Reflexives and Reciprocals in Synchronous Tree Adjoining Grammar

Cristina Aggazzotti
Department of Linguistics
Harvard University
Cambridge, MA 02138
caggazzotti@g.harvard.edu

Stuart M. Shieber
Paulson School of Eng'g. and Applied Sciences
Harvard University
Cambridge, MA 02138
shieber@seas.harvard.edu

Abstract

An attractive feature of the formalism of synchronous tree adjoining grammar (STAG) is its potential to handle linguistic phenomena whose syntactic and semantic derivations seem to diverge. Recent work has aimed at adapting STAG to capture such cases. Anaphors, including both reflexives and reciprocals, have presented a particular challenge due to the locality constraints imposed by the STAG formalism. Previous attempts to model anaphors in STAG have focused specifically on reflexives and have not expanded to incorporate reciprocals. We show how STAG can not only capture the syntactic distribution and semantic representation of both reflexives and reciprocals, but also do so in a unified way.

1 Introduction

In this paper, we present a novel unified analysis of both reflexives and reciprocals in synchronous tree adjoining grammar (STAG). STAG utilizes syntactic-semantic tree pairs that undergo synchronized operations to produce a unified syntactic and semantic analysis of linguistic phenomena. Since anaphors, specifically reflexives (*himself*, *themselves*) and reciprocals (*each other*), require a referential lexical item, or antecedent, to supply their semantic value, they depend on both syntax, in the form of distributional constraints, and semantics, in the form of specific relations with the antecedent. Thus, STAG has the potential to be an effective way of modeling both reflexives and reciprocals. Yet STAG's tight integration of syntax and semantics places strong constraints on the syntax-semantics interface, making anaphors a

challenging and illuminating test case for the formalism.

By way of example, consider the sentences in (1), identical except for the alternation between reflexive and reciprocal.

(1) a. Noah and Emma saw themselves.

  b. Noah and Emma saw each other.

As reciprocals require a plural antecedent, for consistency throughout this paper we use examples with plural antecedents for both reflexives and reciprocals. In the *distributive* reading of the reflexive, the relation (here, *saw*) holds between each atom in the plural antecedent and itself. Similarly, for reciprocals, the core reading (so-called *strong reciprocity*) is one in which the relation holds between each atom in the antecedent and each other distinct atom.[1]

The variant of STAG that we assume, following other recent work on STAG for natural-language semantics, is based on set-local multicomponent TAG (MCTAG). In synchronous set-local MCTAG (henceforth, simply "STAG"), a lexical item is represented by a set of syntactic and semantic elementary trees, all of which substitute or adjoin at the same time into another tree set (thus "set-locally") (Weir, 1988). This formalism has been shown to handle a range of phenomena at the syntax-semantics interface, including nested quantifiers (Nesson and Shieber, 2006), extraction

---

[1]There are additional readings for reflexives and reciprocals as well. For instance, in a *cumulative* reading of the reflexive, Noah and Emma both see the pair containing both of them (as, perhaps, in a mirror). The relation holds of the entire plurality and itself. Similarly, reciprocals can display weaker readings than strong reciprocity (Langendoen, 1978; Dalrymple et al., 1998). Incorporation of these readings into the present framework is enabled by our abstracting out the meanings of the reflexive and reciprocal into separable relations REFL and RECP below, but is well beyond the scope of this paper.

phenomena (Nesson and Shieber, 2007), prepositions (Nesson, 2009), it-clefts (Han and Hedberg, 2006), pied-piping in relative clauses (Han, 2006), and clitic climbing (Bleam, 2000).

Previous applications of TAG to anaphors have either appealed to extra facilities, such as recursive semantic features (Kallmeyer and Romero, 2007; Ryant and Scheffler, 2006; Champollion, 2008), or used the more constrained STAG plus adjustments, such as using de Bruijn indices in the semantics (Nesson, 2009), creating multiple reflexive trees (Storoshenko et al., 2008), or operating at multiple links in the derivation (Frank, 2008). No STAG approach to our knowledge has captured both reflexives and reciprocals. Our analysis seeks to fill this void by showing that both kinds of anaphors can be captured uniformly in STAG.

To achieve this, we simplify and generalize one previous analysis of reflexives in STAG, namely that of Frank (2008), so it can apply to reciprocals and a variety of reflexive cases. We simplify Frank's analysis, eliminating the c-command and dominance relations used for proper variable binding by appealing to fundamental syntactic and semantic constraints. We also generalize his analysis to apply to both reflexives and reciprocals. Our full analysis is described in Section 2. We demonstrate the power of this approach in Section 3 using the examples of cataphoric constructions and anaphors as arguments of object control verbs. Analogous derivations capture ditransitive verbs in which the reflexive can be coindexed with the subject or object, though space limitations preclude their inclusion. To handle non-local cases, we avail ourselves of a version of delayed locality, originally proposed by Chiang and Scheffler (2008), and in Section 4 we show how delayed locality accounts for syntactic constructions such as anaphors as arguments of raising verbs and Exceptional Case Marking (ECM) verbs. The analysis also accounts for anaphors in *picture*-DPs and quantificational *picture*-DPs, anaphors in adjuncts, and sentences with multiple anaphors.

1.1 Synchronous tree adjoining grammar

We use set-local feature-structure-based synchronous MCTAG, supplemented with a version of delayed locality for non-local anaphoric cases. Nodes in syntactic trees are notated with syntactic categories and in semantic trees with semantic types, using the notation $\langle \sigma, \tau \rangle$ to express func-

tion types from σ to τ, or the abbreviated version $\sigma\tau$ where no ambiguity results. The feature-structure-based synchronous MCTAG framework we use is exemplified in Figure 1. Elementary tree sets for DPs, as in (a) and (b), contain multiple syntactic trees and semantic trees, two of each, independently motivated for handling quantification and topicalization. The syntax has a TP auxiliary tree (allowing for frontings such as topicalization, following Nesson and Shieber (2007)) in addition to the "in situ" DP tree; the synchronous semantics has a t auxiliary tree (used for quantifier scope, following Shieber and Schabes (1990) and Williford (1993)) and an e-rooted reference tree. Non-quantificational DPs like *Noah* have a degenerate scope tree t_* that does not modify the derived tree, so merely serves as a placeholder to maintain structural consistency.

