Latvian FrameNet: kick-off

Normunds Grūzītis

University of Latvia, Institute of Mathematics and Computer Science

National information agency LETA

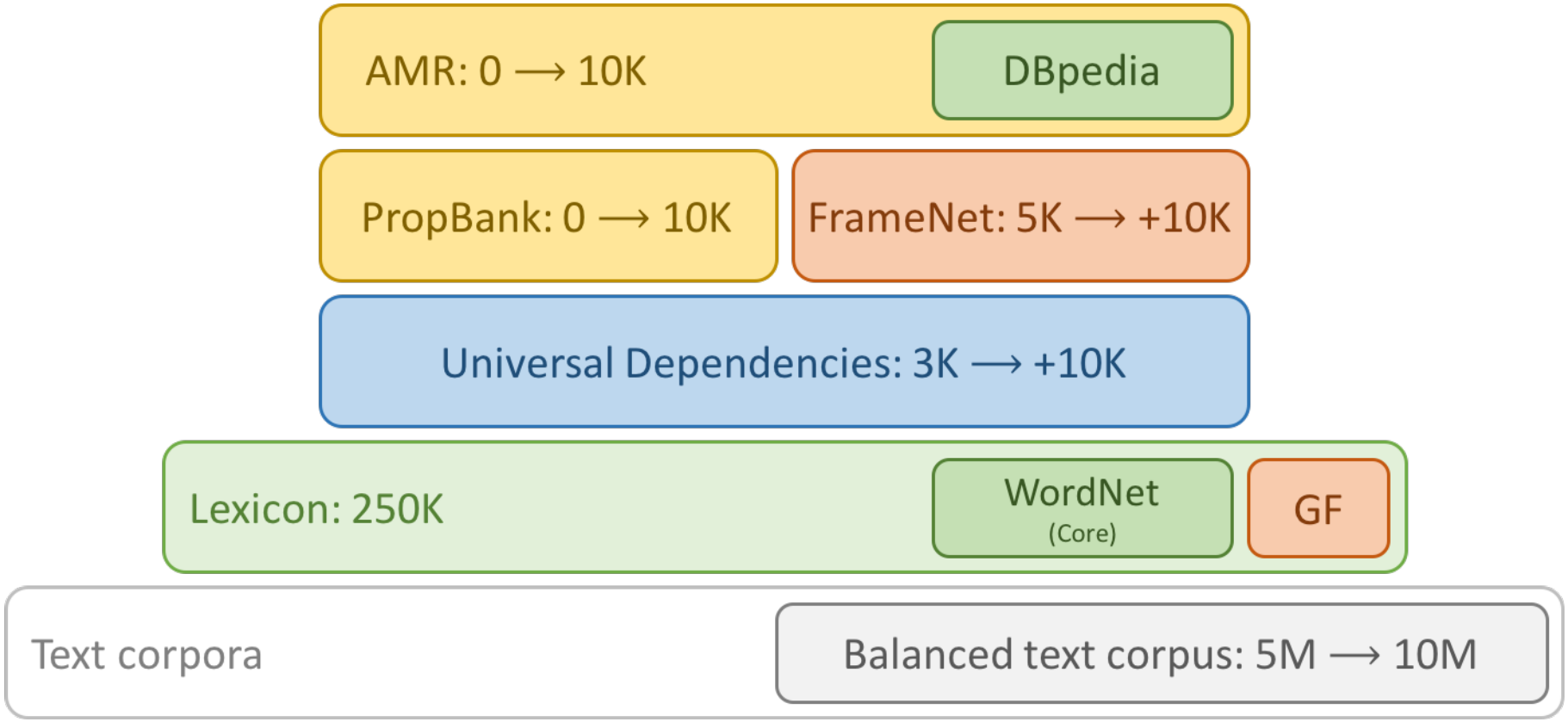
Grant agreement No. 1.1.1.1/16/A/219

Multilingual FrameNet meeting
at ACL 2017

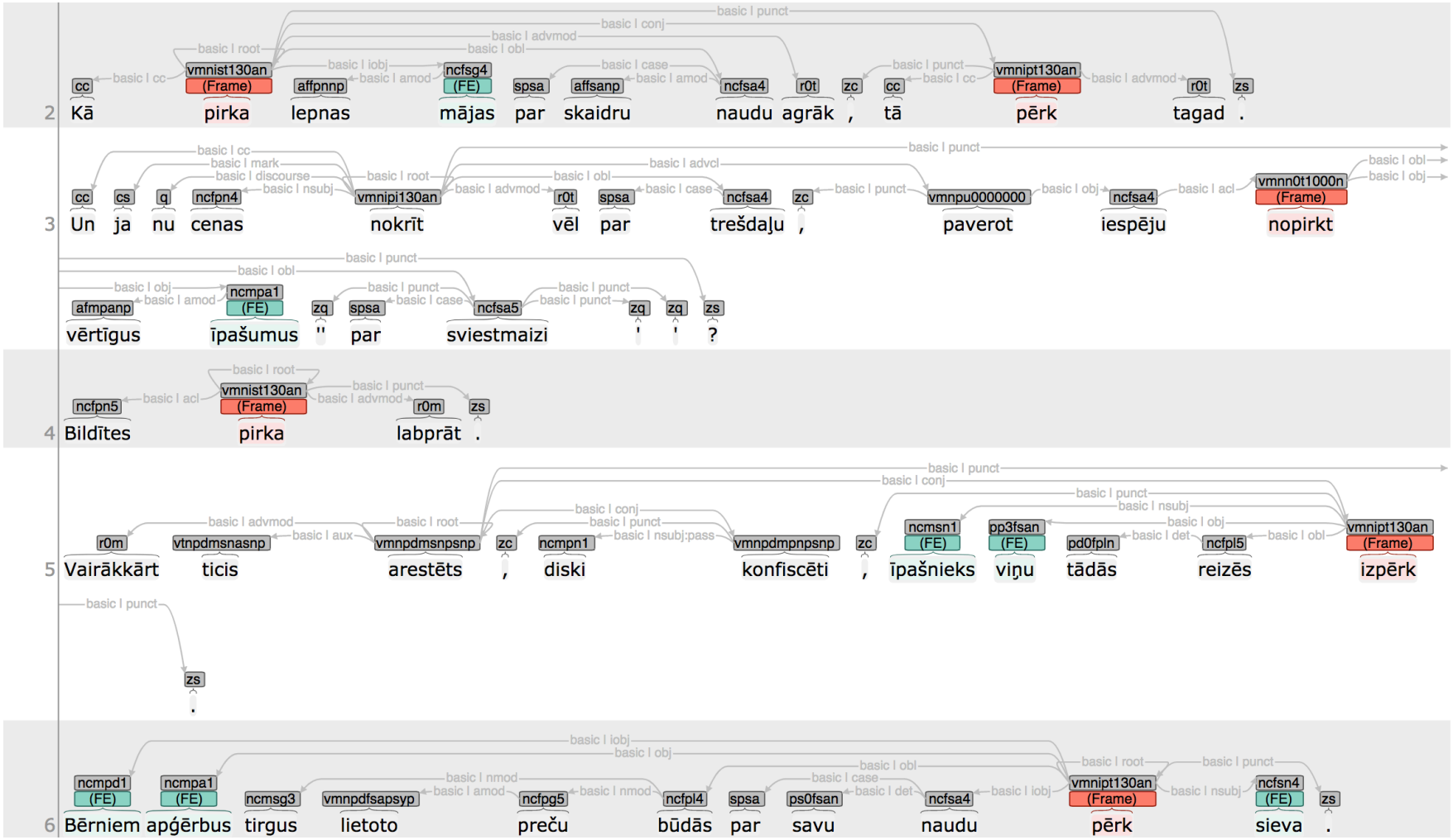
Latvian FrameNet

- **Integrated**: a part of a multi-layered corpus
- **Balanced**
  - Covering at least 2000 common **verbs** with at least 10 examples for each of the 1000 most common verbs
  - A set of paragraphs (vs. docs, or scrambled sentences)
- **Core** FEs only (BFN frames)
  - With exceptions…
  - Skipping the annotation of null instantiations…
- Manually **verified**
  - Including the syntactic layer
- **Accessible** (open data)

