



Prague Dependency Treebank: Annotation of Surface Syntax

Markéta Lopatková

Institute of Formal and Applied Linguistics, MFF UK

lopatkova@ufal.mff.cuni.cz

afun

(main clause)

AuxV



auxiliary verbs

morphology

John is preparing his homework.

Byl by přišel včas.

John has been attending the school since 2011.

Bude chodit včas.

syntactic core
of (surface) sentence

copula

Pred

light verbs

modal verbs
phase verbs

Pred

semantic core

full/lexical/autosemantic verbs

Pred

Copula

- a word used to link the subject of a sentence with a predicate
- semantically empty, but syntactic center

*Mary and John **are** my friends.*

*The sky **was** blue.*

*I **am** taller than most people.*

*The birds and the beasts **were** there.*



Copula



- a word used to link the subject of a sentence with a predicate
- semantically empty, but syntactic center
 - Mary and John **are** my friends.*
 - The sky **was** blue.*
 - I **am** taller than most people.*
 - The birds and the beasts **were** there.*
- it expresses
 - identity ... *The Morning Star is the Evening Star.*
 - membership of a class or a subset relationship ... *She was a nurse.*
 - express some property, relation or position, permanent or temporary:
 - The trees are green.*
 - I am your boss.*
 - John was in England.*
 - The children are confused.*
- main/principal copula typically (equivalent of) "to be"
 - plus "verbs equivalent in function to the principal copula, the verb "be"*
 - no clear cut between copular verbs and full verbs

Copula



Randolph Quirk: *A comprehensive grammar of the English Language*

“A verb is said to have a copular complementation when it is followed by a *subject complement* or a *predicational adjunct*, and when this complement cannot be dropped without changing the meaning of the verb.

The verb in such a clause is a copular (or linking) verb, and is equivalent in function to the principal copula, the verb *be*.”

current copulas (typically stative):

- (i) **be** (friendly) (my friend)
- (ii) appear (happy) (the only solution)
- feel (annoyed) (a fool)
- look (pretty) (a fine day)
- seem (very restless)(a genius)
- smell (sweet)
- sound (surprised) (a reasonable idea)
- taste (bitter)
- (iii) remain (uncertain) (good friends)
- keep (silent)
- stay (motionless)

resulting copulas:

- (iv) **become** (older) (an expert)
- come (true)
- end up (happy) (her slave)
- get (ready)
- go (sour)
- grow (tired)
- prove (rather useful)(his equal)
- turn (cold) (traitor)
- turn out (fortunate) (a success)
- wind up (drunk) (a millionaire)

Copula in Czech



Czech ling. tradition:

Sponová slovesa se využívají v přísudku slovesně-jmenném (přísudek jmenný se sponou), kde jsou doplněna jménem.

Mezi sponová slovesa patří **být**, **bývat**, **stát se**, **stávat se** a v záporu **nebýt**, **nebývat**, **nestat se** a **nestávat se**.

Marta je nemocná.

Simona bývala učitelkou.

Petr se stal lékařem.

(jinde též mít, zůstat, zdát se)

