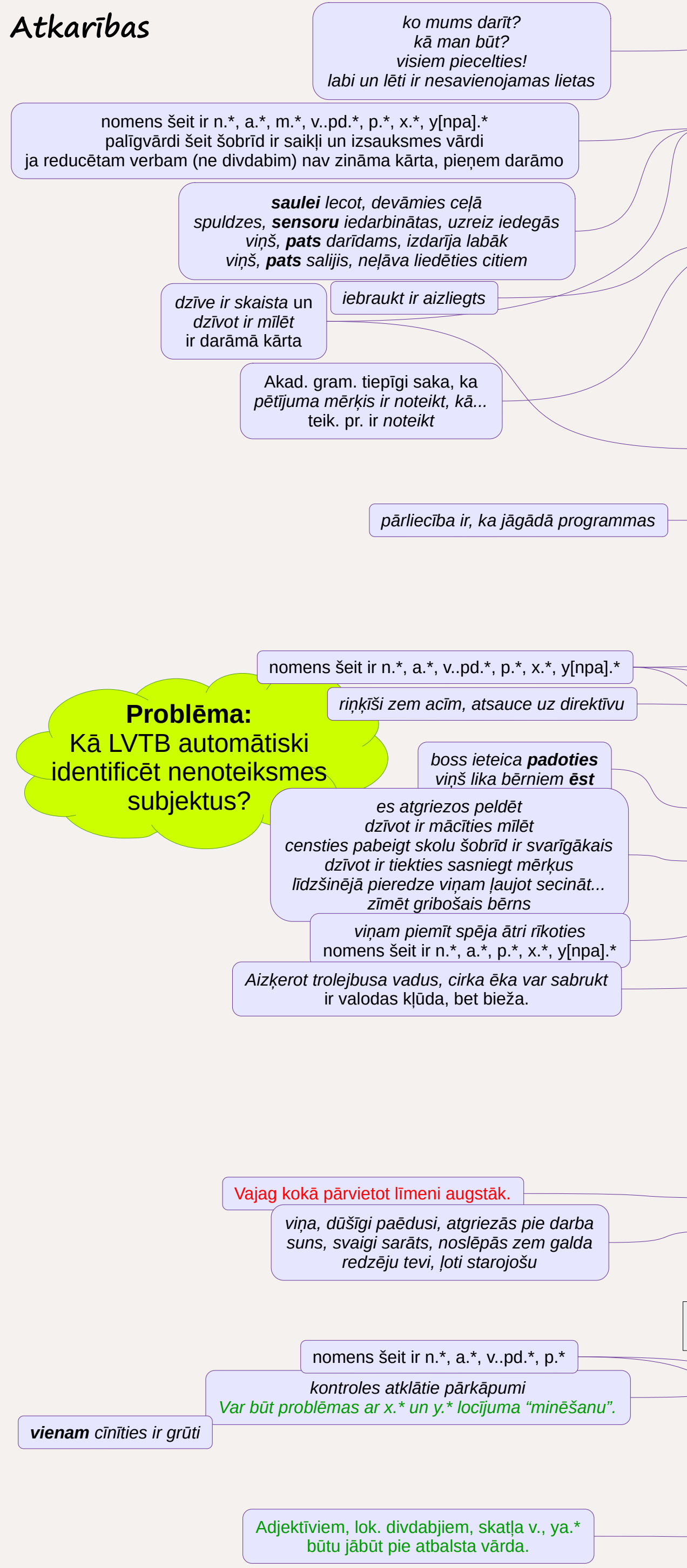


# LVTB un UD lomu atbilstme

## Atkarības



**Nākotnei:** Varētu pārskatīt ar xx.\* un xf.\* izteiktos attr un novērtēt, kas skaitliski tie biežāk ir.

kaut kas notiks 22.00 pantu numuri juridiskos dokumentos

xSimile: vairāk kā/nēkā [pieci]

