

## SÕNALIIK KUI RAKENDUSLIK JA LINGVISTILINE PROBLEEM: SÕNALIIKIDE MÄRGENDAMINE VANA KIRJAKEELE KORPUSES

Külli Habicht, Pille Penjam, Külli Prillop

**Ülevaade.** Artikkel käsitleb sõnade liigitamisega seotud probleeme, lähtudes eesti vana kirjakeele korpuse poolautomaatse märgendamise pikaajalisest praktikast. Tutvustatakse kasutusel olevat märgendus-süsteemi ning põhilisi raskusi selle tegelikul rakendamisel. Universaalsed probleemid on seotud sõnaliikide piirialadega, kuhu kuuluvate sõnade liigi määratlemine saab toimuda vaid üksuse funktsiooni arvestades ja konteksti põhjal. Spetsiifilisemad probleemid on seotud vana kirjakeele tekstide eripäraga – tõkelisuse ja sellega seotud võõr-mõjude, arhailisuse ning välja kujunemata grammatikatradiitsiooniga. Sõnaliikide piirialade, eriti adverbide üle otsustamine on välja toonud huvitavaid leksikaliseerumis- ja grammatiseerumisjuhtumeid, mis avavad kirjakeele leksikaalsete üksuste varasemaid arenguetappe. Funktsiooni alusel otsustamine võimaldab arvestada konteksti ning üksuste universaalse taaskasutatavuse seaduspärasust (üks ja sama üksus on ökonoomiaprintsiibile vastavalt kasutusel eri funktsioonides). Artikkel tutvustab korpuse kasutaja jaoks paljude probleemsete sõnaliigiotsustuste tagamaid ja toob esile just vana kirjakeele spetsiifikaga seotud keerukaid juhtumeid.\*

**Võtmesõnad:** morfosüntaktiline märgendamine, korpuslingvistika, vana kirjakeel, eesti keel

### 1. Sissejuhatus

Sõnade liigitamiseks sobivate aluste otsimine, lähtudes sõnade tähendusest, muutmisviisist ja süntaktilise kasutamise võimalustest, on probleem, millega keelesüsteemi kirjeldades on tegeldud antiikajast alates, ent mis tekitab endiselt

\* Artikli valmimist on toetanud Haridus- ja Teadusministeerium (sihtfinantseeritav uurimisteema SF0180084s08 "Eesti keele morfosüntaktiline ehitus ja areng" ning riikliku programmi "Eesti keel ja kultuurimälu" projekt "Eesti vana kirjakeele sõnastikud").

küsimusi, kui uurija peab määratlema sõnade liigikuuluvust konkreetsetes tekstis ja kindlal rakenduslikul otstarbel.

Hulk spetsiifilisi probleeme kaasneb sõnaliikide määramisega ka vana kirjakeele tekstide märgendamise käigus, sest siin lisanduvad tänapäevastele ja enamjaolt juba varemgi täheldatud probleemidele veel tekstide arhailisuse ja tõlkelisusega seotud küsimused, samuti tekstide kirjapanijate erinevast keeleoskuse tasemest sõltuvad subjektiivsed ümbertõlgendused morfosüntaktiliste struktuuride edasiandmisel olukorras, kus grammatikakirjeldustraditsioon puudus. See küsimuste ring on seotud sõnaliigi määratlemise kui teoreetilise probleemiga.