Syntactic nodes have an associated feature structure containing finite feature values. The feature structure must unify with the feature structure of any substituting or adjoining node in order for the operation to take place; if any features conflict, the unification fails (Vijay-Shanker and Joshi, 1988). In particular, we can mark DP substitution nodes in their feature structure with their case requirements (which can be thought of as a manifestation of their being assigned abstract Case (Polinsky and Preminger, 2014)), while lexical item trees rooted at DP that exhibit morphological case will have that case depicted in the root feature structure as well. This built-in feature checking system will play a role in several aspects of our STAG framework, including matching phi-features (number, person, gender) of anaphors and antecedents, making c-command and dominance constraints unnecessary, and accounting for cataphoric constructions. For reasons of readability and succinctness, we do not show feature structures explicitly in subsequent examples.

Reciprocals reciprocate over plural entities. For our purposes, we do not require a sophisticated semantic notion of plurality (such as Scha (1981), van Bentham (1989), or Westerståhl (1989)). We notate the type of sets of entities of type σ as $\overline{\sigma}$ and the plural entity combining a and b as $a + b$; we further optionally identify singular entities a with singleton plural entities $\{a\}$. Plural DPs will sometimes denote plural entities (like *Noah and Emma* of type $\overline{e}$), and certain verbs (of type $\overline{e}t$, like *met*) will require plural entities. Certain quantified

(a) TP$_*$ DP $\begin{bmatrix} cat: & DP \\ agr: & \begin{bmatrix} num: & sg \\ pers: & 3rd \\ gend: & masc \end{bmatrix} \end{bmatrix}$ t_* e

Noah $noah$

(b) TP$_*$ DP $\begin{bmatrix} cat: & DP \\ case: & acc \\ agr: & \begin{bmatrix} num: & sg \\ pers: & 3rd \\ gend: & fem \end{bmatrix} \end{bmatrix}$ t_* e

her her

(c) TP$_{\boxed{1}\boxed{2}}$

DP$_\downarrow$$_{\boxed{1}}$[case : nom] T'

T VP

V DP$_\downarrow$$_{\boxed{2}}$[case : acc]

saw

$t_{\boxed{1}\boxed{2}}$

et $e_\downarrow$$_{\boxed{1}}$

eet $e_\downarrow$$_{\boxed{2}}$

saw

(d) saw (e) T

$\boxed{1}$ $\boxed{2}$ DP T'

noah her Noah T VP

V DP

saw her

t
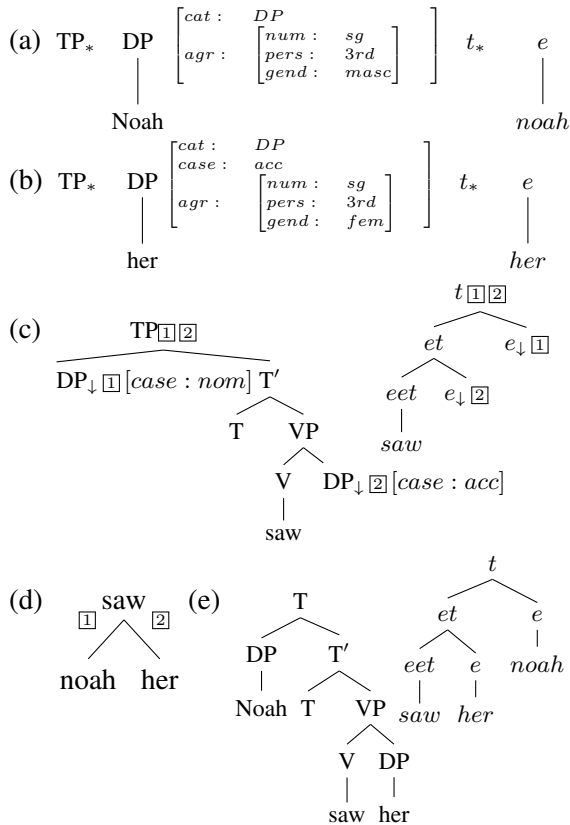
et e

eet e $noah$

saw her

Figure 1: Elementary STAG trees and feature structures for (a) DPs, (b) pronouns, and (c) transitive verbs; (d) is the derivation tree and (e) the derived trees for the simple transitive sentence *Noah saw her.*

DPs (like *everyone*) may involve both interpretations – the quantifier meaning (as in *everyone left*) of type $\langle et, t \rangle$ and the plural interpretation (as in *everyone met*) of type $\overline{e}$.

Nodes that can undergo operations are designated by links, shown in the trees with numbered boxes ($\boxed{}$), that ensure the syntactic and semantic trees accept synchronous operations at corresponding nodes, as in (c). Multiple boxes marked with the same numbered link specify the sites of operation of a set of trees.

1.2 Previous work on reflexives in TAG

In the literature, there are six main applications of TAG to capture anaphors. Three are non-synchronous, building a semantic representation using recursive feature structures (Ryant and Scheffler, 2006; Kallmeyer and Romero, 2007; Champollion, 2008),[2] which can be more pow-

---

[2]Steedman (2000) also provides an account for binding but uses a combination of LTAG and combinatory categorical

erful than TAG (indeed Turing-equivalent under some usages), and look at both reflexives and reciprocals. The other three use the more restrictive STAG framework for both the syntax and semantics but only look at reflexives (Nesson, 2009; Storoshenko et al., 2008; Frank, 2008).

For the non-synchronous TAG approaches, Ryant and Scheffler (2006) employ tree-local multicomponent lexicalized TAG (LTAG) with semantic feature structures and a flat compositional semantics for each elementary tree. The multicomponent tree set for reflexives and reciprocals contains two trees: an NP tree with the lexical anaphor that is c-commanded by a degenerate NP tree that composes with its antecedent through flexible composition (FC) (Joshi et al., 2003), an extension of LTAG. This approach captures reflexives and reciprocals but requires extra subject intervention and c-command constraints to prevent overgeneration.

Kallmeyer and Romero (2007) use a similar approach, but replace Ryant and Scheffler's degenerate anaphor NP tree with a degenerate VP tree. This change does not require the FC extension (except for adjuncts) or stronger c-command constraints, but does require a dominance relation between the degenerate VP and the lexical anaphor as well as a procedure for passing antecedent features. Only with both of these additions do the locality and c-command restrictions of classic binding theory (Chomsky, 1981) then follow.

