Combining Predicate-Argument Structure and Operator Projection: Clause Structure in Role and Reference Grammar

Laura Kallmeyer & Rainer Osswald

Heinrich-Heine-Universität Düsseldorf

13th International Workshop on Tree Adjoining Grammars and Related Formalisms (TAG+13)

September 4–6, 2017

Umeå University, Sweden

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

CRC 991

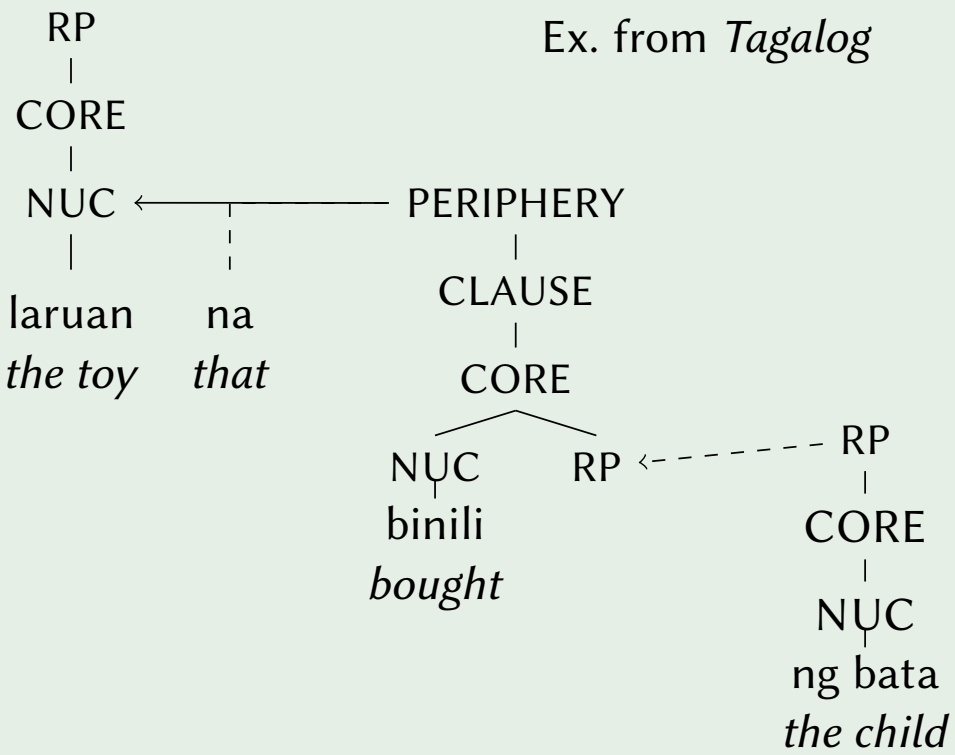
Introduction

Role and Reference Grammar (RRG; Van Valin & LaPolla 1997; Van Valin 2005) is a typologically rich grammar theory.

Introduction

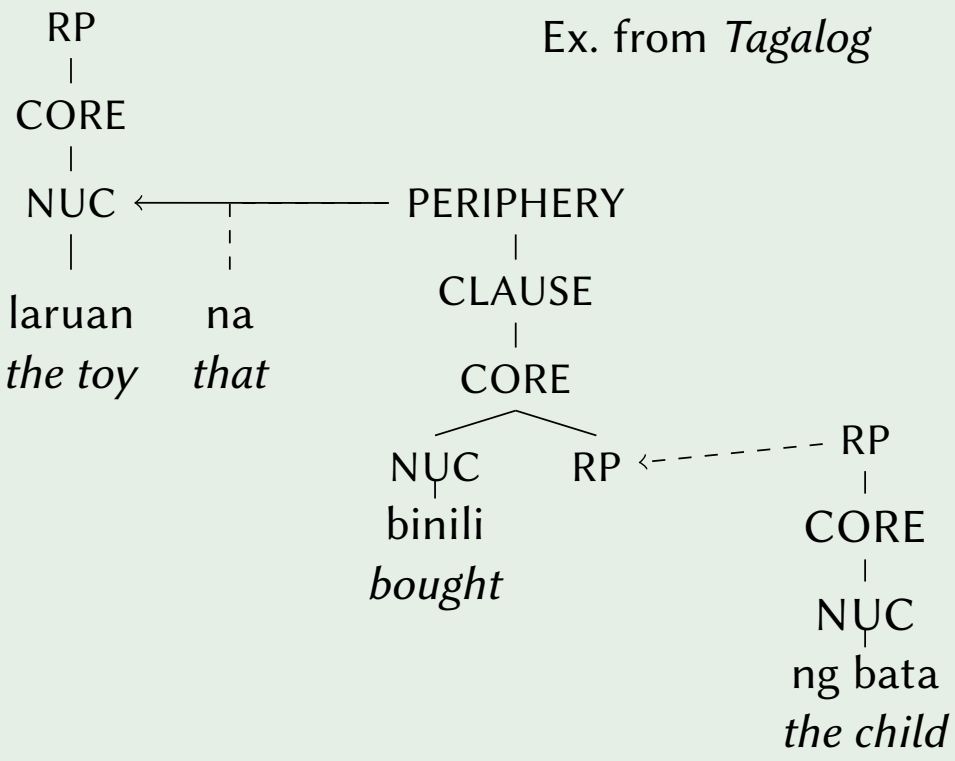
Role and Reference Grammar (RRG; Van Valin & LaPolla 1997; Van Valin 2005) is a typologically rich grammar theory.

Ex. from *Tagalog*



Introduction

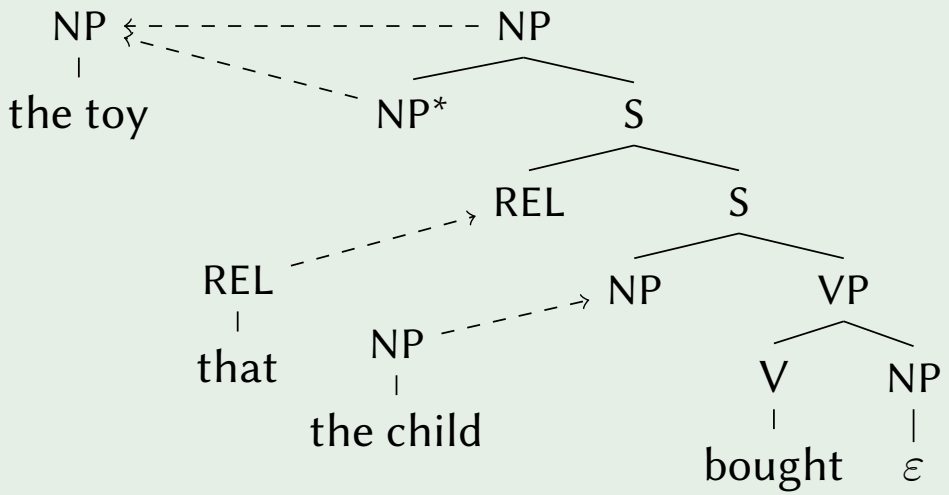
Role and Reference Grammar (RRG; Van Valin & LaPolla 1997; Van Valin 2005) is a typologically rich grammar theory.



