we present an STAG analysis of English reflexives. In the spirit of Ryant and Scheffler (2006) and Kallmeyer and Romero (2007), reflexives are represented as a multi-component set in the syntax, with a degenerate auxiliary tree controlling the  $\phi$  feature agreement between a reflexive and its antecedent. On the semantics side, the reflexive is a valence-reducing  $\lambda$ -expression, identifying two arguments of a single predicate. We then demonstrate that with minimal modifications, our analysis can be extended to capture raising and ECM cases. Finally, we argue that Condition A of Chomsky's binding theory can be derived as a consequence of our treatment of reflexives.

### 1 Introduction

Synchronous Tree Adjoining Grammar (STAG) provides an isomorphic mapping of derivations between a pair of TAG grammars. This mapping can be exploited to map a source syntactic derivation to an isomorphic semantic derivation, which derives a semantic representation for a sentence by combining semantic elementary trees (Shieber, 1994). As a result, STAG is a useful tool for analyzing natural language phenomena at the syntax/semantics interface (Han and Hedberg, 2006; Nesson and Shieber, 2006; Han, 2007; Nesson and Shieber, 2007). We extend that research by presenting an STAG analysis for reflexive pronouns in English, augmented with syntactic feature unification as defined in Vijay-Shanker and Joshi (1988). For the semantic elementary trees, we follow Han

(2007) in using unreduced  $\lambda$ -expressions. This allows  $\lambda$ -conversion to apply in the semantic derived tree, producing the final logical form. Our approach uses three different forms of the reflexive, T'-form, V'-form and TP-form, each represented as a multi-component set in syntax, following Ryant and Scheffler (2006) and Kallmeyer and Romero (2007), and as a reflexive function in semantics. With this, we capture all the core verbal argument cases of reflexive use. We further show how only one of the three forms is acceptable in a given sentence and how Condition A of Chomsky's (1981) binding theory can be derived as a consequence of our analysis.

While we adopt the same basic syntax as Ryant and Scheffler and Kallmeyer and Romero, semantically our approaches are quite different. The previous approaches employ semantic feature unification in the derivation structure (Kallmeyer and Romero, 2008), with composition taking place in a flat, conjunction-based semantics. Our approach uses  $\lambda$ -calculus on the semantic derived tree, which is constructed using the derivation structure on the semantics side that is isomorphic to the derivation structure on the syntax side. Through this, we are more readily able to capture the insights of Reinhart and Reuland (1993), representing our reflexive as a function upon predicates, rather than a relationship between two nominals, the reflexive and its antecedent. As a consequence of this, we make use of different forms of the reflexive depending upon where it appears in a predicate's argument structure. By choosing this approach in which the reflexive works upon a predicate, we are able to capture instances of reflexives occurring in both mono- and multi-clausal

environments within the lexical entry of the reflexive itself.

In section 2, we present our analysis of reflexive binding in mono-clauses. We then extend our analysis to reflexive binding in raising sentences in section 3 and then to instances of exceptional case marking (ECM) sentences in section 4.

## 2 Mono-clausal Reflexives

In the simplest cases, a reflexive appears in the same clause as its antecedent.

- (1) Jim<sub>4</sub> introduces himself<sub>4</sub> to Bill<sub>5</sub>.
- (2) Jim<sub>4</sub> introduces Bill<sub>5</sub> to himself<sub>5</sub>.
- (3) Jim<sub>4</sub> introduces Bill<sub>5</sub> to himself<sub>4</sub>.

Elementary trees for (1) are in Figure 1. In ( $\alpha$ introduces), each DP argument substitution site is specified with an unvalued  $\phi$  feature, which will unify with a  $\phi$  feature from the substituted DP. We adopt the feature structures proposed in Vijay-Shanker and Joshi (1988) and the conception of feature unification defined therein. Each node has a Top feature (notated as  $t$  :), and a Bottom feature (notated as  $b$  :). At the end of a derivation, the Top and Bottom features at each node must unify; incompatible feature values will cause a derivation to crash. In ( $\alpha$ introduces), the  $\phi$  features from the DP subject and the DP direct object are passed over as Top features on the sister bar-level node, and Bottom features on the next highest maximal projection. When adjoining takes place, the Top features of the adjoining site must unify with the Top features of the adjoining auxiliary tree's root node, and the Bottom features of the adjoining site unify with the auxiliary tree's foot node Bottom features. ( $\alpha$ introduces) is paired with a semantic elementary tree ( $\alpha'$ introduces). In the semantic tree, F stands for formula, R for relation and T for term. We will assume that T can host reflexive functions as well as argument variables and constants. Boxed numerals indicate links between the syntactic and semantic elementary tree pairs; if an operation is carried out at one such node on the syntax side, a corresponding operation is carried out at the linked node(s) in the semantics. For simplicity, we only indicate links which are required in the derivation of the example sentences.