Sõnaliik kui informatiivne morfosüntaktiline kategooria on korpuse sõnedele otstarbekas lisada ennekõike päringute kitsendamise ja täpsustamise huvides – see on sõnaliikide märgendamise põhiline rakenduslik eesmärk. Selline märgendus hõlbustab kasutaja jaoks eri tüüpi üksuste, nt suletud sõnaklassidesse kuuluvate grammatiliste sõnade uurimist, sest paljud spetsiifilist pädevust ja vana kirjakeele teatud perioodi tundmist eeldavad keerukad otsused on märgendaja uurija eest ära teinud. Kuna korpuse lemmatiseerimine toimub sõnehaaval ja konteksti arvestades poolautomaatselt, suudab kogenud märgendaja üldjuhul hõlpsasti otsustada, missuguses funktsioonis sõna lauses esineb. Sõnaliikide märgendamine võimaldab seega täpsustada sõnade tekstilise kasutuse eripära ning eristada ühtlasi pärisnimesid, partitsiipe jm eri tunnustega sõnaklasse. Sõnaliik on oluline kategooria ka süntaktilise analüüsi alusena. Eri keelte sõnaliigisüsteeme uurides on selgunud universaalne nähtus, et üksused kalduvad uutes funktsioonides taaskasutamise poole. (Vt täpsemalt Anward 2000, Hengeveld 1992) Seda tendentsi on esile toodud ka eesti keele sõnaliikide süntaktilist ja ennekõike semantilist transpositsiooni analüüsid (Erelt 1980). Taaskasutusprotsessi on võimalik hästi jälgida just diakroonilisest vaatenurgast. Niisiis pakub vana kirjakeele tekstide problemaatiline ja sõnaliikide piirialadele kuuluv aines selgitusi ka üksuste funktsioonide ajas hargnemise kohta.

Vana kirjakeele korpust märgendades tuli esmalt lahendada probleem, misugusest liigitussüsteemist lähtuda, tagamaks mõistliku täpsusega eristust, mis ei suruks vanale materjalile peale mõnd välist või sobimatut süsteemi, kuid ei oleks samal ajal ka liiga arhailine ja sellisena tänapäeva uurijale läbipaistmatu. Põhikriteeriumiks sõnaliikide praktilisel eristamisel on grammatilised karakteristikud: sõna muutuvus/muutumatus ja süntaktiline funktsioon. Ideaalne sõnaliikide märgendamise süsteem peaks võimaldama samal ajal arvesse võtta nii sõna üldkategoriaalset kuuluvust ja morfoloogilist käitumist kui ka süntaktilist ümbrust. Sageli tuleb aga tegeliku märgendustöö käigus üks kriteerium teis(t)ele allutada.

Siinse käsitluse eesmärgiks on tutvustada vana kirjakeele korpuse morfosüntaktilise märgendamise tarbeks kohandatud sõnaliigisüsteemi. Tähelepanu keskmes on sõnaliikide piirialasid puudutavad praktilise määratlemise probleemid, mis loodetavasti aitavad vana kirjakeele korpuse kasutajal paremini mõista üksikute sõnaliigiotsustuste tagamaid. Selliste probleemsete piirialade üle on eesti lingvistikas ka varem põhjalikult arutatud (vt nt KKSM 1977, Karelson 1972, 1985, Erelt 1980; EKK 2007: 168–200, Vare 2006). Näitematerjal pärineb TÜ vana kirjakeele korpusest<sup>1</sup>, põhiliselt 17. sajandi esimese poole tekstidest, mille märgendamisega on seni kõige rohkem tegeldud. Näitelausest on kõnealuste probleemsete sõnade

<sup>1</sup> <http://www.murre.ut.ee/vakkur/korpused/> (25.09.2010).

järele nurksulgudesse lisatud praegu korpusel olev sõnaliigimärgend. Osa seni märgendatud korpusetekstidega saab tutvuda ka internetis<sup>2</sup>.