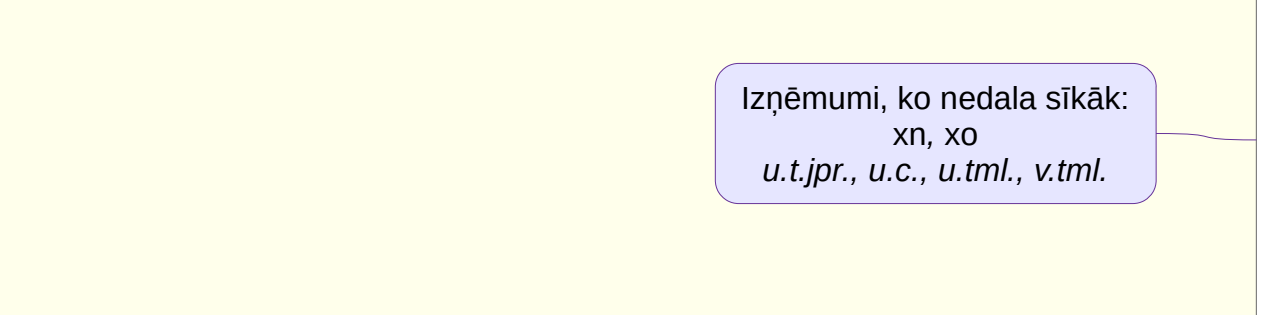
"Washington (CNN) nevaram identificēt"

Pēc līdzības ar daļu etc. angļu valodā un daļu o.v. vācu valodā

**xPred**

- Vispārējā izteicēju principi:
  - \* fāzes un modāle (sākšana, beigšana, varēšana) izteicēji ir izteicējs + xcomp, kur izteicējs = sākšana/beigšana/varēšana, bet xcomp = basElem.
  - \* saistītas funkcija UD lieto tikai (ne)būt, pārējos gadījumos izteicējs + xcomp.
  - \* ciešāmā kārtā ir (ne)tkrapt/būt + -ts divdabjs.
  - \* darāmā kārtā ir -is divdabjs. Te ciešāmā kārtā nemēdz būt.
  - \* -ošs, -ams divdabjs saikto laikus neveido, -ts divdabjs var veidot nominātu izteicēju.
  - \* reizēm ir grūti atšķirt -ts divdabja izteicējus - ciešāmā kārtā vai sastata - tad lieto tagu.
- Sk. LV-UD-annotations.doc, bet AUX ir (ne)būt, (ne)tkr, (ne)tkrpt ar tagiem vc.\* un va.\*

