

LVTB un UD lomu atbilstme

Atkarības

ko mums darīt?
kā man būt?
visiem piecēlies!
labi un lēti ir nesavienojamas lietas

nomens šeit ir n.* a.* m.* v.* pd.* p.* x.* y[npa].*
palīgvardi šeit šobrīd ir saiki un izsaukšanas vārdi
ja reducētam verbam (ne divdabim) nav zināma kārtā, pieņem darāmo

saulei lecot, devāmies ceļā
spulzēs, sensoru iedarbinātās, uzreiz iedegās
viņš, pats durīdams, izdarīja labāk
viņš, pats salījis, neļāva ledieties cietim

dzīve ir skaista un
dzīvot ir mīļot
ir darāmā kārtā

iebraukt ir aizliegts

Akad. gram. tiepiši saka, ka
pēģinājumā mērķis ir noteikt, kā...
teik. pr. ir noteikt

Problēma:
Kā LVTB automātiski
identificēt nenoteiksmes
subjektus?

nomens šeit ir n.* a.* v.* pd.* p.* x.* y[npa].*
rīnīši zem acīm, aisaucē uz direktīvu

boss ieteica padoties
viņš lika bērniem ēst
es atgriezos pašdi
dzīvot ir mācīties mīļot
censies pabeigt skolu šobrīd ir svarīgāks
dzīvot ir iekties sasniegt mērķus
līdzšinējā pieredze viņam ļaujot secināt...
zīmēt grībošais bērns

viņam piemīt spēja ātri rīkoties
nomens šeit ir n.* a.* p.* x.* y[npa].*
Aizkerot troļņbūsa vadus, cirka ēka var sabrukt
ir valodas kļūda, bet bieža.

Vajag kokā pārvietot līmeni augstāk
vīna, dūšņi paēdusi, atgriezās pie darba
suns, svaiji sarāts, noslēpās zem galda
redzēju līvi, toti starojošu

nomens šeit ir n.* a.* v.* pd.* p.*
kontrolēs atklātie pārkāpumi
Var būt problēmas ar x.* un y.* locījuma "minēšanu".

panākumi zinātnēs, seminārs Rīgā (kad lok. nav apstākļi)
šie ļauts pašā bija pārdomāto zemu audzētāji

Adjektīviem, lok. divdabjiem, skaitļa v. ya.*
būtu jābūt pie atbalsta vārda.

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

mātes grāmata, viedoklis par to

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

advmod.emph nav obligāts,
izvēlēts pēc līdzības ar PDT

xPred

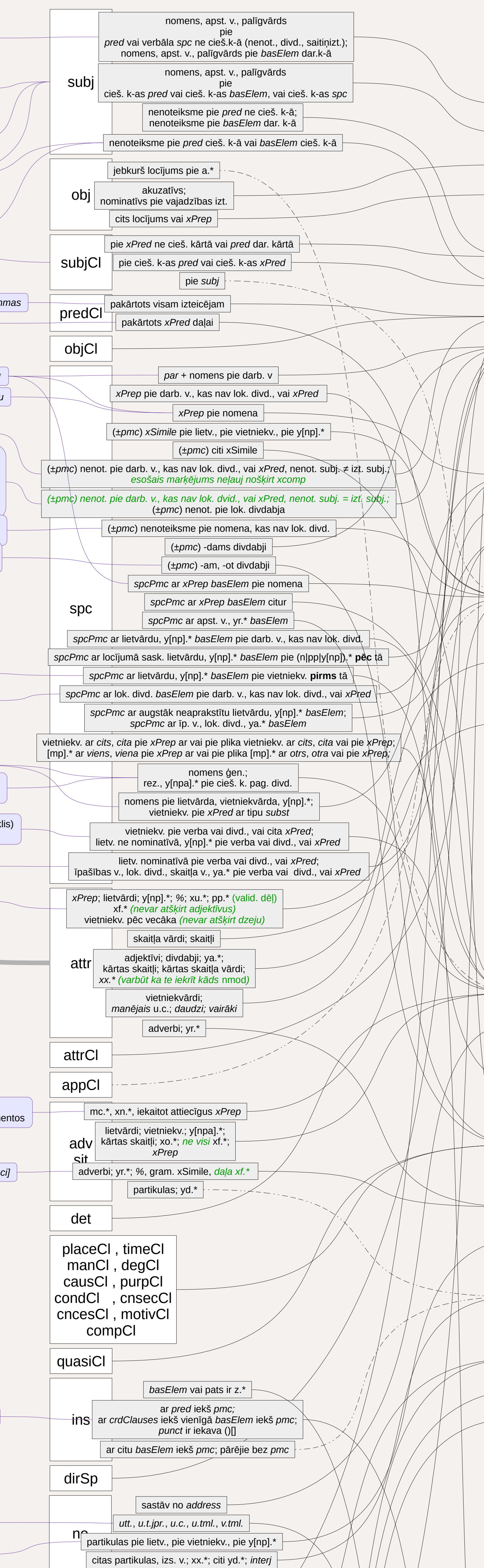
Vispārējie izteicēju principi:
* fāzes un modālie (sākšana, beigšana,
varēšana) izteicēji ir izteicējs + xcomp,
kur izteicējs =
sākšana/beigšana/varēšana, bet xcomp =
basElem.
* sāltinas funkcijā UD lieto tikai (ne)būt,
pārējos gadījumos izteicējs + xcomp.
* ciešamā kārtā ir (ne)tkāpt/būt + ts
divdabīs.
* darāmā kārtā ir -is divdabīs. Te ciešamā kārtā nemēdz būt.
* -oš-, -ams divdabīs saliktos laikus
neveido, -ts divdabīs var veidot nominātu
izteicēju.
* reizēm ir grūti atšķirt -ts divdabja
izteicējus – ciešamā kārtā vai sastata –
tad lieto tagu.