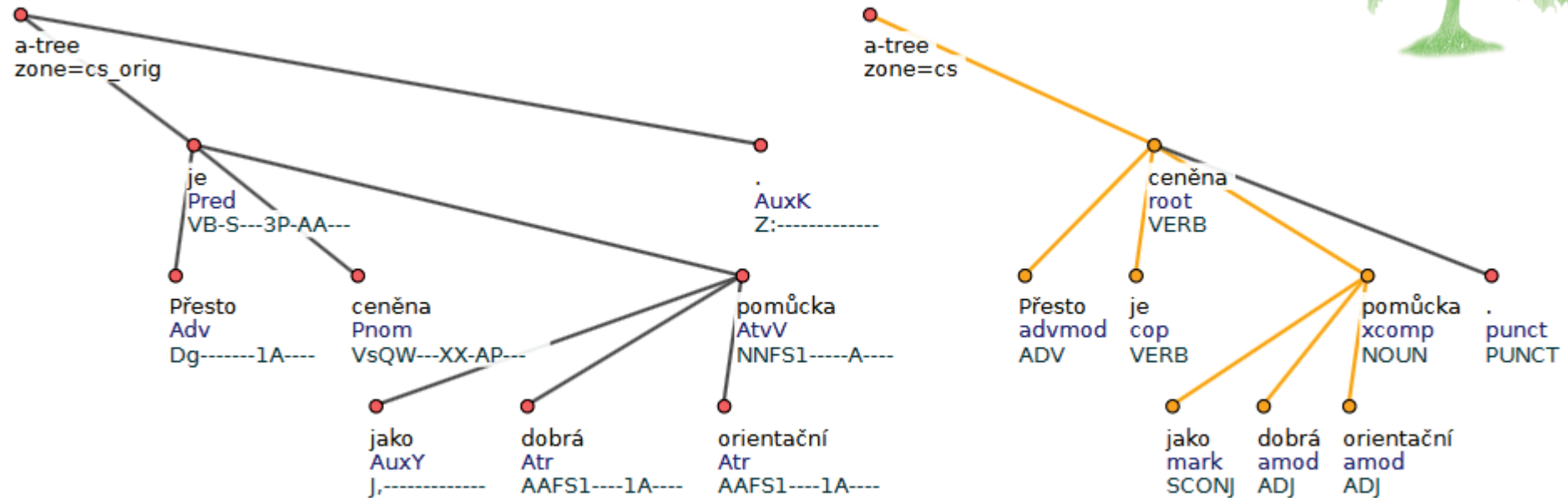
Copula in data (Czech and English)



LINDAT/CLARIN ... provides technical background and assistance
... to share, create and modernise their *tools* and *data* used for research in linguistics
... also provides an open digital repository and archive open to all academics

<http://lindat.mff.cuni.cz/services/pmltq/#!/home>

Copula in HamleDT for Czech (PDT vs. UDs, version 1)



<http://lindat.mff.cuni.cz/services/pmltq/#!/home>

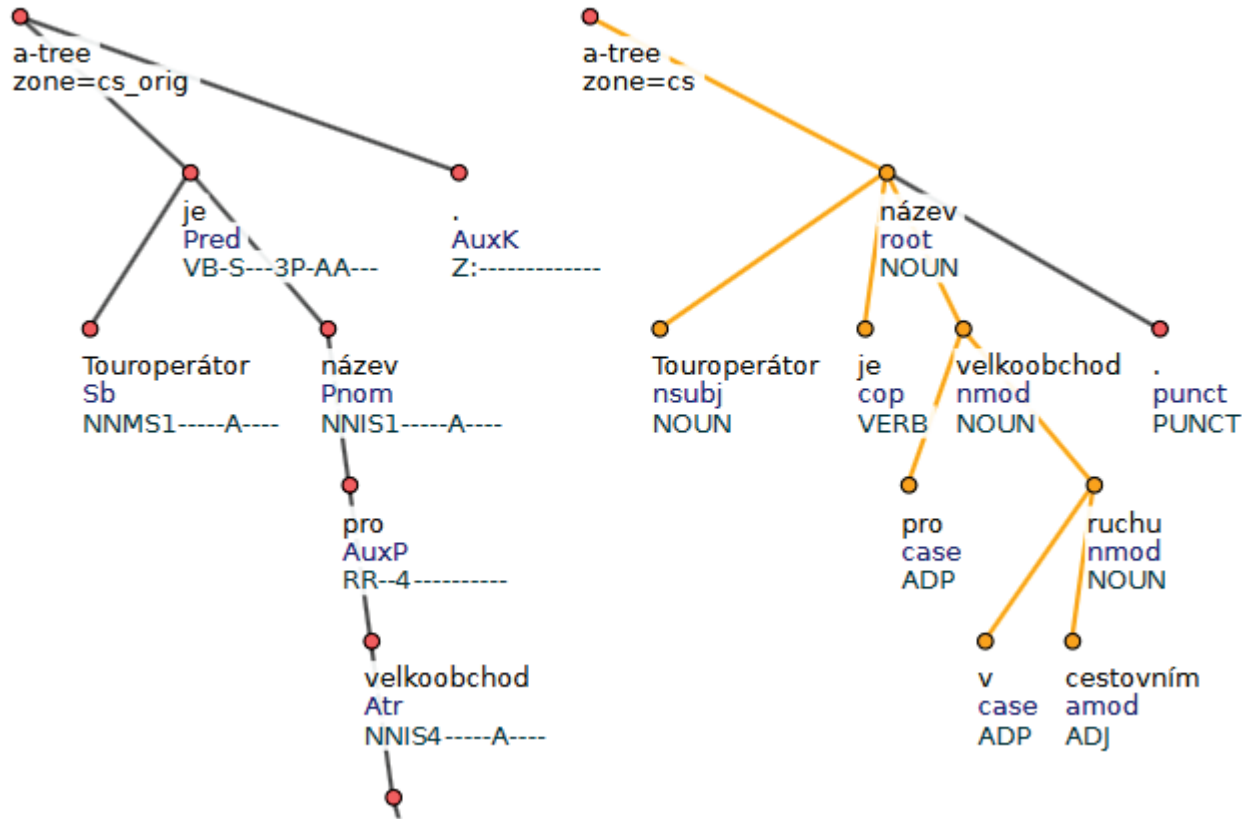
PML-TQ query:

```
a-node $pred :=
[id ~ "prague", depth() = "1",
 1x a-node [ afun = "Pnom",
 ],
]
>> for $pred.lemma give $1, count()
>> sort by $2 desc
```