But: not fully formalized, no implementation framework.

Introduction

Observation: RRG shares with TAG its extended domain of locality and the assumption that a predicate and its arguments are realized within the same elementary tree.

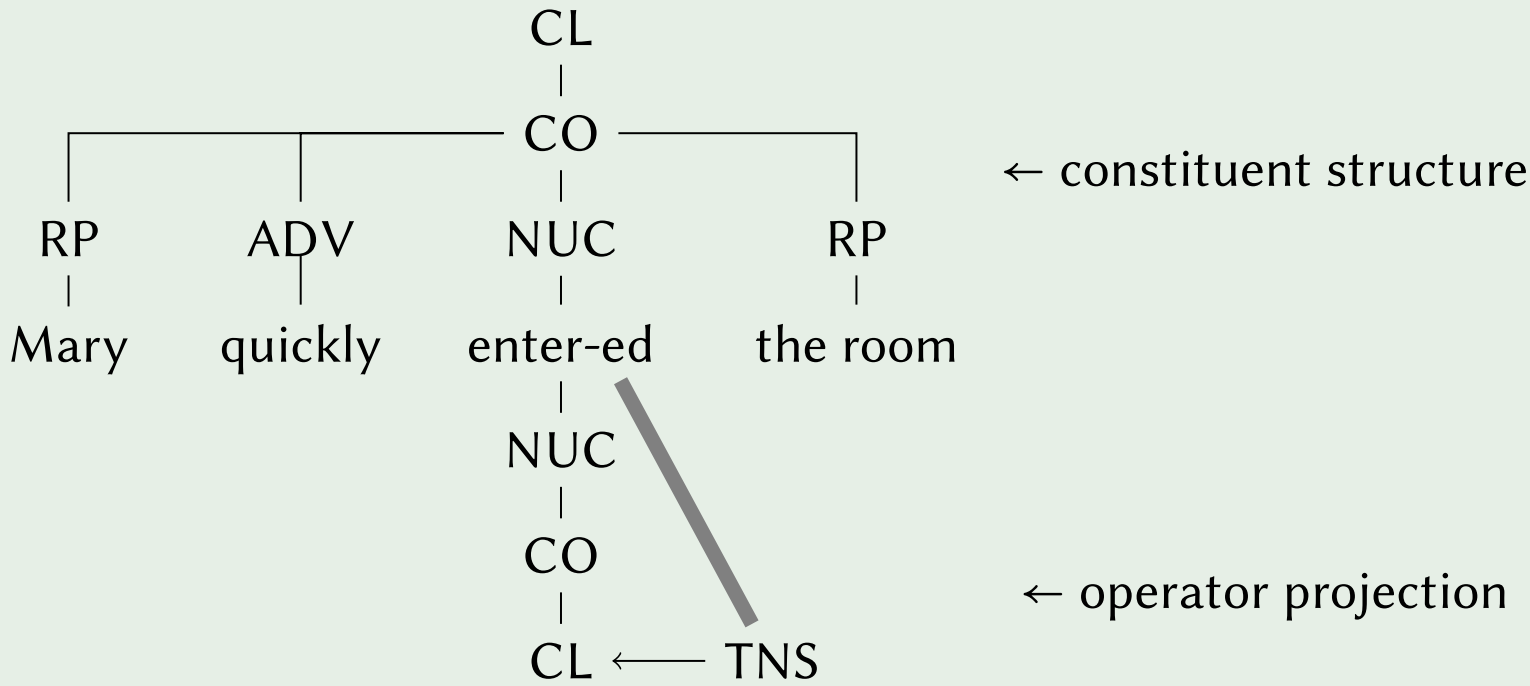


Introduction

- **Overall goal:** formalizing the theory of Role and Reference Grammar (RRG; Van Valin & LaPolla 1997; Van Valin 2005) as a tree-rewriting system
- **Contribution of this paper:** integrate RRG's rather flat constituent structure and its operator projection, which reflects the scopal properties of functional operators, in a single tree

Introduction

- **Overall goal:** formalizing the theory of Role and Reference Grammar (RRG; Van Valin & LaPolla 1997; Van Valin 2005) as a tree-rewriting system
- **Contribution of this paper:** integrate RRG's rather flat constituent structure and its operator projection, which reflects the scopal properties of functional operators, in a single tree



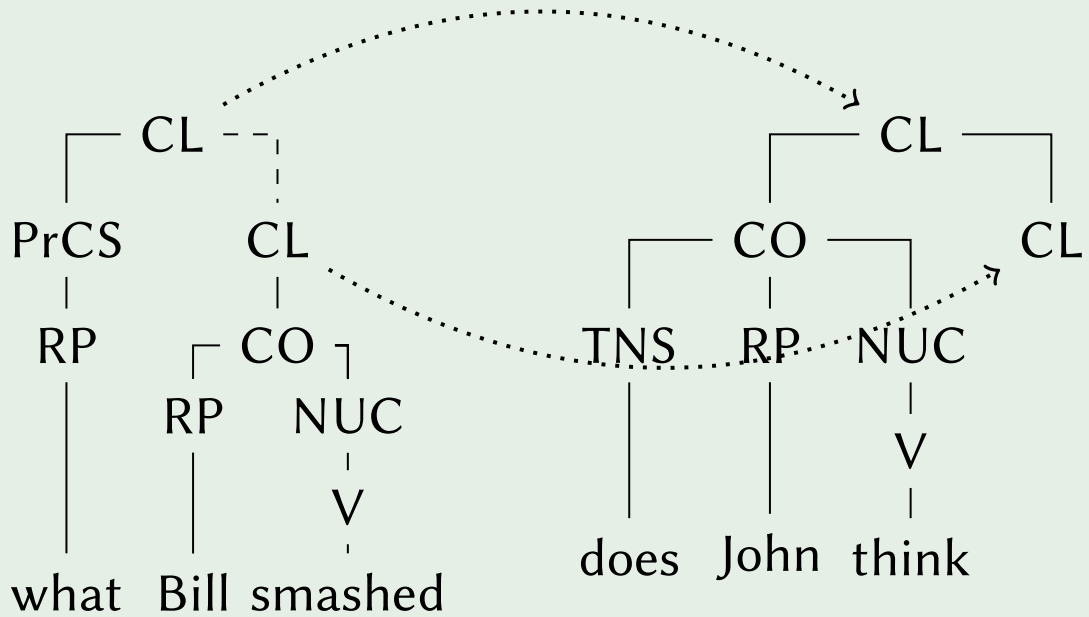
Tree Composition in RRG

Formalization of RRG as a tree-rewriting grammar with two operations for combining elementary trees (Kallmeyer et al., 2013; Osswald & Kallmeyer, in press):