The reflexive employed for (1) is a T'-form, identified as  $himself_{T'}$ . In the syntactic multi-component set, ( $\alpha$ himself $_{T'}$ ) bears a  $\phi$  feature

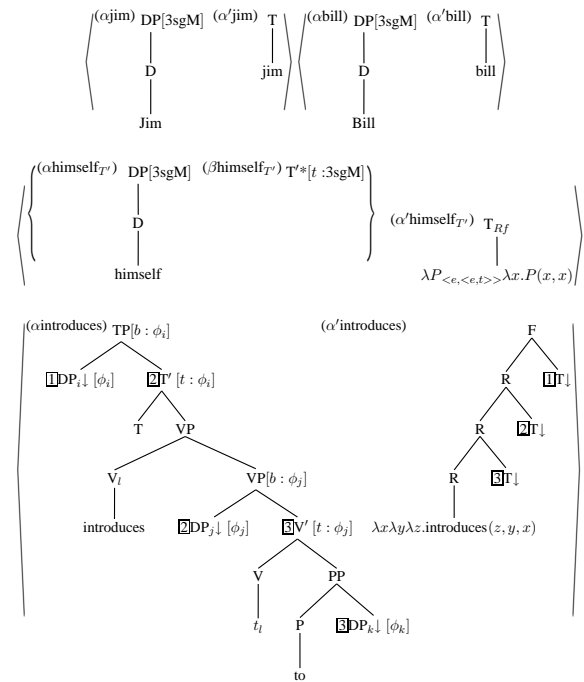


Figure 1: Elementary trees for *Jim<sub>4</sub> introduces himself<sub>4</sub> to Bill<sub>5</sub>*

and will substitute into  $DP_j$  in ( $\alpha$ introduces), and ( $\beta$ himself $_{T'}$ ) is a degenerate T' auxiliary tree, specified with a Top  $\phi$  feature. As in Kallmeyer and Romero (2007), our ( $\beta$ himself $_{T'}$ ) ensures the agreement between the reflexive and its antecedent, the subject DP in [Spec,TP], by adjoining at T' in ( $\alpha$ introduces). The Top  $\phi$  feature of ( $\beta$ himself $_{T'}$ ) must unify with the Top  $\phi$  feature of T', which in turn must agree with the Bottom  $\phi$  feature of TP and the  $\phi$  feature of the subject DP in ( $\alpha$ introduces) through coindexation. Crucially, this is the only syntactic constraint at work. In the semantics, ( $\alpha'$ himself $_{T'}$ ) introduces a function of type  $\langle\langle e, \langle e, t \rangle \rangle, \langle e, t \rangle\rangle$ . This function is labelled as  $T_{Rf}$  (Rf for reflexive), and substitutes into the T node labeled with link  $\boxed{2}$  in ( $\alpha'$ introduces). After  $\lambda$ -conversion, this function returns an  $\langle e, t \rangle$  type predicate where the argument variable corresponding to *himself* and an argument variable corresponding to the antecedent are identified. The isomorphic syntactic and semantic derivation structures are given in Figure 2, and the syntactic and semantic derived trees in Figure 3.

Virtually the same set of trees will derive (2). The only difference is that the form of the reflexive employed here is the V'-type, as defined in

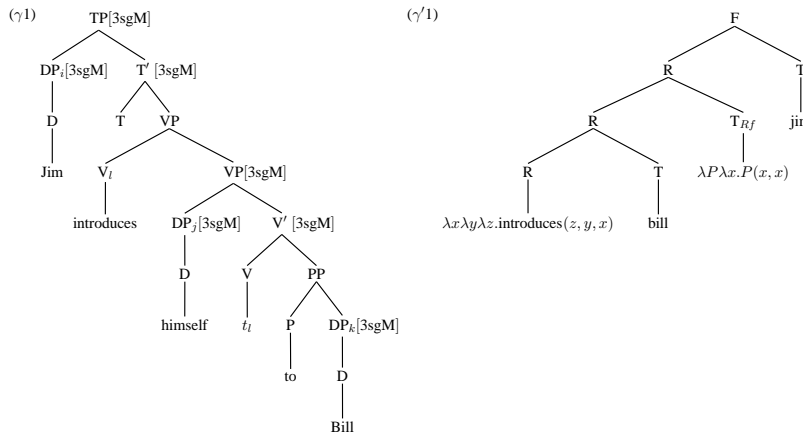


Figure 3: Derived trees for  $Jim_4$  introduces  $himself_4$  to  $Bill_5$

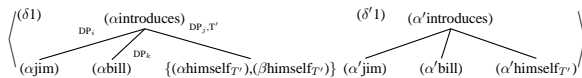


Figure 2: Derivation structures for  $Jim_4$  introduces  $himself_4$  to  $Bill_5$