Copula in HamleDT for Czech (PDT vs. UDs, version 1)

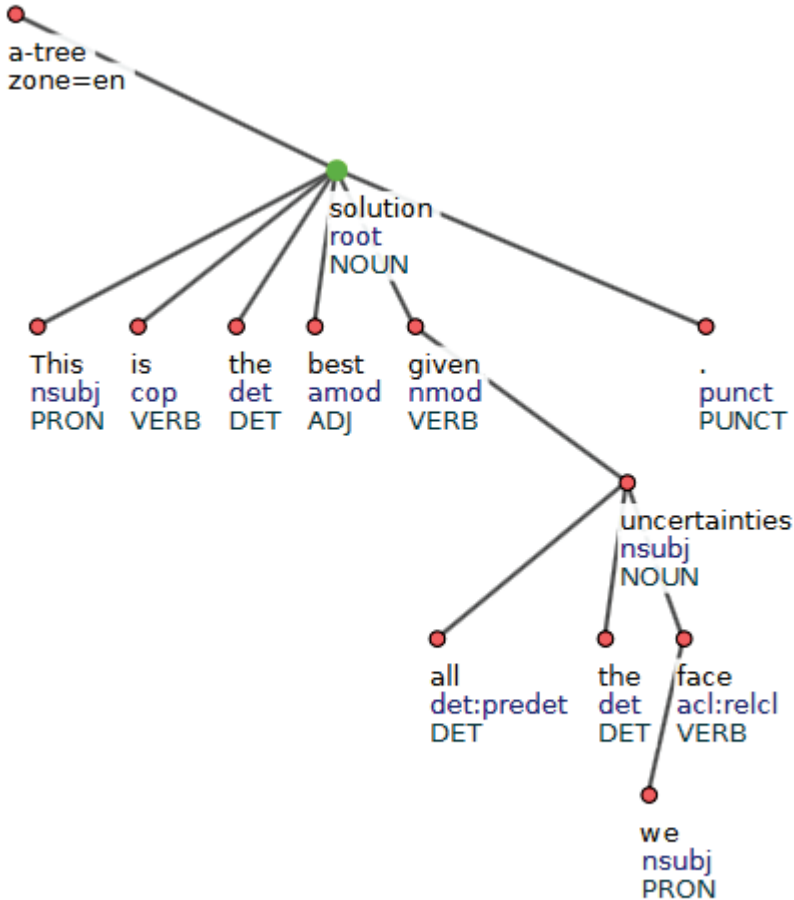


<http://lindat.mff.cuni.cz/services/pmltq/#!/home>



| lemma | count |
|----------------|-------|
| <i>být</i> | 7974 |
| <i>bývat</i> | 41 |
| : | 37 |
| a | 10 |
| - | 5 |
| či | 2 |
| = | 2 |
| <i>bývavat</i> | 2 |
| respektive | 1 |
| , | 1 |
| totiž | 1 |

