

# Harmonization of Dependency Trees

Daniel Zeman, Rudolf Rosa

May 9, 2024



EUROPEAN UNION  
European Structural and Investment Fund  
Operational Programme Research,  
Development and Education

Charles University  
Faculty of Mathematics and Physics  
Institute of Formal and Applied Linguistics

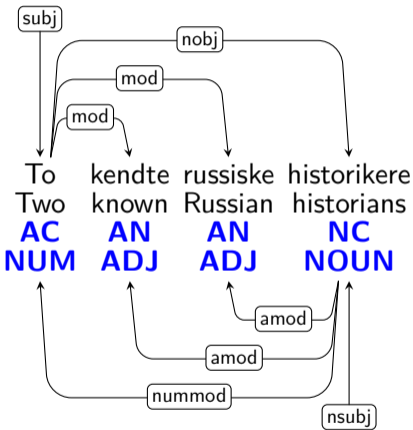


unless otherwise stated

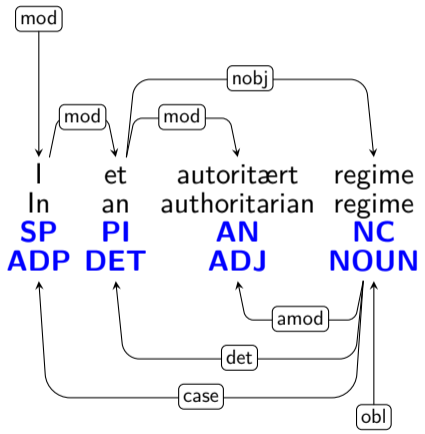
# Harmonization of Dependency Tree Style

- ▶ Universal Dependencies
  - ▶ Solved problem?
  - ▶ Almost. But not completely.
- ▶ Transformations sometimes needed for learnability experiments
- ▶ There are still non-UD treebanks that need to be converted
- ▶ **SUD** (Surface-syntactic Universal Dependencies)
- ▶ Transformations trivial?
  - ▶ Not necessarily...