## 2. Sõnaliikide eristamise teoreetilised lähtekohad

Enamik nüüdiskeelte puhul kasutatavaid sõnade liigitamise süsteeme lähtub kreeka grammatiku Dionysios Thraxi klassikalise kreeka keele kohta loodud kaheksa sõnaliigi käsitlusest (nimisõna, tegusõna, asesõna, eessõna, määrsõna, sidesõna, kesksõna, artikkel). Sõnaliikide hulk keelekirjelduses sõltub ühelt poolt keele iseärasustest ja teiselt poolt liigitamise eesmärgist (nt tüpoloogiline *vs.* üksikkeele seisukohast ammendav; automaatrakenduseks *vs.* käsitsi märgendamiseks mõeldud). Siiski seob eri käsitlusi arusaam, et ühte sõnaliiki kuuluvad üksused jagavad omavahel semantiliste, süntaktiliste ja morfoloogiliste omaduste ainulisi kimpe (Anward 2000: 3). Universaalne on seegi, et sõnaliigisüsteemidele on omane läbipõimumine ja koguni teatav läbipaistmatus: ükskõik missugust kriteeriumi (tähendust, süntaktilist funktsiooni või muutmise eripära) ka sõnaliigi määratlemisel ei arvestataks, on suhe üksikkriteeriumi ja -sõnaliigi vahel ikkagi mitu mitmele, s.t ei kehti ideaalne vastavus “üks tähendus – üks funktsioon – üks vorm”. Nii seostubki enamik sõnaliike mitme semantilise kategooriaga, nt nimisõnad ei väljenda mitte ainult isikuid ega asju, vaid ka sündmuse-, koha-, aja-, hulga- jm mõisteid. Niisama mitmetine on suhe ka süntaktilise funktsiooni ja sõnaliigi vahel, nt saab nii nimisõnu, tegusõnu kui ka omadussõnu kasutada nii argumendi, predikaadi kui ka modifikaatorina. Isegi morfoloogiline kriteerium, sõnamuutmine, ei ole sageli ühemõtteline otsustusalus (Anward 2000: 3–4), sest sõnaliikide piirialadel võib esineda juhtumeid, kus verbil on käändelisi muuteid või mõni käändsõna tegelikult ei käändu.

Keeled kalduvad leksikaalsete ressursside optimaalse kasutamise poole, mis tähendab kalduvust taaskasutada samu üksusi mitmes funktsioonis. Üksuse uus funktsioon peab olema eristatav, mistõttu tuleb keeles saavutada tasakaal paindlikkuse (taaskasutuse) ja eristuvuse (identifitseeritavuse) vahel. Selline tasakaal tõkestab ühtlasi kõikide üksuste täielikku taaskasutatavust ning seetõttu on keeltes kasutusel erinevaid sõnaliigisüsteeme. (Anward 2000: 38–39)

Eesti grammatikakirjeldustes on pikka aega käibinud üheksa sõnaliigi süsteem, mis oli kasutusel juba esimeses eestikeelses eesti keele grammatikakäsiraamatus, Karl August Hermannini grammatikas (Hermann 1884) ning jäi käibe peaaegu kogu 20. sajandi vältel. Alates 1990. aastate algusest kirjeldatakse aga süsteemi 12-liikmelisena (EKG I: 18). Sõnaliikide hulga suurenemine on toimunud adverbide klassi funktsionaalse täpsustamise arvelt, st varasemale üheksale sõnaliigile (nimisõna, omadussõna, asesõna, arvsõna, tegusõna, määrsõna, kaassõna, sidesõna ja hüüdsõna) on lisandunud rõhumäärsõna, asemäärsõna ja abimäärsõna. Sõnaliigiteemalist teoreetilist arutelu on edendanud 1977. a ajakirjas Keel ja Kirjandus ilmunud diskussioon, milles osalesid Ülle Viks, Mati Ereht ja Tiit-Rein Viitso (KKSM 1977: 520–537) ning kust jäid kõlama mõnedki hilisemate käsitluste jaoks tausta loovad seisukohad, nt see, et ühte sõnaliiki kuuluvatel sõnadel on ühesugused tekstilise kasutamise võimalused (KKSM 1977: 524); iga sõnaliiki iseloomustab teatav

<sup>2</sup> [http://www.murre.ut.ee/vakkur/Korpused/Otsi/mrg\\_paring.htm](http://www.murre.ut.ee/vakkur/Korpused/Otsi/mrg_paring.htm) (25.09.2010).

