

## Checking rules

### Afun not preposition

- uzel má na morfologické rovině tag pro předložku (R na první pozici), ale na a-rovině není jako předložka anotován (tj. nemá afun AuxP).

Je-li chyba v morfologickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář)

### <afun1> undef <afun2>

- AuxP a AuxC nemohou viset na sobě. Výjimky.

### AuxP without preposition

- AuxP mohou být jen morfologické předložky (tj. v tagu je na první pozici R) nebo víceslovné předložky (tj. na uzlu visí uzel s afunem AuxY)

Je-li chyba v morfologickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář)

### Cannot be extra dep

- Pred a Denom nemohou být *is\_extra\_dependency* (tj. Pred\_E a Denom\_E je nedovolená kombinace)

### Cannot have children

- AuxY, AuxT, AuxR, AuxV nemá děti kromě AuxZ, AuxG

- AuxK, AuxG, AuxX nemohou mít děti (bez výjimky)

### Deprecated afun

- ExD a kombinované funkce AttrAttr, AttrAdv, AdvAttr, AttrObj, ObjAttr jsou zrušené afuny

### Invalid afun for lemma

- hlídá povolené afuny pro interpunkci:

' => Coord|Apos|AuxX

;' => Coord|Apos|AuxG

(' => Apos|AuxG

:' => Pred|Coord|Apos|AuxG

-' => Pred|Coord|Apos|AuxG

### Invalid AuxK

- AuxK mohou dostat jen tečka, vykřičník a otazník (a trojtečka).

### Invalid AuxV parent

- AuxV závisí na slovese (tag má na první pozici V)

Kontrola nehlásí chybu, pokud uzel má *is\_extra\_dependency*: AuxV\_E

Je-li chyba v morfologickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář)

### Invalid \_Co \_Ap \_P \_E

- afuny AuxK, AuxG, AuxX, AuxS nemohou mít žádnou příponu (tj. např. AuxG\_E je nedovolená kombinace).

### Invalid Sb case

- Podmět v jiném než 1. pádě (2. pád se ukazuje jen jako "Warning")

(Pád je v morfologickém tagu na 5. pozici)

Je-li chyba v morfologickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář).

Jedná-li se o pravidelnou změnu (*přišlo deset lidí* -> *kolem deseti lidí*), je třeba do komentáře poznámky *Other* napsat „sb“.

### Invalid Sb parent

- Rodičem uzlu s afunem Sb je sloveso (tj. uzel s tagem začínajícím na V) nebo - : =.
- Kontrola nehlásí chybu, pokud rodič má *is\_extra\_dependency*.
- Je-li chyba v morfoloickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář).

### Invalid tag/aux combination

- hlídá odpovídající tagy pro afuny AuxC, AuxX, AuxG, AuxK, AuxV, AuxT, AuxR
- AuxC => J,  
AuxX => Z :  
AuxG => Z :  
AuxK => Z :  
AuxV => V  
AuxT => P7--[34]  
AuxR => P7--[34]
- Je-li chyba v morfoloickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář).

### Invalid top node

- jenom Coord, Apos, Pred, Denom, Vocat, a fun\_E mohou být pod AuxS

### Must be top node

- Pred a Denom nemohou být hlouběji ve stromě (pouze v případě "obráceného syntaktického vztahu" (př. *jak je řekl Klaus, je to nesmysl*) visí Pred pod AuxC\_P). Hlouběji zanořené Pred nebo Denom musí mít příponu \_P (*is\_parenthesis root*)

### Noun parent

- pod substantivem (tag NN) by neměl viset Adv, Obj, Sb (rozvití substantive je Attr).

### parent not být/bývat

- Pnom by neměl viset na ničem jiném než na *být* a jeho vidových derivátech (rozeznává i "*bývávat*")

### Parent too far

- mezi uvozovku a jejího rodiče se v povrchovém pořadí nemůže dostat nic kromě interpunkce (počáteční uvozovky visí na slově bezprostředně následujícím, koncové na slově bezprostředně předcházejícím).

### S2 not AuxY

- uzel s tagem S2 má afun AuxY.

### Tag J, – should be AuxC

- morfoloická podřadicí spojka (tag J, ) by měla být AuxC
  - (výjimky jsou *aniž, jakož, takže, jako, než-2, li-2*)
- Je-li chyba v morfoloickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář)

### Tag J^ - should be Coord/Apos/AuxZ

- morfoloická souřadicí spojka (tag J^ ) by měla být Coord/Apos/AuxZ
- Je-li chyba v morfoloickém tagu, je třeba dát poznámku *M-layer* (není třeba komentář)

### Verb parent

- pod slovesem (tag V) by neměl viset Attr

### Warning: no children

- AuxP a AuxC by měly mít nějaké dítě

**Warning: No punctuation**

- uzel s příponou `_P` (*is\_parenthesis\_root*), na kterém nevisí interpunkce (na uzlu s příponou `_P` by měla viset interpunkce (závorka, pomlčka apod.) oddělující parentezi od základního textu. Je však celá řada výjimek (nespojité parenteze, parenteze na začátku nebo na konci věty)

---

Slovnědruhové kategorie v morfologickém tagu:

N noun

A adjectivum

P Pronoun

C číslovka

V verb

D aDverb

R pReposition,

J, podřadicí spojka

J^ souřadicí spojka

T částice

I citoslovce

F cizí slovo

Q písmeno

B zkratka

S segment