

Anotace rematizátorů a discourse connectives v PEDT

Kateřina Veselá

Horní Mísečky 11. 2. 2009

Obsah

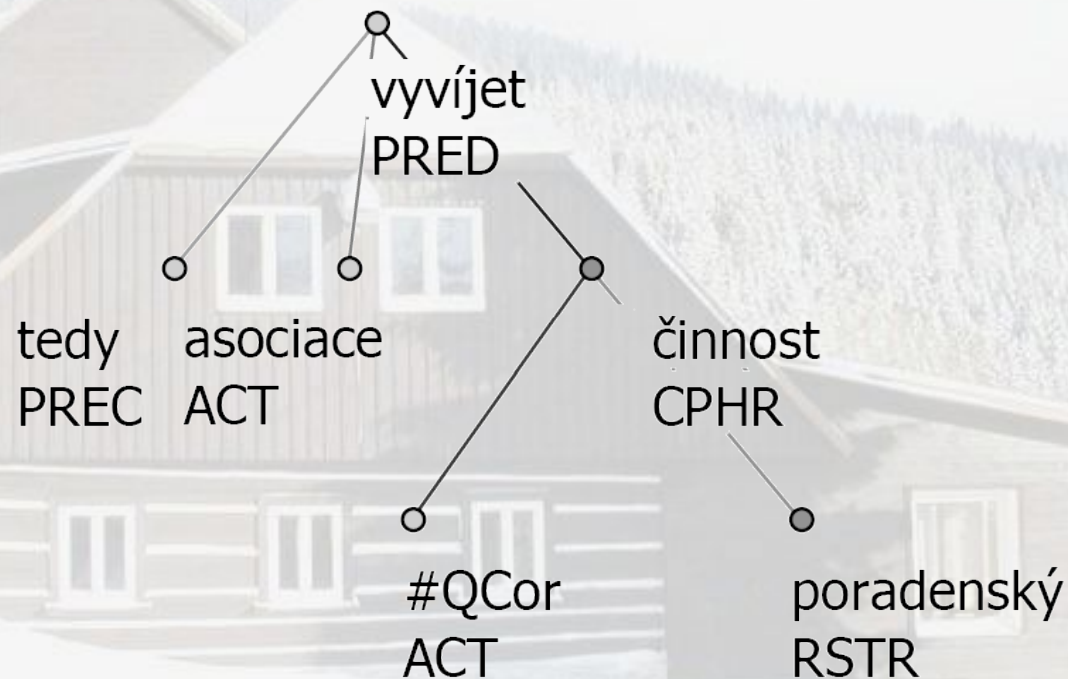
- Úvod
- Anotace rematizátorů v PEDT
- Specifické problémy v anglické anotaci
- Co jsme se naučili o rematizátorech na angličtině
- Co s tím dál

Úvod: rematizátory – co, proč a jak

- Počátek společného plánu anotace rematizátorů: K. Veselá a M. Mikulová při druhém průchodu PDT 2.0
- Rematizátory spojují obě roviny, řídí se TFA a mají vliv na strukturu
- S rematizátory úzce souvisí i anotace spojovacích výrazů, PREC (discourse connectives) a ATT
- PEDT: pokračování anotace na základě výstupů z PDT, nové problémy

Vliv rematizátorů na strukturu

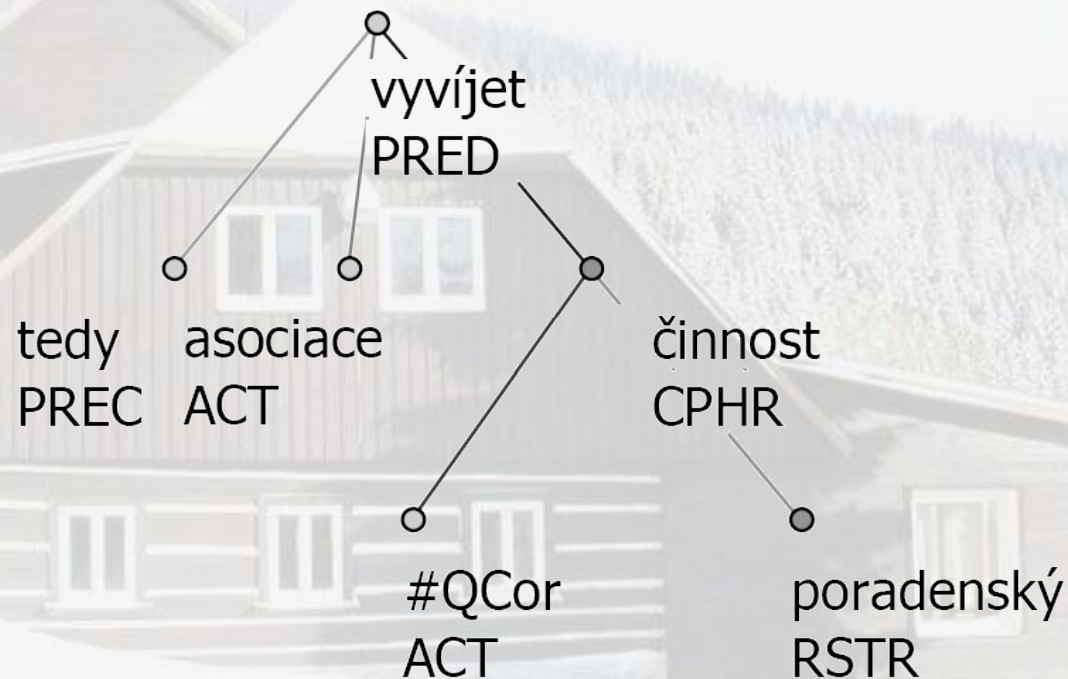
- Rematizátor visí na té hladině stromu, na jejíž uzly má dosah



Bude tedy asociace vyvíjet i poradenskou činnost?