diferentsiaaltunnuste kimp (KKSM 1977: 525); sõnaliikide kirjeldamiseks sobib tsentri-perifeeria põhimõte (vt ka Erelt 1980); ebamäärasus, mis on keelesüsteemi omadus, on omane ka sõnaliigisüsteemile (KKSM 1977: 526); morfoloogiline sõnaliik ei tingi süntaktilist (KKSM 1977: 530). Tiit-Rein Viitso on mõttearenduse tulemusena välja pakkunud originaalse sõnaliigimudeli, mis teeb vahet käänd-, pöörd-, kõrval- ja püsisõnadel (KKSM 1977: 535).

Eesti keele sõnaliikide problemaatikat on viimase kümne aasta jooksul käsitlenud nii eri allkeelte arvutirakendustega seoses, nt kirjakeele korpuse (Kaalep jt 2000), suulise keele korpuse (Hennoste 2002) ning murdekorpuse märgendamise aspektist (Lindström jt 2006), kui ka sõnaliigilist transpositsiooni ja konversiooni puudutavalt (Vare 2004, 2006).

### 3. Vana kirjakeele korpuses eristatud sõnaliigid

Vana kirjakeele tekstide morfosüntaktilisel märgendamisel on praeguse grammatikakirjelduse 12 sõnaliigi süsteemi laiendatud mõne vanade tekstide analüüsil asjakohaseks osutunud sõnaliikide alamklassiga (sõnaklassiga). Lisaks EKG sõnaliikidele on lisatud vanade tekstide eripära arvestavad või täpsustavad kategooriad *artikkel*, *nimisõna (nimi)*, *partitsiip* ja (*muutumatu*) *täiendsõna*. Eristatavate sõnaliikide hulga üle otsustamisel võtsime arvesse ka poolautomaatse märgendamise võimalusi, mis lähtuvad korpuse märgendamisel abiks olevast Filosoofi morfoloogiamoodulist ESTMORF<sup>3</sup>.

Artikliteks võib märgendaja määrata selged definiitsuse väljendamisega seotud näited, kus tõlketekstide mõjul on eesti keeles kasutatud määravat artiklit *see* või umbmäärast artiklit *üks*, aga kahtluse korral ei pea seda tegema. Nii ongi praegu enamik sõnade *see* ja *üks* kasutusjuhtudest märgendatud pronoomeniteks. (Vt ka jaotis 7.)

Pärisnimede eraldi märgendamine võimaldab nimeuurijatel edaspidi kergemini nimematerjalile ligi pääseda. Pärisnimesid on vaja eristada ka seetõttu, et nende vormistik erineb vana kirjakeele tekstides muude noomenite omast, nt on nimedes kasutatud ladina käändelõppe.

Partitsiipide omaette sõnaklassiks nimetamine on põhjendatud sellega, et sageli on tekstides raske vahet teha, kas autor on mõnd partitsiibivormi tõlgendanud verbilisena või pigem atributiivse/predikatiivse (vt ka jaotis 4). Partitsiibi märgendi alusel on uurijatel võimalik edaspidi selline materjal hõlpsasti kätte saada ja vajaduse korral endal otsustada, missuguse funktsiooniga on tegemist. Sellise kategooria kasutuselevõtt aitab lahendada olemusliku probleemi kesksõna kui adjektiivi ja verbi sõnaliigi keskel paikneva kategooria määramisel ning võiks sobida ka eesti üldisesse sõnaliikide süsteemi.

Muutumatu täiendsõna omaette sõnaklassiks pidamise on tinginud taas vanade tekstide eripära, s.t raskus eristada, kas mõne täiendsõna näol on tegemist muutumatu sõna, substantiivi või adjektiiviga, ja otsustada, missuguses käändes see vorm on (nt *viigipuu*). Niisiis on tegemist vana kirjakeele spetsiifilise probleemiga.