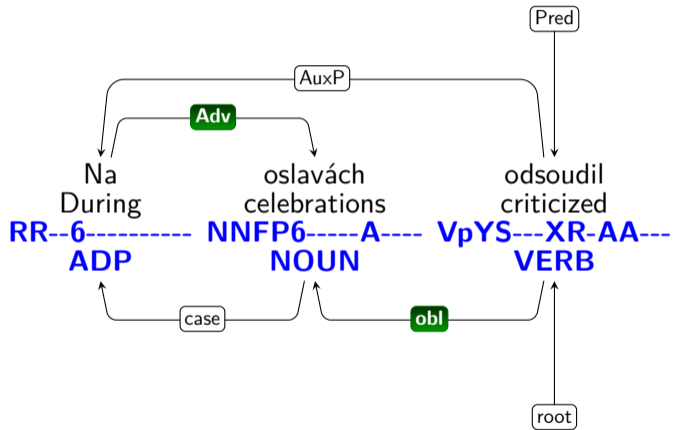
# Danish Dependency Treebank



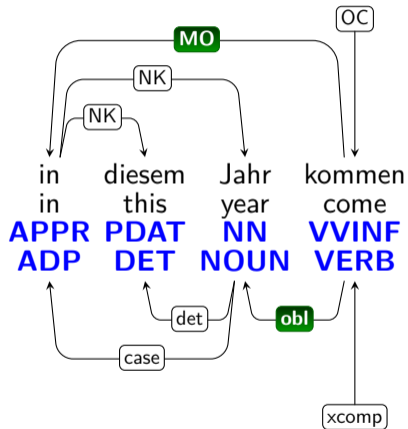
# Danish Dependency Treebank



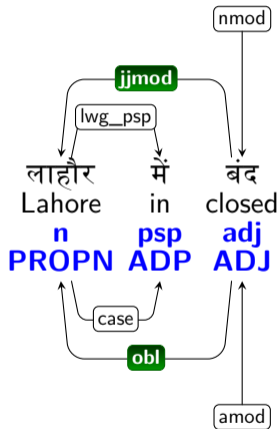
# Adpositions



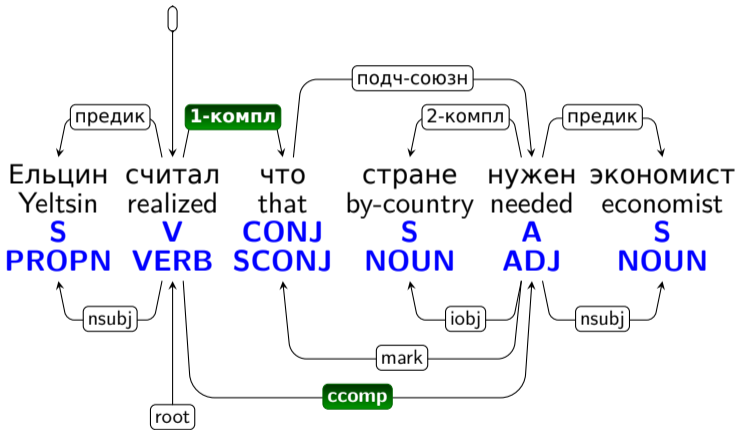
# Adpositions



# Adpositions

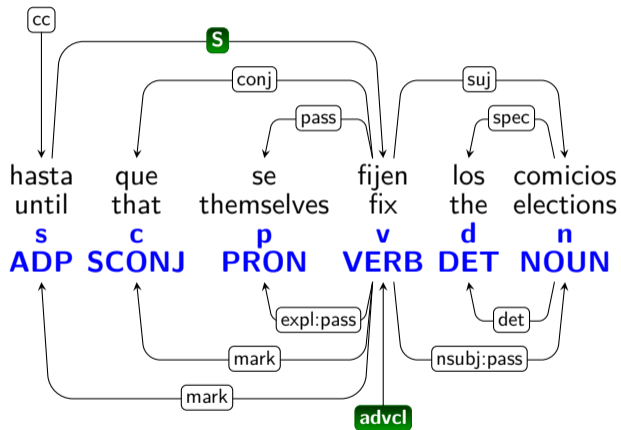


# Subordinate Clauses

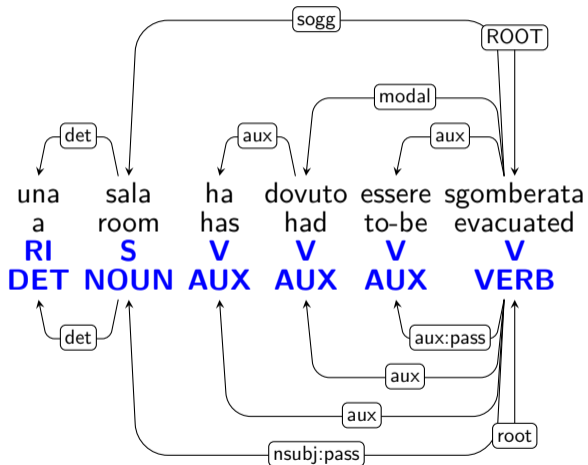




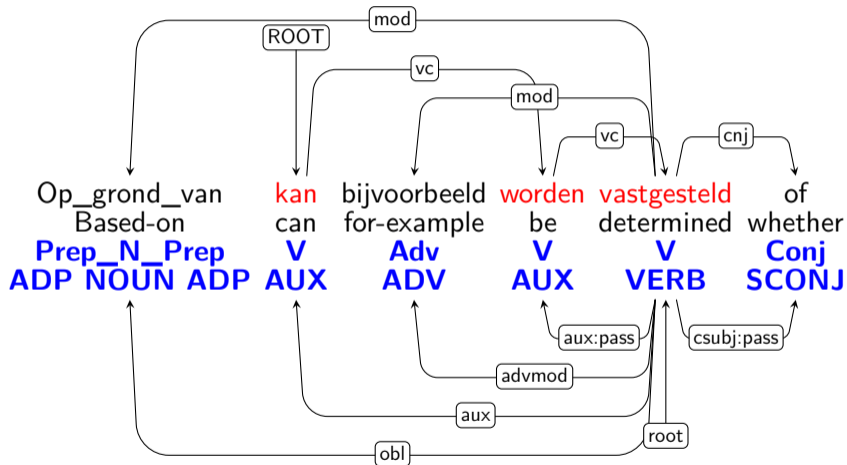
# Subordinate Clauses



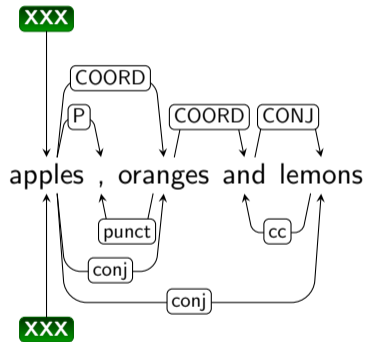