Vliv rematizátorů na strukturu

- Rematizátor visí na té hladině stromu, na jejíž uzly má dosah



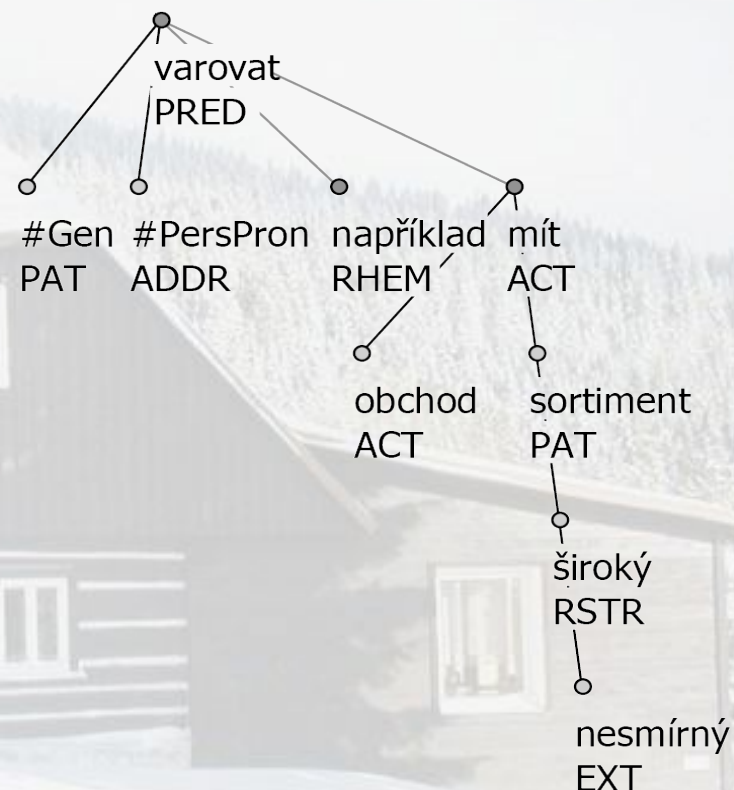
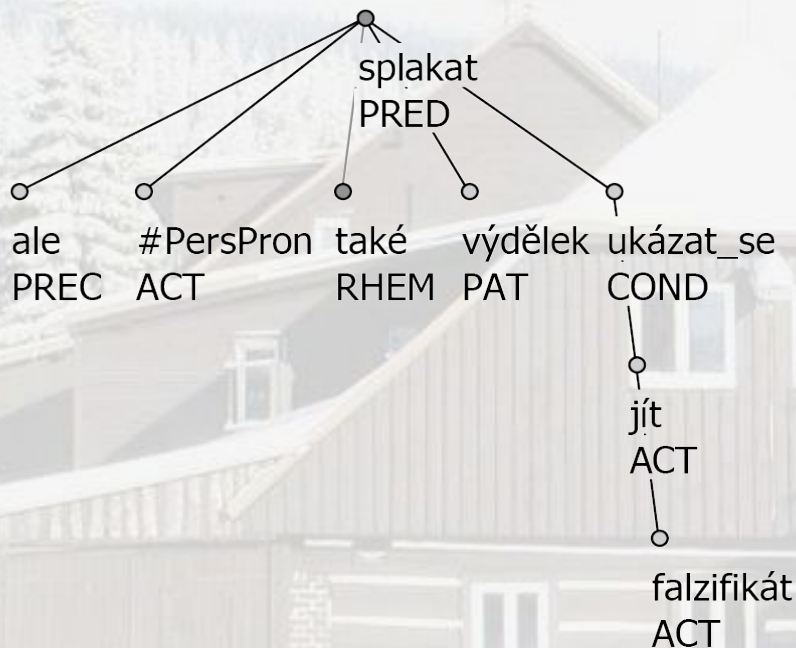
Bude tedy asociace vyvíjet i poradenskou činnost?

Problémy s rematizátory v PDT

- Zavěšování rematizátorů ve stromě (hlavní predikáty, vedlejší větvy)
Řešení: Striktní dodržování pravidla levé sestry s výjimkou hlavního predikátu
- Rematizátory v koordinacích
Řešení: Zdeňkův funktor CM
- Homonymie s PREC, ATT a příslovci
Řešení: Seznamem lemmat s příklady v dodatku

Zavěšení rematizátoru ve struktuře

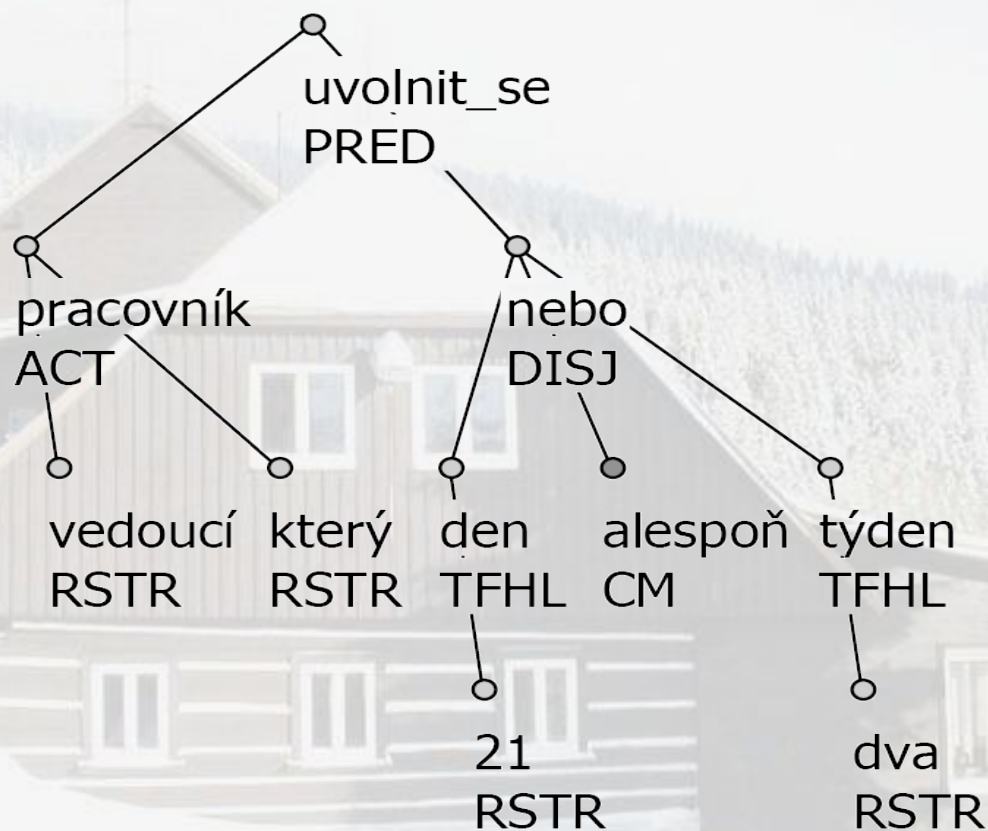
Dosah na hlavní predikát x dosah na vedlejší větu



Ale můžete také splakat nad výdělkem, ukáže-li se, že jde o falzifikát.