Figure 4. The  $(\beta himself_{V'})$  adjoins at the  $V'$  node in  $(\alpha introduces)$  in Figure 1, ensuring the agreement between the reflexive and its antecedent, the direct object DP in [Spec,VP]. On the semantics side,  $(\alpha' himself_{V'})$  introduces a function of type  $\langle\langle e, \langle e, \langle e, t \rangle \rangle \rangle, \langle e, \langle e, t \rangle \rangle \rangle$ , performing essentially the same operation as  $(\alpha' himself_{T'})$  in Figure 1. The derivation structures for (2) are given in Figure 5 and the derived trees in Figure 6. After  $\lambda$ -conversion has taken place on the semantic derived trees, the respective formulas for (1) and (2) are (4) and (5).

- (4) introduces(jim, jim, bill)
- (5) introduces(jim, bill, bill)

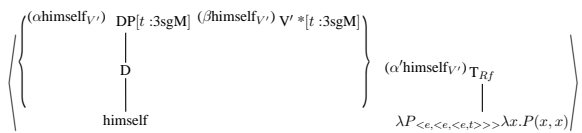


Figure 4: New elementary trees for  $Jim_4$  introduces  $Bill_5$  to  $himself_5$

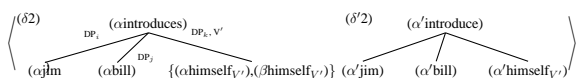


Figure 5: Derivation structures for  $Jim_4$  introduces  $Bill_5$  to  $himself_5$

To derive (3),  $T'$ -type reflexive must be employed but with a different semantic elementary tree from the one in Figure 1. The new  $T'$ -type reflexive tree pair is given in Figure 7.  $(\alpha' himself_{T'})$  in Figure 7 ensures that the variable corresponding to the indirect object  $himself$  and the variable corresponding to the subject antecedent are identified. The isomorphic syntactic and semantic derivation structures are given in Figure 8 and the syntactic and semantic derived trees in Figure 9. After  $\lambda$ -conversion has taken place on the semantic derived tree, the formula for (3) is (6).

- (6) introduces(jim,bill,jim)

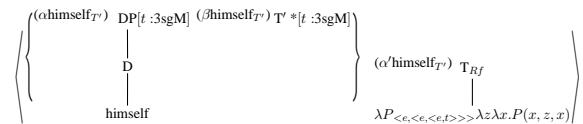


Figure 7: New elementary trees for  $Jim_4$  introduces  $Bill_5$  to  $himself_4$

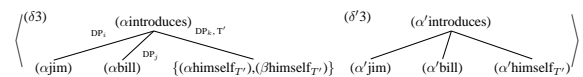
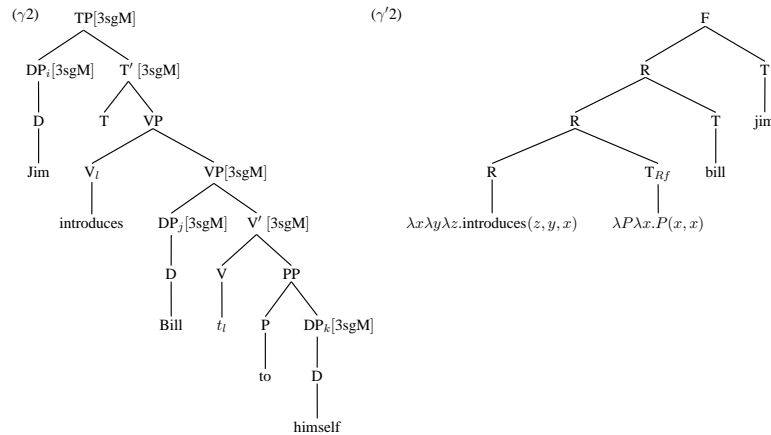
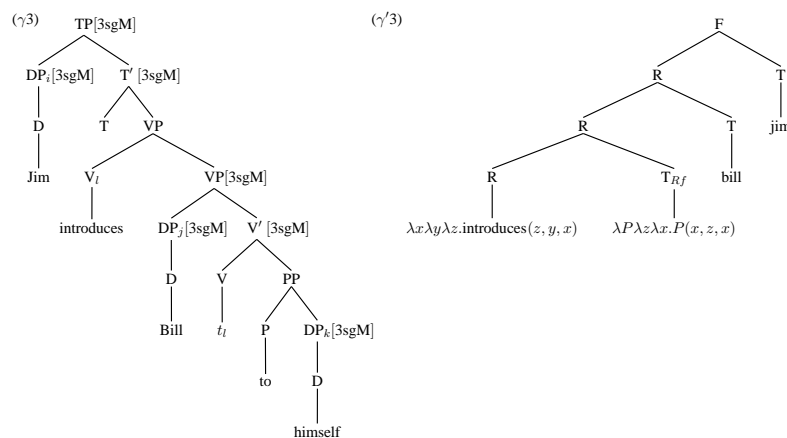