# Verb Groups



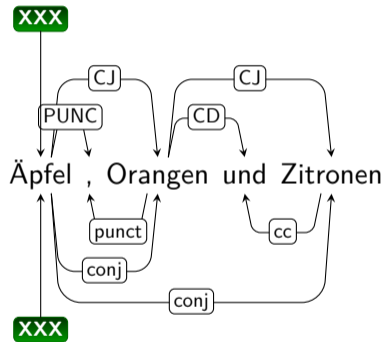
# Verb Groups



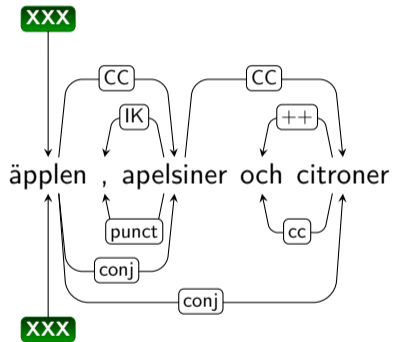
# Coordination: Mel'čuk [en]



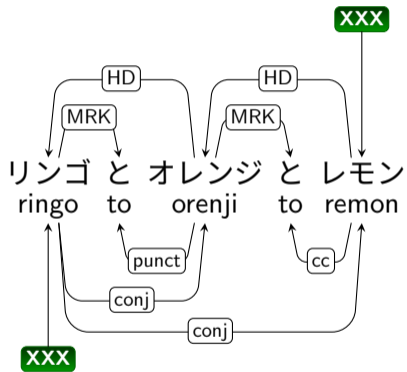
# Coordination: Mel'čuk [de]



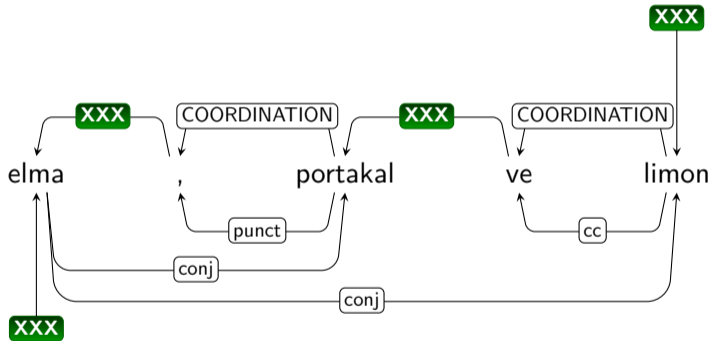
# Coordination: Mel'čuk [sv]



# Coordination: Mel'čuk [ja]

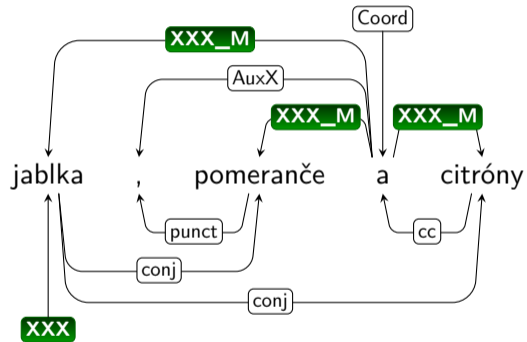


# Coordination: Mel'čuk [tr]

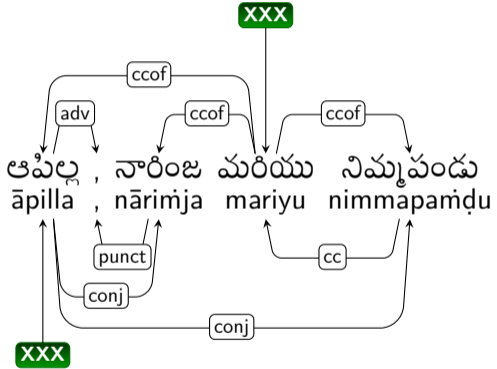




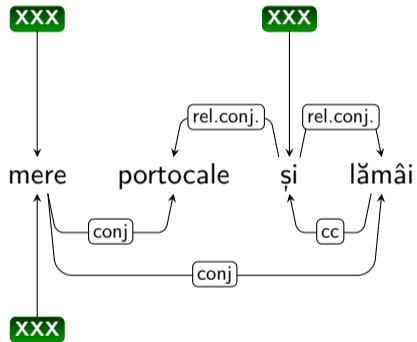
# Coordination: Prague [cs]