Instead of compositional semantics, Champollion (2008) uses the feature-based LTAG formalism of Vijay-Shanker (1987) extended by the use of lists as values of features, as in HPSG (Pollard and Sag, 1992), and list operations, such as appending lists together. Champollion (2008) improves upon the previous non-synchronous approaches in several ways, such as by capturing ECM verbs, adjuncts, and all conditions in binding theory with no further additions to the framework; however, the analysis does not include reciprocals and requires recursive features.

For the STAG approaches, Nesson (2009) uses MCTAG but extends the lambda calculus notation for semantic representation with de Bruijn indices. The de Bruijn notation uses integer indices – instead of explicitly-named free variables – to indicate how many enclosing λ terms away the variable's binding λ is. Although this approach pro-

---

grammar (CCG).

vides more flexibility for locality constraints and can successfully account for a variety of reflexive sentences, it does not allow the differentiation needed for reciprocals because the indices allow specification only of coindexation.

Storoshenko et al. (2008) take a different MC-TAG approach by positing three separate reflexive syntactic trees, whose use depends on the reflexive's binding option as a verbal argument. The semantics relies on dynamically varying what is a function and what is an argument. For a sentence containing a reflexive, the reflexive plays the function role, taking its sister node as its argument; however, if an entity fills that position in the sentence instead of a reflexive, the entity would be an argument and its sister node the function. Although the analysis captures ditransitives, raising verbs, and ECM verbs, why these three reflexive tree sets are the (only) possible options and why each reflexive has its specific semantic type is not well motivated.

It may be possible to extend this analysis to reciprocals: following reflexives, the semantics could be separately defined for each case as needed and agreement could be handled in the syntax through a clever use of features. However, this approach seems to lack a unifying story behind the choice of tree set configuration. We thus turn to the final application of STAG to reflexives, which we show can be extended to both reflexives and reciprocals in a more straightforward way.

1.3 The analysis of Frank (2008)

Frank (2008) uses tree-local MCTAG to capture simple reflexive cases but does not attempt to capture reciprocals and does not definitively extend the analysis to more complicated cases, such as raising and ECM verbs. The analysis is illustrated in Figure 2 for the sentence in (2).[3]

---

[3]We have diverged from Frank's elementary trees (Frank, 2008, Figure 1) slightly, modifying them by clarifying and making explicit two notational issues, according to our understanding. First, we include an explicit t_* tree in the tree set for *John* based on the ① link at the root of the reflexive tree, which we assume indicates the adjunction of a *John* tree at the same time its e-rooted tree substitutes. Second, we removed the ② links on the NP *himself* and on the second semantic x variable tree (see Frank's Figure 1(a)) because they seem to serve as labels, rather than as operable sites like the ① links. We leave the ② link on the root of the semantic verb tree (even though it does not correspond to an explicit adjoining tree) because it may be intended for an extra scope tree in the reflexive tree set; however, it is unclear how that extra scope tree would fit into the included dominance relations so we do not explicitly add it to the reflexive tree set here.
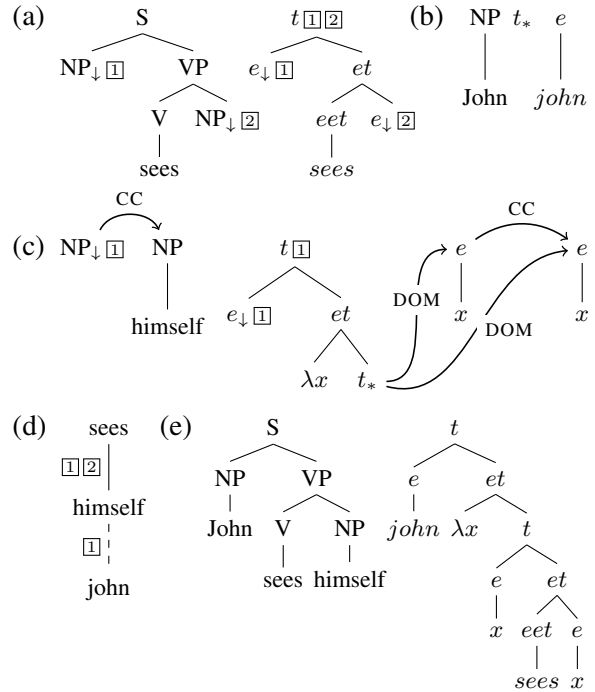
Figure 2: Frank's elementary trees for (a) transitive verbs, (b) type e NPs, and (c) reflexives; (d) is the derivation tree and (e) the derived trees for sentence (2). Extra constraints are indicated with labeled arrows: CC for c-command, DOM for dominance.

(2) John sees himself.

Frank's analysis is novel in two ways: the structure of the derivation and the use of multiple links. First, derivations of reflexive sentences (Figure 2(d)) diverge from derivations of non-reflexive sentences (Figure 1(d)). The derivation tree in Figure 2(d) has the antecedent (subject) first substitute into the reflexive (object), which then as a whole composes into the verb tree at the respective links.[4]

The structure of this derivation is unusual in the TAG literature in not paralleling the non-reflexive derivation tree, in which the subject and object separately substitute directly into the verb tree. However, there may be cross-linguistic evidence for this type of derivation. In languages such as Finnish, which represent reflexivization with a verbal affix that detransitivizes the verb into an intransitive verb, Büring (2005) explains that this verbal reflexive marker is not a syntactic argument

---

[4]Ryant and Scheffler (2006) use a similar derivation tree in that the antecedent composes with the anaphor before both compose into the verb tree, but their use of flexible composition allows composition of trees in either direction, so the derivation tree is not actually equivalent.

or clitic, providing support for a derivation tree in which the verb accepts just one argument, the subject. An analysis of clitics along these lines may be apposite as well.

Second, Frank's analysis crucially relies on allowing a tree set to operate at multiple links. The reflex of this innovation is the multiple links decorating edges in the derivation tree, where in standard STAG, only a single link would appear. In particular, Frank's derivation tree in (d) portrays the reflexive *himself* going into the verb tree at both links ☐ and ☐. Implicitly, Frank is appealing to a novel generalization of MCTAG, in which multiple components of a tree set can apply at multiple links.