Figure 8: Derivation structures for  $Jim_4$  introduces  $Bill_5$  to  $himself_4$

Syntactic constraints on derivation emerge when considering cases where there is no agreement between the reflexive and its antecedent, as in (7).

- (7) \* Jim<sub>4</sub> introduces herself<sub>4</sub> to Gillian<sub>5</sub>.

Here, the reflexive would come with a degenerate  $T'$  tree  $(\beta herself_{T'})$  carrying a feature specifi-

Figure 6: Derived trees for  $Jim_4$  introduces  $Bill_5$  to  $himself_5$ Figure 9: Derived trees for  $Jim_4$  introduces  $Bill_5$  to  $himself_4$ 

cation of [3sgF]. However, substitution of ( $\alpha$ Jim) into ( $\alpha$ introduces) will transfer the value [3sgM] onto the T' node of that tree. This would block the adjoining of ( $\beta$ himself<sub>T'</sub>), as there would be a feature clash preventing unification.

Note also that (1) cannot be derived with the V'-type reflexive. While nothing in the syntax prevents the use of  $\{(\alpha$ himself<sub>V'</sub>), ( $\beta$ himself<sub>V'</sub>)\}, following the links through to the semantics would result in an illegal derivation. ( $\alpha'$ himself<sub>V'</sub>), which takes an argument of type  $\langle e \langle e \langle e, t \rangle \rangle \rangle$  would be substituted at a node where its sister is of semantic type  $\langle e \langle e, t \rangle \rangle$ . The semantic derivation would crash at this point, as functional application cannot be applied.

Thus, both the syntax and semantics work in concert to obviate spurious derivations. What is worth considering here is that illegal derivations have been blocked without any recourse to a constraint such as Condition A. At its core, Condi-

tion A consists of two stipulations: a locality requirement, and a structural relationship between a reflexive and its antecedent. Under our approach, the locality requirement is provided by the formalism, in that the composition of the multi-component set must remain local to a single elementary tree. A binding domain is thus naturally defined. Similar to Kallmeyer and Romero (2007), the c-command relationship between the reflexive and its antecedent is also a consequence of our analysis. The difference is that our analysis accomplishes this without stipulating a dominance relationship between the two members of the reflexive set in the syntax. As shown above, the semantic type of the reflexive's tree governs the location where it can be substituted in semantics. Following the links from the semantics back to the syntax, this translates into a constraint upon the structural relationship between the  $\alpha$  and  $\beta$  trees in the reflexive set. Only the derivation that pro-

duces a syntactic derived tree where the  $\beta$  tree of the reflexive set dominates the  $\alpha$  tree can be mapped onto a fully composable semantic derived tree. As in the case of Kallmeyer and Romero, this necessary dominance easily translates into the c-command constraint embedded within Condition A, as the  $\beta$  tree of the reflexive must be adjoined at a sister node to a potential antecedent. As a result, both portions of Condition A are consequences of the present analysis and constraints upon semantic well-formedness.

### 3 Raising

Our analysis of English reflexives is extendable to instances of raising, as in (8) and (9).

- (8) Jake<sub>4</sub> seems to himself<sub>4</sub> to be happy.
- (9) Julian<sub>4</sub> seems to Miles<sub>5</sub> to love himself<sub>4</sub>.

In the first raising case, (8), the reflexive is an argument of a different predicate than its antecedent. The elementary trees required for (8) are given in Figure 10. We use the *seems to* tree presented in Storoshenko (2006), extended with a matching semantic tree.<sup>1</sup> Following the derivation in Figure 11, in syntax,  $(\beta\text{himself}_{T'})$  adjoins to the  $T'$  root of  $(\beta\text{seems\_to})$ , unifying with its Top  $\phi$  feature. This feature must then unify with the Top  $\phi$  feature of  $T'$  in  $(\alpha\text{ happy})$ , the adjunction site for  $(\beta\text{seems\_to})$ , and agree (through coindexation) with the Bottom  $\phi$  feature of TP and the  $\phi$  feature of the subject DP in  $(\alpha\text{happy})$ . In semantics,  $(\alpha\text{himself}_{T'})$  substitutes into  $(\beta'\text{seems\_to})$ , which adjoins to  $(\alpha'\text{ happy})$ . Derived trees are shown in Figure 12. After  $\lambda$ -conversion on  $(\gamma'8)$  is complete, the formula for (8) is (10).