Tabelist 1 on näha, missugustel sõnaliikidel on teiste sõnaliikidega kõige rohkem probleemseid kokkupuutealasid. Probleemseim sõnaliik on adverb, mis avatud sõnaliigina võtab vastu uusi leksikaliseerunud üksusi igat liiki noomenite, samuti

verbide kivinenud vormide hulgast. Määratlemisprobleeme on ka adjektiividel ja substantiividel omavahel ning mõlemal neist piirialadel adverbi, adpositsiooni ja afiksaaladverbiga.

**Tabel 1.** Ülevaade sõnaliikide piirialadest. Numbrid märgivad artikli jaotisi, kus vastavat probleemi käsitletakse

	V	S	Adj	Num	Konj	Adv	AfAd	PrAd	MdAd	Pron	Adp	Int
V		4	4			4			4			
S	4		5			5	5				5	
Adj	4	5				6	6				6	
Num						7				7		
Konj						8		8				8
Adv	4	5	6	7	8		10	10	10	9		
AfAd		5	6			10					9	
PrAd					8	10						
MdAd	4					10						
Pron				7		9						
Adp		5	6				9					
Int					8							

#### 4. Verbide (sh partitsiipide) eristamine teistest sõnaliikidest

Verb on üldjuhul kõige lihtsamini eristuv sõnaliik, mis muutub pöördes, ajas, kõneviisis, tegumoes ja kõneliigis ning esineb lauses predikaadina (EKG I: 19). Probleeme tekitavad enamasti verbi käändelised vormid, mis oma funktsioonilt ongi pigem käändsõnalised või adverbilised üksused. Kuni 1950. aastateni peeti näiteks *mine-* ja *ja-*tuletisi tegusõna käändelisteks vormideks, mitte substantiivideks (nt Jänes 1942: 68), sest neid on võimalik moodustada enamikust verbitüvedest. Samamoodi on võimalik enamikust verbitüvedest moodustada *nu-* ja *tu-*tuletisi, mis suure vormihomonüümia tõttu jäävad **substantiivide** ja partitsiipide piirialale. Enamasti võib vanades tekstides iseseisva lauseliikmena toimivaid (sageli käändetunnusega) vorme pidada substantiivideks (1), (2).

- (1) SÄhl olli ütZ Inimene/ se .. lähatis oma sullaset welja .. ütlemale neile **palwel-tuile** [S]: Tulcket/ sest kihk om walmis. (Rossihnius 1632: 121)
- (2) Ke agkas leeha himmode sees ellap/ se on ellaw **surnut** [S]. Welche aber in Wollüsten lebet/ die ist lebendig **Tod**. (Stahl 1641: 109)

Ulatuslik mitmeti tõlgendatav rühm on **adjektiivide** ja verbi käändeliste vormide piiriala. Verbi partitsiibivormidel on tüüpiliselt hulk adjektiividele omaseid tunnuseid (esinemine täiendi ja öeldistaitena, oleviku partitsiipide puhul ka ühildumine), samas esinevad mineviku partitsiibi vormid eesti keeles ka liitjavormide koosseisus ja nende laiendid on tüüpilised verbilaiendid. Kuna prototüüpset verbaalsete (3) ja prototüüpset adjektiivsete (4) mineviku partitsiipide kasutuste osakaal on võrdlemisi väike ning suur hulk juhtumeid kuulub adjektiivide ja verbide vahelisse hajusasse üleminekualasse, kus sõnaliigi ühemõtteline määratlemine osutub keerukaks (5), on korpuses mineviku partitsiipide jaotamisest adjektiivideks ja verbideks loobutud – eristatakse iseseisvat partitsiibi sõnaklassi, millesse on paigutatud kõik korpustes esinenud *nud-* ja *tud-*partitsiibi vormid (v.a substantiivsed, vt eelmine lõik).