Sk. LV-UD-annotations.doc, bet
AUX ir (ne)būt, (ne)tkāpt,
(ne)tkāpt ar tagiem vc.* un va.*

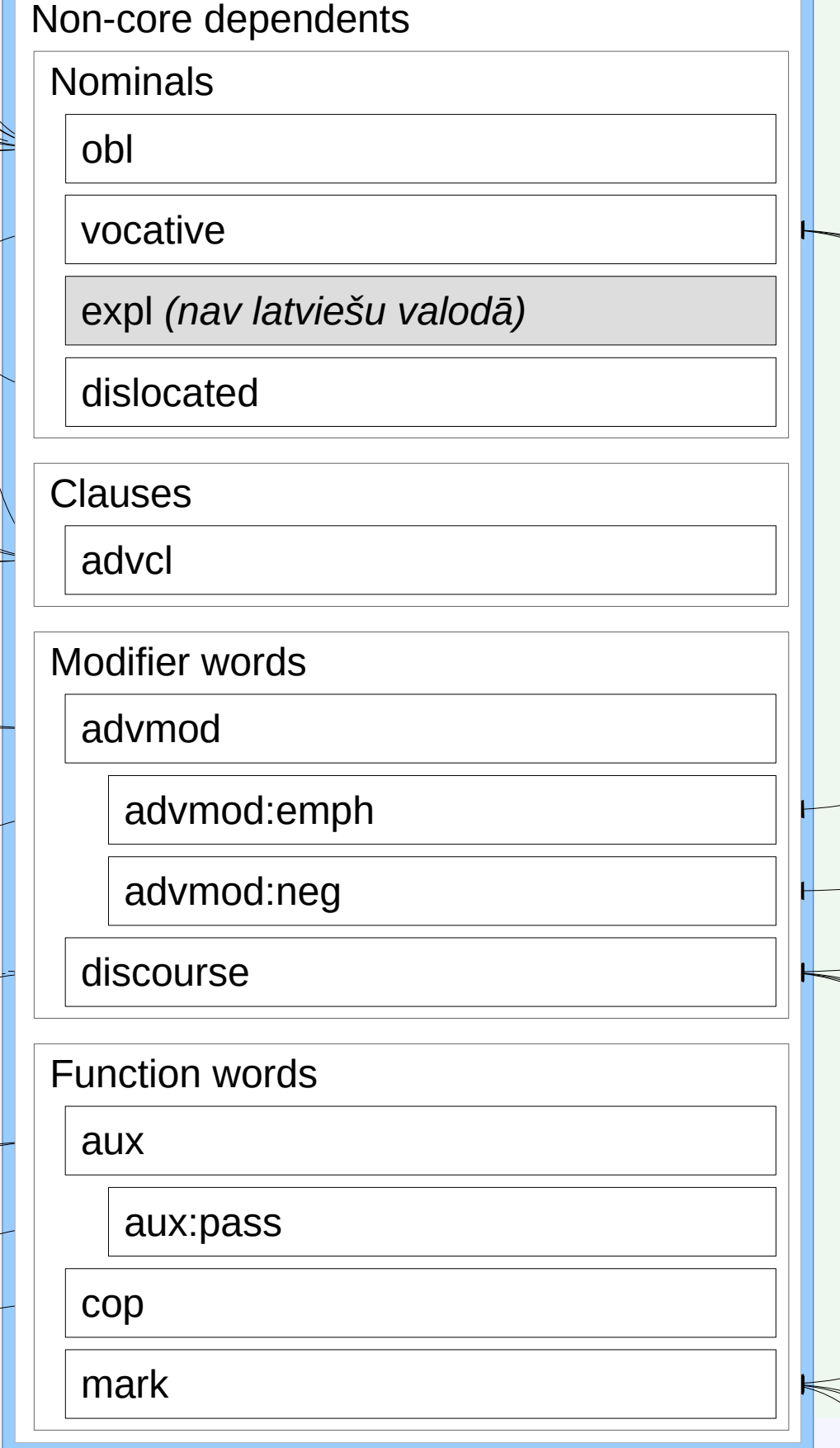