- (10)  $\text{seems\_to}(\text{happy}(\text{jake}), \text{jake})$
- (11)  $\text{seems\_to}(\text{love}(\text{julian}, \text{julian}), \text{miles})$

In the second raising case, (9), both antecedent and reflexive are arguments of the same predicate, to which  $(\beta\text{seems\_to})$  adjoins with a sep-

arate experiencer. The new elementary trees required for (9) are in Figure 13. As shown in Figure 14, in syntax,  $(\alpha\text{Miles})$  is substituted into  $(\beta\text{seems\_to})$ , which is then adjoined into  $(\alpha\text{to\_love})$ . Both components of the  $\text{himself}_{T'}$  set then compose with  $(\alpha\text{to\_love})$ :  $(\alpha\text{himself}_{T'})$  substitutes into  $\text{DP}_k$ , and  $(\beta\text{himself}_{T'})$  adjoins onto  $T'$ . Here, we assume multiple adjunction, as defined in Schabes and Shieber (1994), so that  $(\beta\text{himself}_{T'})$  and  $(\beta\text{seems\_to})$  adjoin to the same  $T'$  node in  $(\alpha\text{to\_love})$ . As  $(\beta\text{himself}_{T'})$  is a degenerate auxiliary tree, the order of adjoining is unimportant, as either order results in the same derived tree. In semantics,  $(\alpha'\text{himself}_{T'})$  substitutes into  $(\alpha'\text{to\_love})$  and  $(\beta'\text{seems\_to})$  adjoins to  $(\alpha'\text{to\_love})$ . The derived trees are in Figure 15.  $(\gamma'9)$  yields the formula in (11) after  $\lambda$ -conversion.

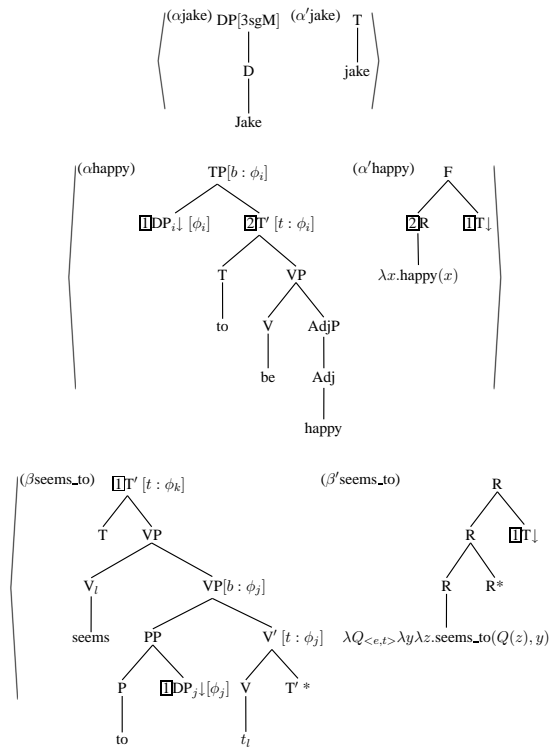


Figure 10: Elementary trees for *Jake<sub>4</sub> seems to himself<sub>4</sub> to be happy*

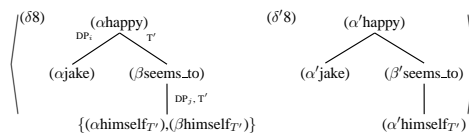


Figure 11: Derivation structures for *Jake<sub>4</sub> seems to himself<sub>4</sub> to be happy*

<sup>1</sup>A reviewer questions why the semantics of *seems to* presented here contains an argument slot for the subject of the embedded clause, when it is not present in the syntactic elementary tree of *seems to*. This is a function of the fact that the semantic elementary tree for *seems to* that we have defined adjoins to the predicate of type  $\langle e, t \rangle$  coming from the embedded clause. As this predicate takes an argument to return a proposition which is one of the arguments of *seems to*, an argument slot for the subject of the embedded clause is necessary in the  $\lambda$ -expression for *seems to*.