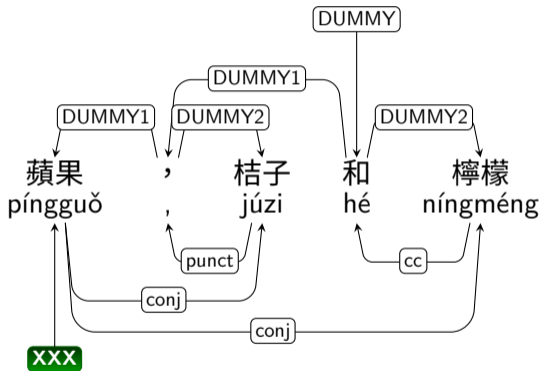
# Coordination: Prague [te]



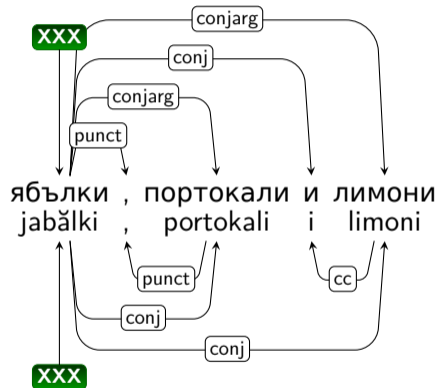
# Coordination: Half-Prague [ro]



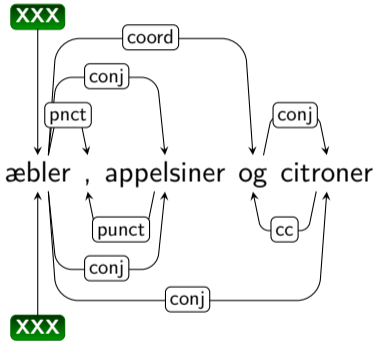
# Coordination: Obligatorily Nested Prague [zh]



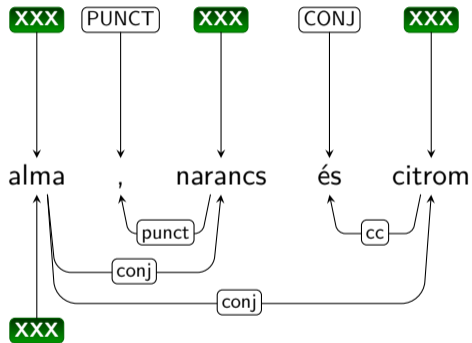
# Coordination: Stanford [bg]



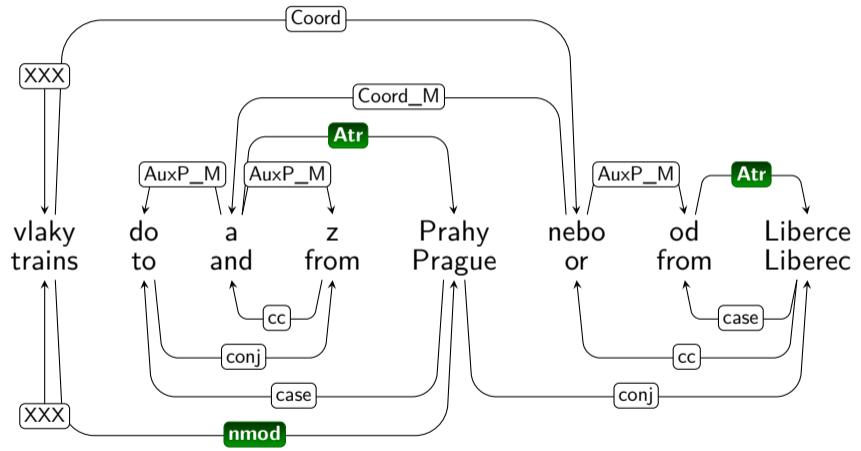
# Coordination: Half-Stanford [da]