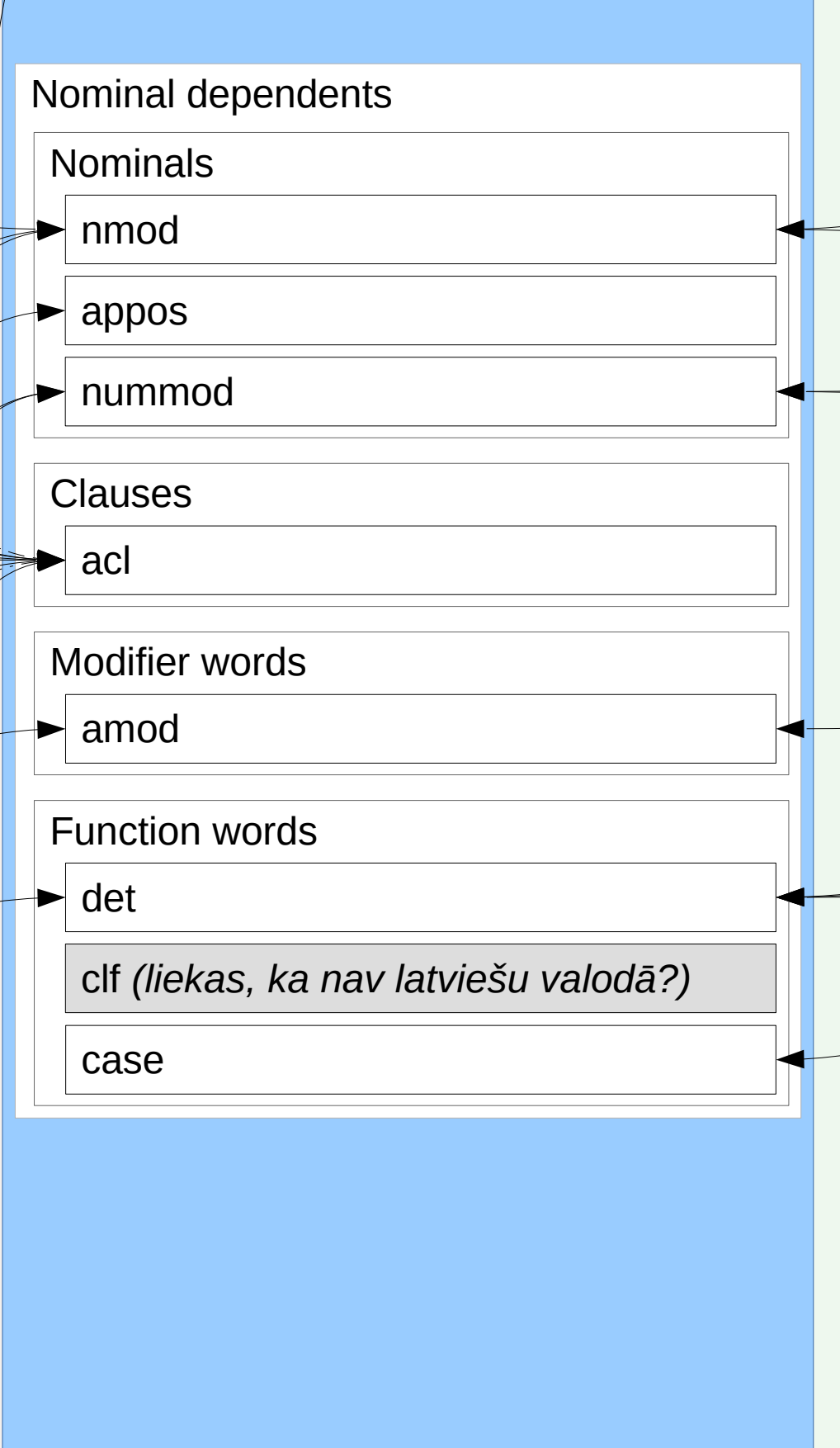
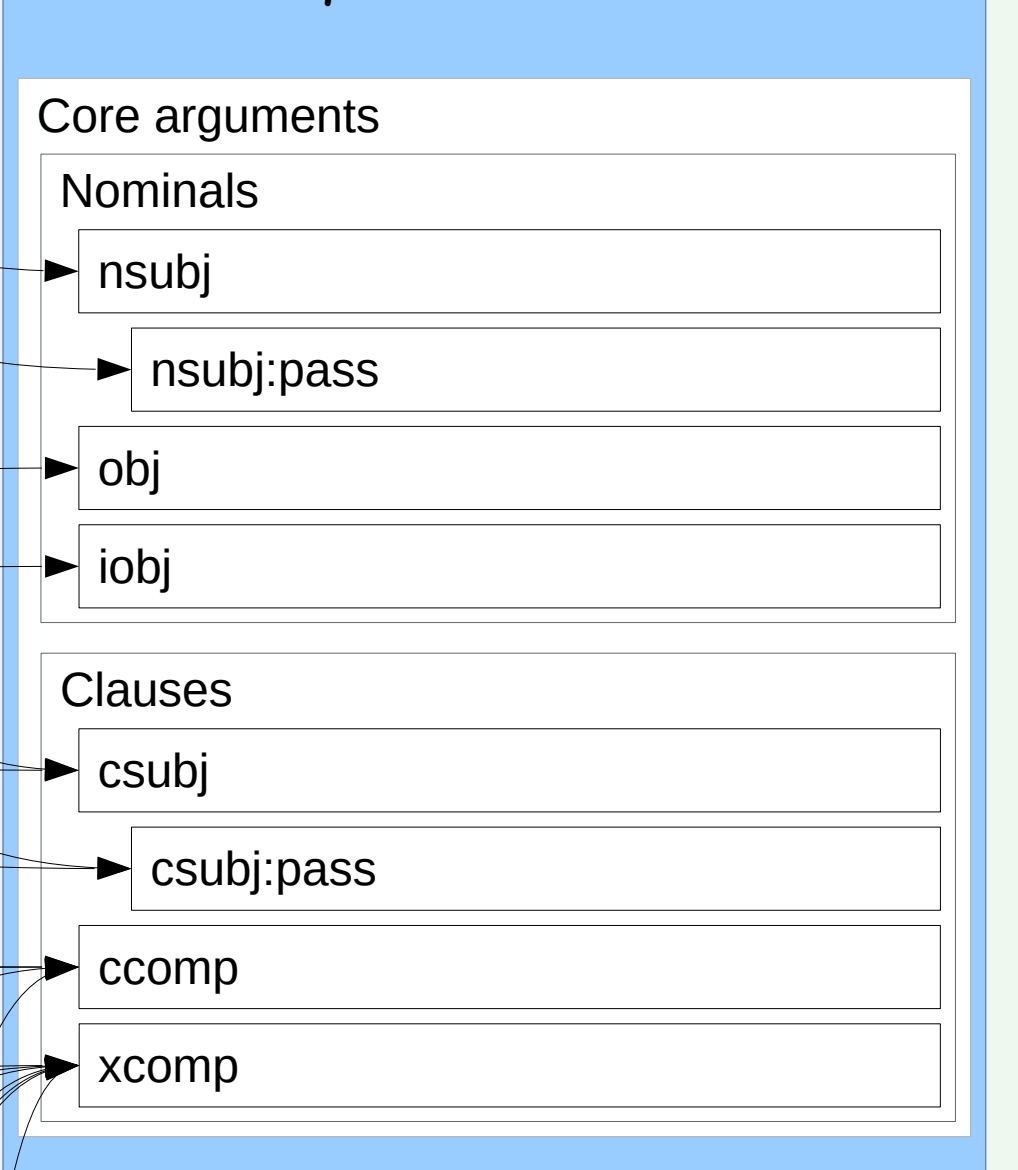
Vārdi ar atstarpēm