## Vārdi ar atstarpēm



**Diskusija:** jāuzmanās, jo UD goeswith ir domāts tikai kļūdaini sadalītiem vārdiem

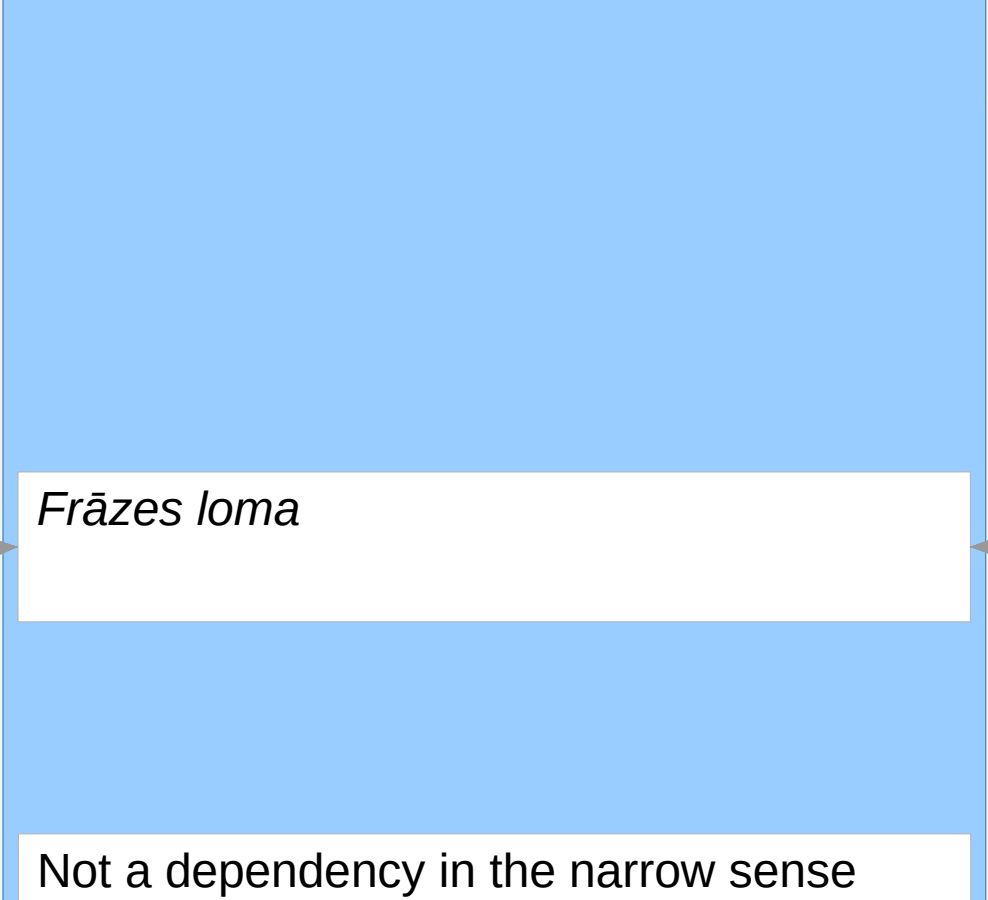
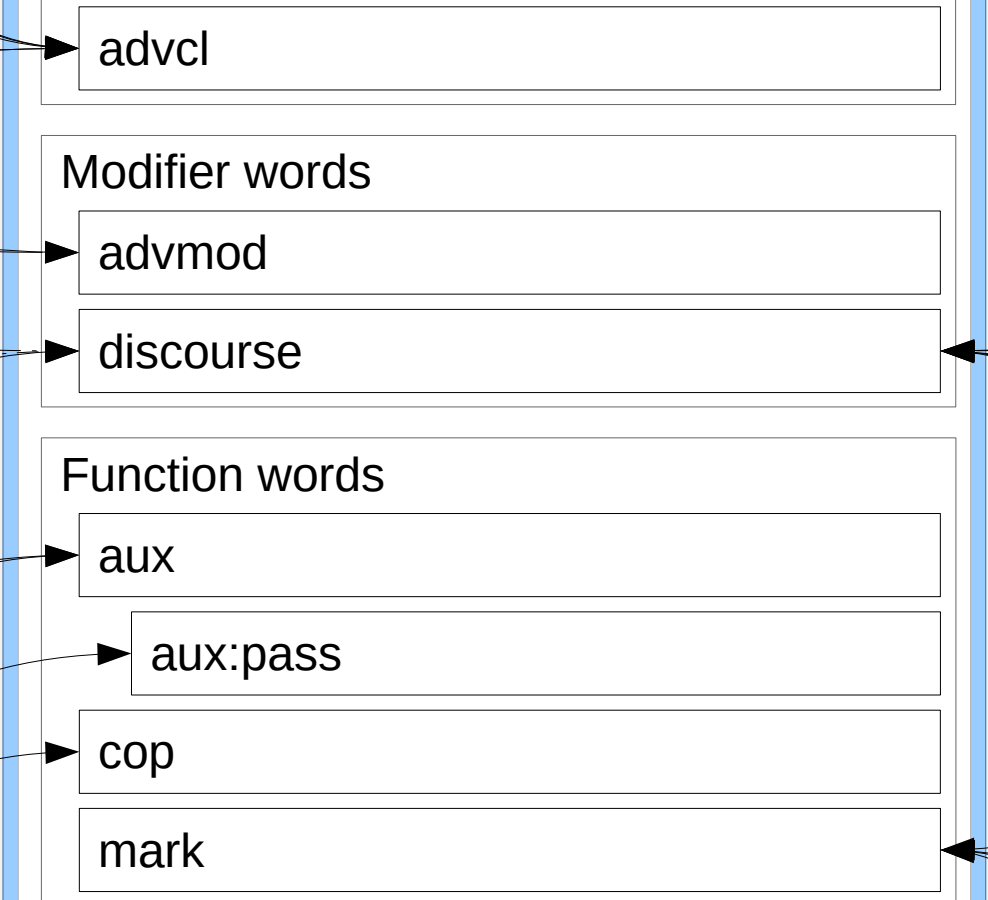
