Bootstrapping Role and Reference Grammar Treebanks via Universal Dependencies

Kilian Evang Tatiana Bladier
Laura Kallmeyer Simon Petitjean

2022-02-16
TreeGraSP Meeting #7

Outline

Introduction

UD to RRG Conversion

Impact on Annotation Effort

Conclusion

Outline

Introduction

UD to RRG Conversion

Impact on Annotation Effort

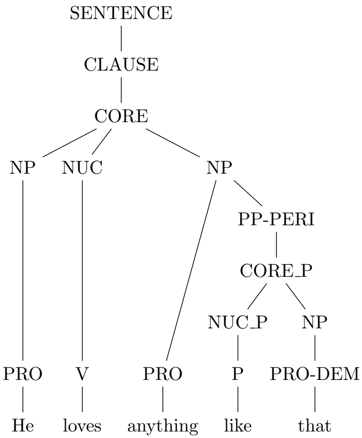
Conclusion

Introduction

- Context: creation of treebanks of syntactic structures (rrgbank, rrgparbank, etc.) based on RRG [Van Valin Jr. and LaPolla, 1997, Van Valin Jr., 2005]
- Problem: tedious process when starting from scratch
- Aim: provide a (reasonable) starting point for annotations
- Machine learning: only possible once enough data is annotated
- Dependency parsers: provide analyses for a large set of languages
- ud2rrg: convert a dependency parse into an RRG structure

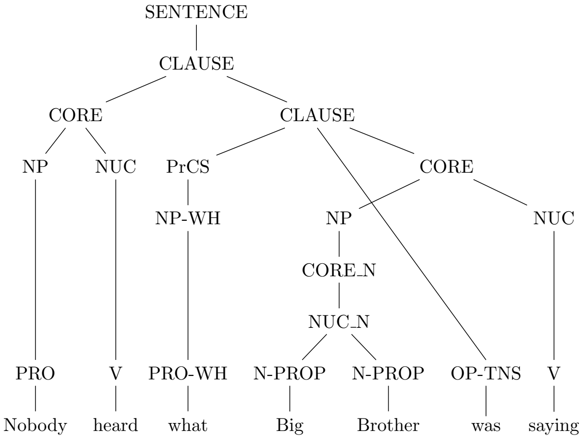
Introduction

- UD [Nivre et al., 2016, Nivre et al., 2020] and RRG:
  - descriptively adequate across typologically diverse languages
  - reflect their commonalities in analyses
- Slighlty adapted representation of operator projection:



Introduction

. . . possibly with crossing branches



Outline

Introduction

UD to RRG Conversion

Impact on Annotation Effort

Conclusion

Auxiliary Formalism

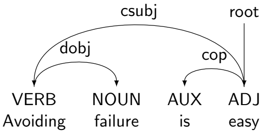
- Custom formalism inspired by LTAG: elementary trees + composition operations
- Operations compose elementary trees following RRG juncture-nexus types: coordination, subordination, cosubordination
- Operations apply at different levels: NUCLEUS, CORE, CLAUSE, PrCS, PrDP

Conversion Rules

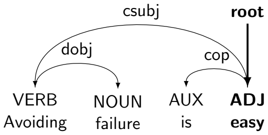
General rules:

- Every node in the dependency tree is converted to a RRG elementary tree
- Every edge in the dependency tree is converted to a composition operation
- Single top-down traversal of the ud tree, ideally one node and its incoming edge at the time