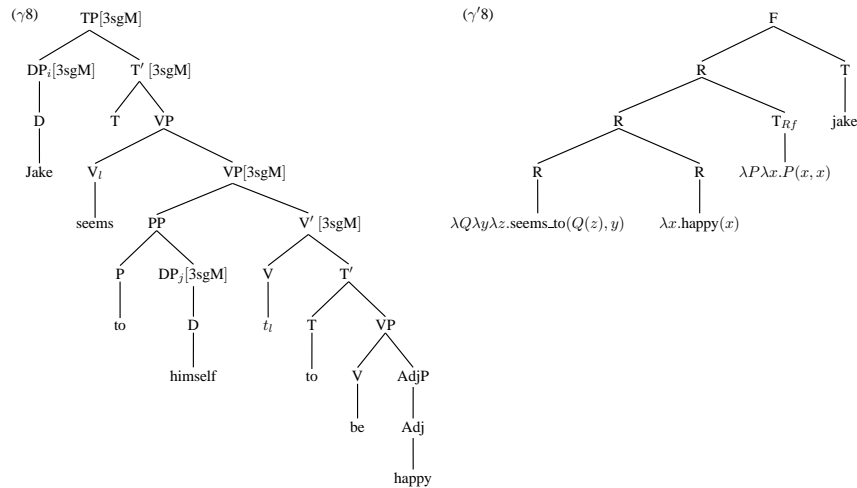


Figure 12: Derived trees for  $Jake_4$  seems to  $himself_4$  to be happy

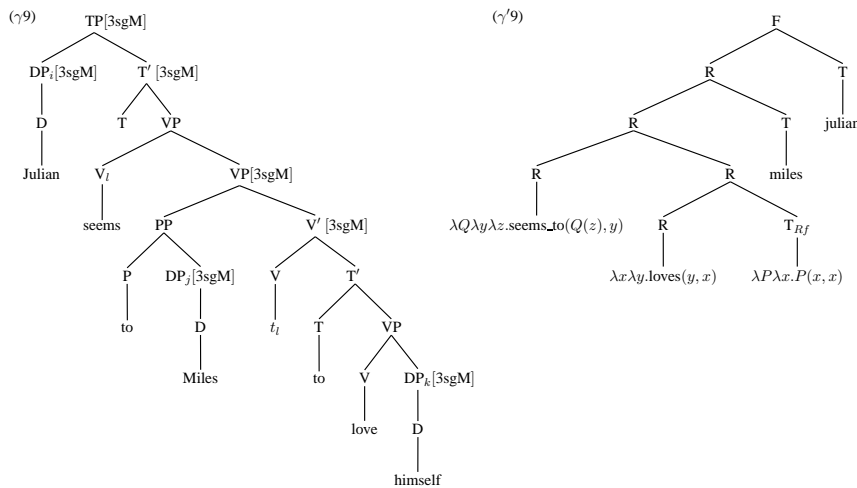


Figure 15: Derived trees for  $Julian_4$  seems to  $Miles_5$  to love  $himself_4$

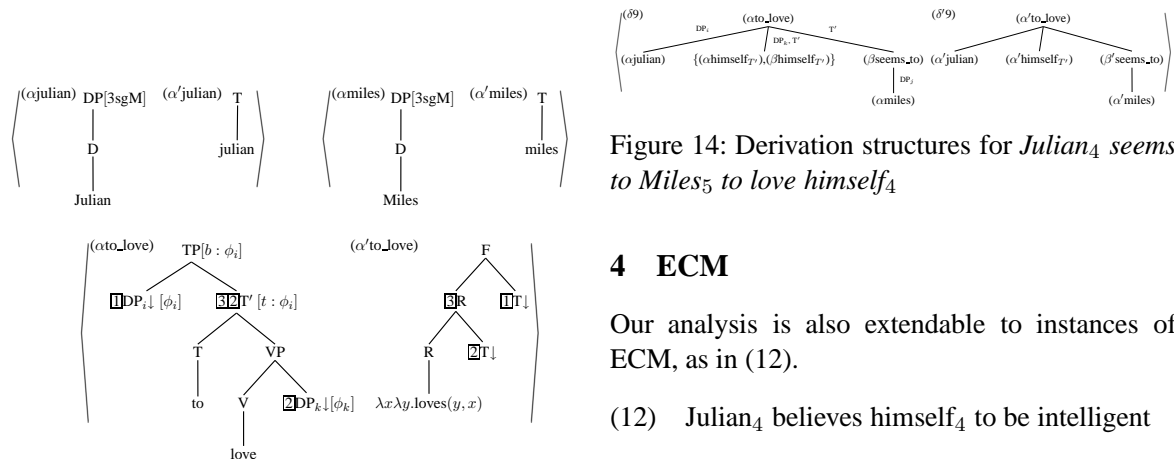


Figure 13: New elementary trees for  $Julian_4$  seems to  $Miles_5$  to love  $himself_4$

#### 4 ECM

Our analysis is also extendable to instances of ECM, as in (12).