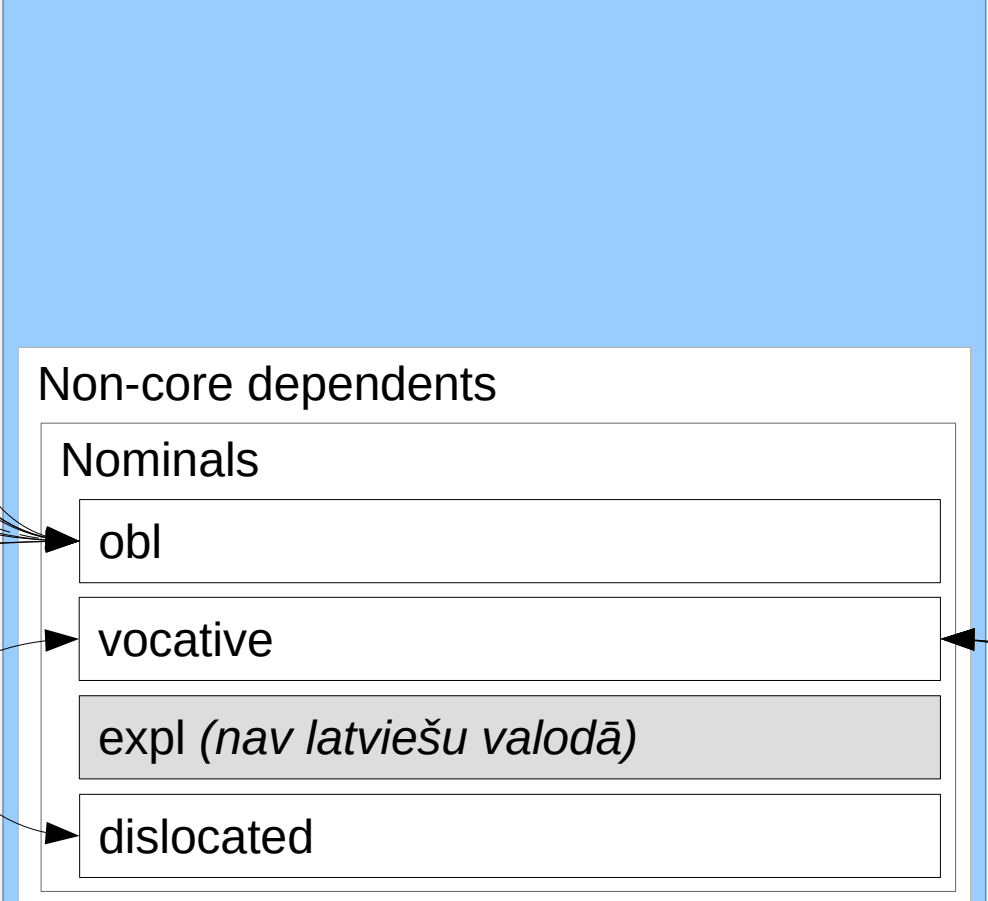
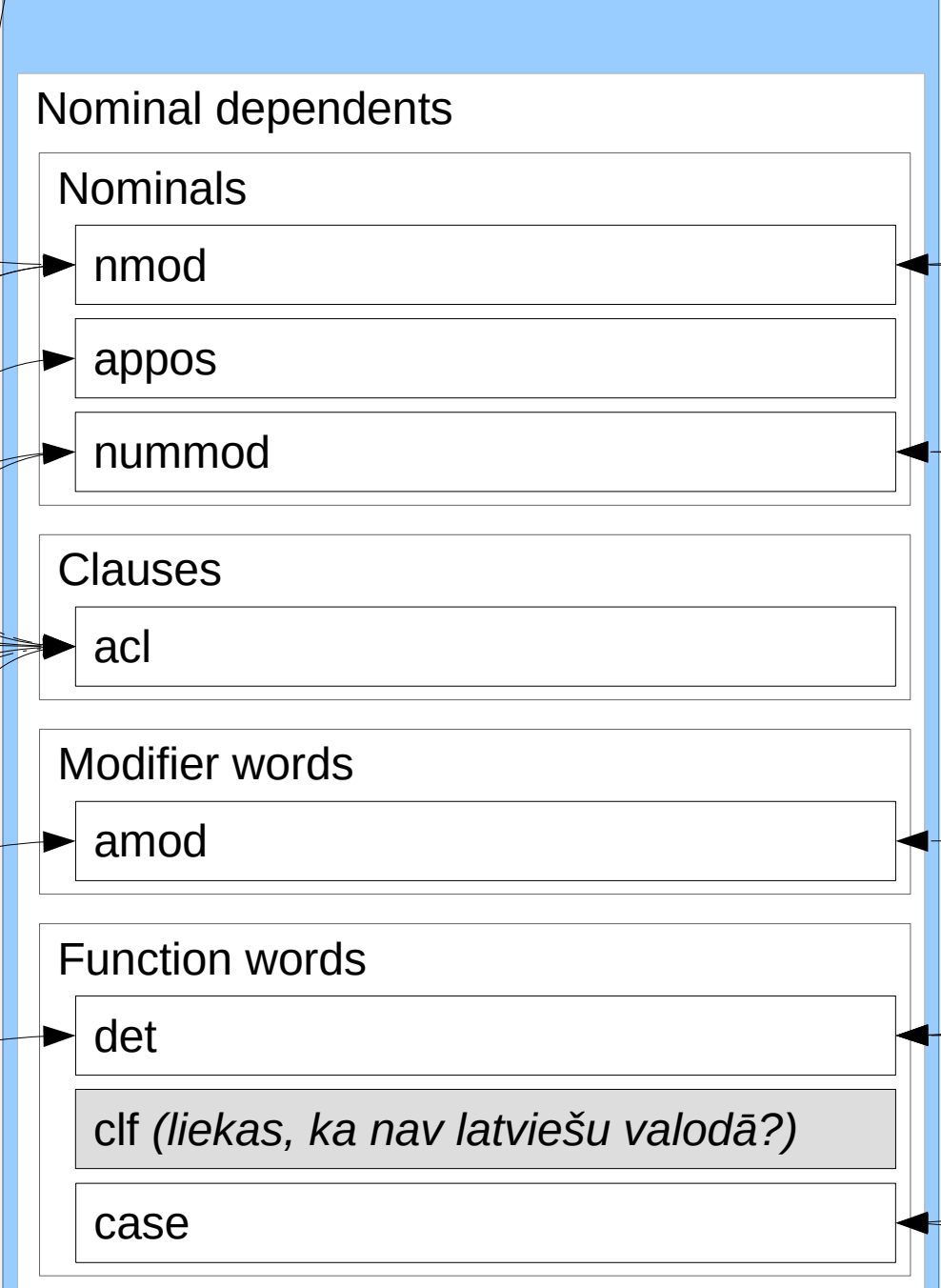