Tree Composition in RRG

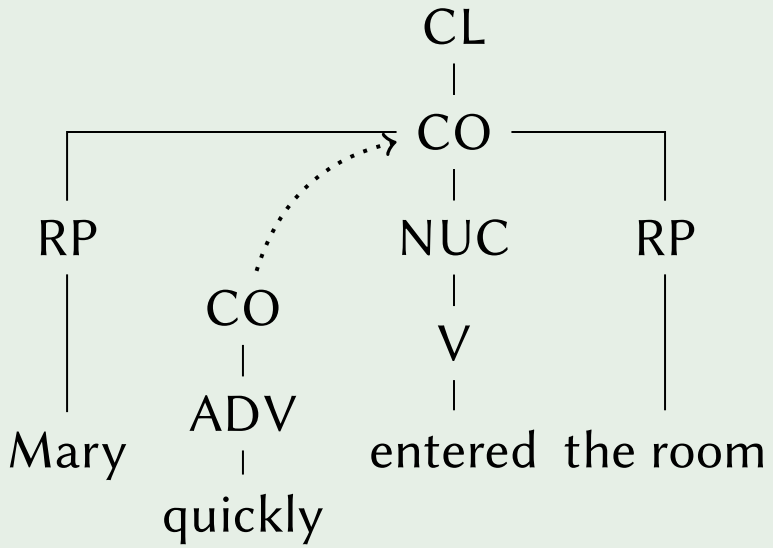
Formalization of RRG as a tree-rewriting grammar with two operations for combining elementary trees (Kallmeyer et al., 2013; Osswald & Kallmeyer, in press):

(Wrapping) substitution for argument composition



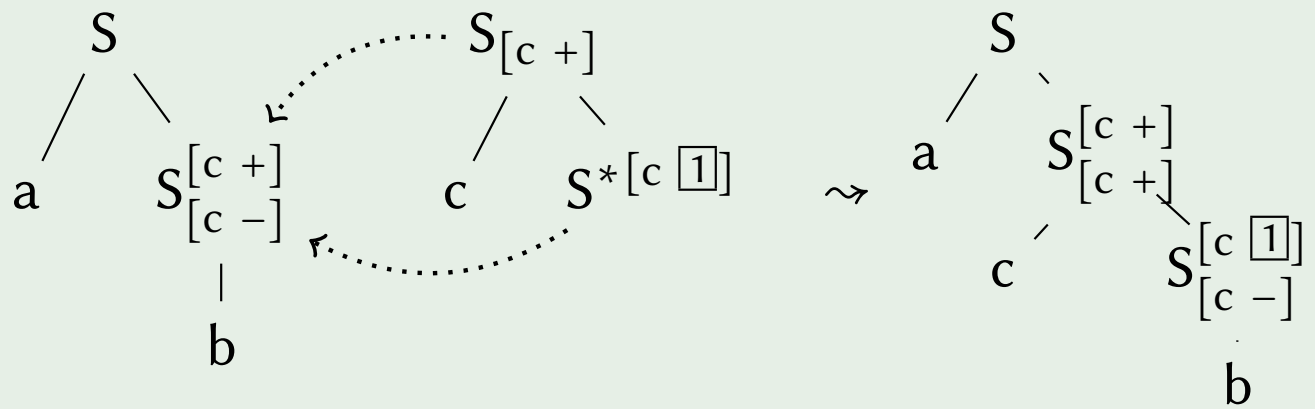
Tree Composition in RRG

Sister adjunction for modification



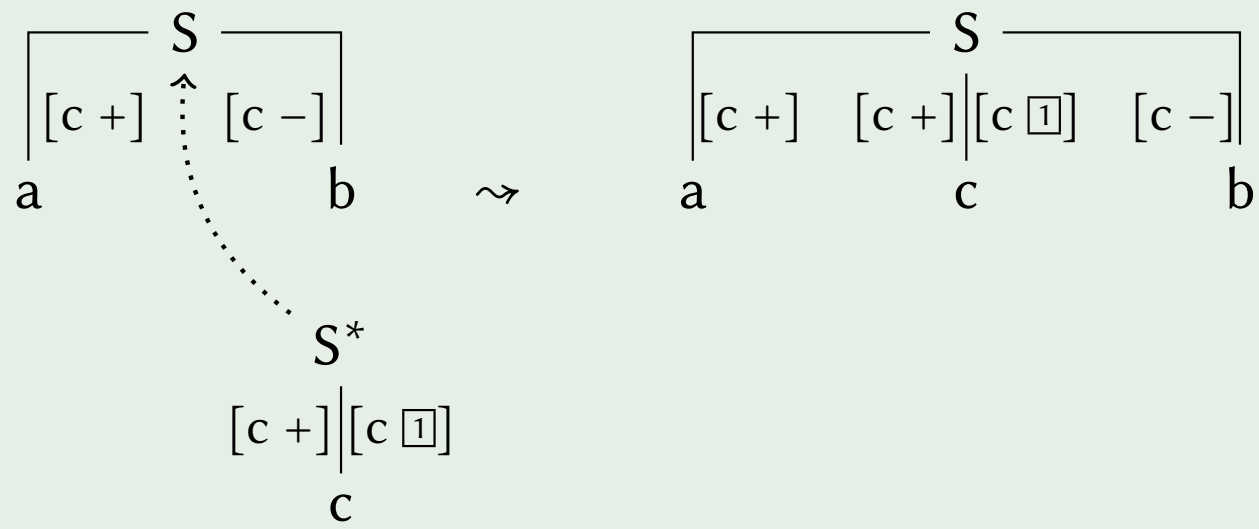
Adding features

In TAG (mostly binary tree structures), we have top and bottom feature stuctures that can constrain adjunction.