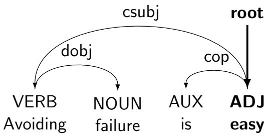
Conversion with General Rules

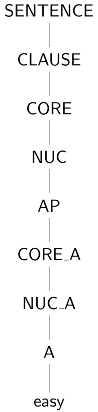


Conversion with General Rules

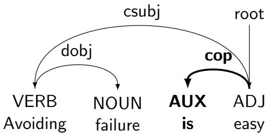


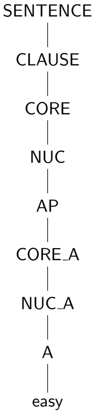
Conversion with General Rules



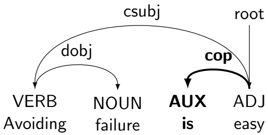


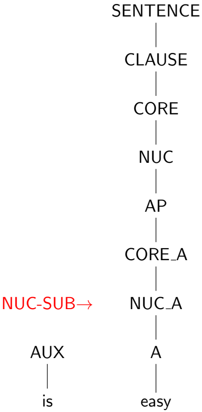
Conversion with General Rules



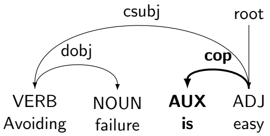


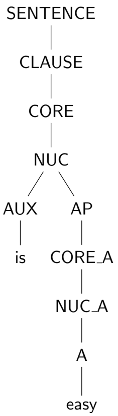
Conversion with General Rules



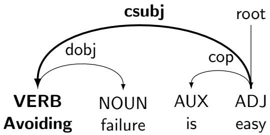


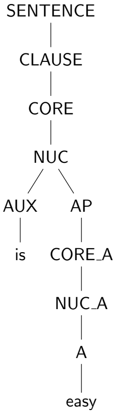
Conversion with General Rules



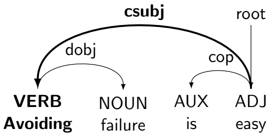


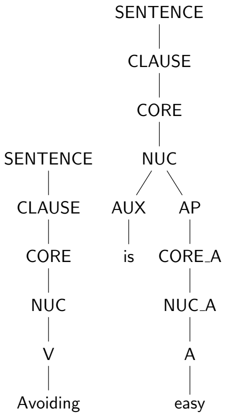
Conversion with General Rules



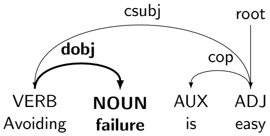


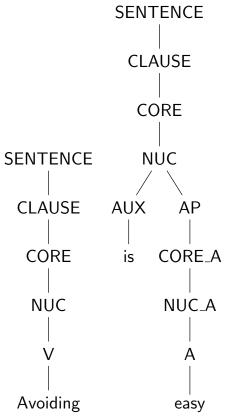
Conversion with General Rules



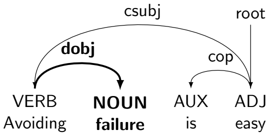


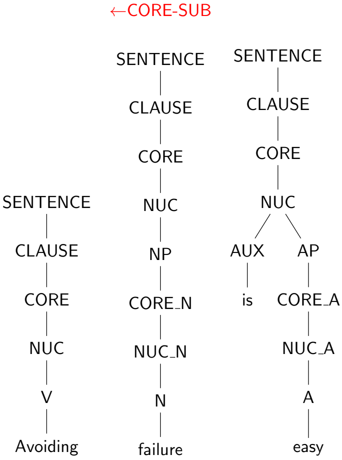
Conversion with General Rules



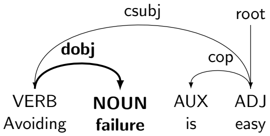


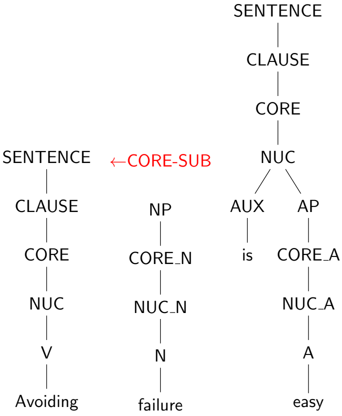
Conversion with General Rules



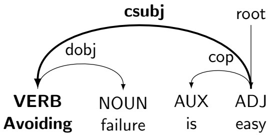


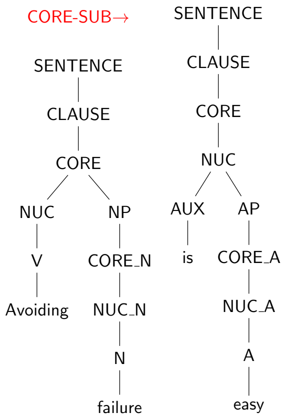
Conversion with General Rules



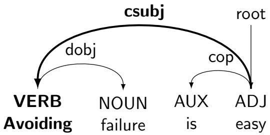


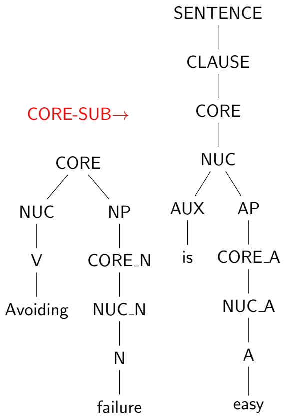
Conversion with General Rules



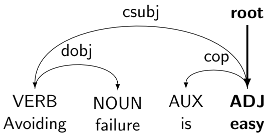


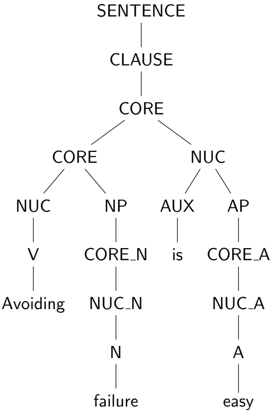
Conversion with General Rules





Conversion with General Rules

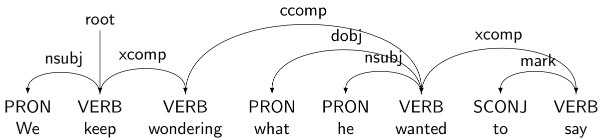




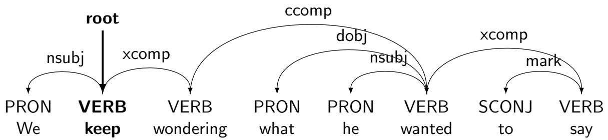
Conversion Rules: Special Rules

- Cases where RRG analyses are more informative than UD trees
- Example: ccomp dependency (clausal complements) →
  - CLAUSE subordination for verbs of cognition and saying
  - CORE subordination in other cases
- Example: xcomp dependency (open clausal complements) →
  - CLAUSE cosubordination for phase verbs (*starts walking*, *keep wondering*)
  - CORE coordination for some raising constructions (*seems*, *scheinen*)
  - CORE cosubordination in other cases
- Need to add lexical / language specific rules