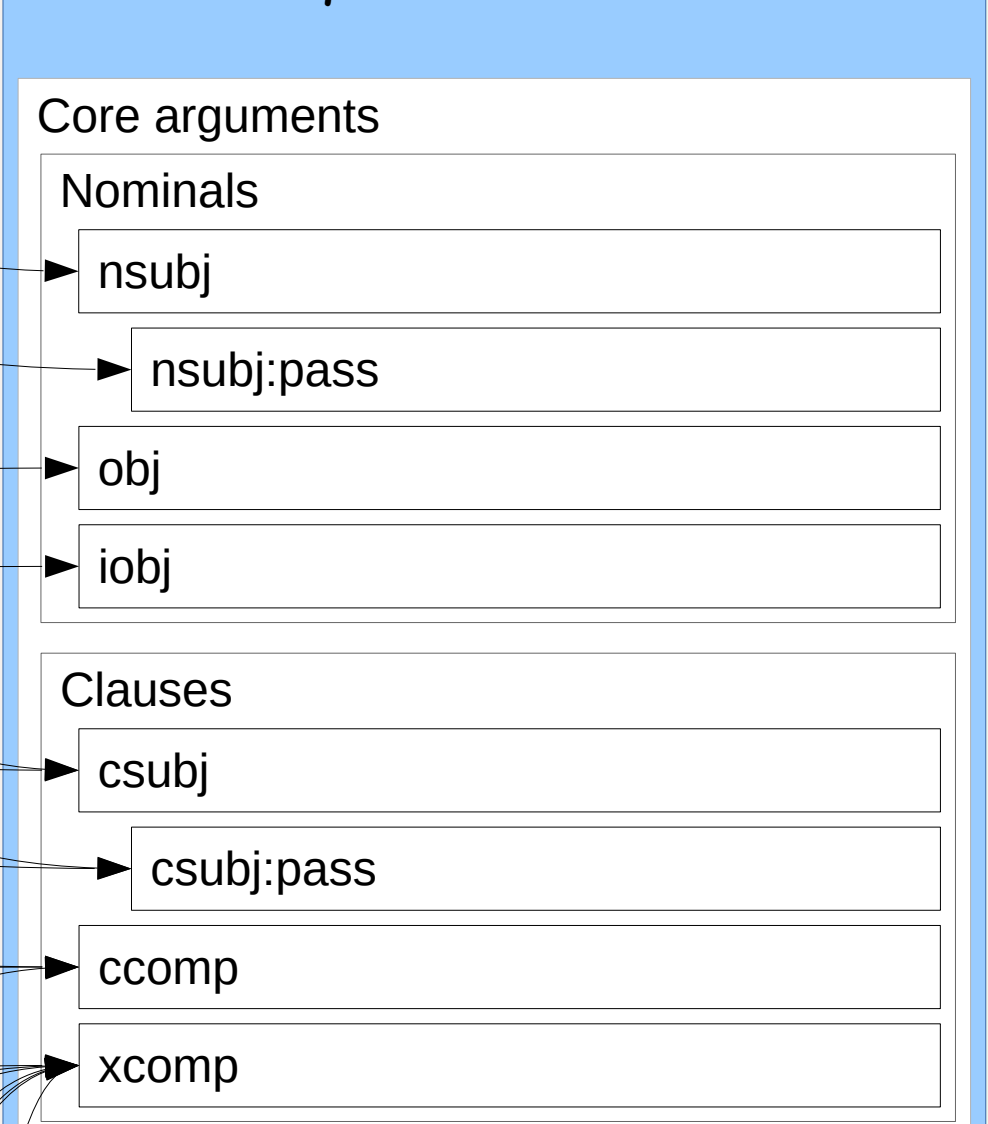
Aprakstīts LV-UD-annotations.docx