Adding features

In our flat structures with sister adjunction, we use left and right edge features to capture adjunction constraints.



Adding features

- Finite set of untyped feature structures with structure sharing within elementary trees (just like TAG, Vijay-Shanker & Joshi, 1988).
- Nodes have a single feature structure while edges have a left one and a right one.
- In a sister adjunction, the feature structure of the root of the adjoined tree unifies with the one of the target node.
- In the final derived tree, the two feature structures between two neighbouring edges have to unify.

  Furthermore, features on the leftmost (resp. rightmost) edge percolate upwards, except if there is a substitution node, which blocks feature percolation.

Integrating operators

Each operator belongs to a certain level of RRG's layered structure:

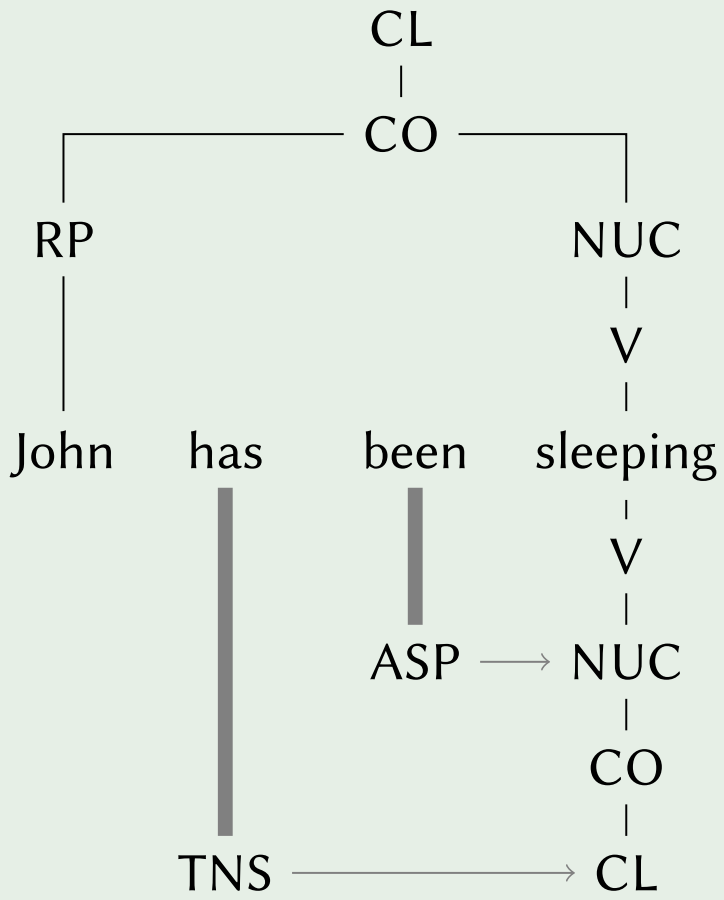
Layer	Nucleus	Core	Clause
Operators	Aspect	Directionals	Status
	Negation	Event quantification	Tense
	Directionals	Modality	Evidentials
		Negation	Illocutionary Force

The operator level explains

- the scope behavior: structurally higher operators take scope over lower ones
- surface order constraints: higher operators are further away from the nucleus of the structure.

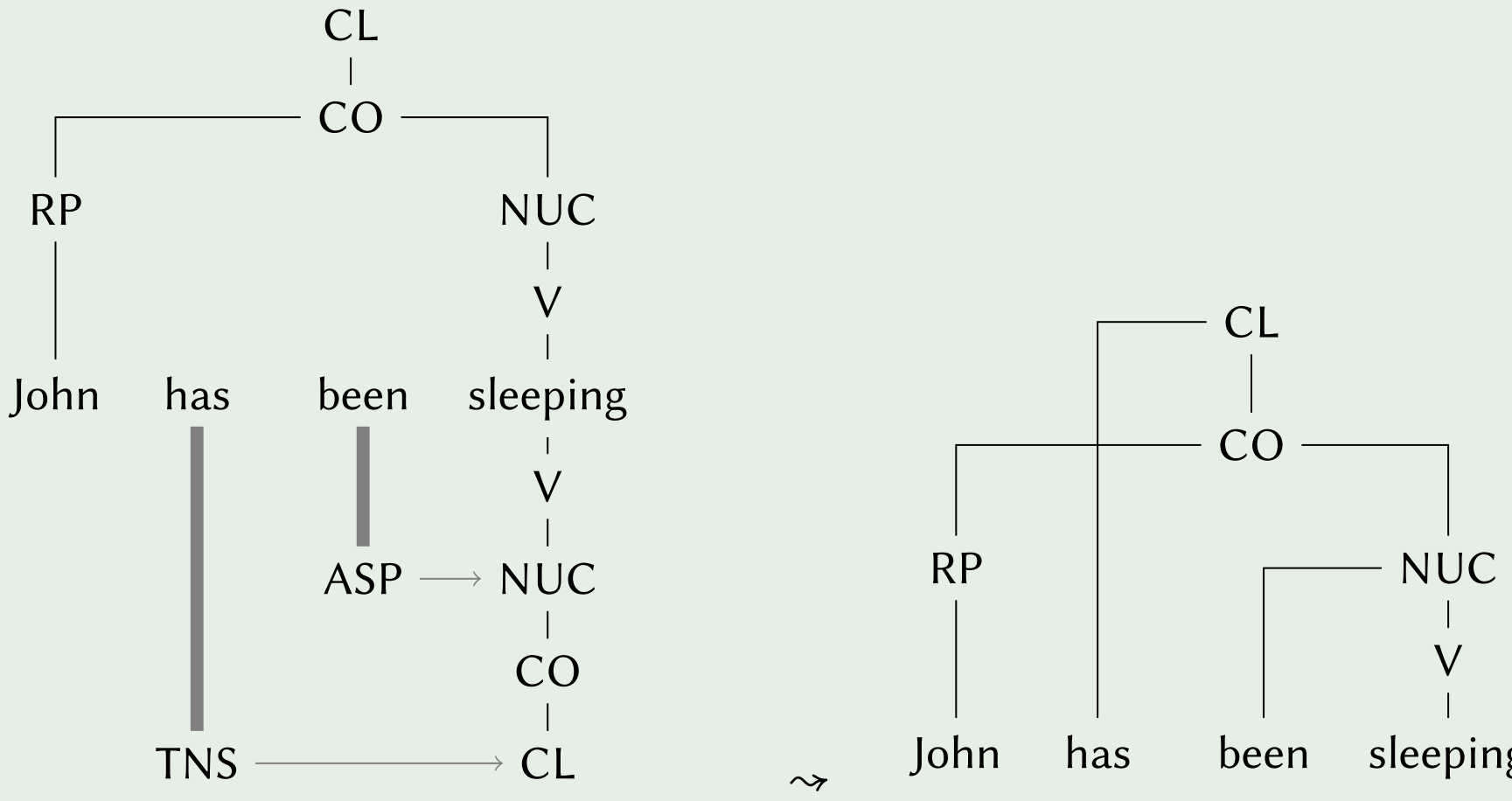
Integrating operators

Problem: constituent and operator structure are not completely parallel. An operator belonging to a specific layer can be surrounded by elements belonging to a lower layer in the constituent structure.



