

Lõputööde teemasid

Kaili Müürisep

ATI keeletehnoloogia õppetooli vanemteadur

kaili.muurisep@ut.ee

Teemad

- Süntaks
 - Morfoloogiline ühestamine mõne statistilise meetodiga
 - Fraasipiiride tuvastamine statistilise meetodiga
 - Määruslike täiendite ja määruste eristamine statistiliste meetoditega
 - Eesti keele süntaksi fragmente HPSG-s või LFG-s
 - Puudepank
- Sisukokkuvõtete tegemine
 - Töö hindamine
 - Statistilised meetodid
- Muu

Morfoloogiline ühestamine

Eesti

Eesti+0 // _S_ prop sg gen #cap //

Eesti+0 // _S_ prop sg nom #cap //

eesti+0 // _G_ #cap //

vanimad

vanim+d // _A_ super pl nom //

asukad

asukas+d // _S_ com pl nom //

saabusid

saabu+sid // _V_ main indic impf ps2 sg ps af #Intr //

saabu+sid // _V_ main indic impf ps3 pl ps af #Intr //

siia

siia+0 // _D_ //

siig+0 // _S_ com sg gen //

pärast

pärast+0 // _D_ //

pärast+0 // _K_ post #gen //

pärast+0 // _K_ pre #part //

pärane+t // _A_ pos sg part //

pära+st // _S_ com sg el //

viimast

viimane+t // _A_ pos sg part //

vii+mast // _V_ main sup ps el #NGP-P //

jääaega

jää_aeg+0 // _S_ com sg adit //

jää_aeg+0 // _S_ com sg part //

Morfoloogiline ühestamine

- Statistilised ühestamise meetodid:
 - TnT tagger
 - Support Vector Machines
 - jmt

Fraasipiiride leidmine

[Teadmata kadunuks tunnistatud isiku taasilmumisel] või [tema viibimiskoha kindlakstegemisel] tähistab [kohus] [isiku enda või huvitatud isiku nõudel] [tema teadmata kadunuks tunnistamise] .

Määruslikud täiendid vrs määrused

- Ta võttis praeahjust panni
- Ta võttis vasest ahjuroobi
- Lapsed läksid koeraga onu juurde
- Nõnda ongi lihtsakoeline Arnold kabinetivõitluses (ADV L <NN NN>) kaotajaks jäänud

Sügavam grammatikakirjeldus

- Pindsüntaktiliselt märgendatud korpuse teisendamine puudepangaks
- LinGo Grammar Matrix (HPSG)
- XLFG või GFU-LAB (LFG)

Sisukokkuvõtted

Kuningannat ootab Tallinnas karu

Suurbritannia kuningannat Elizabeth II ootab Tallinna saabudes ees karukostüümis loomakaitseühenduse PETA liige, kes protestib sel moel brittide kaardiväes kantavate ehtsast karunahkadest mütside vastu.

PETA liikmed kutsuvad üles Suurbritannia kuningannat Elizabeth II ja Suurbritannia kaitseministeeriumi lõpetama mustade karude massimõrv sümboolsete peakatete nimel, on kirjas Delfile saadetud teates.

Suurbritannia parlamendi üleskutse nr 1756, mis kutsub üles kasutama kunstkarusnahka kaardiväe mütsides, on kogunud üle 200 kaastundliku Suurbritannia parlamendisaadiku allkirja.

PETA karu külastab ka Tallinna, et kuninganna külaskäigu ajal tõmmata tähelepanu kaardiväe karunahast mütside kasutamisele.

Sisukokkuvõtted

- Sisukokkuvõtja töö hindamine
- Statistilistel õppimismeetoditel põhinev sisukokkuvõtja
- Sisukokkuvõtja sidumine lingvistilise mooduliga
- Sisukokkuvõtja kasutajaliides ning tekstide eeltöötlus

Suulise keele teemad

- Morfoloogilise ühestaja kohandamine suulisele keelele

*ee ja kui kaua h mi mis ajal peab puh noh enne kaitsmist ütleme
mul on viieteistkümnes millal ma pean selle bakalaureusetöö
esitama*

- Uue meedia keel: jutukad, blogid, foorumid,
portaalikommentaarisid

Grammatikakorrektor

- Reeglite formuleerimine ja testimine

Programmeerimine

- VISL-parseri koodi kaasajastamine (C++)
- Visualiseerimisvahendite ja puuredaktori kirjutamine/parandamine (Java)

Muud

- Vene keele kitsenduste grammatika
- Kogu eelnev jutt mõne teise keele jaoks

Muud

- unistus kirjutada
 - enneolematu masintõlkesüsteem
 - krimiromaanide generaator
 - ja palju muud