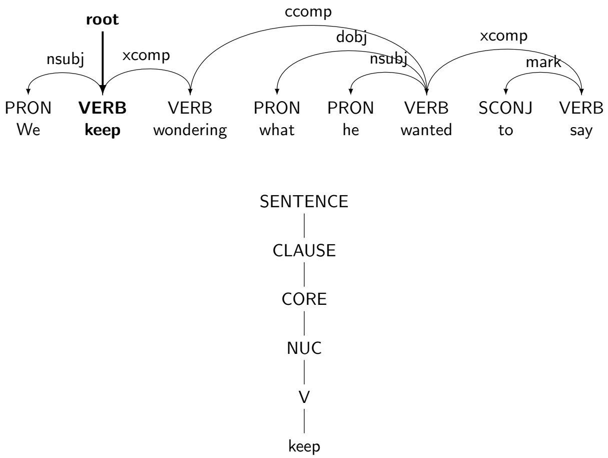
Conversion with Special Rules



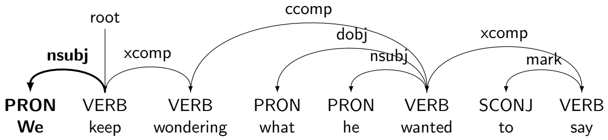
Conversion with Special Rules

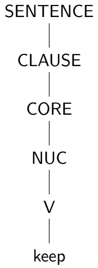


Conversion with Special Rules

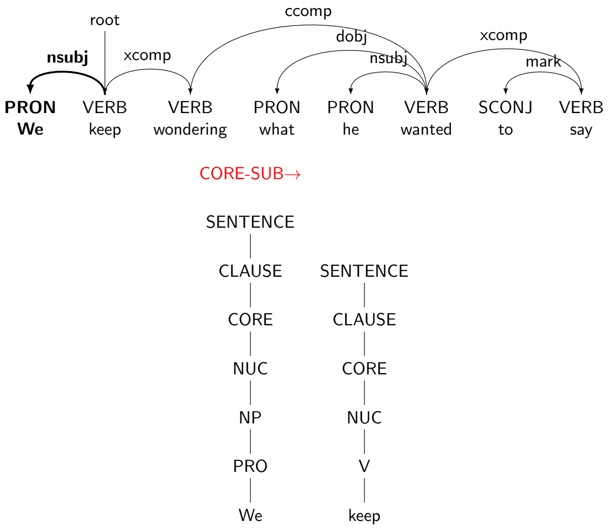


Conversion with Special Rules

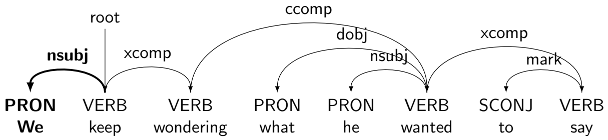


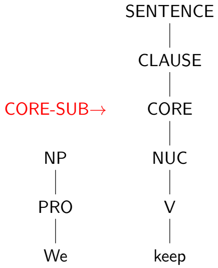


Conversion with Special Rules

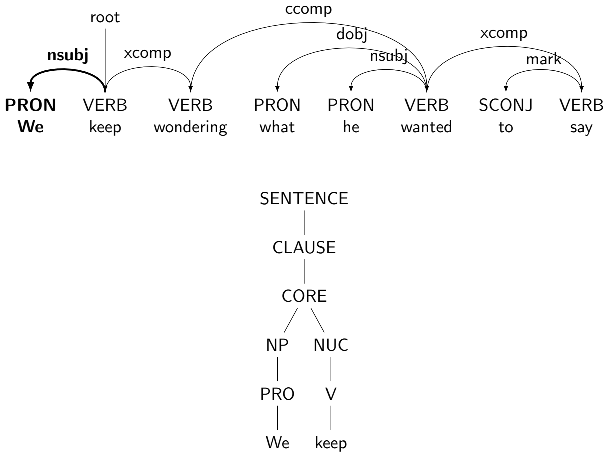


Conversion with Special Rules

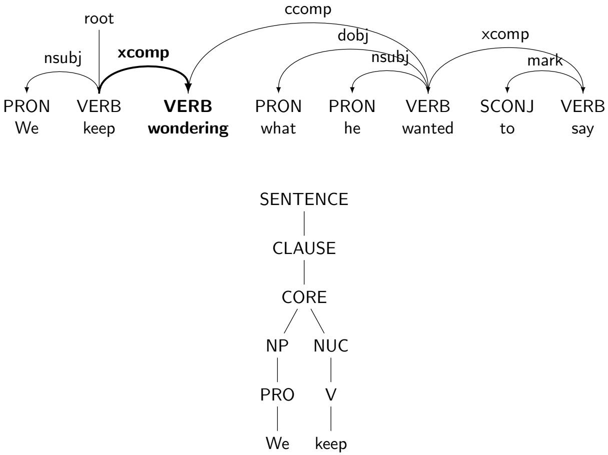




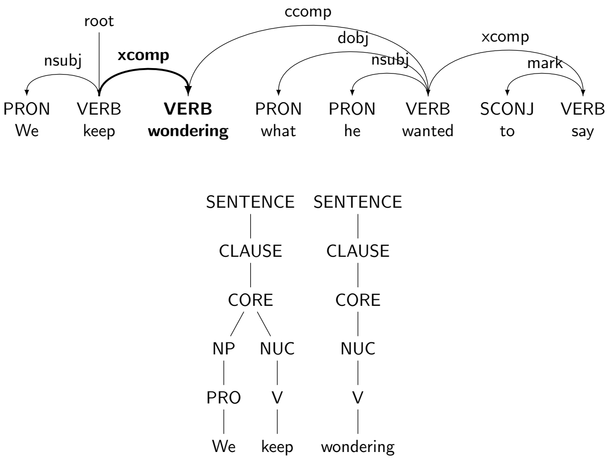
Conversion with Special Rules



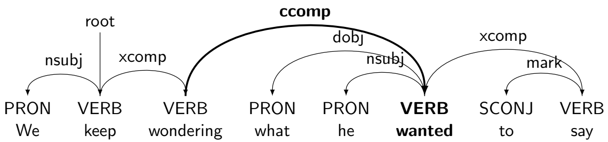
Conversion with Special Rules

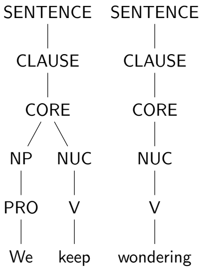