Frank's approach accounts for simple reflexive antecedents, quantifier-bound reflexives, reflexives embedded in a *picture*-DP, and reflexives occurring as the argument of a ditransitive predicate. However, the approach does not directly extend to reciprocals.

Unlike reflexives, reciprocals are not simply inherently coindexed with their antecedent since the antecedent must be distributed into its atomic parts. Frank's approach as it stands cannot account for this. The semantic trees contain only one binder of two instantiations of the same variable and are thus inherently detransitivizing. By maintaining separate binders of the two argument positions, our modifications below not only account for both reflexives and reciprocals, but also do so in a unified and simplified way.

2 Our analysis

In this section, we explain how our analysis builds directly on Frank's. We adjust the analysis to be in line with the framework outlined in Section 1.1.

2.1 Frank's analysis revised

As in Frank's analysis, the reflexives will use *both* subject and object links, and thus will be composed of four syntactic and four semantic trees. The tree set follows Frank's approach with only minor changes, as shown in Figure 3(a). In the syntactic tree set are two TP$_*$ placeholder trees, one for each of the DP trees. The first DP tree is degenerate, accepting the antecedent by substitution, and the second contains the reflexive. The semantic tree set contains a t auxiliary scope tree for each of the e-rooted variables. In the first scope tree, a reflexive operator REFL (described shortly)

has been added as another binary branch in the elementary reflexive tree, along with two binding λ terms (instead of just one). The e-rooted variable trees correspondingly contain two distinct variables. As shown in (b), we use the same derivation tree as Frank, also taking advantage of the multi-link extension of MCTAG.

As described in Section 1.1, our framework makes use of case and feature unification for pronouns, which can additionally ensure the correct configuration of lexical substitution of the antecedent and reflexive, thus making the c-command (CC) and dominance (DOM) constraints redundant. Eliminating these extra constraints greatly simplifies the analysis by relying on the inherent features of the formalism instead of on externally-added restrictions.

On the semantics side, the reflexive operator REFL serves as a formalization of the reflexive relation. For the purposes of this paper, in which we focus on the distributive reflexive reading of plural reflexives, the REFL operator is given as in (3). Abstracting out the reflexive operator allows flexibility in its semantic definition and comparison to alternatives (such as the RECP operator we introduce shortly).

(3) $\text{REFL} \equiv \lambda R \,.\, \lambda Z \,.\, \forall x : x \in Z \,.$
$$\forall y : y \in Z \land y = x \,.\, R\,y\,x$$

Informally speaking, the operator holds of a binary relation R and an antecedent set Z just in case every pair x and y in the set Z, where x and y are not distinct, are in the relation R.[5] (The benefit of the apparent redundancy of the two universal quantifiers will become evident shortly.)

The STAG derivation corresponding to sentence (1a) proceeds as in Figure 3. The resulting logical form can be simplified as shown in (4), demonstrating that the distributive reading is appropriately captured.

(4) $\text{REFL} \,(\lambda a \,.\, \lambda b \,.\, saw\,a\,b)\,(n + e)$
$\quad = (\lambda R \,.\, \lambda Z \,.\, \forall x : x \in Z.$
$\qquad\quad \forall y : y \in Z \land y = x \,.\, R\,y\,x)$
$\qquad\quad (\lambda a \,.\, \lambda b \,.\, saw\,a\,b)\,(n + e)$
$\quad = \forall x : x \in (n + e).$
$\qquad\quad \forall y : y \in (n + e) \land y = x \,.\, saw\,y\,x$
$\quad = saw\,n\,n \land saw\,e\,e$

---

[5]This definition of REFL can also account for singular antecedents by interpreting them as singleton plural entities.

(a)

$\text{TP}_* \boxed{1}$ $\text{DP}_\downarrow \boxed{1}$ TP_* $\quad$ DP
|
themselves

$t \boxed{1}$ $\quad\quad$ e $\quad t_*$ $\quad e$
$\overline{et}$ $\quad\quad \overline{e}_\downarrow \boxed{1}$ $\quad b$ $\quad\quad a$

$\langle eet, \overline{et}\rangle$ $\quad eet$
|
REFL $\quad \lambda a$ $\quad et$
λb $\quad t_*$

(b)
saw
$\boxed{1}\boxed{2}$ |
themselves
$\boxed{1}$ |
noah+emma

$\quad\quad\quad\quad\quad\quad t$
$\quad\quad\quad \overline{et}$ $\quad\quad \overline{e}$
$\quad \langle eet, \overline{et}\rangle$ $\quad eet$ $\quad n+e$
$\quad\quad$ REFL $\quad \lambda a$ $\quad et$
$\quad\quad\quad\quad\quad \lambda b$ $\quad t$

(c)
$\quad\quad$ TP
$\quad$ DP $\quad\quad$ T′
Noah&Emma $\quad$ T $\quad$ VP
$\quad\quad\quad\quad$ V $\quad$ DP
$\quad\quad\quad$ saw themselves

$\quad\quad et \quad e$
$\quad eet \quad e\,b$
$\quad | \quad\quad |$
$\quad saw \quad a$

Figure 3: (a) Elementary tree set for a reflexive; (b) is the derivation tree and (c) the derived trees for sentence (1a).

2.2 Comparison of analyses

There are four differences between our reflexive analysis and Frank's.

1. We use extra placeholder trees to maintain a parallel structure among all DPs. The extra trees are necessitated on the syntax side by the DP tree used in the Nesson and Shieber (2007) fronting analysis and on the semantics side by the quantifier scope tree. This modification is not essential to our reflexive analysis as it arises solely from our incorporation of the independent fronting analysis (as described in Section 3.1).

2. We eliminate binding constraints like c-command and dominance, which permits the flexibility needed for cataphora, since these relations are already captured through case checking.

3. We employ two bindings of distinct variables instead of one binding of a single variable twice, as this allows the appropriate grain needed for reciprocals.

4. We abstract away the reflexivity notion from Frank's trees with an operator REFL, which

generalizes to also be compatible with reciprocals using a parallel operator RECP, as described in the next section.

2.3 Adding reciprocals

Using an operator for both reflexives and reciprocals captures their underlying similarities, creating a unified account of both. It seems logical to group reflexives and reciprocals together syntactically, as structurally interchangeable constructions, and distinguish between them semantically, as differing with respect solely to distribution over the antecedent. This is the motivation behind our proposed approach.