Integrating operators

Problem: constituent and operator structure are not completely parallel. An operator belonging to a specific layer can be surrounded by elements belonging to a lower layer in the constituent structure.

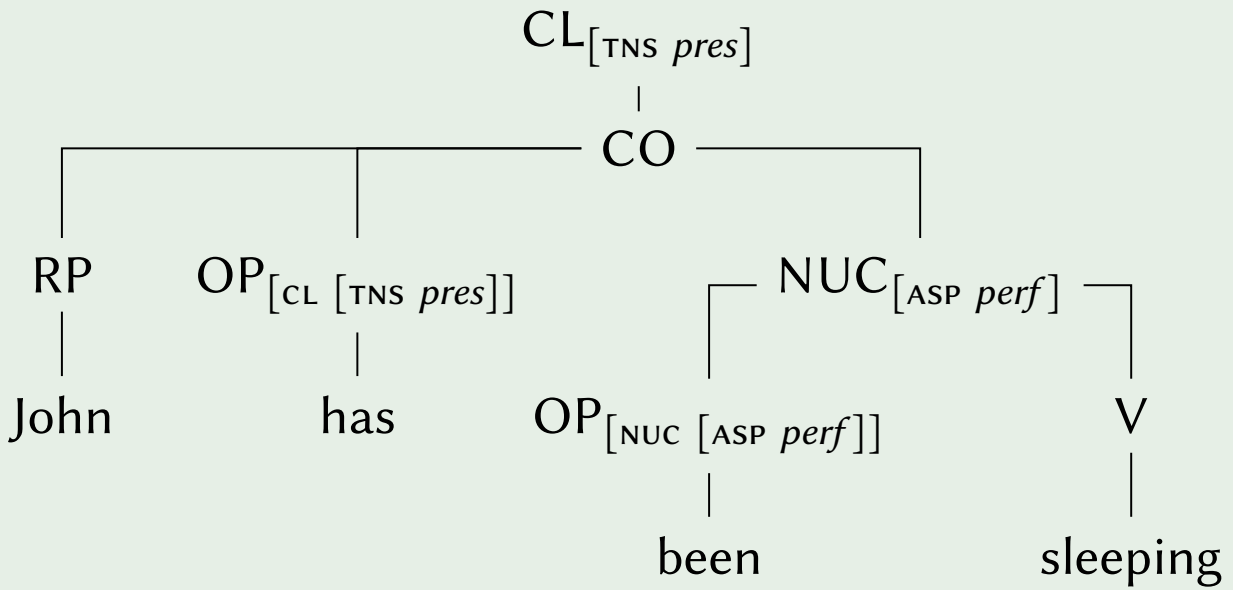


Integrating operators

The following holds:

- The hierarchical order of constituent and operator structure is the same.
- The existence of a layer in the operator projection requires that this layer also exists in the constituent structure.

We model the operator projection within the features while attaching the operators at their surface position.