Izņēmumi, ko nedala sīkāk:
xi, xo
u.t.jpr., u.c., u.tml., v.tml.

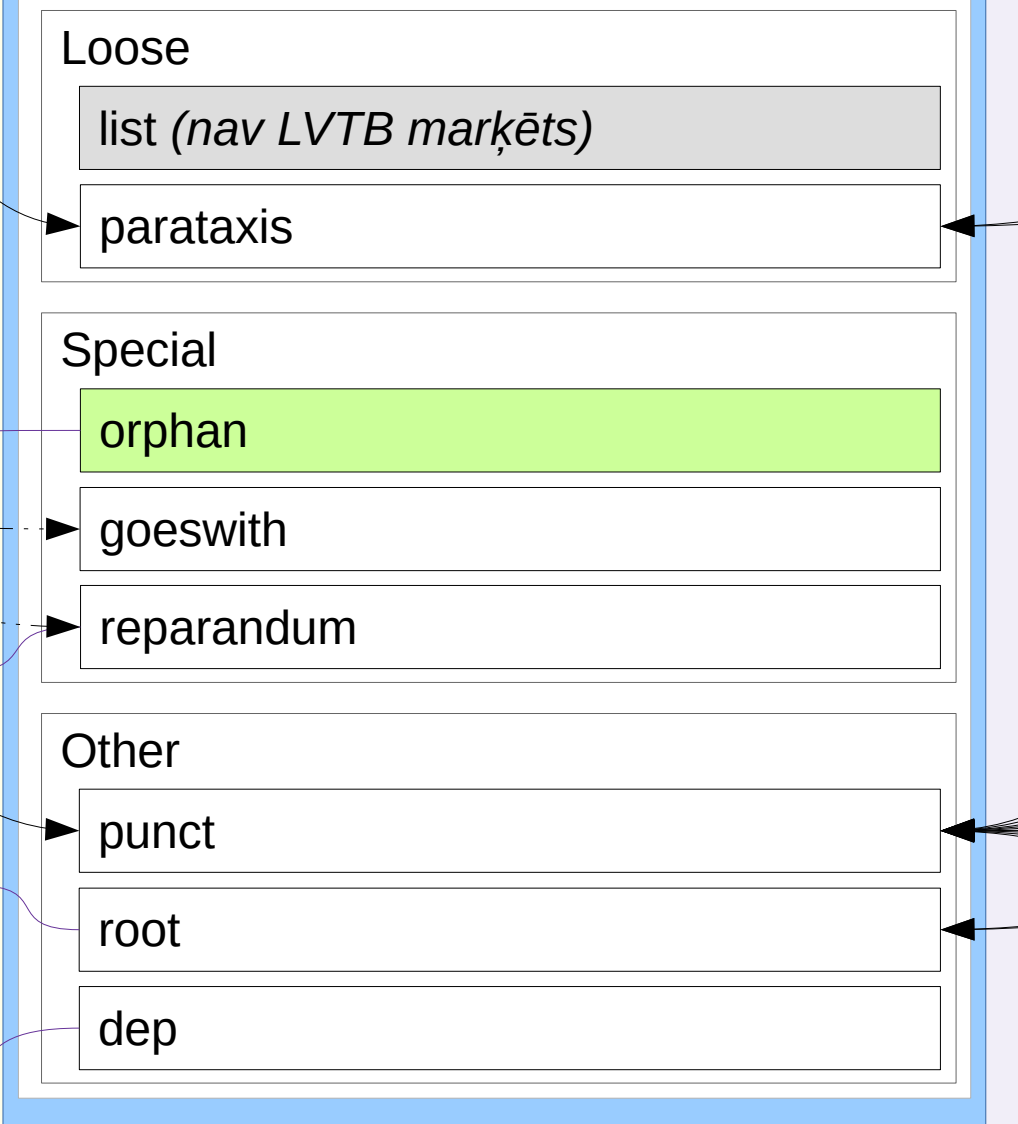
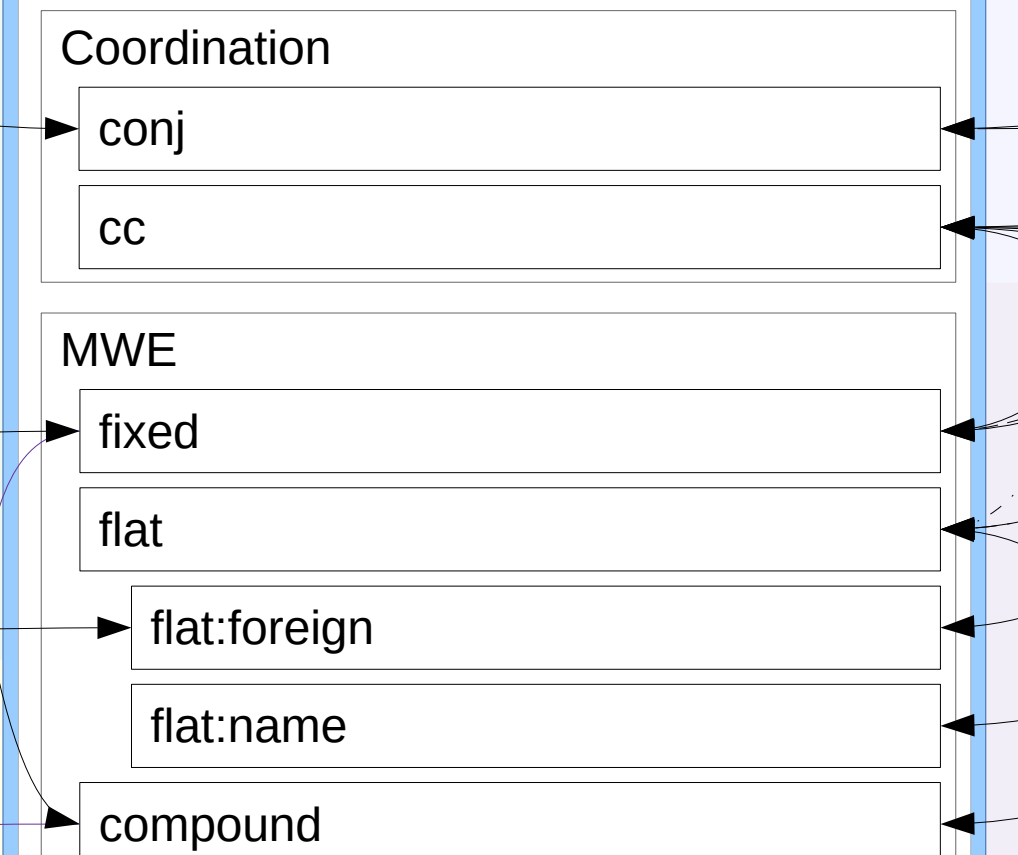