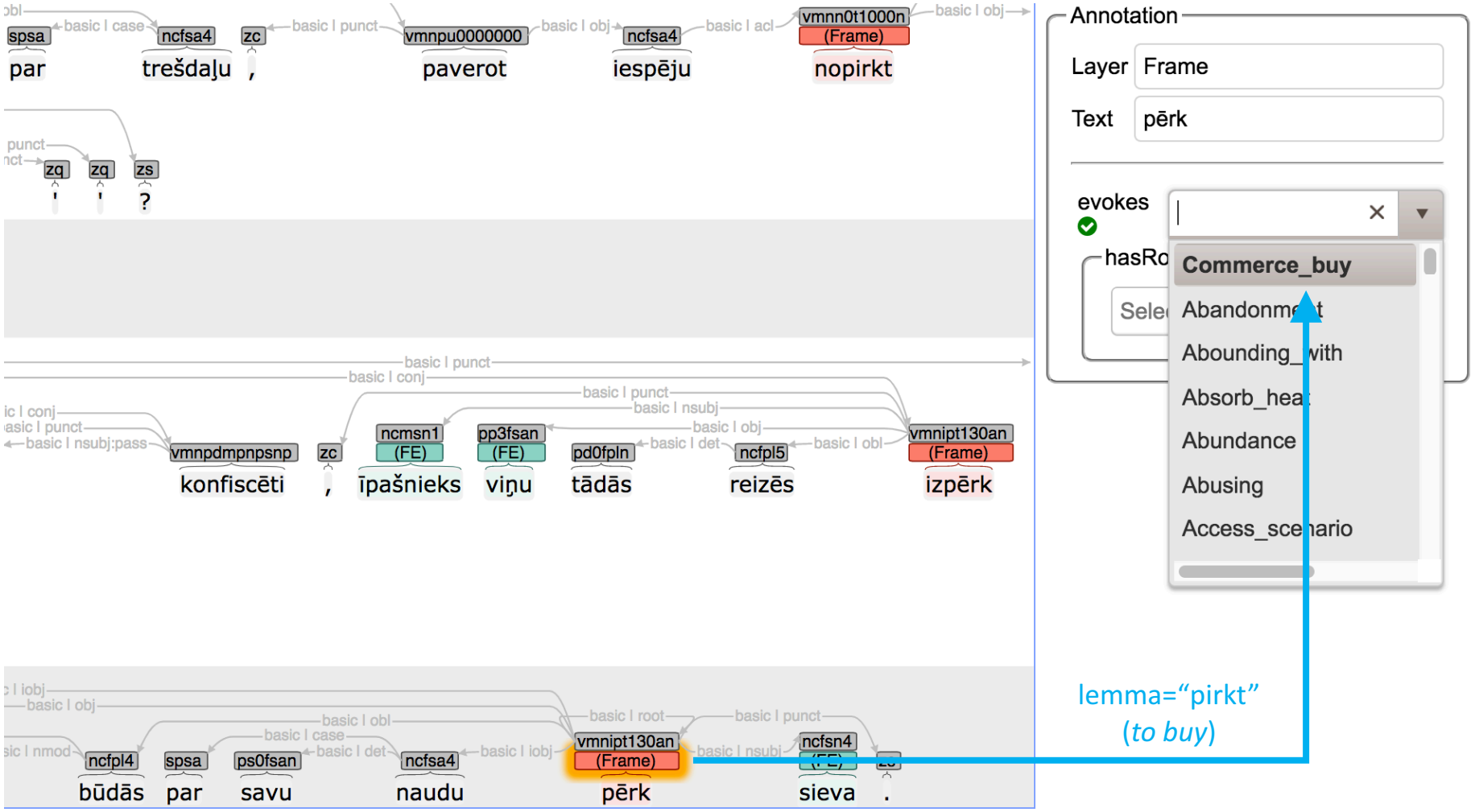
Part of a multilayer corpus for NLU and NLG