Conversion with Special Rules

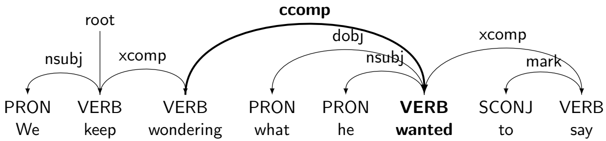


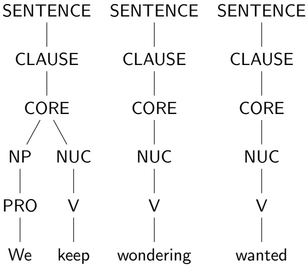
Conversion with Special Rules



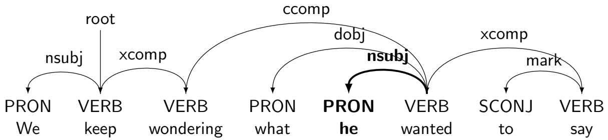


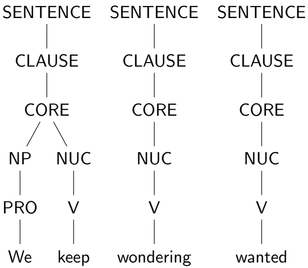
Conversion with Special Rules



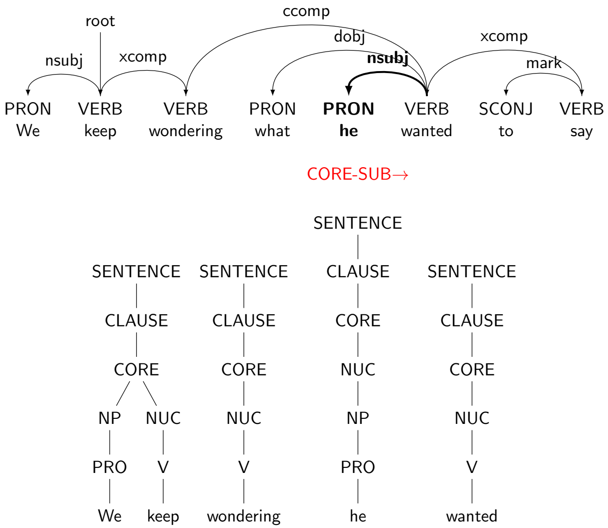


Conversion with Special Rules

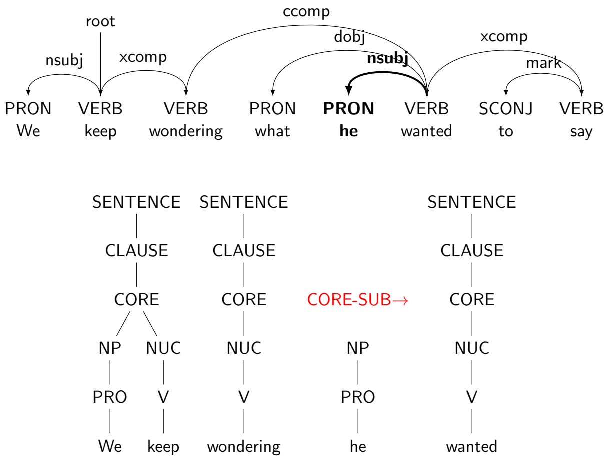




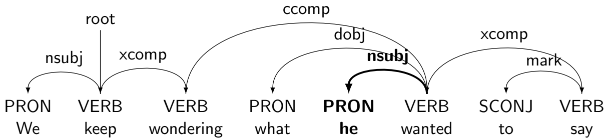
Conversion with Special Rules

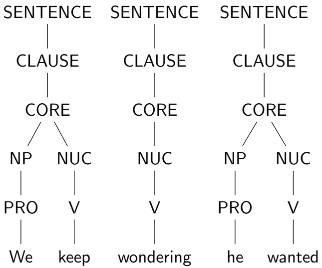


Conversion with Special Rules

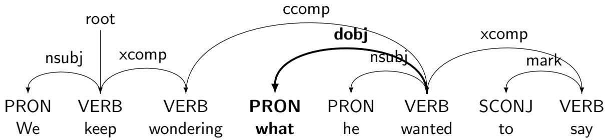


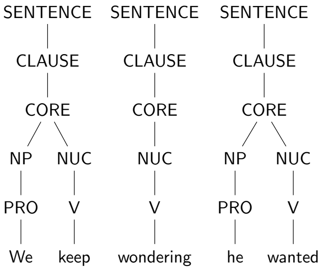
Conversion with Special Rules



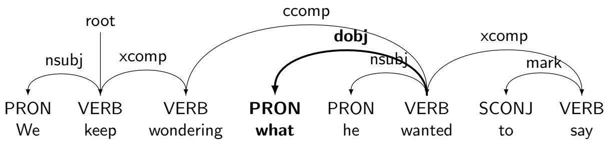


Conversion with Special Rules

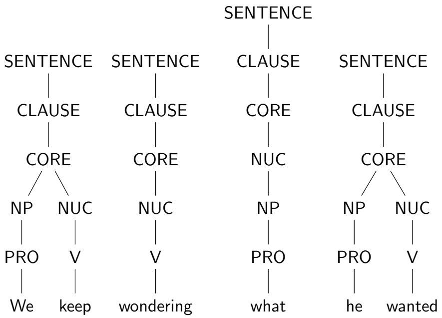