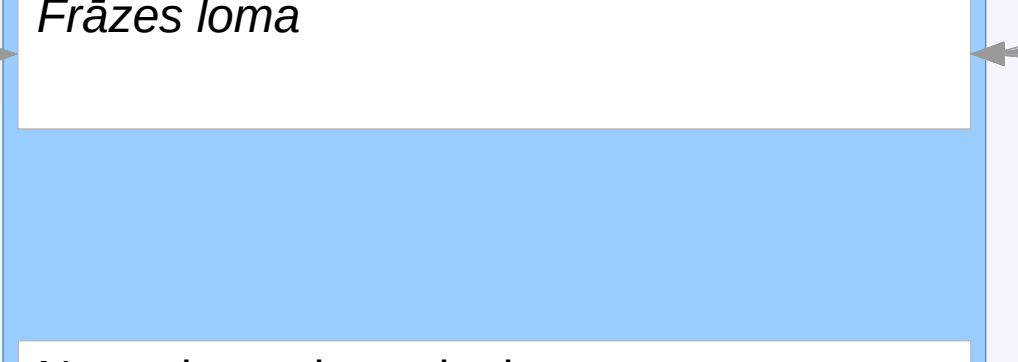
Vārdforma satur ''
pirmais
otrais+
koptags xu.*, xx.*
koptags xf.*
pārējie gadījumi