Situācija, kad cilvēks pārlabo iepriekš teikto. LVTB nav šādu piemēru, mūsu modelī nav izdomāts, kā marķēt.

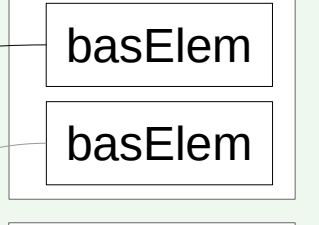
Izteicēja atkarība no mākslīgas koka virsotnes. Šāda loma kokā var būt tikai viena un, ja citādi galīgi nesaskan, tad UD vadlīnijas liek kreisāko no tiem mākslīgi padarīt par vienīgo root.

Kad visi strikti trūkst: kaut ko nevar konvertēt vai parsēris devis nepilnu rezultātu.

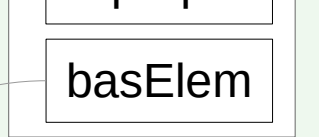
## Universal Dependencies



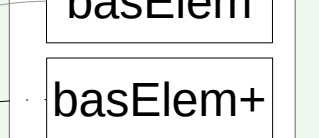
## xApp



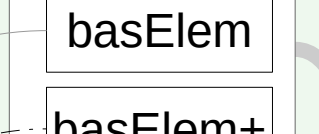
## xPrep



## phrasElem



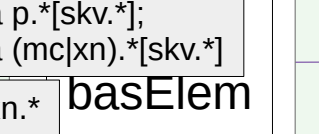
## subrAnal vw



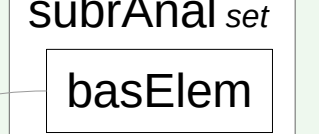
## subrAnal pv



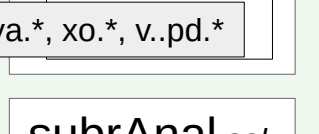
## subrAnal svk



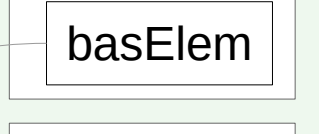
## subrAnal set



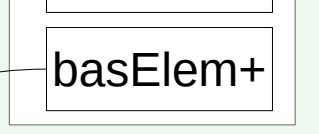