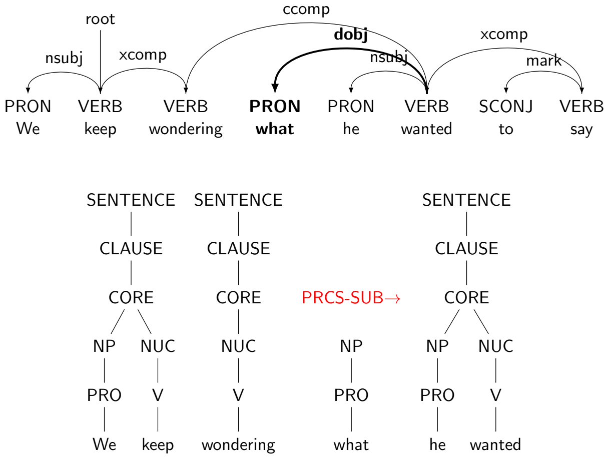
Conversion with Special Rules



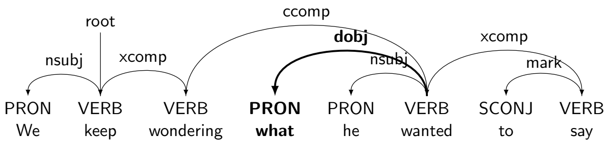
PRCS-SUB→

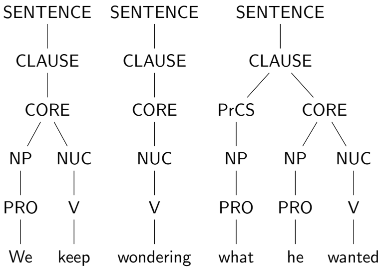


Conversion with Special Rules

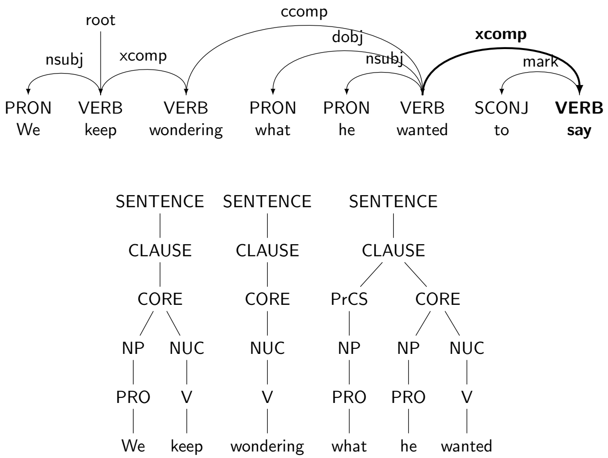


Conversion with Special Rules

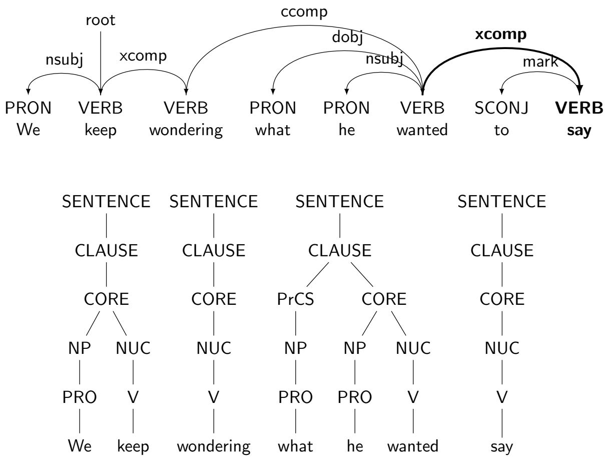




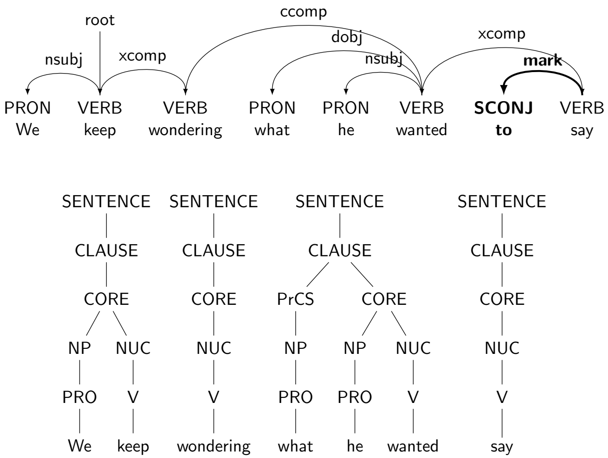
Conversion with Special Rules



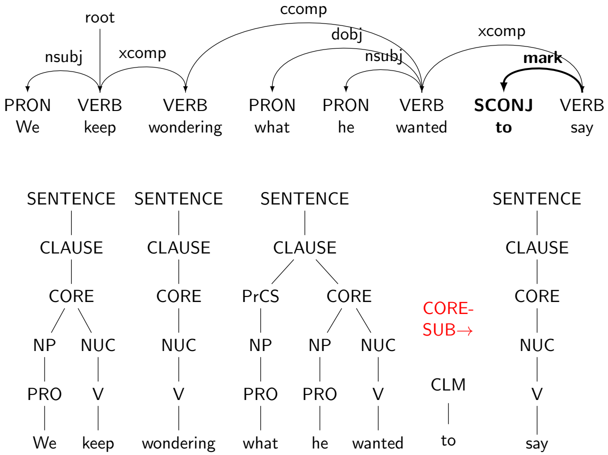
Conversion with Special Rules



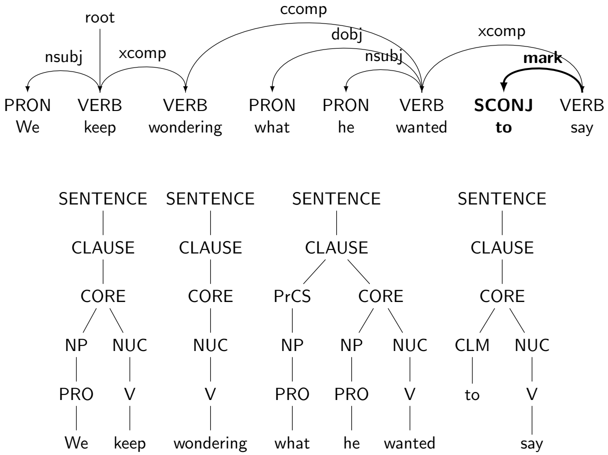
Conversion with Special Rules



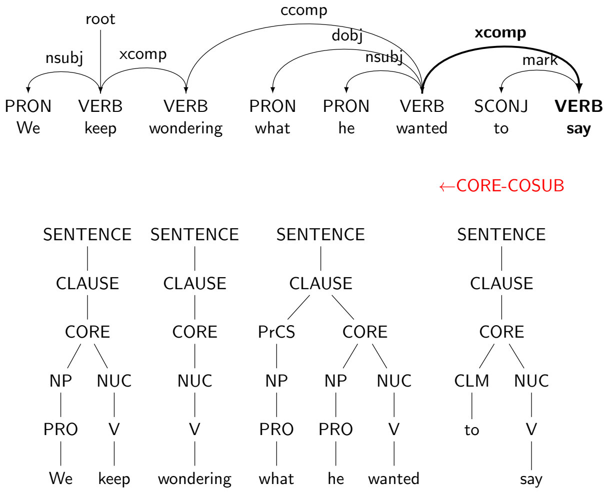
Conversion with Special Rules



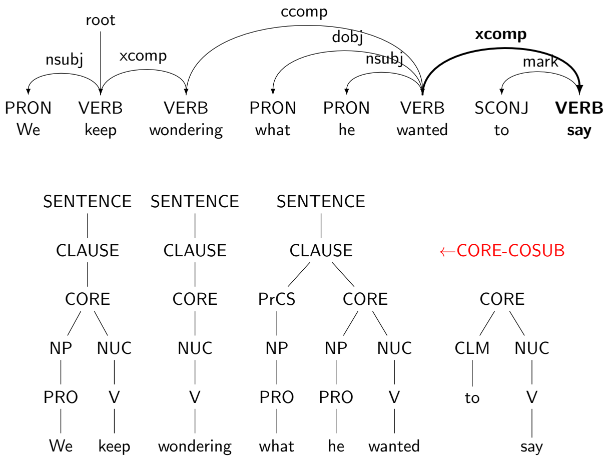
Conversion with Special Rules



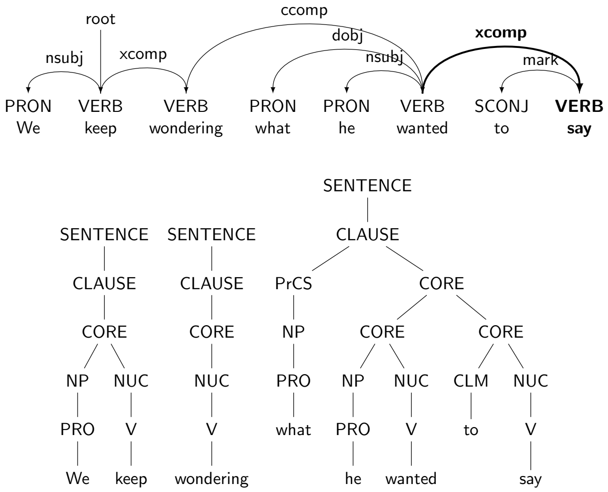