Universal Dependencies



Frāzes loma



UD: fixed ir lsi, grammatikalizējusies elementi
ar fiksētu struktūru, savukārt flat ir "semi-fixed"
vai "headless (exocentric)": compound ir "headed
(endocentric)" struktūra. Savukārt, salīdzinot
compound ar nmod – compound ir stabils (?)

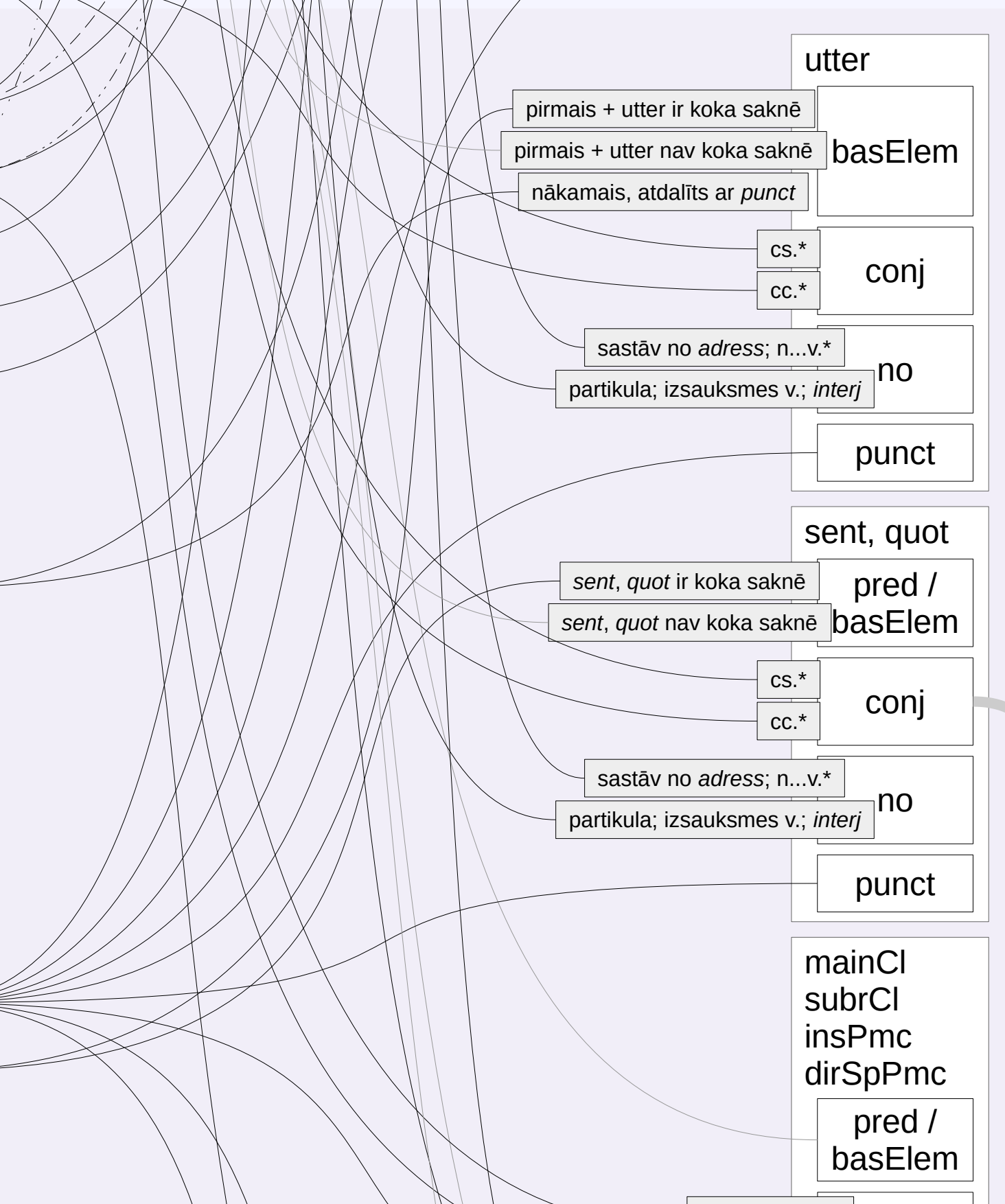
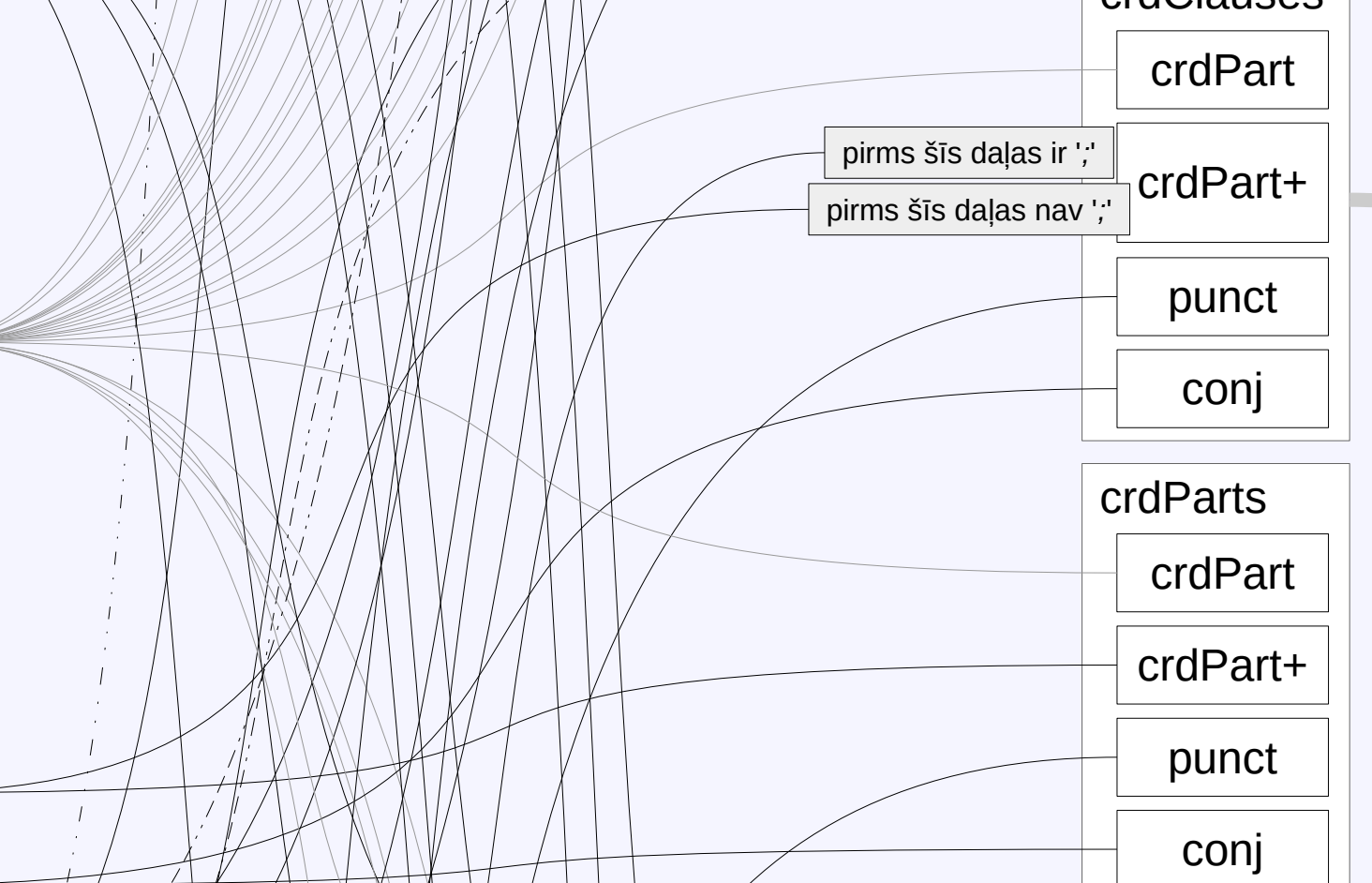
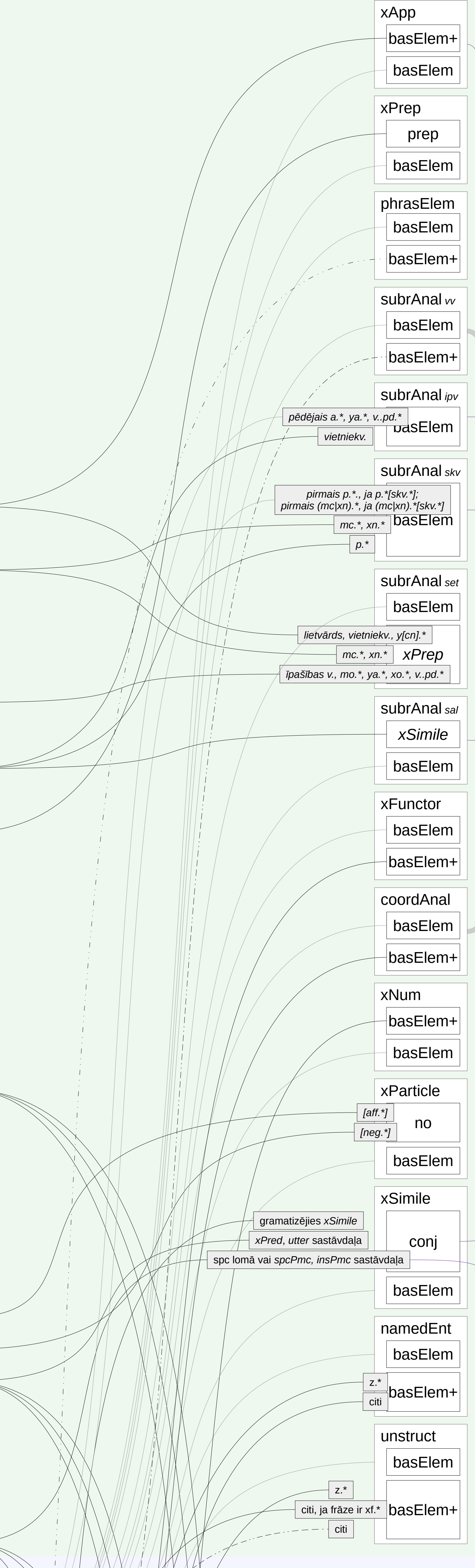
Aprakstīts
LV-UD-annotations.docx