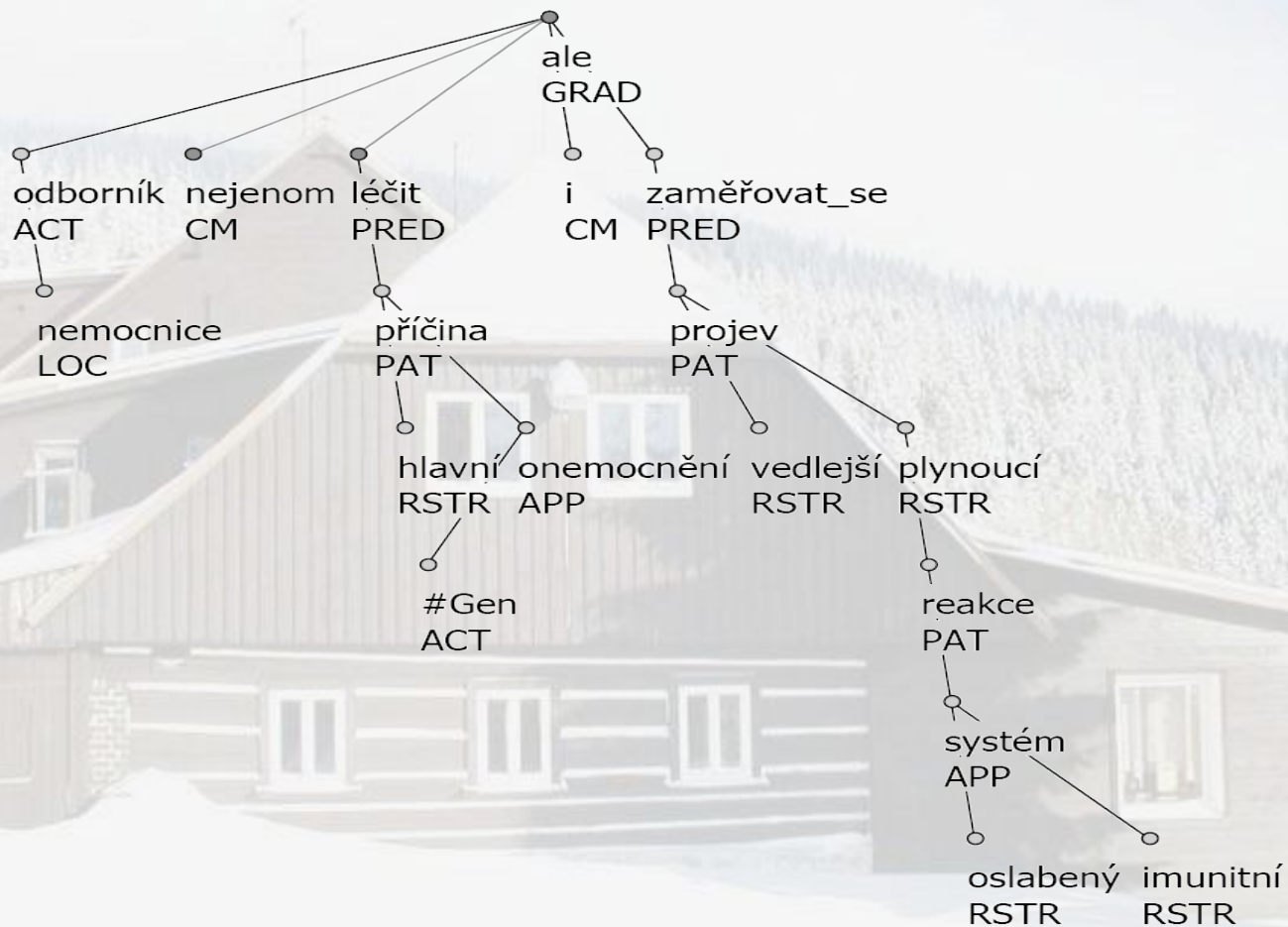
Varovat by ho mohlo například to, že má obchod nesmírně široký sortiment.

Koordinace členská



Málokterý vedoucí pracovník se může uvolnit na 21 dní, nebo alespoň na dva týdny.

Koordinace větná



Odborníci v nemocnicích léčí nejenom hlavní příčiny onemocnění, ale zaměřují se i na vedlejší projevy plynoucí z reakcí oslabeného imunitního systému.

Homonymie

Homonymie s PREC, ATT a příslovci

- Řeší se hlavně podle dosahu – když má dosah na jednu konkrétní věc, kterou rozvíjí, je to příslovečné určení, když na celou větu, je to PREC, když na část věty, jde o rematizátor

Hlavně.PREC na to nezapomeň. / Nezapomeň hlavně.RHEM na to.

Zvlášť.RHEM tohle je velké. / Tohle je zvlášť.EXT velké.