Conversion with Special Rules



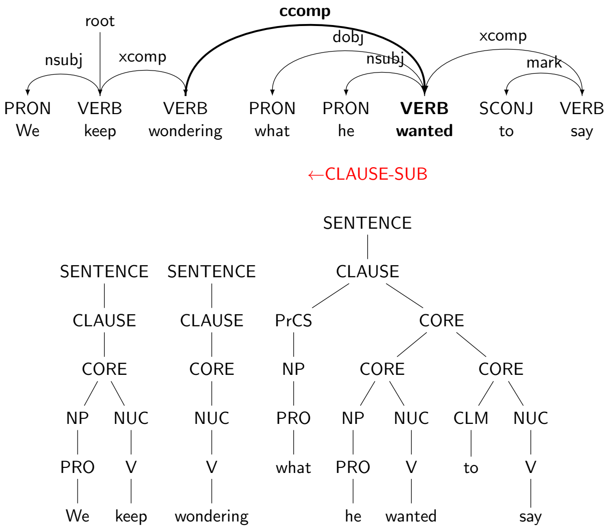
Conversion with Special Rules



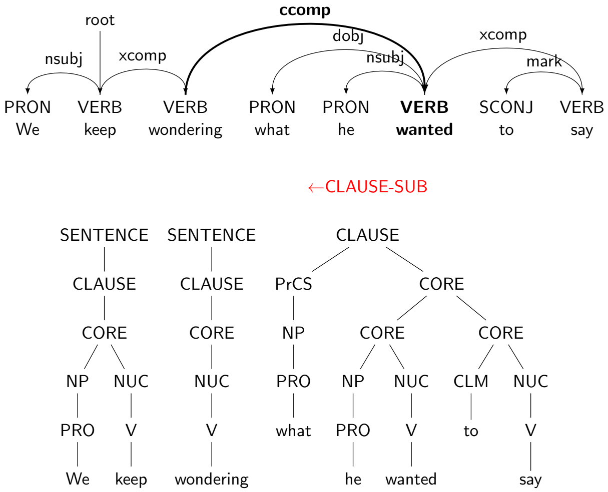
Conversion with Special Rules



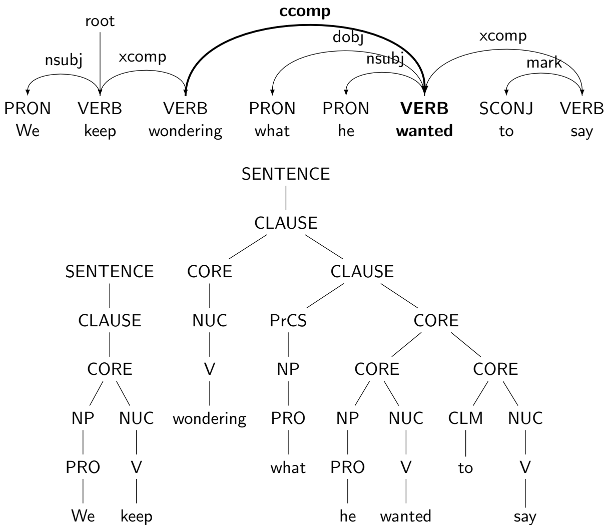
Conversion with Special Rules



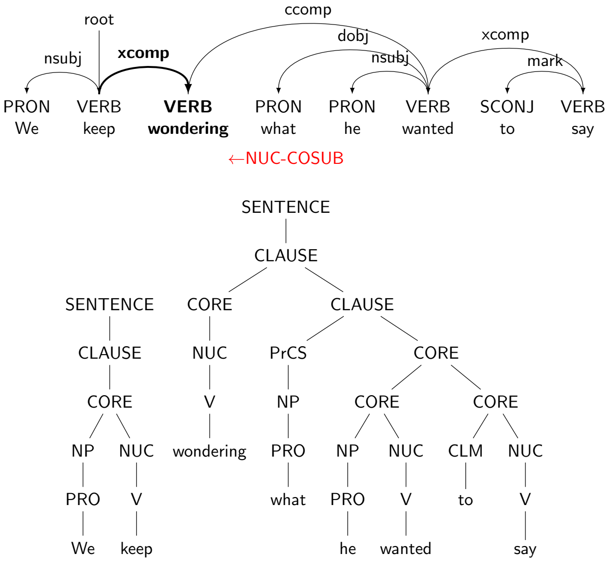
Conversion with Special Rules



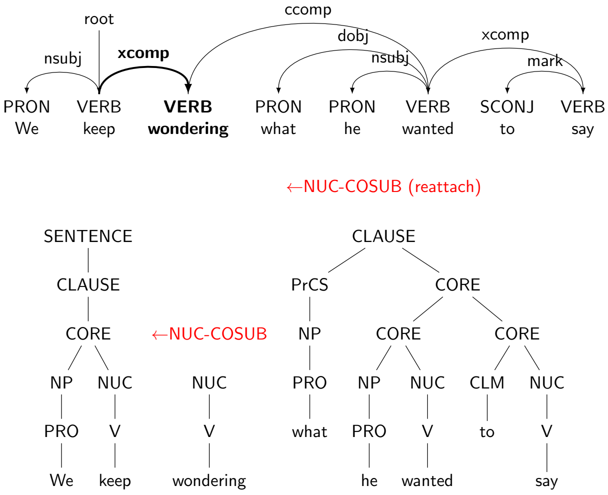
Conversion with Special Rules



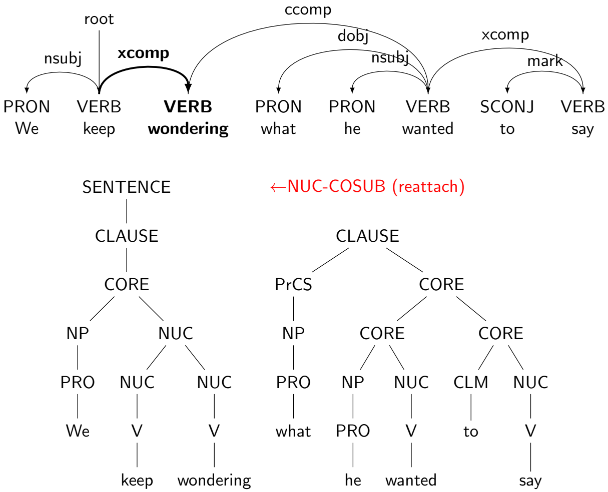
Conversion with Special Rules



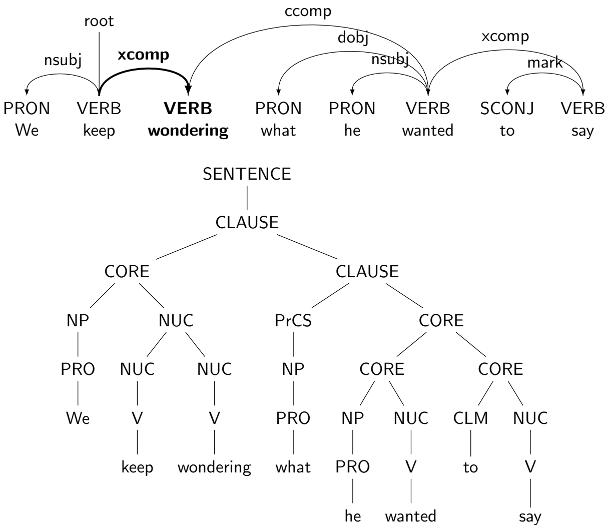
Conversion with Special Rules



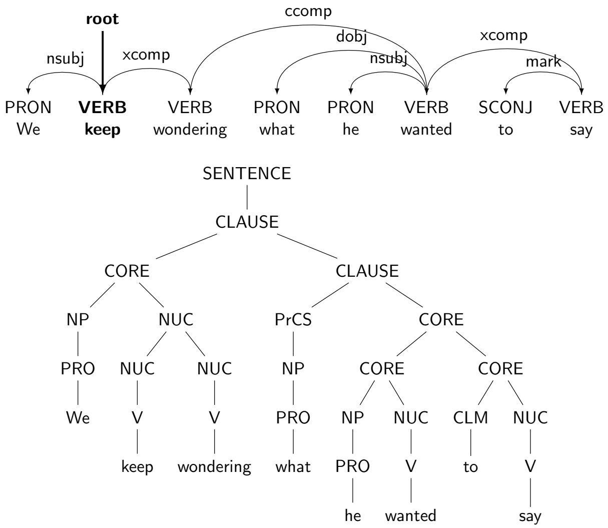
Conversion with Special Rules



Conversion with Special Rules



Conversion with Special Rules



Outline

Introduction

UD to RRG Conversion

Impact on Annotation Effort

Conclusion

Impact on Annotation Effort

- Compare annotation efforts when using different starting points: output of ud2rrg, output of a statistical parser, 'blank' tree
- Evaluate the effort in terms of number of clicks or drag and drops (create / delete node, update label, reattach subtree) in the graphical interface
- Standard measures for tree similarity: tree editing distance (TED), EVALB
- bottom-up replugging (BURP): novel algorithm computing tree similarity consistent with our annotation interface (cost of 1 for reattaching a whole subtree)
- Data (rrgparbank): 4 languages from MULTEXT-East dataset [Erjavec, 2017] (en, de, fr, ru), parsed with UDpipe2 [Straka, 2018]