## subrAnal sai



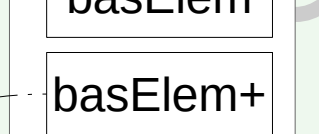
## xNum



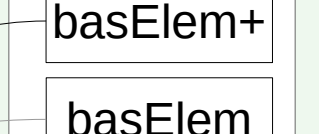
## xParticle



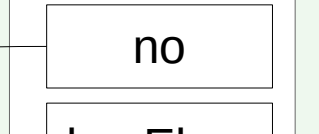
## xSimile



## namedEnt



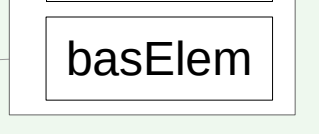
## unstruct



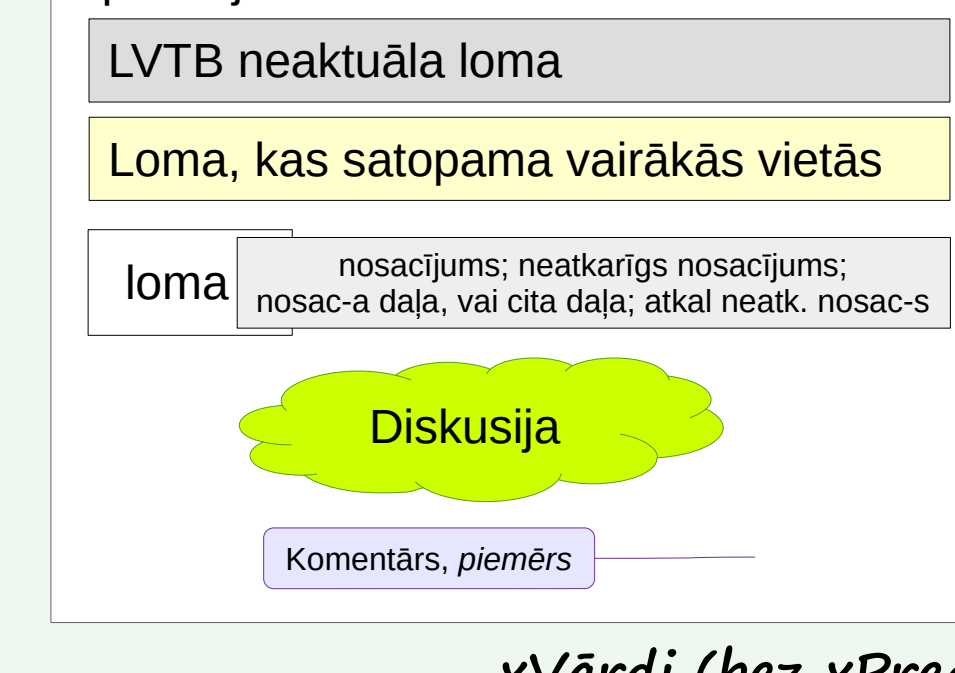
## crdClauses



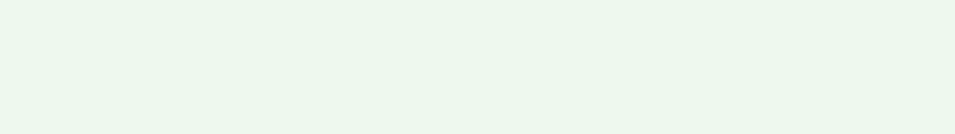
## utter



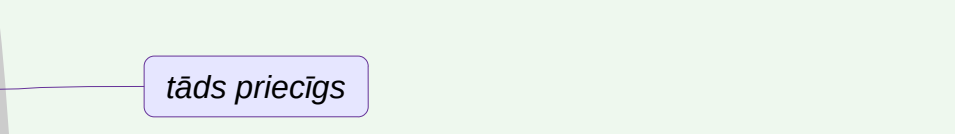
## Apzīmējumi



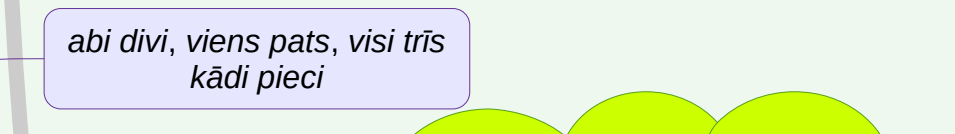
## xVārdi (bez xPred)



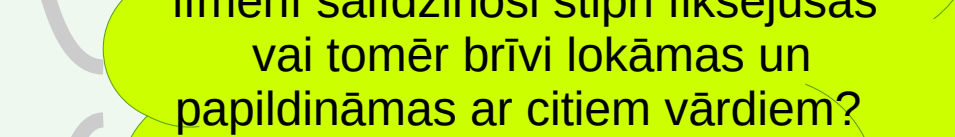
## Koordinācija



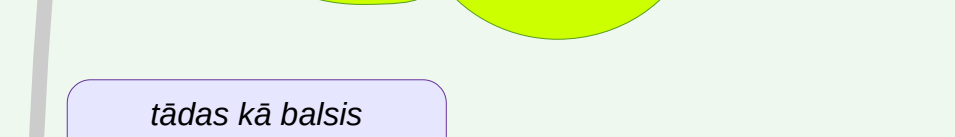
## PMc



## Pārdomas



## Problema: Kā LVTB automātiski identificēt nenoteiksmes subjektus?



## Problema: Mūsu frāzes ir gramatiskā līmeni salīdzinoši stipri fiksējušās vai tomēr brīvi lokāmas un papildināmas ar citiem vārdiem? Atkarībā no tā, tās marķējamas kā flat, fixed vai compound.