In order to incorporate reciprocals into the STAG framework, we simply add the reciprocal counterparts in the same place as reflexives in the multicomponent tree set for reflexives, as in Figure 4. On the syntax side, we replace the lexical item *themselves* with *each other* and on the semantics side, we replace the reflexive operator (REFL) with a reciprocal operator (RECP). We indicate the shared structure by placing corresponding components of reflexives and reciprocals in the same node as interchangeable options.

An attractive property of this analysis is that simply by replacing the $=$ in the semantic representation of REFL with $\neq$, we get the formalization of the reciprocal relation RECP:

(5) $\text{RECP} \equiv \lambda R \,.\, \lambda Z \,.\, \forall x : x \in Z \,.$
$$\forall y : y \in Z \wedge y \neq x \,.\, R\, y\, x$$

Similarly to REFL, the RECP operator holds of a binary relation R and an antecedent set Z just in case every pair x and y in the set Z, where x and y *are* distinct, are in the relation R. For the reciprocal version of sentence (1a), in (1b), RECP provides the correct (and only) reading – the strong reciprocity reading – that Noah saw Emma and Emma saw Noah. The reduction proceeds in parallel fashion to that of reflexives. Comparing these trees to Frank's trees in Figure 2, the reader can confirm that the derivation tree is identical and both methods produce the same result (up to the modification in the logical form). With this example as a foundation, we now show the utility of this representation for a range of increasingly complex reflexive and reciprocal phenomena.

$$\text{TP}_{*}\boxed{1} \quad \text{DP}_{\downarrow}\boxed{1} \quad \text{TP}_{*} \quad \text{DP}$$

themselves/
each other

$t\boxed{1}$ ⏐ e t_* e

$\overline{et}$ $\overline{e}_{\downarrow}\boxed{1}$ b a

$\langle eet, \overline{et}\rangle$ eet

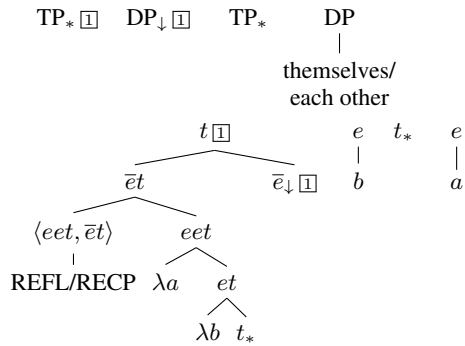REFL/RECP λa et

λb t_*

Figure 4: Anaphor elementary tree set for both a reflexive and a reciprocal

3 Applications

The analysis essentially unchanged accounts for various reflexive and reciprocal phenomena, including cataphora, anaphors with object control verbs, and anaphors as arguments of ditransitive verbs. The analysis also has the potential to apply to reflexives and reciprocals in other languages, but we leave this extension for future work. We show here only the analysis for cataphora and anaphors with object control verbs due to space constraints, but the other applications follow similarly.

3.1 Cataphora

Cataphora, such as in (6a), would appear to present a problem for analyses requiring c-command constraints, as the required c-command relation does not appear to hold overtly in the derived tree. Our approach however is completely consistent with the account of topicalization of Nesson and Shieber (2007), by treating the anaphor as a topicalized item.[6] We illustrate this derivation in Figure 5 for the simplified cataphoric reciprocal sentence in (6b).[7]

(6) a. (Noah and Emma like many people, but) each other, they can't stand.

b. Each other, Noah and Emma saw.

The syntactic tree set for the reflexive, shown in Figure 5(a), simply reflects topicalization of the

---

[6]In this paper, we focus on the core anaphoric cases, excluding logophoric (point of view or emphatic/focus) anaphors, as discussed in, for instance, Reinhart and Reuland (1993). For cataphora in particular, we only consider examples in which the anaphor stands alone as a topicalized item, and do not address anaphors embedded in topicalized adjuncts or adverbial phrases.

[7]Cataphoric reflexives follow similarly so are not shown.

(a) $\text{TP}_{*}\boxed{1}$ $\text{DP}_{\downarrow}\boxed{1}$ TP DP

DP T′ ϵ

each other T TP_{*}

(b) saw (c) TP

$\boxed{1}\boxed{2}$⏐

each other DP T′

$\boxed{1}$⏐ Each other T TP

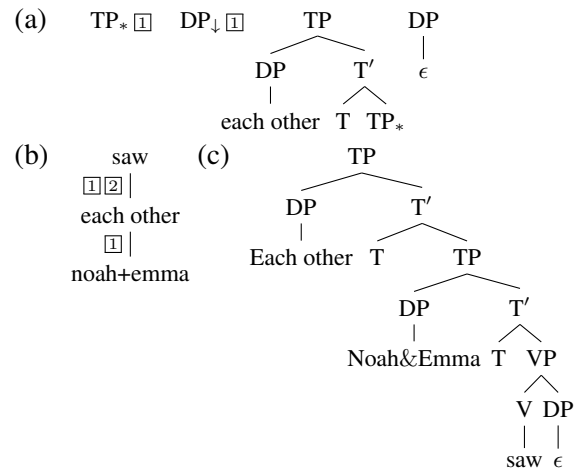noah+emma

DP T′

Noah&Emma T VP

V DP

saw ϵ

Figure 5: (a) Syntactic elementary trees for a cataphoric (topicalized) reflexive; (b) is the derivation tree and (c) the syntactic derived tree for sentence (6b)

reflexive following directly the topicalization analysis of Nesson and Shieber (2007): the TP auxiliary tree now contains the lexical reflexive and the corresponding DP tree contains the empty string; the semantics side remains unchanged so is not shown. The derivation proceeds as usual.

Using a feature-checking system instead of binding principles provides the flexibility needed for capturing cataphora without additional machinery because the topicalized anaphor, instead of the empty DP, receives accusative case and thus no feature conflicts arise.

3.2 Anaphors with object control verbs

Syntactic constructions with object control verbs, such as *persuade* in (7), follow directly from our analysis as put forth so far.

(7) Noah and Emma persuaded themselves/each other to be happy.