# Coordination: Tesnière [hu]

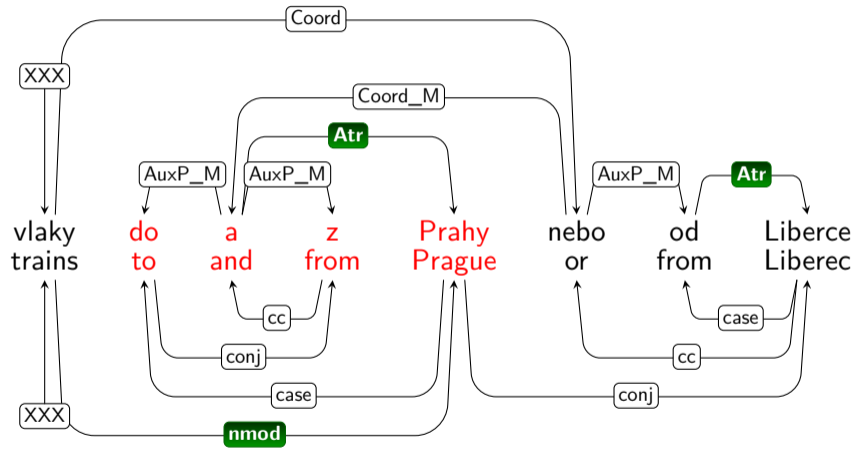


# Interaction of Structures in Prague

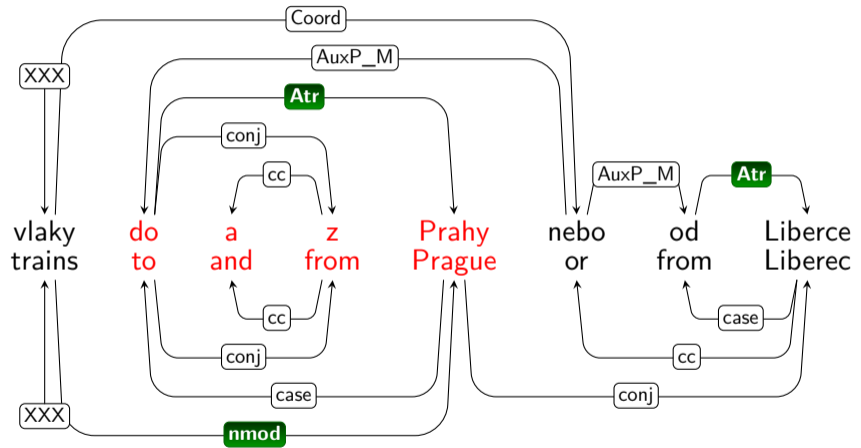




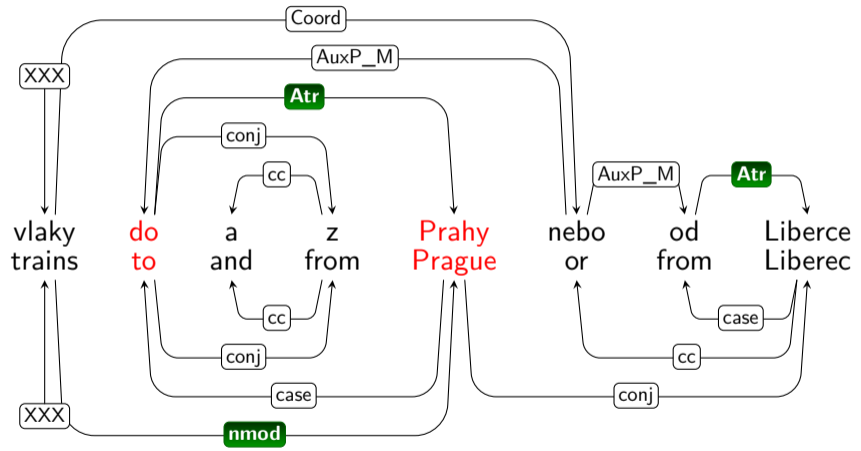
# Interaction of Structures in Prague



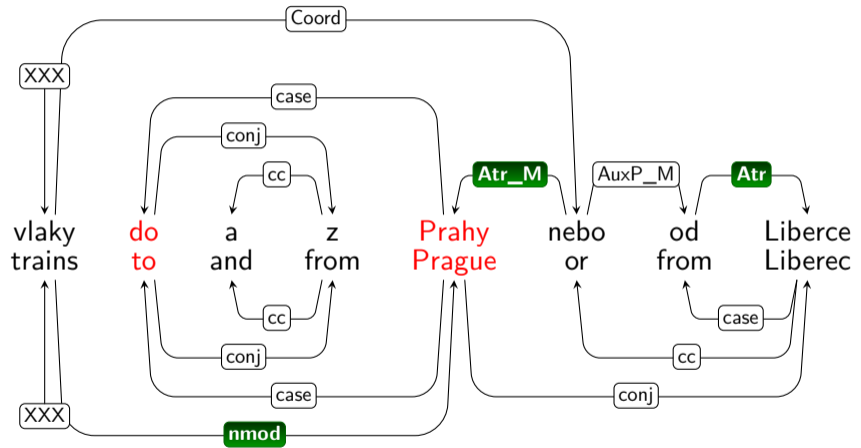
# Interaction of Structures in Prague



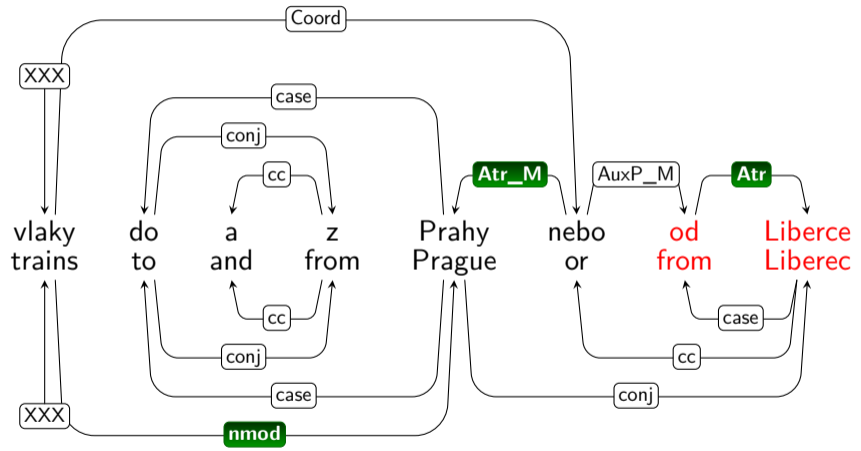