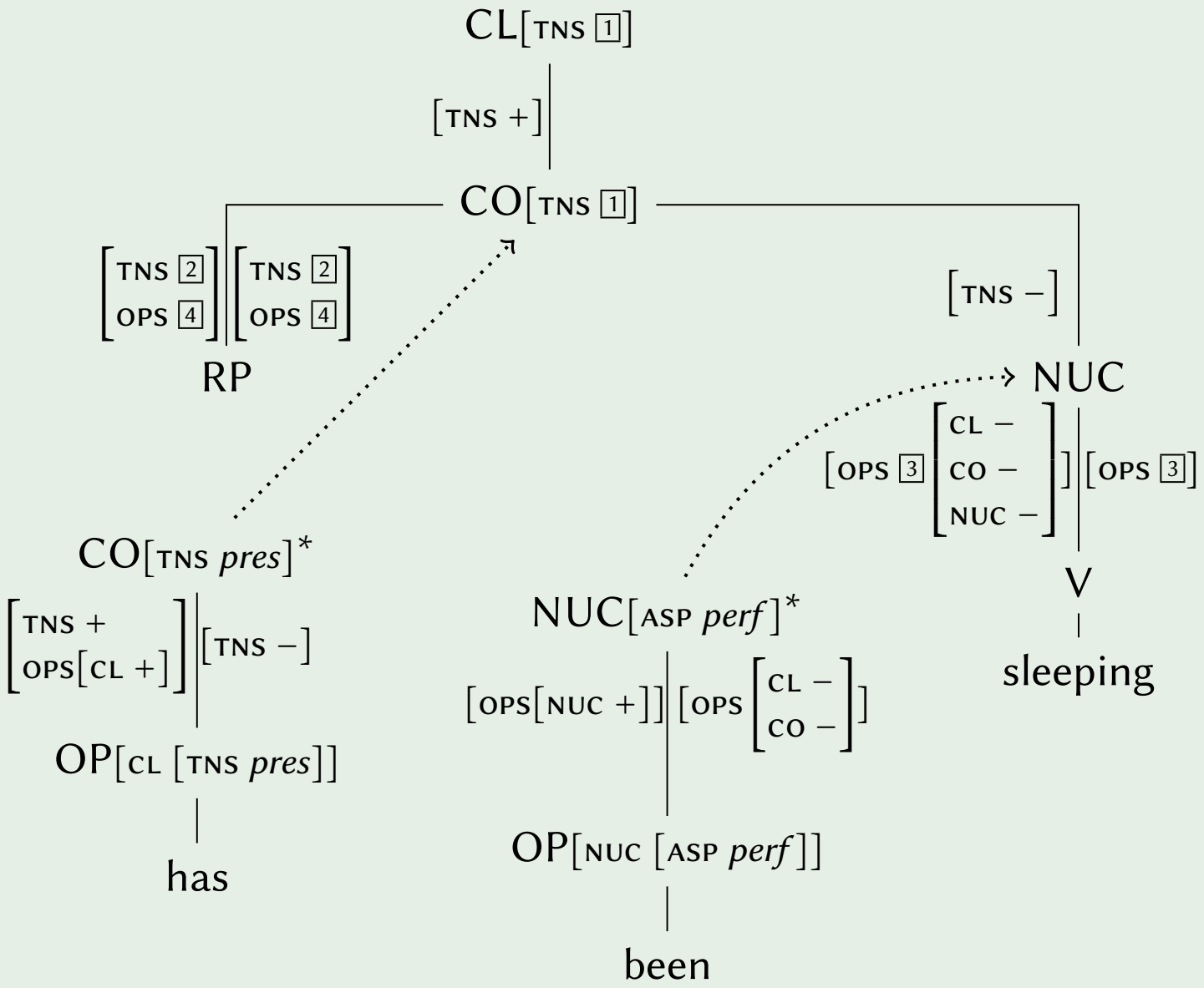


Integrating operators

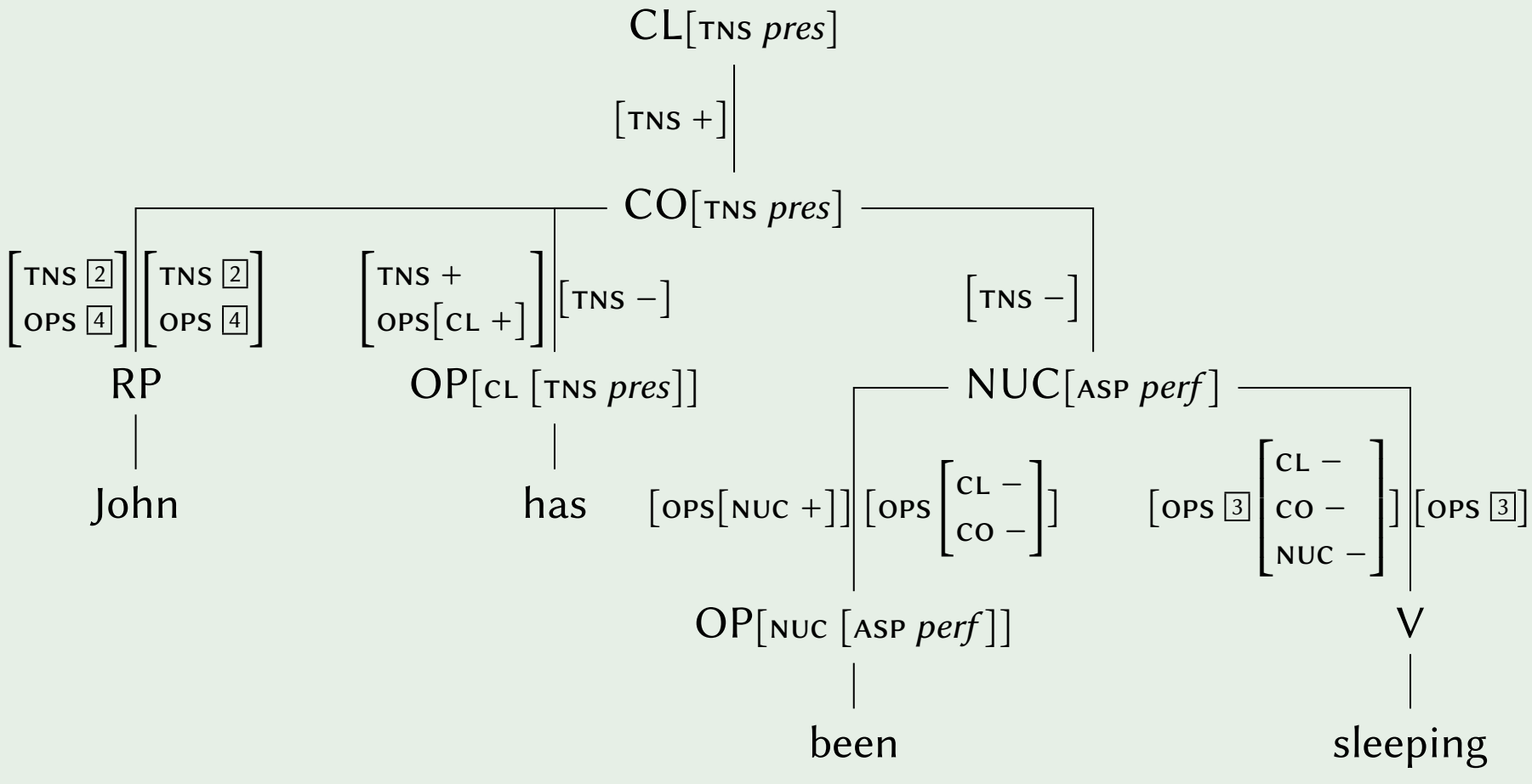
Features for operators (syntactic category OP):

- edge features TNS etc. that express the presence/absence of a specific operator and that can be used to formulate obligatory adjunction constraints.
- edge feature OPS (= operator structure), its value being a feature structure with features CL, CO and NUC with possible values + or −. OPS guarantees that nuclear, core and clausal operators have to appear in this order when moving outwards from the nuclear predicate.
- node features that specify the contribution of the operator, for instance $[\textsc{nuc}\ [\textsc{asp}\ \textit{perf}], \textsc{cl}\ [\textsc{tns}\ \textit{past}]]$ for the operator *had* in "John had slept".

Integrating operators



Integrating operators



Operators in complex sentences: Cosubordination

Cosubordination structures in RRG

- have basically the form $[[\ ]_X\ [\ ]_X]_X$.
- have the characteristic property that X-operators are realized only once but have scope over both constituents.

Examples from Van Valin (2005):

(1) [[Gid-ip]$_{CO}$ [gör-meli-yiz]$_{CO}$]$_{CO}$ (Turkish)
go-LM[1] see-MOD-1PL
'We ought to go and see.'

(2) [[Kim must$_{MOD}$ go]$_{CO}$ [to try]$_{CO}$ [to wash the car]$_{CO}$]$_{CO}$

We assume that it is a general property of cosubordination elementary trees that operator features get passed upwards to the higher X.