Copula for English in UDs, version 1



<http://lindat.mff.cuni.cz/services/pmltq/#!/home>

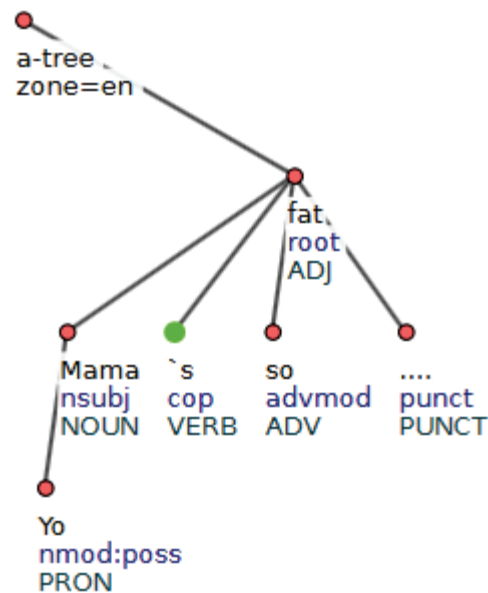
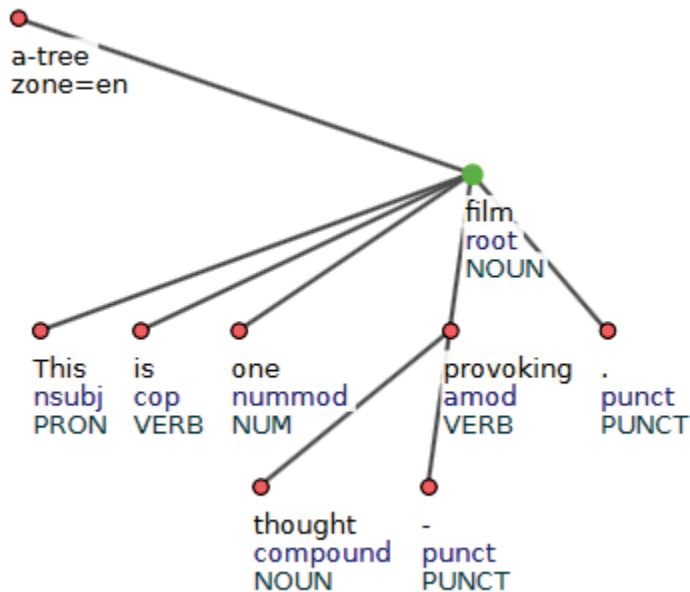
PML-TQ query:

```
a-node $cop := [ deprel = "cop" ]
>> for $cop.lemma give $1, count()
>> sort by $2 desc
```