# Interaction of Structures in Prague



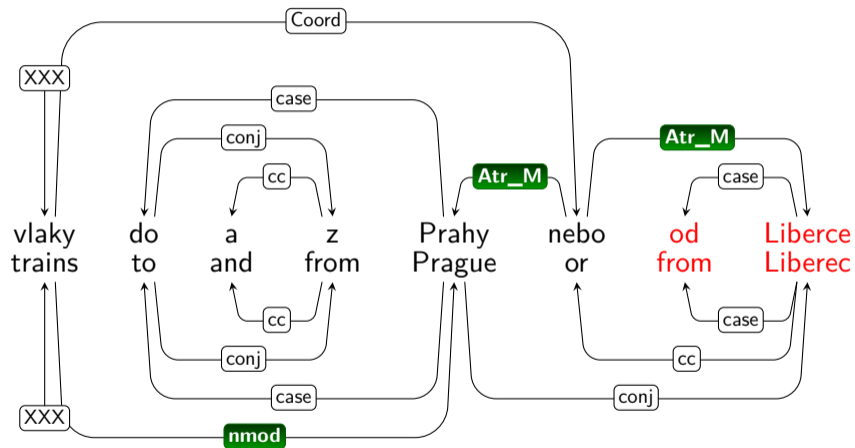
# Interaction of Structures in Prague



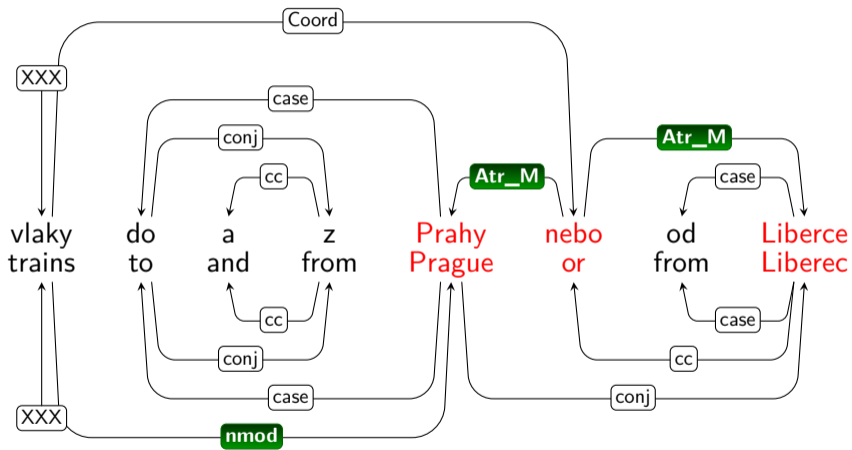
# Interaction of Structures in Prague



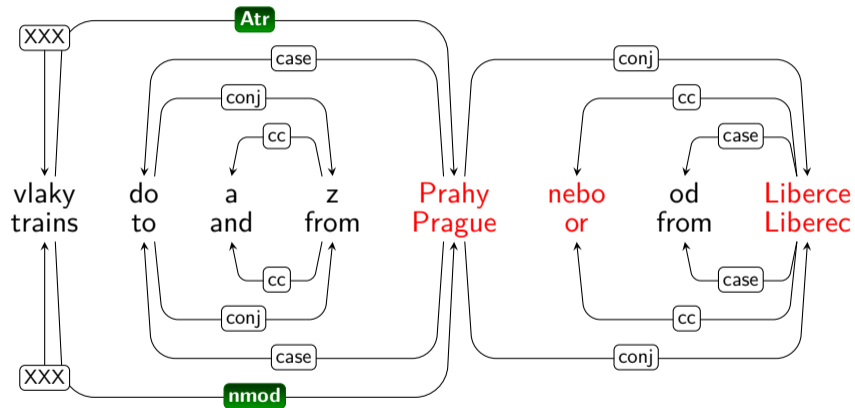
# Interaction of Structures in Prague



# Interaction of Structures in Prague

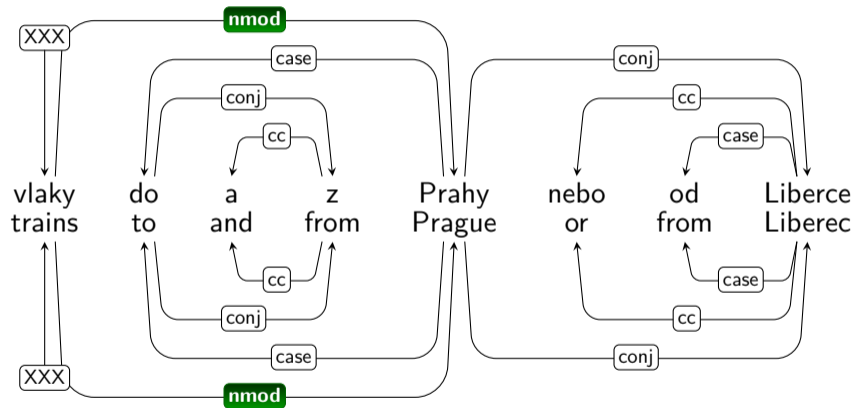