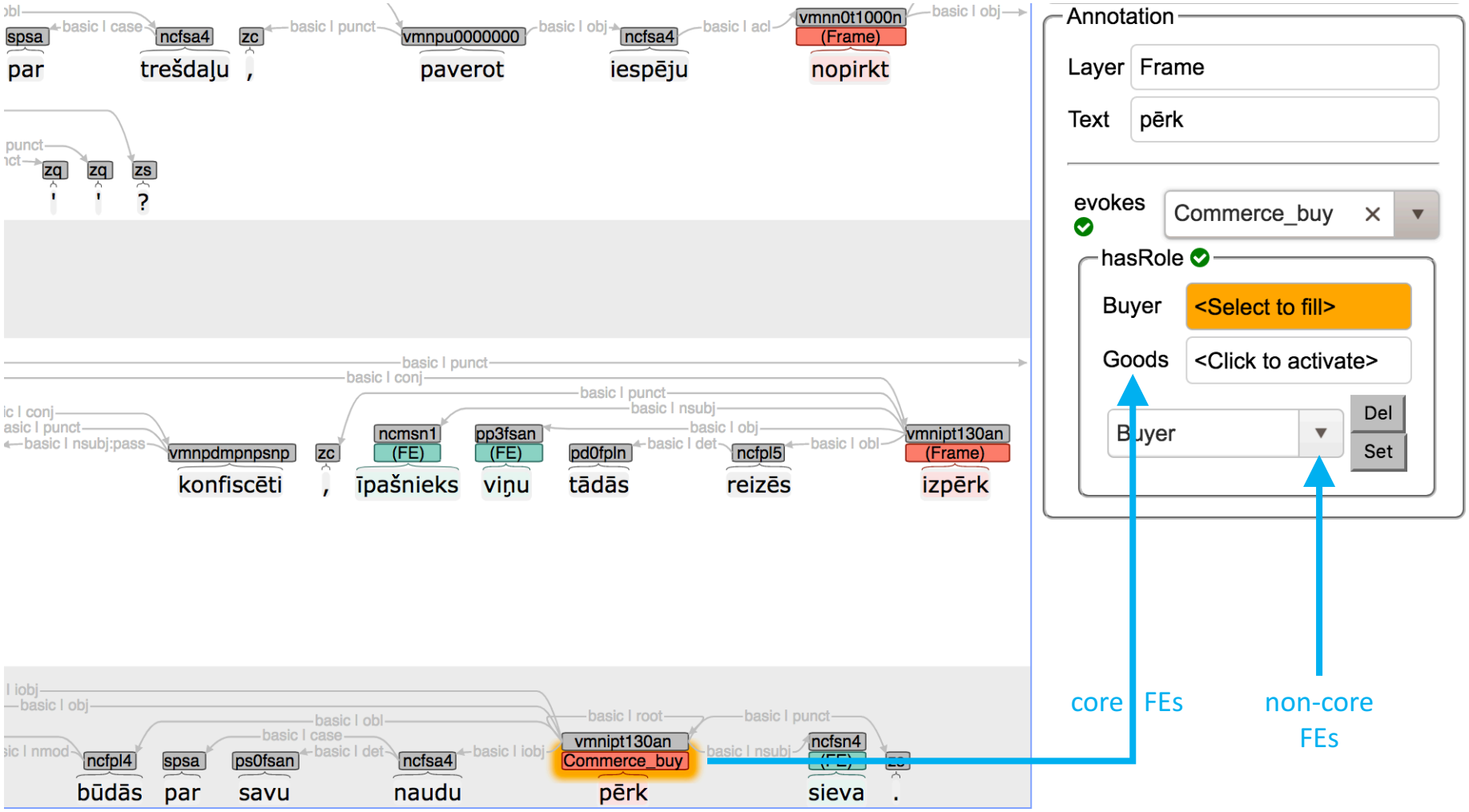
UD-guided annotation in WebAnno

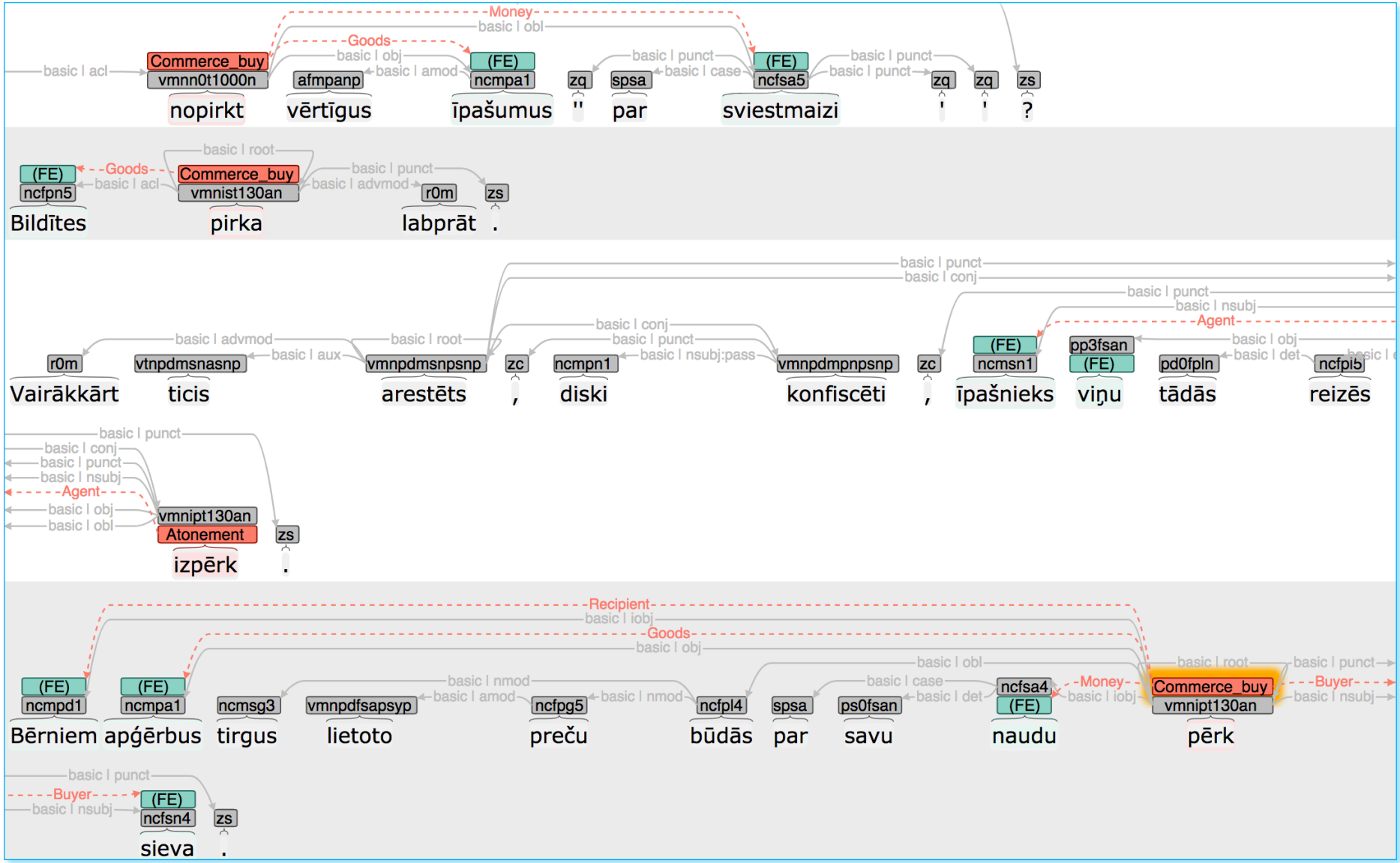


WebAnno constraints: lemma-frame suggestions



WebAnno constraints:
frame-core FE suggestions




Money
basic | obl
Goods
basic | obj
basic | amod
basic | acl
Commerce_buy
vmnn0t1000n
nopirkt
afmpanp
vērtīgus
(FE)
ncmpa1
īpašumus
basic | punct
basic | case
zq
spsa
par
(FE)
ncfsa5
sviestmaizi
basic | punct
zq
'
zq
'
zs
?
basic | root
(FE)
ncfpn5
Goods
basic | acl
Commerce_buy
vmnist130an
basic | punct
basic | advmod
r0m
zs
Bildītes
pirka
labprāt
.
basic | punct
basic | conj
basic | punct
basic | nsubj
Agent
basic | advmod
basic | root
basic | conj
basic | punct
basic | nsubj:pass
basic | aux
r0m
vtnpdmsnasnp
vmnpdmsnpsnp
zc
ncmpn1
vmnpdmpnpsnp
zc
(FE)
ncmsn1
pp3fsan
(FE)
basic | obj
basic | det
pd0fpln
ncfpi5
Vairākkārt
ticis
arestēts
,
diski
konfiscēti
,
īpašnieks
viņu
tādās
reizēs
basic | punct
basic | conj
basic | punct
basic | nsubj
Agent