[1] LM = linkage marker

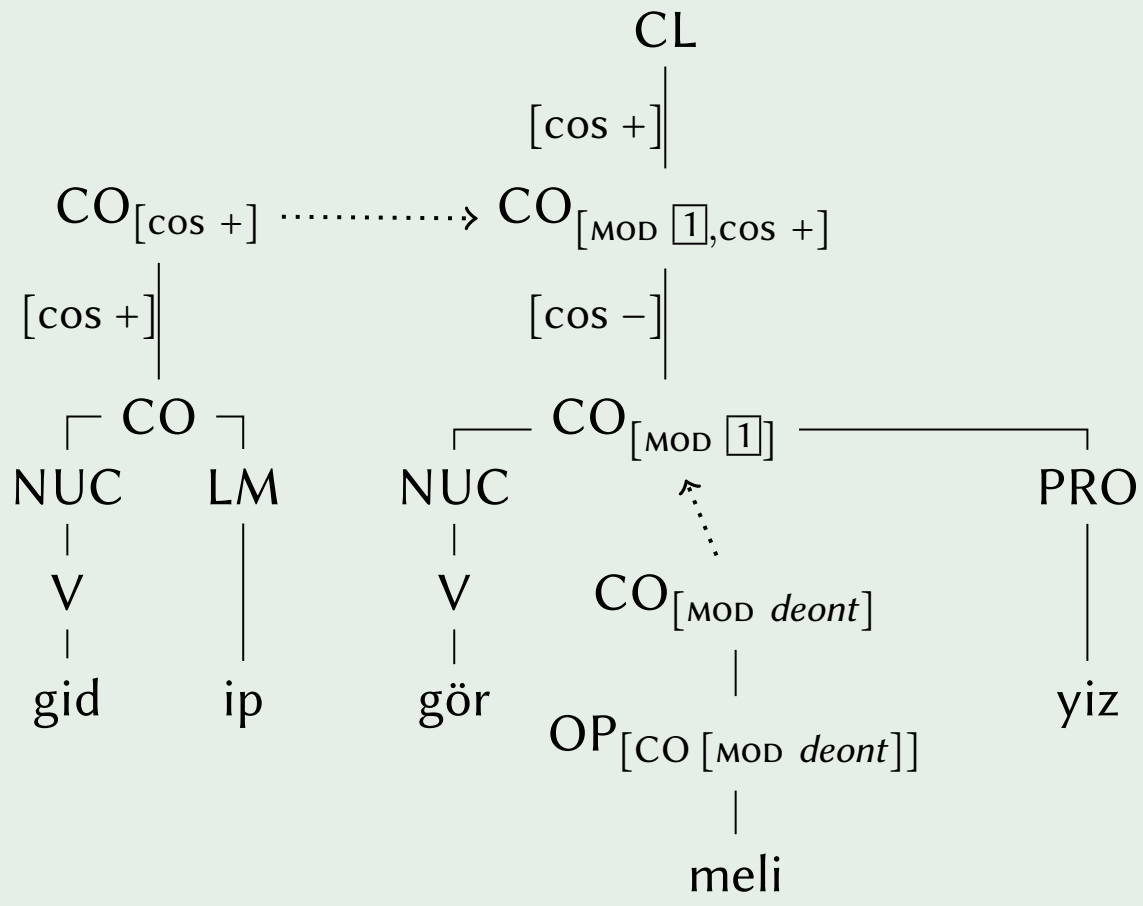
Operators in complex sentences: Cosubordination

[[Gid-ip]$_{\text{CO}}$ [gör-meli-yiz]$_{\text{CO}}$]$_{\text{CO}}$

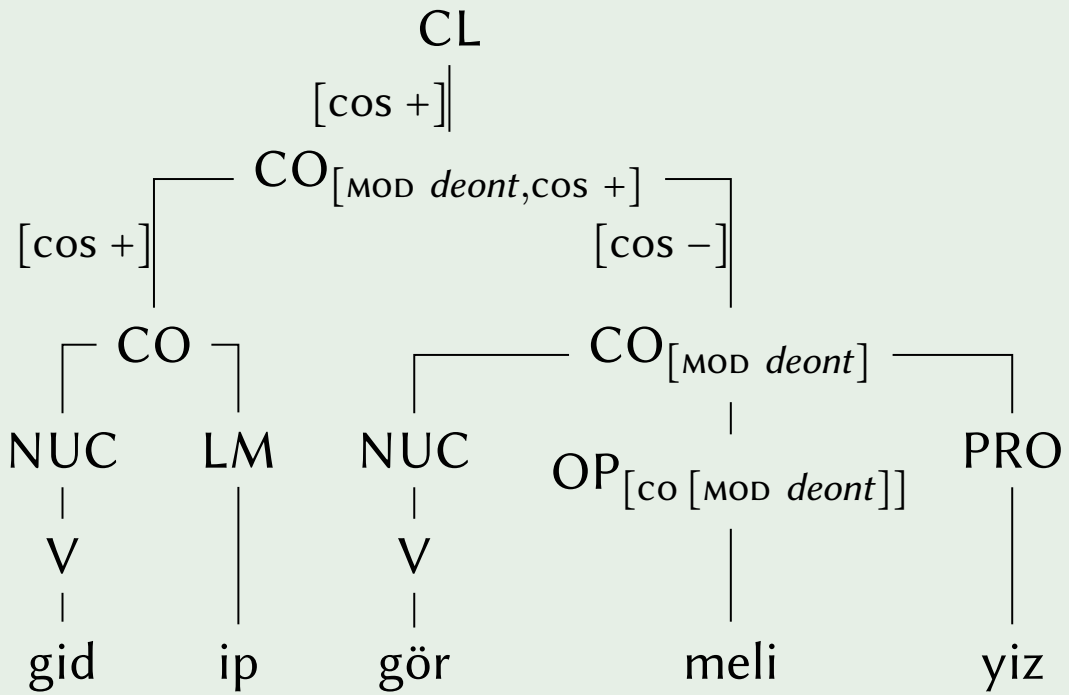
Proposal for the elementary trees:

- Special cosubordination tree for *gör PRO* that provides a lower and a higher CO node.
- CO operator features (e.g., MOD) are shared between the two CO nodes and thereby passed upwards from the lower node.
- *gid-ip* is added by adjunction, targeting the higher CO node, thereby adding a second CO daughter.
- Edge feature cos (values +/-) that indicates that adjunction of at least one more core to the left is obligatory.
- Node feature cos (values +/-) that indicate whether a node is the root of a cosubordination structure.