- (3) Kæ Tajwa ninck Maa on **lohnut** [Pts], se læßib iße kuiwa holkidde pæl .. (Müller 7: 3)
- (4) .. eth temma hend lasckis arwama/ Maria omma **towotut** [Pts] Pruti kahs .. (Stahl 1638: 10)
- (5) .. keick mea sinna mah pehl siddoma saht/ peap kahs taiwa sees **siddotut** [Pts] olla/ ninck keick/ mea sinna mah pehl pehstma saht/ peap kahs taiwas **pehstetut** [Pts] olla. .. Alles was du auff Erden binden wirst/ sol auch im Himmel **gebunden** sein/ Vnd was du auff Erden lösen wirst/ sol auch im Himmel **los** seyn. (Stahl 1638: 189)

Oleviku partitsiibid, mis täiendina oma põhjaga ühilduvad ja on enamasti leksikaliseerunud, saavad pigem adjektiivse tõlgenduse, nt

- (6) Needt lessedat peawat kahs Hanna jallaasemedesse sisse astma/ omma lohtusse pannema se **ellawa** [Adj] Jummala pehle .. Die Witwen sollen auch in Hanna fußstapffen treten/ jhre hoffnung stellen auff den **Lebendigen** GOTT .. (Stahl 1641: 108)

Adjektiivilaadsena on paljudel juhtudel kasutatud ka supiini abessiivivorme, mis niisugustel juhtudel ongi süntaktilise funktsiooni alusel loetud adjektiivideks (7). Paljudel juhtudel ei ole piir adjektiivse ja verbilise tarvituse vahel siiski kuigi selge (8).

- (7) Temma sai toibatut se Paradysi sisse/ ninck kuhlis ilma **weljapajata-matta** [Adj] sönnat .. (Rossihnius 1632: 93)
- (8) Minna ei taha teile/ armat weljet/ sedda mitte **ütlematta** [V] olla/ et meije wannambat ommat kihck se pilwe all olnut .. (Rossihnius 1632: 90)

Päris selget piiri ei saa tõmmata ka teatud sõnadest moodustatud üksikute verbivormide ja **adverbide** vahele, sest eesti vanemas kirjakeeles leidub rohkesti tähenduselt iseseisvunud verbivorme, mis on tekstilise kasutuse seisukohalt adverbid, ehkki kujunenud verbilisest lähtevormist. Sellesegi sõnarühma kuuluvad eelkõige *mata*-vormid, mis on leksikaliseerumise käigus adverbistunud, väljendades viisitähendust, nt (*ilma*) *kogemata*, *ilmalõpmata*, *ilmajällesjätmata* 'järelejätmata, lõpmata', (*ilma*) *teadmata*, *ilmaväljapajata* 'ärarääkimata' jt. Adverbistunud on ka osa *des*-gerundiivi vorme, nt *elades* 'iialgi, eales', *silmanähes* (vt Habicht 2001: 53–67). Kõik sellised juhtumid on korpuses märgendatud adverbideks.

Verbivormist *ei* on tänapäevaks kujunenud **modaaladverb**. 17. sajandil kuulus *e*-tüveline eitusverb arvatavasti veel (vaegmuuteliste) verbide hulka. Tänapäeval esinevad vaegmuutelisena sarjad *pole* : *polnud*..., *ära* : *ärgem* : *ärge* : *ärgu* (EKG I: 20) Vanimates kirjalikes tekstides on üksikutel juhtudel kasutatud pöördvorme *en*, *emme* ja *ewat*. Arvestades nende vormide üsna juhuslikku esinemust, pole meid märgendanud verbideks, vaid negatsiooni väljendavateks modaaladverbideks (9).