Plán anotace v PEDT

- Srpen 2008 S. Cinková a J. Toman vymysleli „jevovou anotaci“ - anotaci jednotlivých jevů v korpusu v už anotovaných i neanotovaných datech
- První z úloh se týká právě rematizátorů
- Lidé: Od září se jí věnují L. Mladová, K. Čermáková a K. Veselá
- Postup: Automatická procedura vyhledá „podezřelé uzly“, které anotátoři kontrolují podle dodatku
- Tempo: anotace se chýlí ke konci, bylo vyhledáno cca 10000 výskytů, rozdělených na 21 sekcí (7 na jednoho)

První anotační „šok“

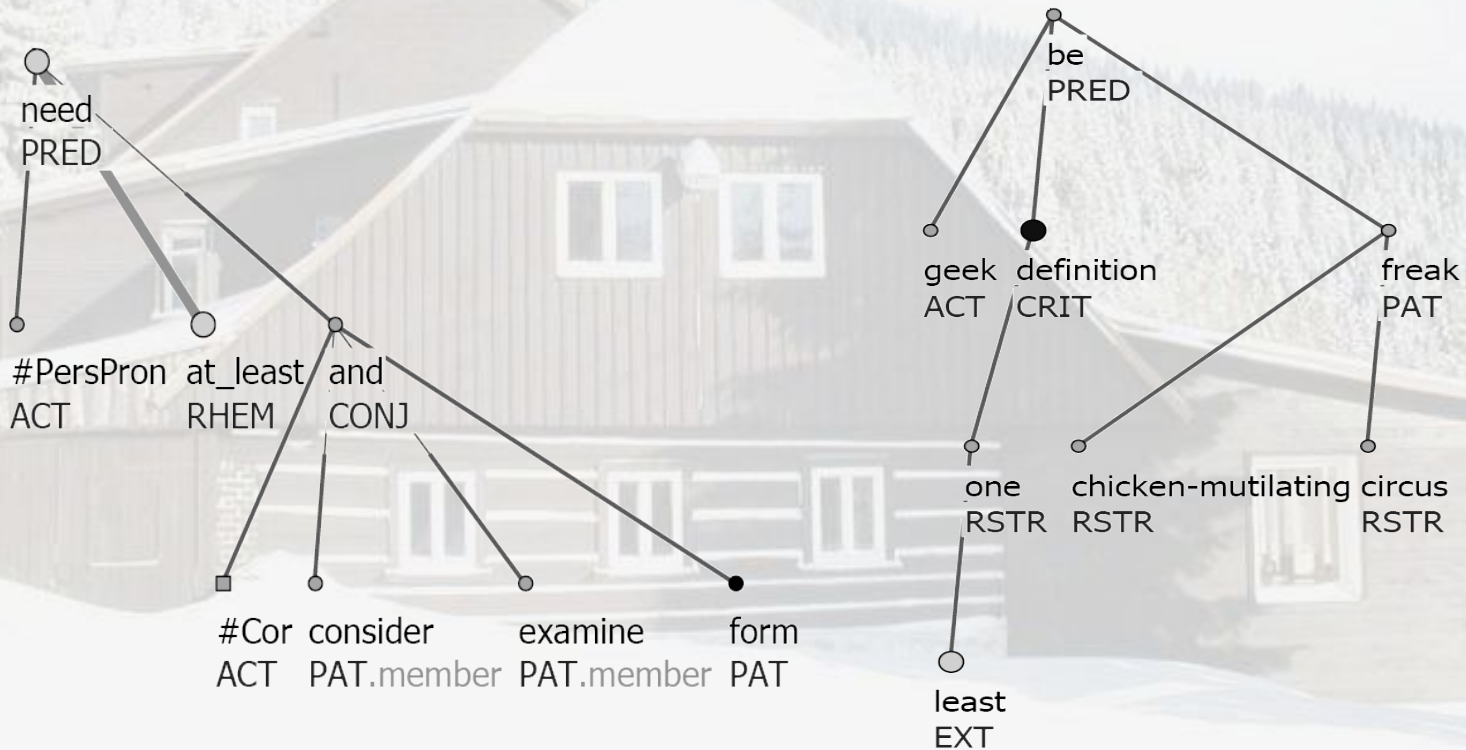
- Anotace rematizátorů bez TFA – práce s kontextem, hloubkový slovosled
- Práce na neanotovaných stromečích – hranice strukturních úprav
- Specifické rysy anglických rematizátorů, na které není připraven manuál
- Neřeší se negace

Specifické jevy a problémy v angličtině

- Složená t_lemmata
- Rematizátory s omezenou volností pozice
- Distantní postavení rematizátoru
- Postpozice rematizátoru
- Víceslovné PREC
- Koordinační struktury

Víceslovná t_lemmata

V angličtině se vyskytují rematizátory s předložkou, např. *at least* - ve funkci rematizátoru/discourse connective zůstává předložka v lemmatu, ve funkci volného doplnění se schová do a/aux.rf



We need at least to consider and examine forms of controlled legalization of drugs.

Geeks, by at least one definition, are chicken-mutilating circus freaks.

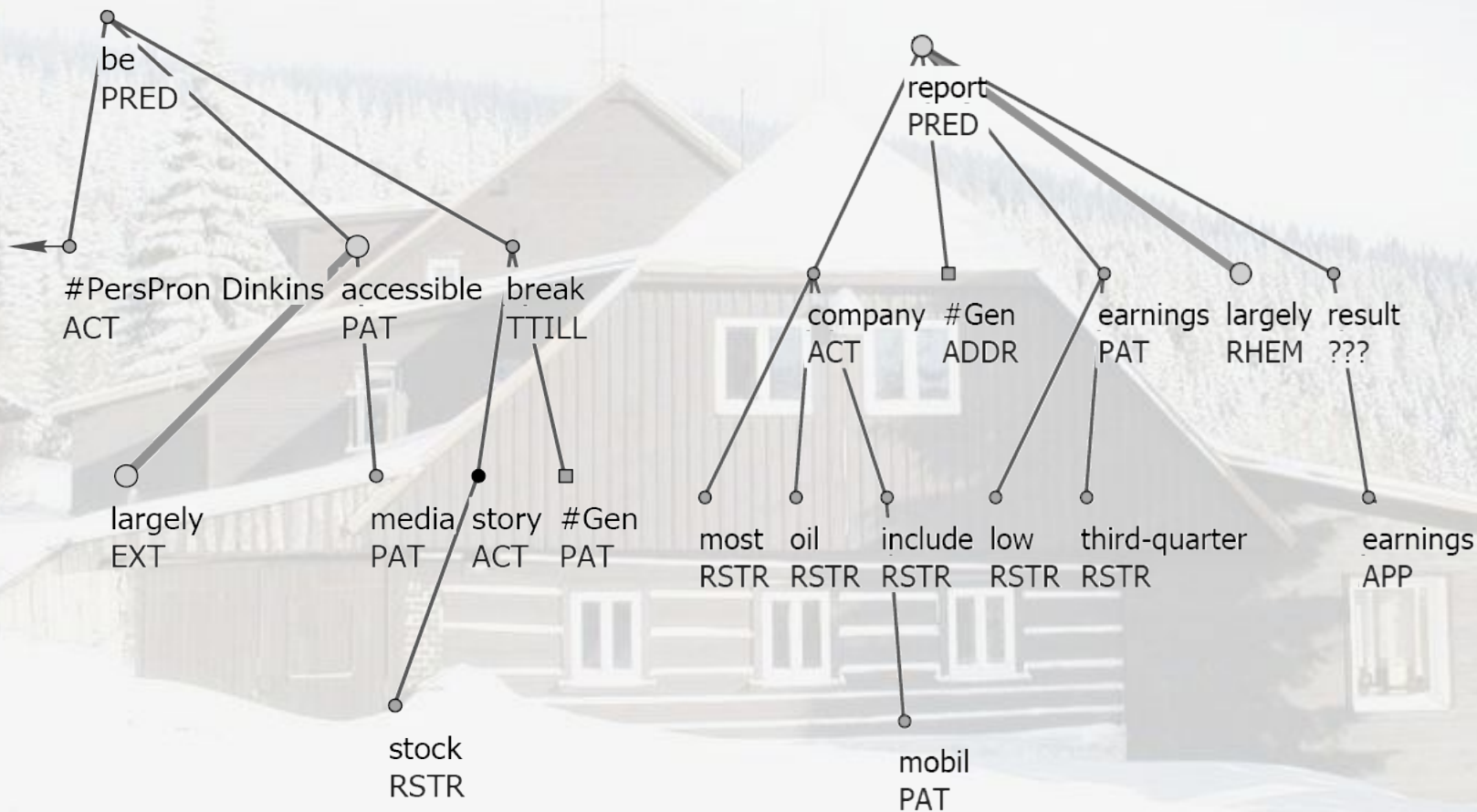
Rematizátory vázané na určité pozice ve větě