Evaluating the Annotation Effort

We would like to answer the following questions:

- How does the addition of new composition rules impact the annotation effort?
- How much does ud2rrg reduce the annotation effort compared to:
  - starting the annotations from scratch
  - using the output of a statistical parser as starting point
- How does ud2rrg perform compared to similar tools

Evolution of the Annotation Effort

- General composition rules: apply to all languages (universal dependencies and POS tags)
- New language → add special rules when new constructions appear in the annotated data
- The annotation effort decreases progressively as the annotated data grows
- Example: performance of ud2rrg on Russian data (4 635 sentences) at different development steps

Timestamp	nTED	nBURP	LF1	failed
#1	0.53	0.66	61.02	1 100
#2	0.49	0.57	64.09	773
#3	0.44	0.47	68.75	355
#4	0.33	0.33	72.51	221
#5	0.22	0.20	79.96	0

Comparison with Annotations from Scratch

- Evaluation of the difference of effort when using a ud2rrg output as a starting point or not
- Baseline: starting from a tree where all words are attached below the root

language		de	fr	ru	fa
nBURP	baseline	1.24	1.22	1.18	1.16
LF1		6.56	8.97	7.64	9.14
nBURP	ud2rrg	0.18	0.21	0.20	0.30
LF1		79.24(926)	79.80(402)	79.96(939)	72.09(211)
# sents (annot.)		5723	2177	4635	1110
∅ len. (annot.)		17.00	12.57	11.76	9.01
failures		9	1	0	37
# sents (entire corpus)		6661	7261	6669	6604

Comparison with Statistical Parsing

- Evaluate the number of annotated trees needed to train a statistical parser [Bladier et al., 2020] which outperforms ud2rrg
- Comparison on English data, using different amounts of training data:

approach	train sz.	failures	nTED	LF1 (exact match)	nBURP
ud2rrg		0	0.34	76.51 (84)	0.21
statist.	500	131	0.42	63.45 (85)	0.63
parser	1 000	1	0.35	70.27 (85)	0.29
	2 000	0	0.27	76.13 (113)	0.21
	3 000	0	0.24	78.73 (133)	0.18
	4 000	0	0.22	80.62 (135)	0.17
	>4 000	0	0.22	80.30 (137)	0.16
# sent.			526		
∅ len.			14.02		

Comparison with Related Work

- [Chiarcos and Fäth, 2019]: RDF/SPARQL-based converter to RRG
- Data: 351 sentences from [Van Valin Jr. and LaPolla, 1997]
- Conversion with ud2rrg without update, after normalization of the data:

nBURP	nTED	LF1	exact matches
0.16	0.18	85.75	15.38%

Outline

Introduction

UD to RRG Conversion

Impact on Annotation Effort

Conclusion

Conclusion

- ud2rrg: conversion tool from dependency trees to RRG structures
- Reduces the annotation effort with minimal need of annotated data
- Language independent general rules + custom rules
- Addition of rules as new constructions appear in the treebank
- When enough annotated data is available ($\sim$2000 sentences), statistical parsing offers better starting points for annotation

Bibliography I

Bladier, T., Waszczuk, J., and Kallmeyer, L. (2020).
Statistical parsing of tree wrapping grammars.
In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 6759–6766, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Chiarcos, C. and Fäth, C. (2019).
Graph-Based Annotation Engineering: Towards a Gold Corpus for Role and Reference Grammar.
In Eskevich, M., de Melo, G., Fäth, C., McCrae, J. P., Buitelaar, P., Chiarcos, C., Klimek, B., and Dojchinovski, M., editors, *2nd Conference on Language, Data and Knowledge (LDK 2019)*, volume 70 of *OpenAccess Series in Informatics*

Bibliography II

(OASIcs), pages 9:1–9:11, Dagstuhl, Germany. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Erjavec, T. (2017).
MULTEXT-East.
In Ide, N. and Pustejovsky, J., editors, *Handbook of Linguistic Annotation*, pages 441–462, Dordrecht. Springer Netherlands.

Nivre, J., de Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajič, J., Manning, C. D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., Tsarfaty, R., and Zeman, D. (2016).
Universal Dependencies v1: A multilingual treebank collection.
In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1659–1666, Portorož, Slovenia. European Language Resources Association (ELRA).

Bibliography III

Nivre, J., de Marneffe, M.-C., Ginter, F., Hajič, J., Manning, C. D., Pyysalo, S., Schuster, S., Tyers, F., and Zeman, D. (2020).
Universal Dependencies v2: An evergrowing multilingual treebank collection.
In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 4034–4043, Marseille, France. European Language Resources Association.

Straka, M. (2018).
UDPipe 2.0 prototype at CoNLL 2018 UD shared task.
In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 197–207, Brussels, Belgium. Association for Computational Linguistics.

Bibliography IV

Van Valin Jr., R. D. (2005).
Exploring the Syntax-Semantics Interface.
Cambridge University Press.

Van Valin Jr., R. D. and LaPolla, R. J. (1997).
Syntax: Structure, meaning and function.
Cambridge University Press.