Copula for English in UD, version 1

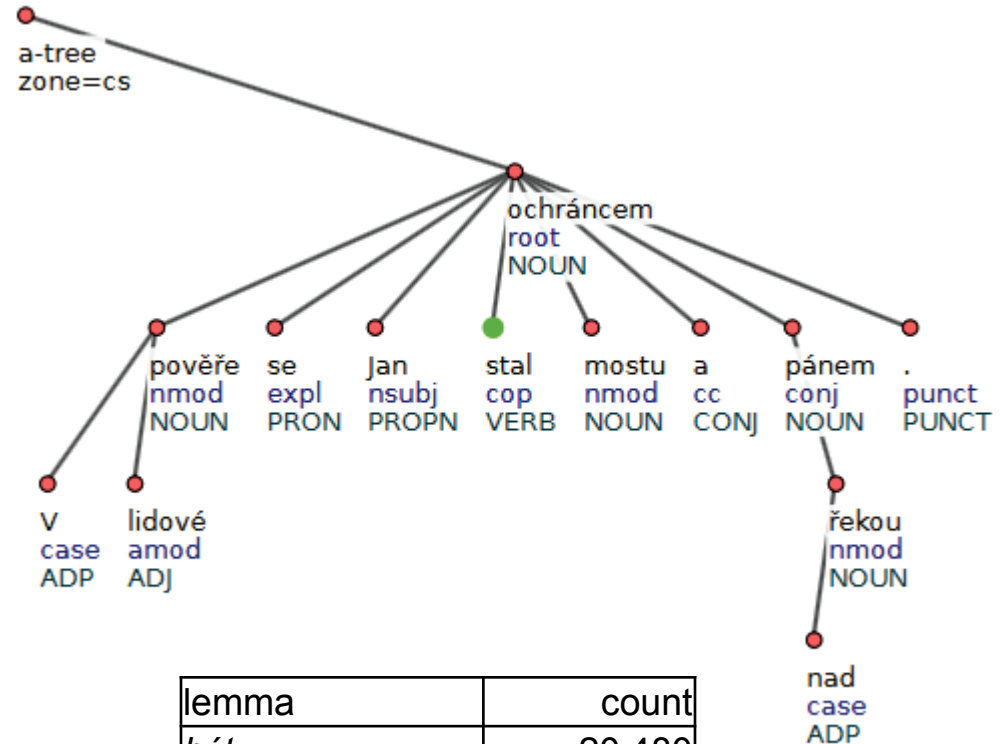
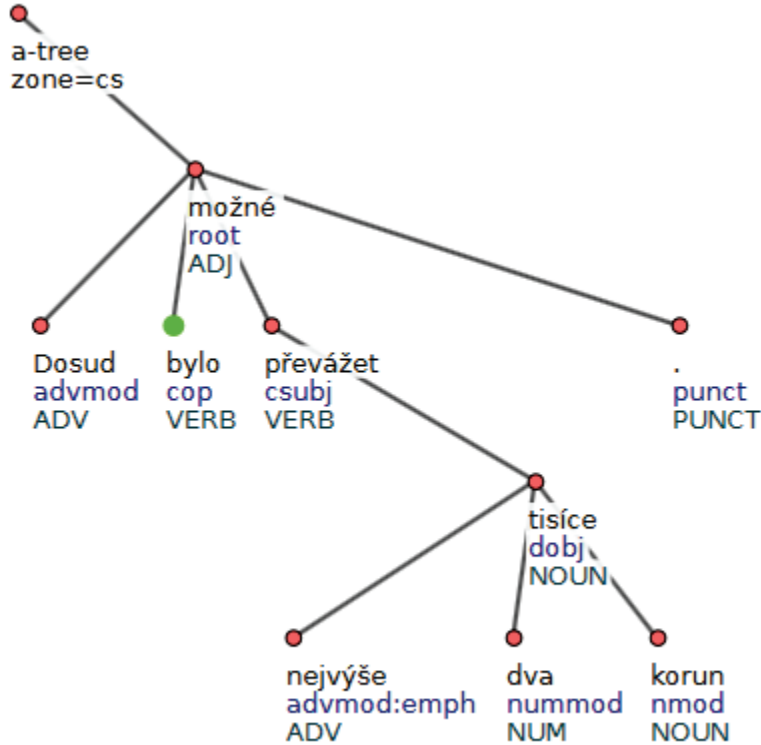
<http://lindat.mff.cuni.cz/services/pmltq/#!/home>



| copula | count | example |
|--------|-------|--|
| be | 5598 | This is one thought-provoking film. |
| 's | 8 | "Yo Mama's so fat" |
| ' | 1 | "so, enemy, what's your point?" |
| b | 1 | "o and the cheaper the better (we are trying to save money now, sooo..thatd b helpful) thanks" |



Copula for Czech in UDs, version 1



| lemma | count |
|----------------|--------|
| <i>být</i> | 20 480 |
| <i>bývat</i> | 110 |
| <i>stát</i> | 3 |
| <i>bývávat</i> | 3 |
| <i>moci</i> | 1 |
| <i>mít</i> | 1 |