Object control verbs have three arguments: an agent (*Noah and Emma*), a theme (*themselves/each other*), and an open proposition (*to be happy*). This configuration is represented in the elementary object control verb tree set in Figure 6(a). The lower verb cannot have its own subject, so the *persuaded* tree set contains a DP tree in the syntax and a corresponding variable tree in the semantics that substitute into the subject position of the lower verb. The derivation proceeds according to the derivation tree in (c), in which the antecedent composes into the anaphor tree set,
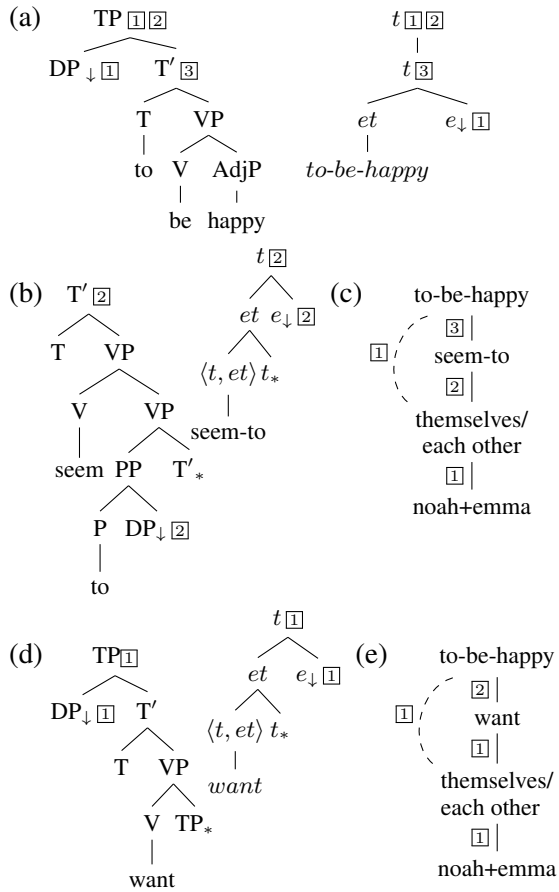
Figure 6: Elementary trees for (a) object control verbs and (b) non-finite predicates with appropriate links; (c) is the derivation tree. The derived tree pair is provided in Figure 8 in the appendix.

which then as a whole composes into the object control tree set in a tree-local fashion. This tree set then composes into the non-finite verb tree.

4 Extensions with delayed locality

Although a wide variety of interactions between anaphors and other constructions are captured by this analysis, there is an entire class of cases that are not expressible under the set-local view of STAG derivation we have been presupposing. In this section, we extend the derivation notion to allow for *delayed locality*, first proposed by Chiang and Scheffler (2008). Delayed locality relaxes the set-locality constraint to allow a delay in composition. Two trees in a multicomponent tree set may compose into (any number of) other trees before eventually composing into the same elementary tree.[8] This differs from the more expressive non-local MCTAG in requiring that the members eventually compose into the same elementary tree (Chiang and Scheffler, 2008). Delayed locality has permitted analyses of non-local right-node raising (Han et al., 2010), bound vari-

---

[8]Storoshenko and Han (2013) propose a slightly different definition of a delay than Chiang and Scheffler (2008); we postpone committing to a particular definition to future work, but recognize that overgeneration is a concern, since without further constraint our analysis could allow, for instance,

(8) * Noah$_i$ thinks that Emma likes himself$_i$.

able pronouns (Storoshenko and Han, 2010), and clitic climbing (Chen-Main et al., 2012).

With this extension, our analysis allows for anaphors in a variety of syntactic constructions, including *picture*-DPs, quantificational *picture*-DPs, adjuncts, raising verbs, ECM verbs, and multiple anaphors in the same sentence, but due to space limitations we again demonstrate only for raising and ECM verbs.

4.1 Anaphors with raising verbs

In contrast to object control verbs, raising verbs, such as *seem* in (9), do not have an inherent subject argument; therefore, the usual representation of *seem* in the TAG literature (with minor variations) does not contain a DP subject node, as shown in Figure 7(b).

(9) Noah and Emma seem to themselves/each other to be happy.

Use of the present anaphor analysis with this configuration violates set-locality because the anaphor would compose into the raising verb tree, but there would not be a position for the antecedent to also compose.[9] However, the relaxation provided by delayed locality allows the lexical anaphor part of the tree set to compose into the raising verb through delay, which then composes into the lower clause verb trees at link 3, while the antecedent part is not delayed and composes directly into the lower clause verb trees, as depicted in Figure 7(c). In order to ensure that all variables are properly bound, the semantic predicate *to-be-happy* tree in Figure 7(a) has the root node split into an upper $t_{1,2}$ node and a lower t_3 node to ensure that the REFL/RECP tree binds the a variable in the raising verb tree.

4.2 Anaphors with ECM verbs

ECM (or "subject-to-object raising") verbs, as in (10), have two arguments: a subject (*Noah and Emma*) and a proposition (*themselves/each other to be happy*). Based on these structural properties, the elementary tree for an ECM verb contains a subject position and adjoins into a predicate to fill its proposition argument, as shown in Figure 7(d).

---

[9]An alternative local derivation would be to simply include a subject position in the elementary raising verb tree. Although this solution solves the locality issue, it has implications for the treatment of raising constructions in general so we do not pursue it here.

(a) TP[1][2]

DP↓[1] T'[3]

T VP

to V AdjP

be happy

t[1][2]

t[3]

et e↓[1]

to-be-happy

(b) T'[2]

T VP

V VP

seem PP T'*

P DP↓[2]

to

t[2]

et e↓[2]

$\langle t, et \rangle$ t_*

seem-to

(c) to-be-happy
[3] |
[1] ⁄   seem-to
[2] |
themselves/
each other
[1] |
noah+emma

(d) TP[1]

DP↓[1] T'

T VP

V TP*

want

t[1]

et e↓[1]

$\langle t, et \rangle$ t_*

want

(e) to-be-happy
[2] |
[1] ⁄   want
[1] |
themselves/
each other
[1] |
noah+emma

Figure 7: Elementary trees for (a) non-finite predicates with appropriate links and configuration for variable binding and (b) raising verbs with an anaphor object; (c) is the derivation tree with delayed locality for sentence (9); (d) elementary trees for ECM verbs and (e) is the derivation tree with delayed locality for sentence (10).

(10) Noah and Emma want themselves/each other to be happy.

In contrast to the previous example, for ECM verbs the antecedent part of the anaphor tree set is the delayed part, first composing into the ECM verb trees and then composing into the non-finite verb trees at link [2]. The derivation tree in (e) reflects this difference through the links shown.