(12)  $Julian_4$  believes  $himself_4$  to be intelligent

The elementary trees required for (12) are shown in Figure 16. Here, we propose a third form of the reflexive, the TP-type, specified for subject positions. Because the reflexive is a subject, it is impossible for the antecedent to be found locally,

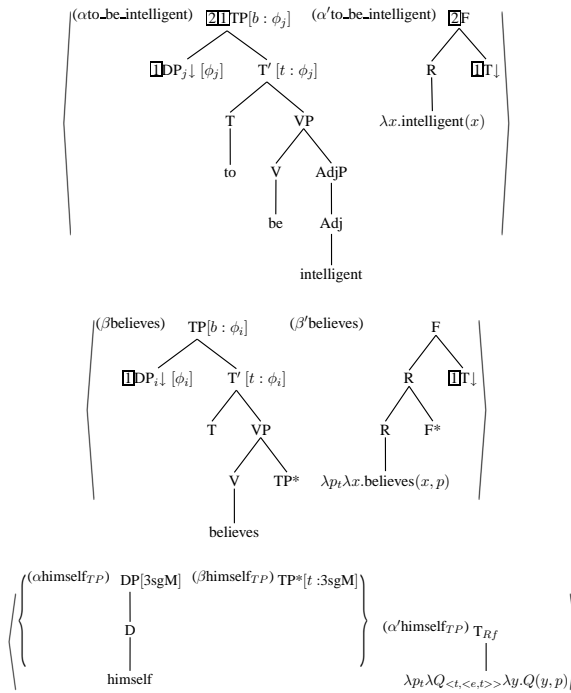


Figure 16: Elementary trees for *Julian<sub>4</sub> believes himself<sub>4</sub> to be intelligent*

motivating a distinct treatment bridging two separate predicates.  $(\alpha\text{himself}_{TP})$  is unchanged from the previous forms, while  $(\beta\text{himself}_{TP})$ , with its Top  $\phi$  feature, is a TP-adjoining auxiliary tree.  $(\alpha'\text{himself}_{TP})$  introduces a function that ensures the identification of the subject argument of the embedded clause and the subject argument of the higher clause. Following the derivation in Figure 17,  $(\beta\text{himself}_{TP})$  and  $(\beta\text{believes})$  multiply adjoin to the TP node of  $(\alpha\text{to\_be\_intelligent})$ . The TP nodes of both  $(\alpha\text{to\_be\_intelligent})$  and  $(\beta\text{believes})$  receive  $\phi$  feature values from DP's substituted at their respective subject positions. Through adjoining  $(\beta\text{himself}_{TP})$  and  $(\beta\text{believes})$  to the TP node of  $(\alpha\text{to\_be\_intelligent})$ , the Top  $\phi$  feature from  $(\beta\text{himself}_{TP})$  and the Bottom  $\phi$  feature from the root TP in  $(\beta\text{believes})$  must unify, as Top features present at an adjoining site must unify with the features of the root of an adjoining tree. This ensures the agreement between the reflexive which is the subject of the embedded clause and the antecedent which is the subject of the higher clause. Note that under Vijay-Shanker and Joshi's definition of feature unification, the Bottom  $\phi$  features of the root TP node of  $(\alpha\text{to\_be\_intelligent})$  would not have to unify with the  $\phi$  features of the root node of

$(\beta\text{believes})$ ; the reflexive's Top feature is responsible for carrying the agreement across clauses. The syntactic and semantic derived trees are in Figure 18. The final formula reduced from  $(\gamma'12)$  is (13).<sup>2</sup>

$$(13) \text{ believes}(\text{julian}, \text{to\_be\_intelligent}(\text{julian}))$$

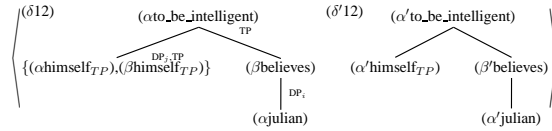


Figure 17: Derivation structures for *Julian<sub>4</sub> believes himself<sub>4</sub> to be intelligent*

In our analysis of ECM, we have required no ECM-specific featural specifications on the predicates, contrary to the ECM derivations in Kallmeyer and Romero (2007). There, the ECM predicate was endowed with special features to permit a variable representing the subject to be passed downward into the embedded clause; our approach limits the differences to the form of the reflexive itself.