# Interaction of Structures in Prague

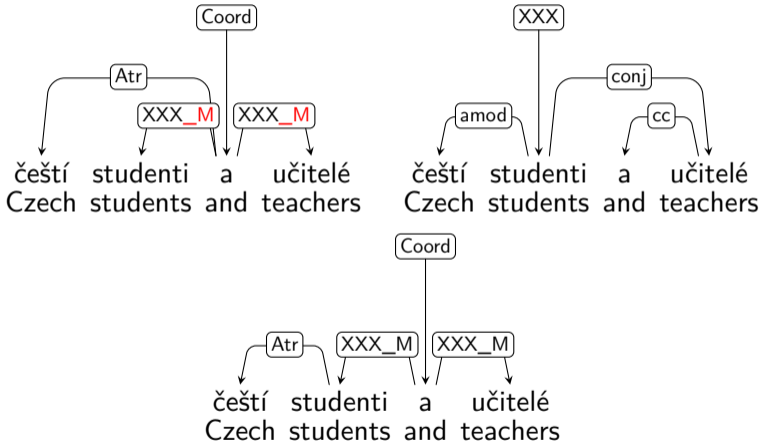




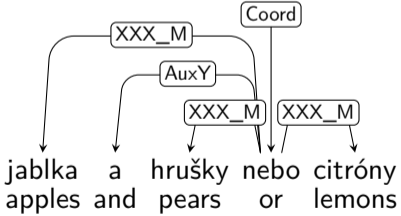
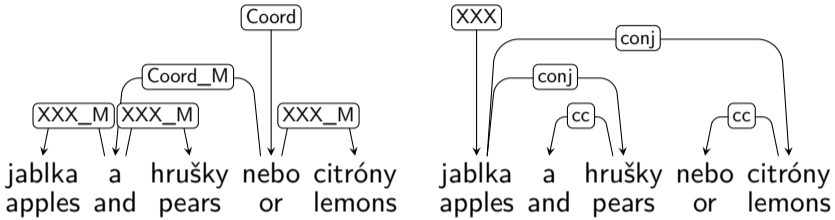
# Interaction of Structures in Prague



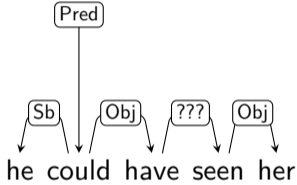
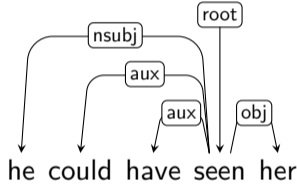
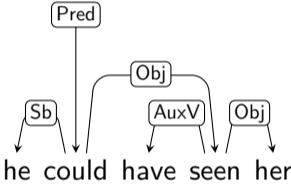
# Round-Trip: Shared Dependents