- (9) Semperrast **en** [MdAd ] kartame meije hendes/ kus se Jlm hucka lehx/ .. (Stahl 1641: 207)

## 5. Substantiivide eristamine teistest sõnaliikidest

Substantiivid on tüüpjuhul käändes ja arvus muutuvad sõnad, mis nimetavad asju ning võivad lauses esineda subjekti ja objekti funktsioonis (EKG I: 20). Piiriala olemasolu on olemuslik ja universaalne substantiivide ning **adjektiivide** puhul (vt Erelt 1980, Hengeveld 1992, Pajunen 1998). Eesti keeles on väga levinud adjektiivide substantiveerumine ja ka mõningate substantiivide adjektiveerumine (vt EKG I: 20–21). Keeleteadlaste hulgas on aga erinevaid arvamusi selle kohta, kas sõna võimalikku kasutamist nii adjektiivi kui ka substantiivina (*haige, pime*) “pidada konversiooniks ja üldse sõnamoodustuseks või on siin tegemist lihtsalt noomenikasutuse suhteliselt avarate võimaluste ja süntaktilise ellipsiga” (Kasik 1996: 10). Seda tüüpi markeerimata sõnaliigimuutus e konversioon on levinud paljudes keeltes. Eesti keeles on seda ulatuslikult kasutatud juba vana kirjakeele perioodil (Habicht 2002). Eesti keele konversiooninähtusi on põhjalikult käsitlenud Silvi Vare (2004, 2005, 2006)

Enamasti taandubki mitmeti tõlgendatavuse probleem sellele, kas vaadelda noomenifraasi väljajäetelisena (sel juhul oleks fraasi põhjaks juurdemõeldav nimisõna ning kaheldava analüüsiga sõna saaks adjektiivi tõlgenduse) või mitte, nt (10).

- (10) Minna ütle teile/ nünda sahb kahn röh m ollema Taiwa sissen/ ülle ütte pattatze Inimesse/ ken omma pattut parrandab/ üttesa ninck üttesa-kümne **euckede** [Adj] ehst/ kelle sedda parrandamist ei mitte tarbis olle. (Rossihnius 1632: 123)

Substantiivse tõlgenduse poole suunab küsitavale sõnale eelnev artikkel või demonstratiivpronoomen (11), (12), samuti paiknemine loendis koos teiste substantiividega (13). Sageli on otsustamisel abi ka saksakeelsest rööptekstist.

- (11) Kus kekit pajatap/ Minna armastan Jum~ala ninck weehap om~a wenda/ se on üx **wallelick** [S]. So jemand spricht/ Jch liebe GOTT/ vnd hasset seinen Bruder/ der ist ein **Lügener**. (Stahl 1638: 91)
- (12) Ach kikeweggeff Iumal ../ ken sina need **üllembat** [S] ninck **Wañambat** [S] essi sädnut/ ninck neine Kunnigatte söamet omma käghe sissen peat .. (Rossihnius 1632: 49)
- (13) Seperrast küssi mina sinust/ kabs sinna tahat sesinatze N. Oma Abbi-kahsas wötta/ temma kahn kannata/ mea Iumal teije pähle panneb/ **hä** [S] ächk **Kurja** [S]/ Waisusse ninck Rickusse .. darumb/ wolt jhr diese N. zum Ehelichen Gemahl haben/ mit jhr leben vnd tragen/ was GOTT euch aufferleget/ **Glück** vnd **Vnglücke** Armuht vnd Reichthumb .. (Rossihnius 1632: 38)

Tänapäeva eesti keelele on iseloomulik *ja*-liiteliste noomenite rohkus substantiivide ja adjektiivide vahelises perifeerses alas (EKK 2007: 177). Mitut moodi käituvaid *ja*-tegijanimesisid esineb ka 17. sajandi esimese poole tekstides. Sageli määrab kontekst nende tõlgenduse üheselt (14), kuid esineb ka märgendaja jaoks probleemsemaid lausestruktuure (15). Sellistel puhkudel tuleb teha otsus iga konkreetse

juhtumi korral eraldi, arvestades konteksti ja eelistades kahtluse korral prototüüpset sõnaliiki (*ja*-liiteliste noomenite korral substantiivi).