basic | obj
basic | obl
vmnipt130an
Atonement
zs
izpērk
.
Recipient
basic | iobj
Goods
basic | obj
basic | obl
basic | root
basic | punct
(FE)
ncmpd1
(FE)
ncmpa1
ncmsg3
basic | nmod
vmnpdfsapsyp
basic | amod
ncfpg5
basic | nmod
ncfpl4
spsa
basic | case
ps0fsan
basic | det
ncfsa4
(FE)
Money
basic | iobj
Commerce_buy
vmnipt130an
Buyer
basic | nsubj
Bērniem
apģērbus
tirgus
lietoto
preču
būdās
par
savu
naudu
pērk
basic | punct
Buyer
basic | nsubj
(FE)
ncfsn4
zs
sieva
.

Comments

- Verb-by-verb annotation (vs. doc-by-doc)
- Non-core FEs:
  - Location, Time (important to our downstream apps)
  - *iobj* and other *obj*, which are not core FEs
- Semantic types of FEs (selectional restrictions)
  - FrameNet-style vs. WordNet-style vs. VerbNet-style vs. …
- Data format: CoNLL
- Automatic generation of PropBank annotations (?)
- UD and FrameNet annotations are essential (!)
  - Basic vs. enhanced UD dependencies
  - P.S. FrameNet annotation unveils bugs in the UD layer