### 5 Conclusion

Using STAG mechanisms including links and isomorphic syntactic and semantic derivations, we have shown that different binding possibilities for verbal argument reflexives are captured within the definition of the reflexive itself. Furthermore, we have shown that Condition A can be derived from constraints upon STAG derivation. We have not provided a treatment of 'picture' noun phrase cases here, preferring to see these as logophors (Pollard and Sag, 1992; Reinhart and Reuland, 1993), and we defer cases of non-argument reflexives, such as *Jim did it himself*, to future work.

### Acknowledgment

We are extremely indebted to the three anonymous reviewers of TAG+9. Their insightful comments and criticisms were crucial in reshaping our paper. All remaining errors are ours. This work was supported by NSERC RGPIN/341442 to Han.

<sup>2</sup>Nothing in our analysis so far rules out (i).

- (i) \* John believes that himself is intelligent.

An independent fact of the grammar that *himself* cannot receive accusative case from the subject position of a finite clause accounts for the ill-formedness of (i).

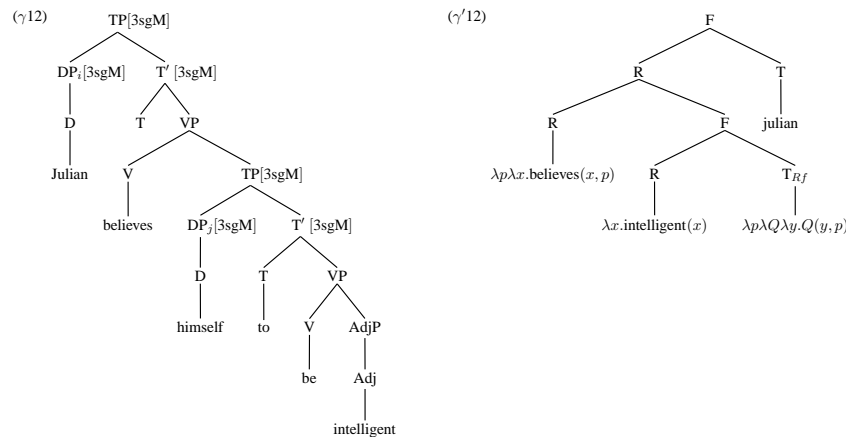


Figure 18: Derived trees for  $Julian_4$  believes  $himself_4$  to be intelligent

## References

- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- 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)*, pages 33–40, COLING-ACL workshop, Sydney, Australia.
- Han, Chung-hye. 2007. Pied-piping in relative clauses: Syntax and compositional semantics using Synchronous Tree Adjoining Grammar. *Research on Language and Computation*, 5(4):457–479.
- Kallmeyer, Laura and Maribel Romero. 2007. Reflexive and reciprocals in LTAG. In Geertzen, Jeroen, Elias Thijssse, Harry Bunt, and Amanda Schiffrin, editors, *Proceedings of the Seventh International Workshop on Computational Semantics IWCS-7*, pages 271–282, Tilburg.
- Kallmeyer, Laura and Maribel Romero. 2008. Scope and situation binding in LTAG using semantic unification. *Research on Language and Computation*, 6(1):3–52.
- Nesson, Rebecca and Stuart M. Shieber. 2006. Simpler TAG Semantics through Synchronization. In *Proceedings of the 11th Conference on Formal Grammar*, Malaga, Spain. CSLI.
- Nesson, Rebecca and Stuart Shieber. 2007. Extraction phenomena in Synchronous TAG syntax and semantics. In Wu, Dekai and David Chiang, editors, *Proceedings of the Workshop on Syntax and Structure in Statistical Translation*, Rochester, New York.
- Pollard, Carl and Ivan Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry*, 23(2):261–303.
- Reinhart, Tanya and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry*, 24(4):657–720.
- Ryant, Neville and Tatjana Scheffler. 2006. Binding of anaphors in LTAG. In *Proceedings of the 8th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+ 8)*, pages 65–72, COLING-ACL workshop, Sydney, Australia.
- Schabes, Yves and Stuart M. Shieber. 1994. An Alternative Conception of Tree-Adjoining Derivation. *Computational Linguistics*, pages 167–176.
- Shieber, Stuart. 1994. Restricting the weak-generative capacity of Synchronous Tree-Adjoining Grammars. *Computational Intelligence*, 10(4).
- Storoshenko, Dennis Ryan. 2006. Reconsidering raising and experiencers in English. In *Proceedings of the 8th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+ 8)*, pages 159–166, COLING-ACL workshop, Sydney, Australia.
- Vijay-Shanker, K. and Aravind Joshi. 1988. Feature structure based Tree Adjoining Grammars. In *Proceedings of COLING'88*, pages 714–719, Budapest.