Výrazy, které mají menší variabilitu v umístění ve větě a jen v konkrétních pozicích se chovají jako rematizátory – např. *mostly, largely, partly, particularly...*

He has canceled numerous campaign appointments and was largely inaccessible to the media until the stock story broke.

Most oil companies, including Mobil, have been reporting lower third-quarter earnings, largely as a result of lower earnings from chemicals as well as refining and marketing businesses.

Rematizátory vázané na určité pozice ve větě - stromy



He was largely inaccessible to the media until the stock story broke.

Most oil companies, including Mobil, have been reporting lower third-quarter earnings, largely as a result of lower earnings from chemicals as well as refining and marketing businesses.

Distantní postavení rematizátoru

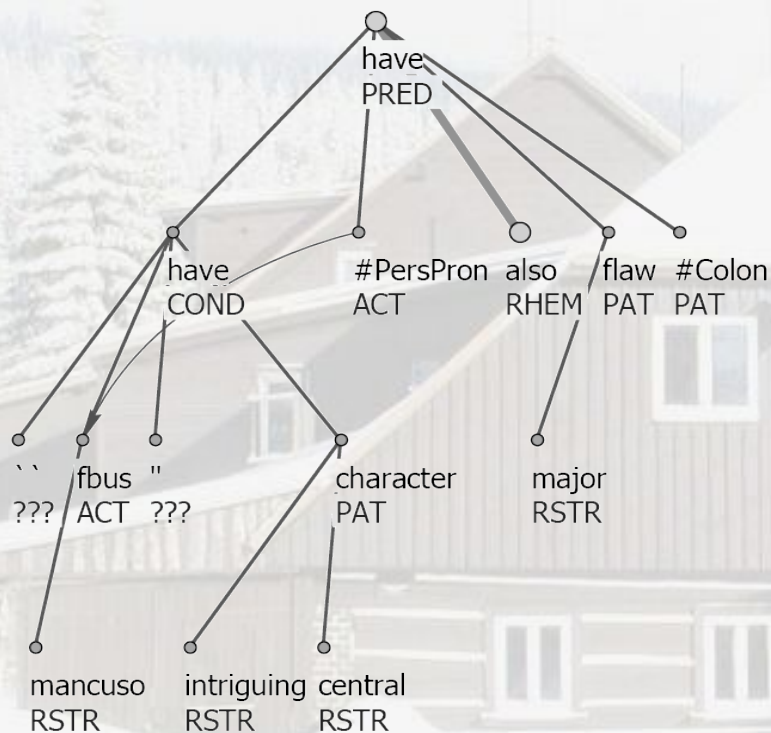
- Vyskytuje se i u rematizátorů s úplně volnou pozicí, jako např. *also, just, only...*
- Anglické rematizátory mají svou „oblíbenou“ pozici ve větě před slovesem

If “Mancuso FBI” has an intriguing central character, it also has a major flaw:

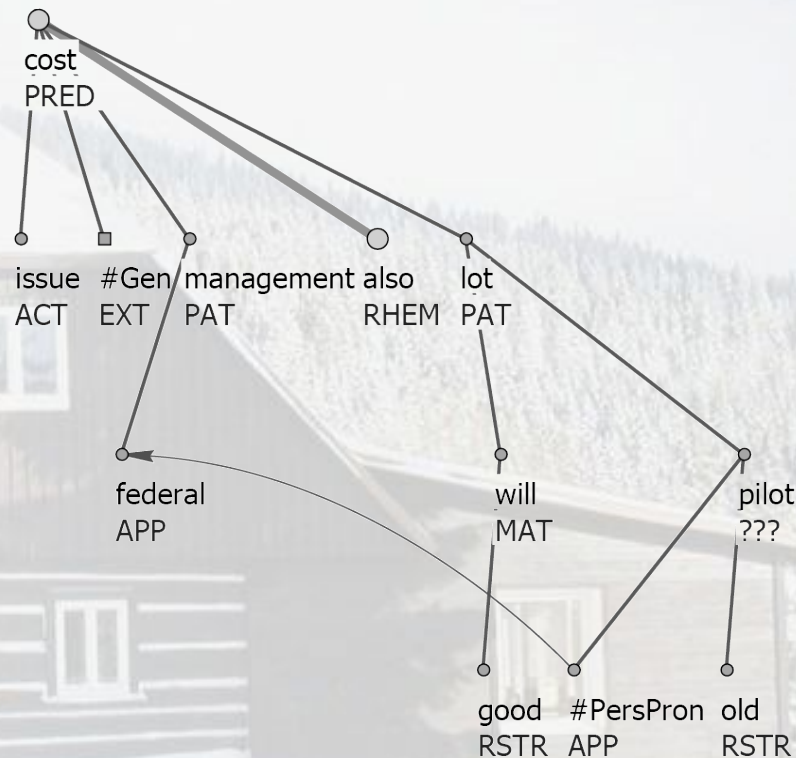
(Meanwhile, the company is operating with two separate pilot groups and seniority lists, and that is costing Federal „a big number“, says James Barksdale, executive vice president and chief operating officer.)