# Round-Trip: Nested Coordination



# Round-Trip: Auxiliary Verbs



- ▶ Udapi

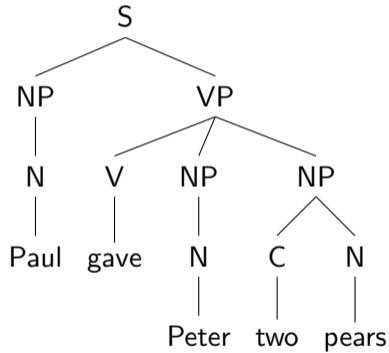
- ▶ <http://udapi.github.io/> (main page)

- ▶ <https://ufal.mff.cuni.cz/~zeman/vyuka/deptreebanks/NPFL075-working-with-UD.pdf>  
(tutorial)

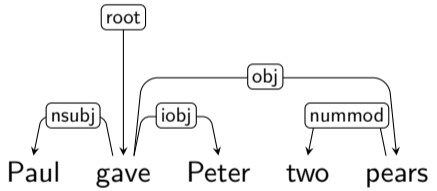
- ▶ Grew

- ▶ <https://grew.fr/>

# Phrase vs. Dependency Trees



↔



# Phrase vs. Dependency Trees

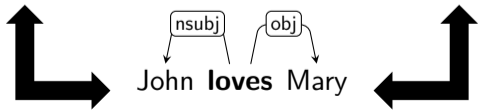
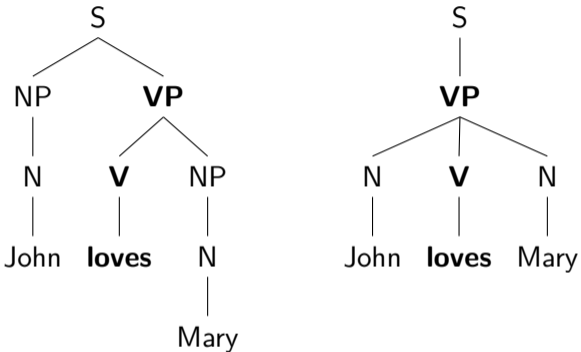
- ▶ Phrase trees
  - ▶ Usually do not mark the **head**
  - ▶ May not mark the **function** of the constituent in the superordinate constituent

# Phrase vs. Dependency Trees

- ▶ Phrase trees
  - ▶ Usually do not mark the **head**
  - ▶ May not mark the **function** of the constituent in the superordinate constituent
- ▶ Dependency trees
  - ▶ Do not show **nonterminals** (phrase types)
    - ▶ Nor any other phrase-level features
  - ▶ Do not show “how the sentence is generated” (order, recursion, **proximity** of constituents)

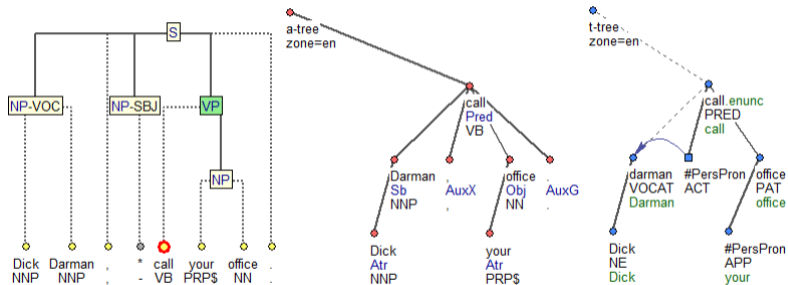


# Example

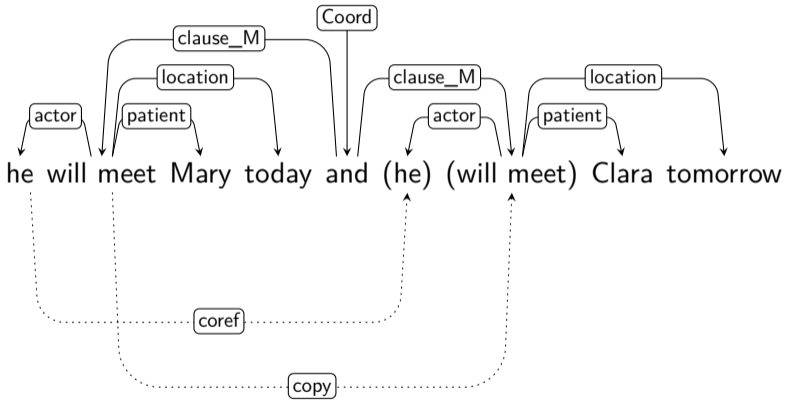


# Prague Tectogrammatical Layer

*Dick Darman, call your office.*



# Prague Tectogrammatical Layer: Ellipsis



# Basic vs. Enhanced Universal Dependencies