For cases in English with multiple (surface accusative) objects, such as in the ECM construction in (11a), appealing to case is not sufficient to account for the ungrammaticality of (11b). A more nuanced case analysis, in which the equational constraint on case (that the antecedent's case is nominative and the anaphor's case is accusative) is replaced by an inequational constraint over a set of cases ordered by obliqueness (that the antecedent's case is less oblique than the anaphor's case) suffices to cover these as well, predicting the grammaticality of (11a) and ungrammaticality of (11b).

(11) a. Emma wants him to love himself.

b. * Emma wants himself to love him.

5 Conclusion

In this paper, we have shown how the formalism of STAG can not only handle both reflexives and reciprocals, but also provide a unified account of both, founded on the idea that these anaphors share a syntactic distribution but differ slightly and uniformly in their semantics. To accomplish this, we provide STAG tree sets for reflexives and reciprocals that differ only in their lexical presentation and their interpretation through operators REFL and RECP that capture the parallel semantic nature of reflexives and reciprocals. It is, to our knowledge, the first STAG analysis to provide for reciprocals as well as reflexives. The analysis is consistent with earlier STAG analyses accounting for such syntactic phenomena as topicalization and semantic phenomena as quantification, while building on the previous STAG account by Frank (2008) of reflexives alone, making anaphoric notions more explicit, eliminating the need for c-command and dominance constraints, and generalizing the analysis to capture reciprocals as well.

Areas for future work include investigating appropriate further limits on delayed locality to prevent overgeneration, expanding our preliminary application of the operators crosslinguistically, and refining the operators' semantic definitions to account for additional anaphoric interpretations.

References

Tonia Bleam. 2000. Clitic climbing and the power of tree adjoining grammar. In Anne Abeillé and Owen Rambow, editors, *Tree Adjoining Grammars: Formalisms, Linguistic Analysis, and Processing*, CSLI Publications, pages 193–219.

Daniel Büring. 2005. *Binding Theory*. Cambridge University Press, UK.

Lucas Champollion. 2008. Binding theory in LTAG. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+9)*. Tübingen, Germany, pages 1–8. http://aclweb.org/anthology/W08-2301.

Joan Chen-Main, Tonia Bleam, and Aravind K. Joshi. 2012. Delayed tree-locality, set-locality, and clitic climbing. In *Proceedings of the 11th International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+11)*. Paris, France, pages 1–9. http://aclweb.org/anthology/W12-4601.

David Chiang and Tatjana Scheffler. 2008. Flexible composition and delayed tree-locality. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+9)*. Association for Computational Linguistics, Tübingen, Germany, pages 17–24. http://aclweb.org/anthology/W08-2303.

Noam Chomsky. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Mary Dalrymple, Makoto Kanazawa, Yookyung Kim, Sam McHombo, and Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21(2):159–210. https://doi.org/10.1023/A:1005330227480.

Robert Frank. 2008. Reflexives and TAG semantics. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+9)*. Association for Computational Linguistics, Tübingen, Germany, pages 97–104. http://aclweb.org/anthology/W08-2313.

Chung-hye Han. 2006. Pied-piping in relative clauses: Syntax and compositional semantics based on synchronous tree adjoining grammar. In *Proceedings of the Eighth International Workshop on tree Adjoining Grammars and Related Formalisms (TAG+8)*. Association for Computational Linguistics, Sydney, Australia, pages 41–48. http://aclweb.org/anthology/W06-1506.

Chung-hye Han and Nancy Hedberg. 2006. A tree adjoining grammar analysis of the syntax and semantics of it-clefts. In *Proceedings of the Eighth International Workshop on tree Adjoining Grammars and Related Formalisms (TAG+ 8)*. Association for Computational Linguistics, Sydney, Australia, pages 33–40. http://aclweb.org/anthology/W06-1505.

Chung-hye Han, David Potter, and Dennis Ryan Storoshenko. 2010. Non-local right-node raising: An analysis using delayed tree-local mc-tag. In *Proceedings of the 10th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+10)*. Linguistics Department, Yale University, New Haven, Connecticut, pages 9–16. http://aclweb.org/anthology/W10-4402.

Aravind K. Joshi, Laura Kallmeyer, and Maribel Romero. 2003. Flexible composition in LTAG: Quantifier scope and inverse linking. In *Proceedings of the International Workshop on Compositional Semantics*. Tilburg, The Netherlands.

Laura Kallmeyer and Maribel Romero. 2007. Reflexives and reciprocals in LTAG. In Harry Bunt, Jeroen Geertzen, Elias Thijsse, and Amanda Schiffrin, editors, *Proceedings of the Seventh International Workshop on Computational Semantics ICWS-7*. Tilburg, pages 271–282.

D. Terence Langendoen. 1978. The logic of reciprocity. *Linguistic Inquiry* 9(2):177–197. http://www.jstor.org/stable/4178051.

Rebecca Nesson. 2009. *Synchronous and Multicomponent Tree-Adjoining Grammars: Complexity, Algorithms and Linguistic Applications*. Ph.D. thesis, Harvard University.

Rebecca Nesson and Stuart M. Shieber. 2006. Simpler TAG semantics through synchronization. In *Proceedings of the 11th Conference on Formal Grammar*. Malaga, Spain. https://perma.cc/JM6Y-6QBY.

Rebecca Nesson and Stuart M. Shieber. 2007. Extraction phenomena in synchronous TAG syntax and semantics. In *Proceedings of SSST, NAACL-HLT 2007 / AMTA Workshop on Syntax and Structure in Statistical Translation*. Association for Computational Linguistics, pages 9–16. http://aclweb.org/anthology/W07-0402.

Maria Polinsky and Omer Preminger. 2014. Case and grammatical relations. In Andrew Carnie, Dan Siddiqi, and Yosuke Sato, editors, *Routledge Handbook of Syntax*, Routledge, chapter 8.

Carl Pollard and Ivan A. Sag. 1992. *Head-driven Phrase Structure Grammar*. University of Chicago Press.

Tanya Reinhart and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4):657–720. http://www.jstor.org/stable/4178836.

Neville Ryant and Tatjana Scheffler. 2006. Binding of anaphors in LTAG. In *Proceedings of the Eighth International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+8)*. Association for Computational Linguistics, Sydney, Australia, pages 65–72. http://aclweb.org/anthology/W06-1509.