"Doplněk" ~ verbal attribute





"Doplněk" ~ verbal attribute

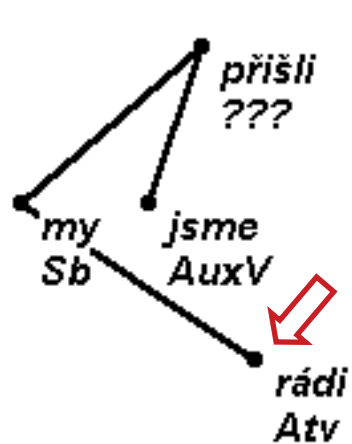
Complement (verbal attribute) ... analytical functions: **Atv, AtvV**

- modifies two sentence members, verb and noun

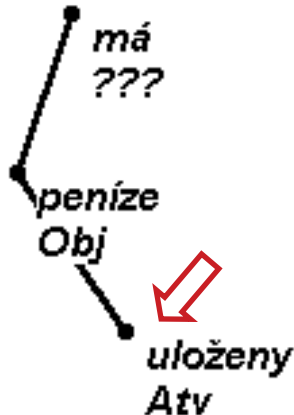
chlapec ležel nemocen, viděl ho nemocného

[boy – lay – ill]

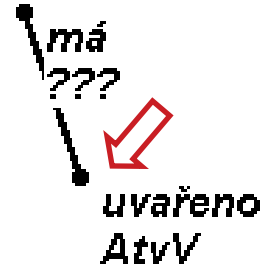
[he – saw – him – (being) ill]



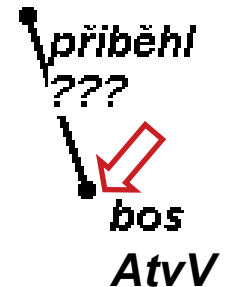
we - Aux - came - glad
we were glad to come