Situācija, kad cilvēks pārļābo iepriekš
teikto. LVTB nav šādu piemēru, mūsu
modelī nav izdomāts, kā markēt.

Izteicēja atkarība no mākslīgās
koka virsotnes. Sāda loma kokā
var būt tikai viena un, ja citādi galīgi
nesanāk, 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.

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



Apzīmējumi

LVTB neaktuāla loma

Loma, kas satopama vairākās vietās

loma nosacījums; neatkarīgs nosacījums;
nosac-a daļa, vai cita daļa, atkal neatk. nosac-s

Diskusija

Komentārs, piemērs

xVārdi (bez xPred)

padomnieks leitnans Priedītis

tāds precīgs

abi divi, viens pats, visi trīs
kādi pieci

Problēma:
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.

tādas kā balsis
tāds kā precīgs
xSimile ir gramatizējis

Gramatizējamies xSimile – pirmais vārds
saņem frāzes lomu, pārējie ir fixed.
Pieņemums, ka gram. xSimile pirmais
vienmēr ir basElem

neļūties ne kā bērns, ne kā filozofs

Koordinācija

Virsteikumi, kas atdalīti ar komatiem - kā vienlīdzīgi izteicēji.
Virsteikumi, kas atdalīti ar semikolu – parataxis.

Problēma:
Vai parataxis ir tikai
šie gadījumi?

PMC

Pārdomas:
Nez vai šis ir adekvāti arī
vientuļiem sakārtojuma saikļiem
teikuma vidū?