The issue has also cost Federal management a lot of good will among its old pilots .

Distantní postavení rematizátoru - stromy



If `` Mancuso FBI " has an intriguing central character, it also has a major flaw:



The issue has also cost Federal management a lot of good will among its old pilots.

Postpozice rematizátoru

- V angličtině může mít rematizátor dosah i na výrazy, které mu předcházejí, jak bezprostředně, tak distantně
- Typické rematizátory s dosahem doleva jsou *too*, *as_well*, *either*
- Dosah doleva mohou mít i ostatní rematizátory, většinou mají potenciálně dosah na obě strany
- Tento dosah se dá určit jedině z kontextu

Postpozice rematizátoru - příklady

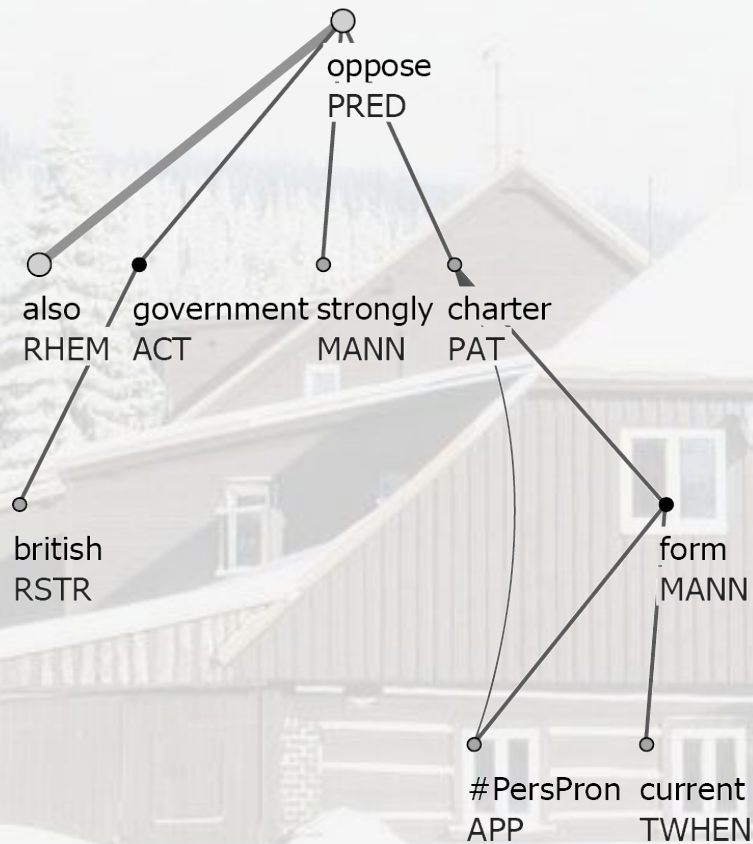
(European Community employers fear that the EC Commission 's plans for a „charter of fundamental social rights" is a danger to industrial competitiveness.)

The British government also strongly opposes the charter in its current form .

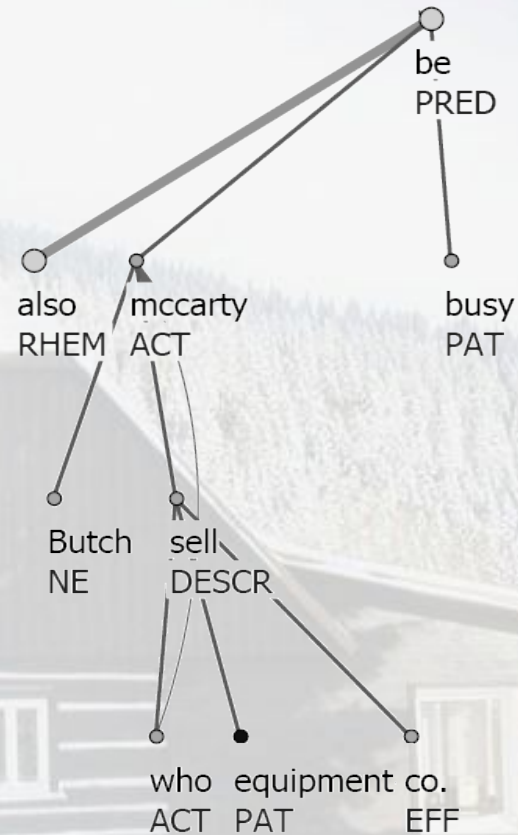
(Norman Young , a „mud-logger" at the Sniper well , has worked all but about nine days of this year.)

Butch McCarty , who sells oil-field equipment for Davis Tool Co. , is also busy.

Postpozice rematizátorů - stromy



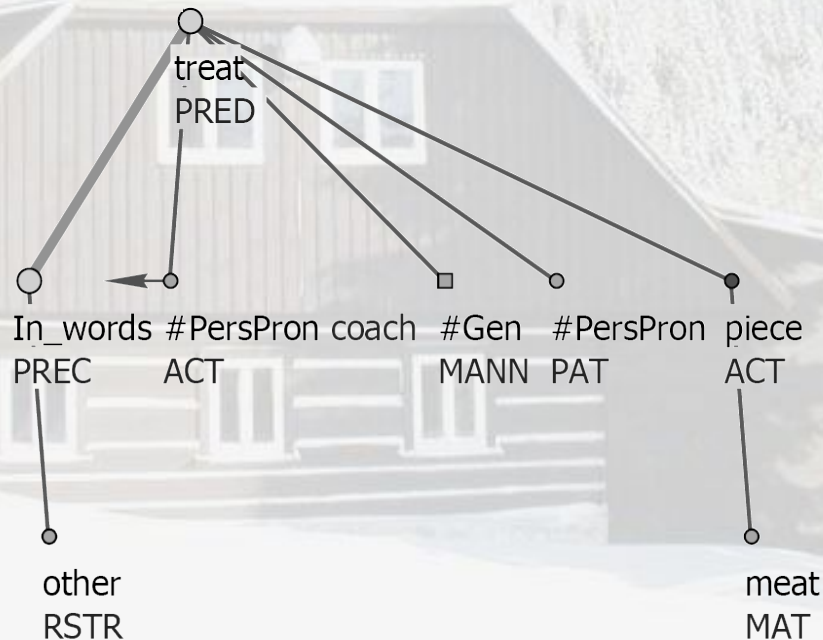
The British government also strongly opposes the charter in its current form.



