Harmonization of Dependency Trees

Daniel Zeman, Rudolf Rosa

May 12, 2022
• 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…
Two known Russian historians

Danish Dependency Treebank
In an authoritarian regime
During celebrations criticized

Adpositions

Na

During

oslavách

celebrations

odsoudil

criticized

Pred

Adv

AuxP

case

obl

root

RR--6--------
ADP

NNFP6------A----
NOUN

VpYS---XR-AA---
VERB

4/33
Adpositions

in diesem Jahr kommen

APP R

ADP

DET

NOUN

VERB

MO

OC

NK

NK

det

case

obl

xcomp
Lahore in closed
Ельцин считал что стране нужен экономист
Yeltsin realized that by-country needed economist
hasta que se fijen los comicios

until that they fix the elections
Una sala ha dovuto essere sgomberata.

**Verb Groups**

**Diagram:**
- **Verb:** essere sgomberata
- **Auxiliary:** ha dovuto
- **Determiner:** una
- **Noun:** sala
- **Root:** essere
- **Modal:** ha
- **Auxiliary:** dovuto
- **Determiner:** una
- **Noun:** sala
- **Auxiliary:** essere
- **Verb:** sgomberata

**Dependency Structure:**
- **Det:** una
- **Noun:** sala
- **Aux:** ha
- **AUX:** dovuto
- **Verbs:** essere, sgomberata
Op grond van kan bijvoorbeeld worden vastgesteld of
Based-on can for-example be determined whether
apples, oranges and lemons
Äpfel, Orangen und Zitronen
ändran, apelsiner och citroner
Coordination: Mel’čuk [ja]

リンゴ と オレンジ と レモン
ringo to orenji to remon

conj

punct

cc

conj

HD MRK

HD MRK

XXX

XXX

XXX
Coordination: Mel’čuk [tr]
jablka, pomeranče a citróny
Coordination: Prague [te]

అపిలల్, నారింజ మరియు నిమమ్పండు

apilla, nārimja mariyu nimmampāmpu
Coordination: Half-Prague [ro]
苹果，橘子和柠檬

pingguo, juzi he ningmeng
ябълки, портокали и лимони
jabălki, portokali i limoni
æbler, appelsiner og citroner
Coordination: Tesnière [hu]

alma, narancs és citrom
Interaction of Structures in Prague

vlaky do a z Prahy nebo od Liberec
trains to and from Prague or from Liberec
Interaction of Structures in Prague

Trains to and from Prague or from Liberec.
Interaction of Structures in Prague

trains to and from Prague or from Liberec

Coord

AuxP_M

Atr

conj

cc

case

nmod
Interaction of Structures in Prague

Trains to and from Prague or from Liberec.
interaction of structures in Prague

trains to and from Prague or Liberec
Interaction of Structures in Prague
Interaction of Structures in Prague

trains to and from Prague or from Liberec
Interaction of Structures in Prague

Trains to and from Prague or from Liberec.
Interaction of Structures in Prague

vlaky do a z Prahy nebo od Liberec
trains to and from Prague or from Liberec
Interaction of Structures in Prague

vlaky do a z Prahy nebo od Liberce
trains to and from Prague or from Liberec
Round-Trip: Shared Dependents

[Diagram showing the structure of sentences in Czech, with labels for Atr, Coord, amod, conj, and cc, translating to Czech students and teachers.]

- čeští studenti a učitelé → Czech students and teachers
- čeští studenti a učitelé → Czech students and teachers
Round-Trip: Nested Coordination

- jablka a hrušky nebo citróny
  - apples and pears or lemons
- jablka a hrušky nebo citróny
  - apples and pears or lemons
- jablka a hrušky nebo citróny
  - apples and pears or lemons
Round-Trip: Auxiliary Verbs

he could have seen her

he could have seen her

he could have seen her
Tree-Rewriting Toolkits

- **Udapi**

- **Grew**
  - [https://grew.fr/](https://grew.fr/)
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
Dick Darman, call your office.
he will meet Mary today and (he) (will meet) Clara tomorrow
he will meet Mary today and (he) (will) (meet) Clara tomorrow