Operators in complex sentences: Cosubordination



Operators in complex sentences: Cosubordination



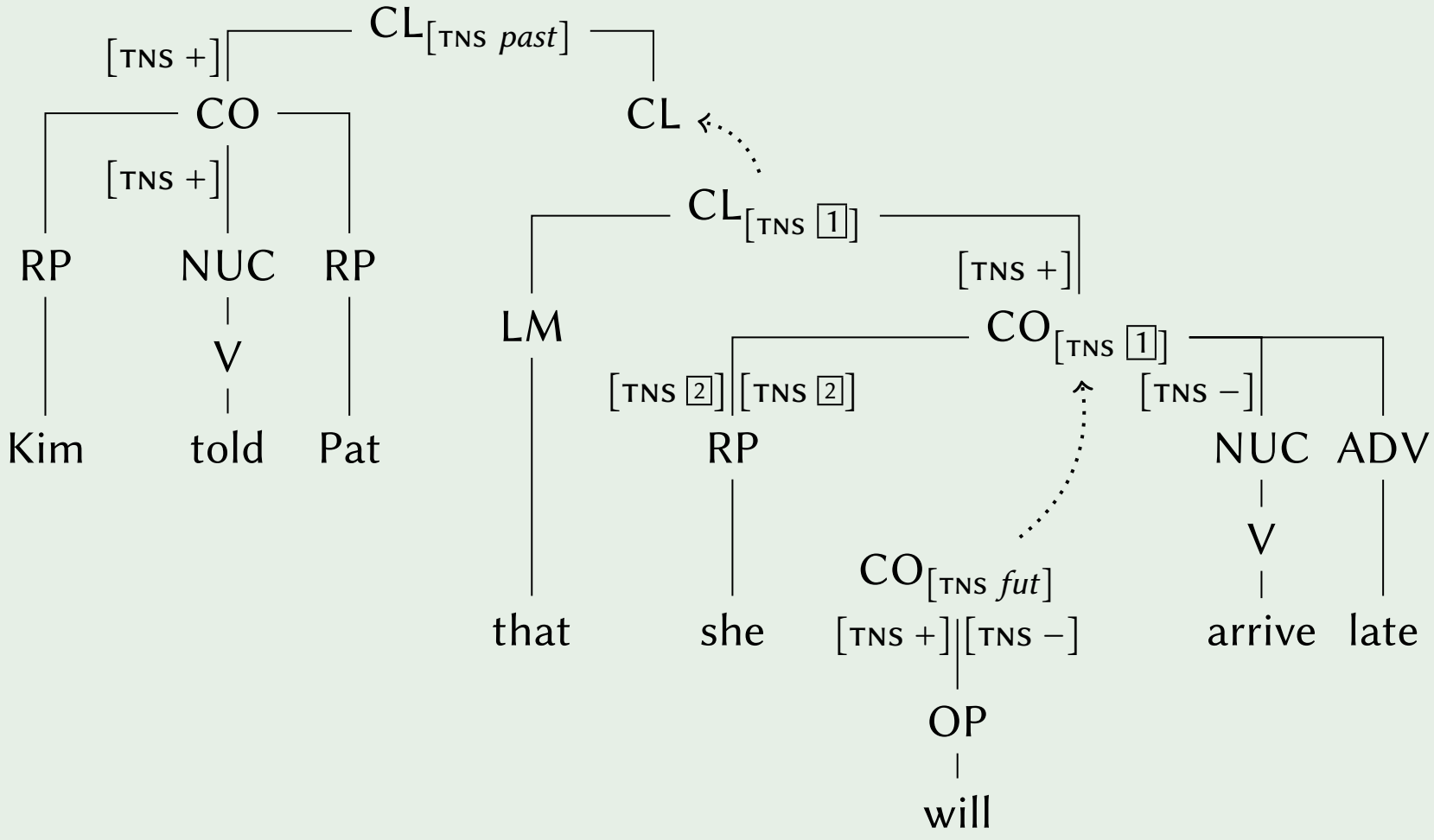
Operators in complex sentences: Subordination

In **subordination** structures, operator projections are built locally. The composition operation is substitution, which means that edge feature percolation is blocked.

(3) [[Kim told Pat]$_{\text{CO}}$ [that [she will arrive late]$_{\text{CO}}$]$_{\text{CL}}$]$_{\text{CL}}$

The two CL nodes in this structure have different TNS values, provided by *told* and *will* respectively.

Operators in complex sentences: Subordination



Conclusion: Summary

- Background: TAG-inspired formalization of RRG with (wrapping) substitution for complement insertion and sister adjunction for adding modifiers.
- We introduced features and proposed to use left and right edge features in order to model adjunction constraints.
- Given this architecture, RRG's operator projection can be integrated into the constituent structure, modeling the operator hierarchy and its interaction with the constituent structure within the features.
- The difference between cosubordination (lower operators take wide scope) and subordination (operator scope is blocked) can be accounted for by appropriate elementary trees and by treating substitution nodes as islands concerning edge feature percolation.

Conclusion: Future work

- Inspect further cases of complex sentences.
- Model the scopal structure of periphery modifiers (e.g., adverbs).
  The assumption is that this can be done in a similar way as in the case of the operator scope.
  Example of a nuclear modifier that attaches at the core node:

  (4) Leslie immersed herself completely in the new language.

- Integrate this formalization of RRG into XMG in order to enable grammar implementation.
- Integrate RRG parsing into TuLiPA in order to enable grammar parsing for testing.
- Long-term goal: full formalization of RRG and integrated framework for RRG-based grammar development.

References

Bohnemeyer, Jürgen & Robert D. Van Valin. 2017. The macro-event property and the layered structure of the clause. *Studies in Language* 41(1). 142–197.

Kallmeyer, Laura, Rainer Osswald & Robert D. Van Valin, Jr. 2013. Tree wrapping for Role and Reference Grammar. In Glyn Morrill & Mark-Jan Nederhof (eds.), *Formal Grammar 2012/2013*, vol. 8036 LNCS, 175–190. Springer.

Osswald, Rainer & Laura Kallmeyer. in press. Towards a formalization of Role and Reference Grammar. In *Proceedings of the 2013 conference on role and reference grammar*, .

Van Valin, Robert D. 2005. *Exploring the syntax-semantics interface*. Cambridge University Press.

Van Valin, Robert D. & Randy LaPolla. 1997. *Syntax: Structure, meaning and function*. Cambridge University Press.

Vijay-Shanker, K. & Aravind K. Joshi. 1988. Feature structures based tree adjoining grammar. In *Proceedings of coling*, 714–719. Budapest.