Butch McCarty, who sells oil-field equipment for Davis Tool Co., is also busy.

Víceslovné PREC

- Výrazy typu *on the other hand*, podobně jako v češtině *v první řadě* apod. - mají vnitřní strukturu, přestože jsou atomické, mají děti

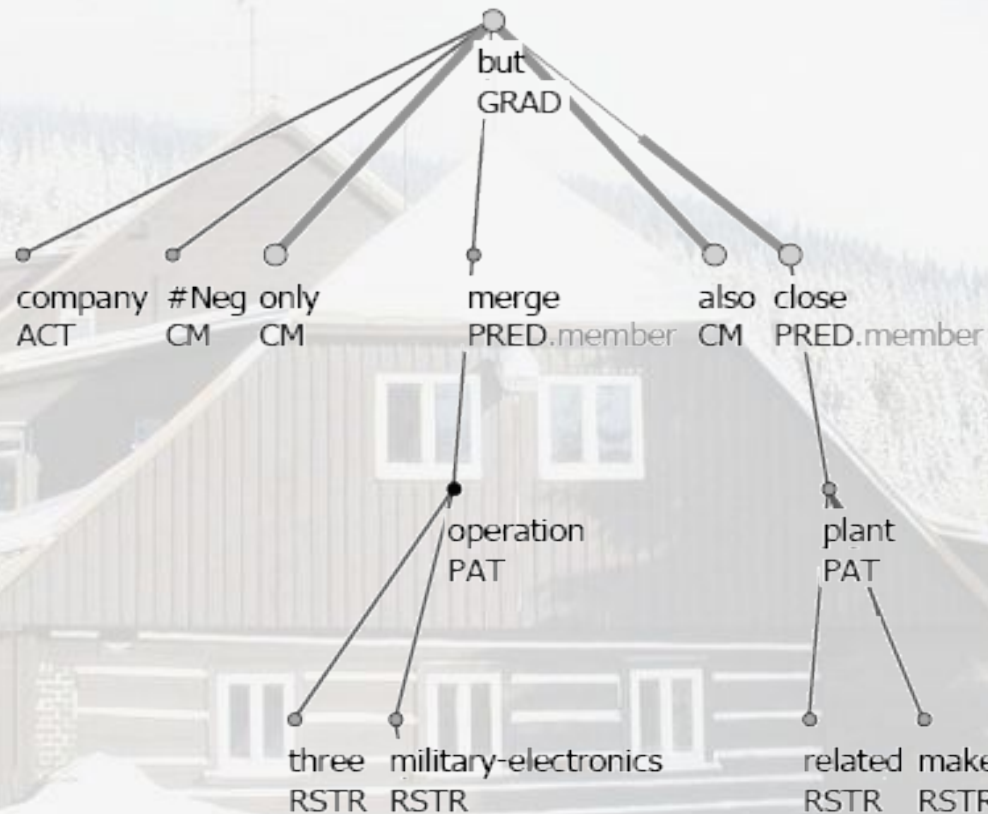


In other words, they treat you like a piece of meat . . .

Koordinační struktury

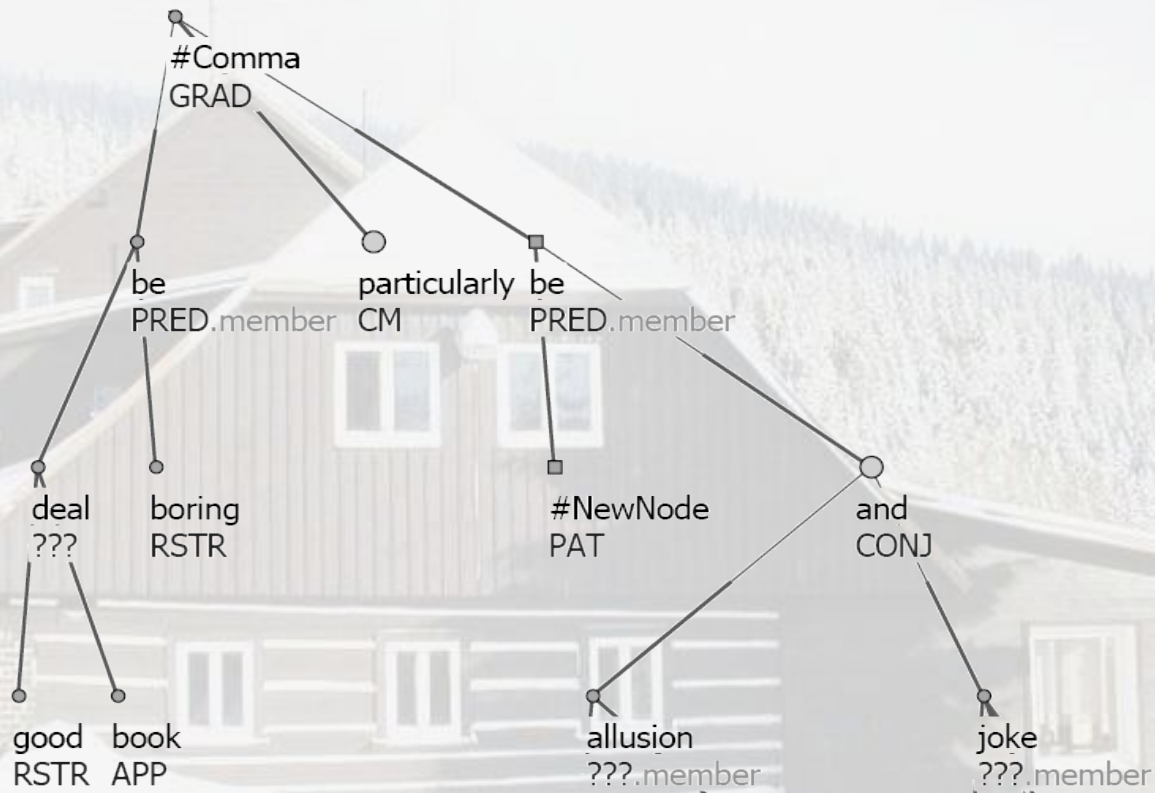
- V PEDT se na rozdíl od PDT anotují koordinace tak, že visí „co nejnižší“
 - Vzrůstá podíl členských koordinací, kde se vyskytují CM
 - Pokud to TFA vyžaduje, musí se koordinace posunout výš
 - Velký počet „prozatímních“ neprojektivit
- Nejčastější složené spojovací výrazy:
 - *Not only but (also), either or, neither nor, and not, and especially*
 - Velice podobné typy jako v češtině