money - he-has - deposit
he has put the money on a deposit

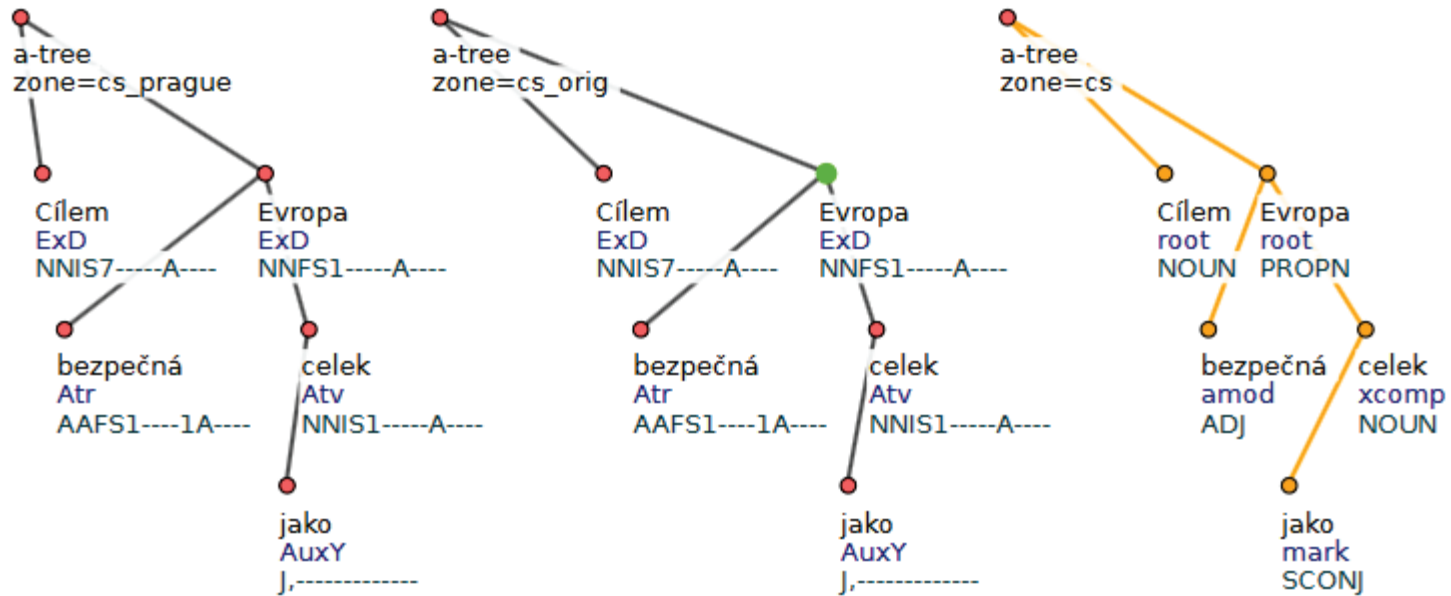


she-has - cooked
he has done cooking

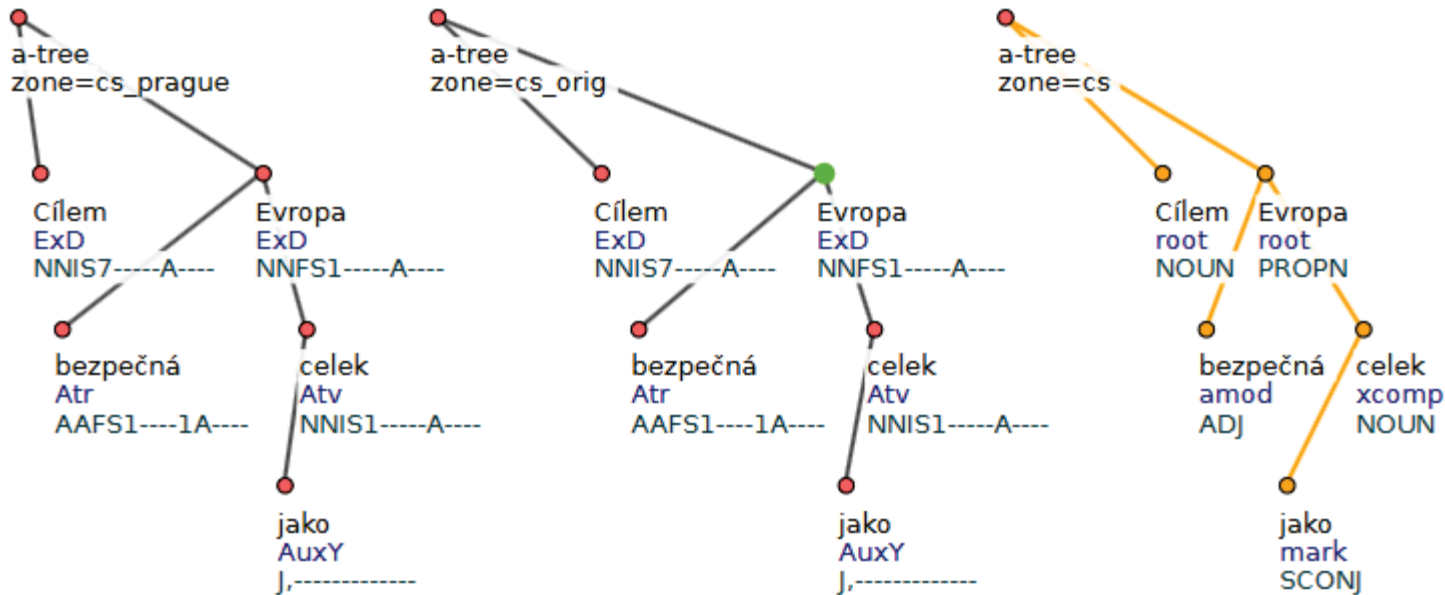


he-arrived - barefooted

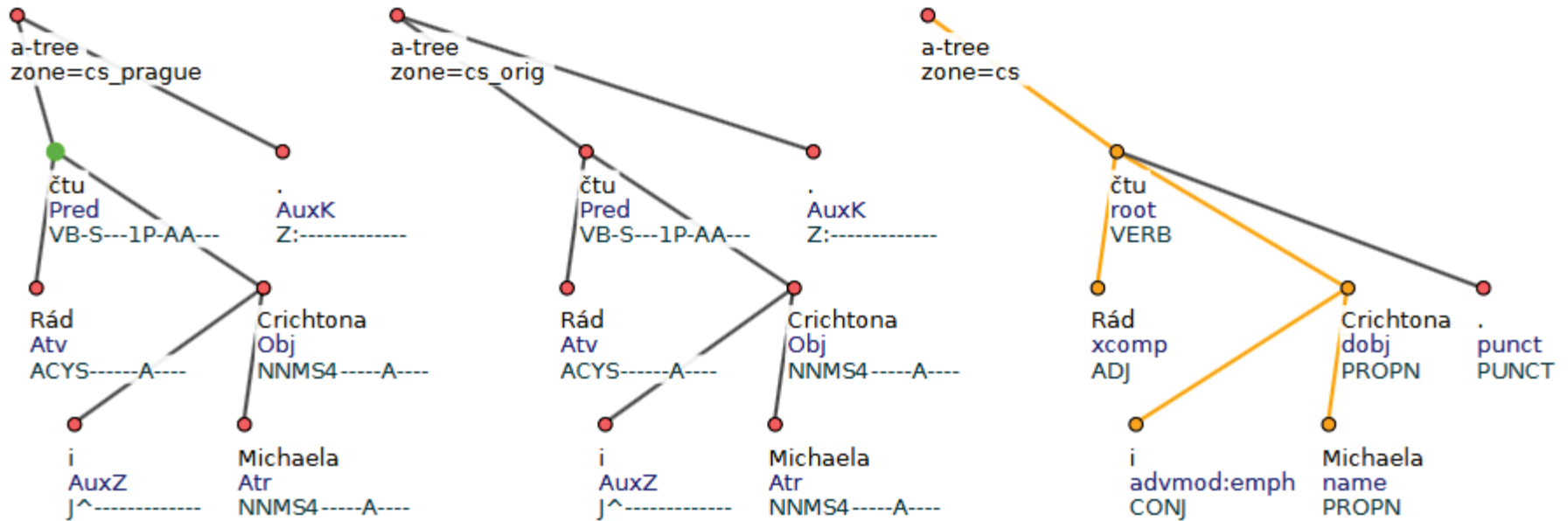
"Doplněk" in HamleDT for Czech (PDT vs. UDs, version 1)



"Doplněk" in HamleDT for Czech (PDT vs. UDs, version 1)



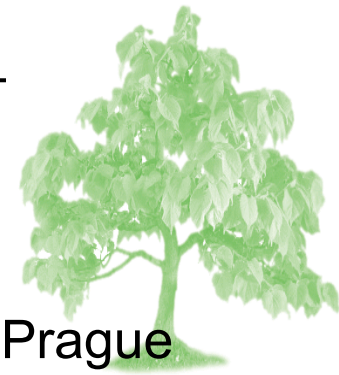
"Doplněk" in HamleDT for Czech (PDT vs. UDs, version 1)



Complex sentences



- **main clause** (= independent clause) vs. **non-main clause**
- **relative clause** ... syntactic function: Attribute
 - The Ferrari [which I can't afford _].*
 - I imagined the house where I want to live.*
 - starting with a relative pronoun or relative adverb
- **dependent clause** ... either modify the main clause of a sentence or serve as a component of it (plus recursivity!!)
 - ... syntactic function: Object or Adverbial
 - starting with a subordinating conjunction (that, if, whether, in order that)
I know (that) he is here.
 - starting with an interrogative pronoun or relative adverb
I know who said that.
I know where he wanted to go.



References

- Hajič, J. (1998) Building a Syntactically Annotated Corpus: The Prague Dependency Treebank". In E. Hajičová (ed.): *Issues of Valency and Meaning. Studies in Honour of Jarmila Panevová*, Karolinum, Charles University Press, Prague, Republic, pp. 106-132
- Manual for Analytical Annotation
<http://ufal.mff.cuni.cz/pdt2.0/doc/manuals/en/a-layer/html/index.html>
- PML-TQ and repository
<http://lindat.mff.cuni.cz/services/pmltq/#!/home>