Remko Scha. 1981. Distributive, collective and cumulative quantification. In J. A. G. Groenendijk, T. M. V. Janssen, and M. B. J. Stokhof, editors, *Formal Methods in the Study of Language, Part 2*, Mathematisch Centrum, Amsterdam, pages 483–512.

Stuart M. Shieber and Yves Schabes. 1990. Synchronous tree-adjoining grammars. In *Proceedings of the 13th International Conference on Computational Linguistics*. Helsinki, Finland, volume 3, pages 253–258. http://aclweb.org/anthology/C90-3045.

Mark Steedman. 2000. Implications of binding for lexicalized grammars. *Tree Adjoining Grammars: Formalisms, Linguistic Analysis, and Processing* pages 283–301.

Dennis R. Storoshenko, Chung-hye Han, and David Potter. 2008. Reflexivity in English: an STAG analysis. In *Proceedings of the Ninth International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+9)*. Association for Computational Linguistics, Tübingen, Germany, pages 149–156. http://aclweb.org/anthology/W08-2320.

Dennis Ryan Storoshenko and Chung-hye Han. 2010. Binding variables in english: An analysis using delayed tree locality. In *Proceedings of the 10th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+10)*. Linguistics Department, Yale University, New Haven, Connecticut, pages 143–150. http://aclweb.org/anthology/W10-4418.

Dennis Ryan Storoshenko and Chung-hye Han. 2013. Using synchronous tree adjoining grammar to model the typology of bound variable pronouns. *Journal of Logic and Computation* 25(2):371–403. https://doi.org/10.1093/logcom/exs064.

Johan van Bentham. 1989. Polyadic quantifiers. *Linguistics and Philosophy* 12(4):437–464. https://doi.org/10.1007/BF00632472.

K. Vijay-Shanker. 1987. *A study of Tree Adjoining Grammars*. Ph.D. thesis, University of Pennsylvania.

K. Vijay-Shanker and Aravind K. Joshi. 1988. Feature structures based tree adjoining grammars. In *Coling Budapest 1988 Volume 2: International Conference on Computational Linguistics*. http://aclweb.org/anthology/C88-2147.

David Jeremy Weir. 1988. *Characterizing mildly context-sensitive grammar formalisms*. Ph.D. thesis, University of Pennsylvania.

Dag Westerståhl. 1989. Quantifiers in formal and natural languages. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic: Volume IV: Topics in the Philosophy of Language*, Springer Netherlands, Dordrecht, pages 1–131.

Sean Michael Williford. 1993. *Application of Synchronous Tree-Adjoining Grammar to Quantifier Scoping Phenomena in English*. Bachelor's thesis, Harvard College. http://nrs.harvard.edu/urn-3:HUL.InstRepos:10951941.

A Appendix: Derived trees

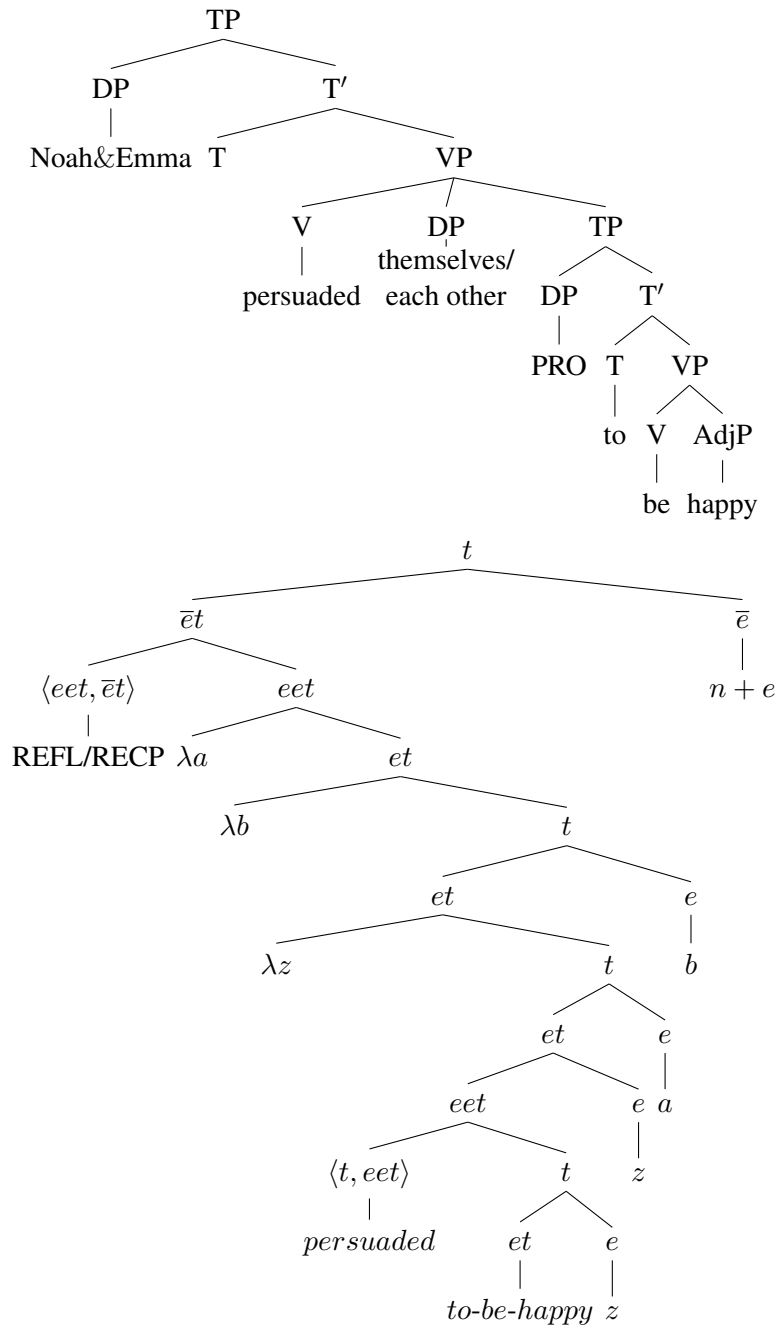The derived trees for the object control example (7) using the elementary trees and derivation of Figure 6 are provided in Figure 8.

Figure 8: Derived trees for the object control example (7)