Koordinační struktury – větná koordinace



The company not only merged three military-electronics manufacturing operations, but also closed an unrelated plant that *T*-1 makes ordnance devices used * in fighter planes and missiles.

Koordinální struktury – větná koordinace



A good deal of the book *ICH*-1 is boring, particularly the endless allusions to high and pop culture and the frequent jokes festooning the text.

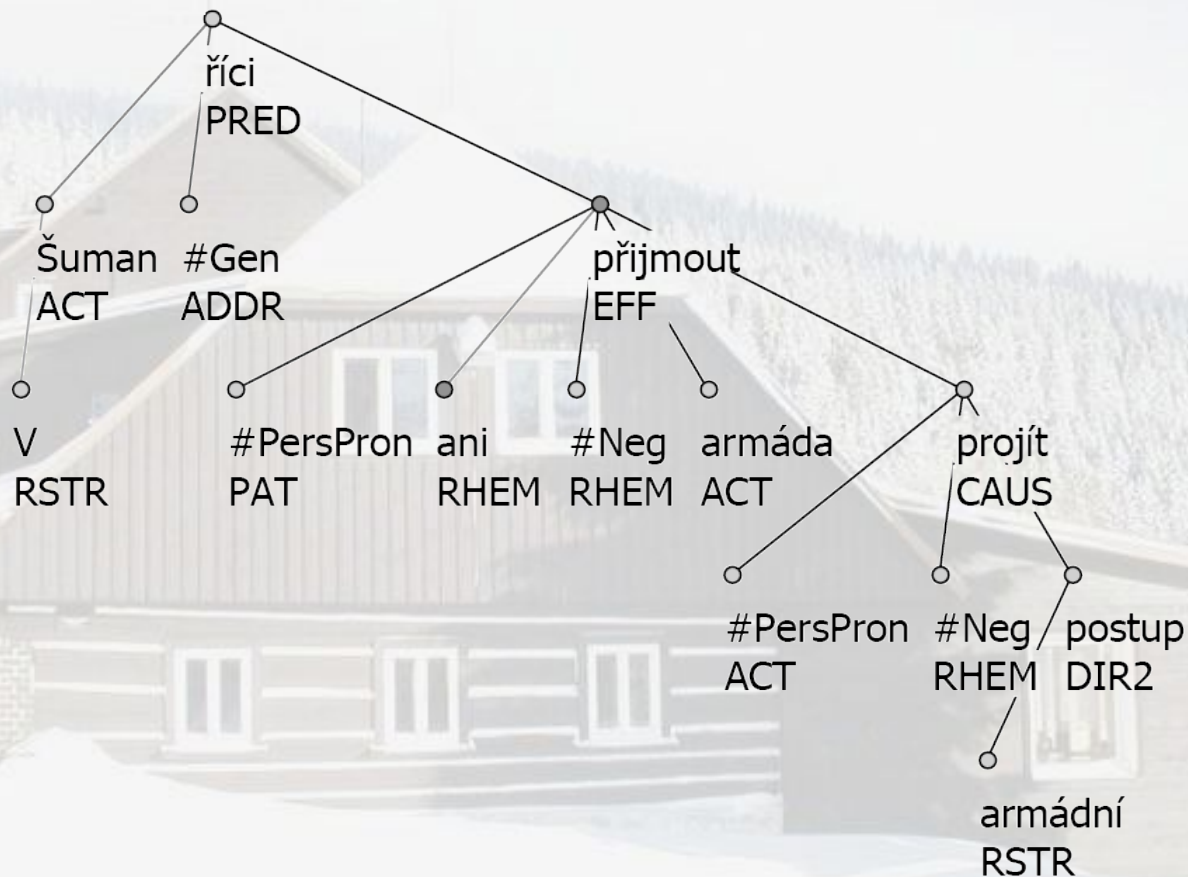
Co jsme se na angličtině naučili

- Otázka dosahu některých rematizátorů znepokojivě koresponduje s jejich českými ekvivalenty (also – také, rovněž)
- Je třeba specifikovat zavěšování rematizátorů v přímé řeči a vedlejších větách
- Dále se problematizuje hypotéza rematizátoru na hranici dvou hodnot kontextové zapojenosti
- Rematizátor většinou kumuluje několik funkcí (sémantickou, aktuálněčlenskou, diskurzivní)

Dosah doleva v češtině?

- Při pohledu na anglické věty nás napadne, zda to v češtině nefunguje podobně
- Přes všechny komplikace to vypadá, že ano, věty s také mohou mít víc významů:
 - (Karel má kočku.) Karel má taky psa. - kromě kočky má Karel psa.
 - (Já mám psa.) Karel má taky psa. - kromě mě má psa i Karel.

Zavěšení ve vedlejších větách a přímé řeči



Nebyl by ani přijat armádou, protože neprošel armádním postupem, řekl V. Šuman.

Co s tím dál

- Uzavřít anotaci DC a spočítat všechna data
- Anotovat TFA na PEDT
- Ověřit na zvukových datech
- Zjistit, jak promlouvají rematizátory do diskurzu