- (14) Olcket söhmata ninck johmata/ ninck walwket/ sest teye wasto saibja/ se kurrat keub ümber/ kui ützt **mürriseja** [Adj] löuw. (Rossihnius 1632: 122).
- (15) Olcket **maja wötjat** [S] teine teise sean ilmanürrisematta/ ninck the-nitzeket teine teist .. (Rossihnius 1632: 114)

**Adpositsioonide** sõnaklassi osaline produktiivsus ja pidev täienemine noomenite käändevormide arvel on eesti keelele iseloomulik protsess (Sepper 2006, Habicht, Penjam 2007: 52), mis muudab hägusaks substantiivide ning adpositsioonide piiriala. Näiteid grammatiseeruvatest, substantiivide ja adpositsioonide piirialale jäävatest noomenitest võib rohkesti leida ka vanema kirjakeele tekstidest, nt

- (16) SIhn om ützt par raghwast N. N. need tachtwat hendas Iumala sädusse perra se abbi ello saissusse sisse anda/ ninck pallewat/ et meije Risti Palwe neine ehst/ peame teggema/ et nemat Iumala **Nimmel** [Adp] woisit nackada/ ninck neine kessi hästi woiß keuwa; ES wollen sich nach dem willen GOTTes in den H. Ehestand begeben/ N. N. begehren das gemeine Christliche Gebett für sich/ das sie es **in** GOTtes **Nahmen** anfahen vnd wolgerahte. (Rossihnius 1632: 38)
- (17) Pahstuda ninck **ihuliko kombel** [Adp] hendas walmista/ om küll ützt hä weljalick kassindus. Fasten vnd sich **leiblich** bereiten/ ist wol eine feine eusserliche Zucht. (Rossihnius 1632: 21)

Konkreetse tekstisõna märgendamisel räägib kaassõnalise kasutuse poolt see, kui abstraktsemaks muutunud tähendusega vorm on kasutusel süntaktilises ümbruses *Noomen genitiivis + /\_\_\_/*, väljendades mingit üldkategoriaalset tähendust, nt märkides koha-, aja-, seisundi-, asendus- või omandussuhet (vt ka Habicht 2000).

Samamoodi täieneb substantiivivormide leksikaliseerumise kaudu **adverbide** sõnaklass. Tänapäeva eesti keeles peetakse adverbideks hulka erandlikke, üldistunud tähendusega, arhailisi või murdelisi käändevorme. Sageli pole nende seos algse käändsõnaga enam ilmne. Osa niisuguseid vorme on tänapäevastes käsiraamatutes liigitatud kord nimi-, kord mäarsõnadeks. Nt *kodunt* on 2006. aasta ÕS-is märgitud seestütleva käände vormiks, samuti on tavaliseks käändevormiks peetud vormi *külla*. 1976. aasta ÕS ja EKG nimetavad neid vorme aga adverbideks (EKG I: 23). Kirjakeele seletussõnaraamatus on *kodunt* adverb ja *külla* substantiiv, mille juures on viide “adverbilaadsele” kasutusele. Niisiis on tegemist probleemse piirkonnaga, mille suhtes puudub uurijatel üksmeel.

Vana kirjakeele puhul on võrdlusmaterjali vähesuse tõttu veelgi raskem selgitada, kas vaadeldav sõnavorm kuulus teksti ilmumise ajal mõne noomeni paradigmasse või oli juba kujunenud muutumatuks sõnaks. Kui niisugustel juhtudel eelistada alati nimisõnaks määramist, on korpuse kasutajal hiljem keeruline uurida adverbide kujunemist, kui aga eelistada adverbiks määramist, on keeruline uurida käändevorme ja leksikaliseerumise astmeid. Sel põhjusel oleme oma märgendusotsused teinud üksiksõnuti.



Probleemsed on näiteks ajamäärused *öösel* ja *ööseks*. Tänapäeva ühiskeeles on vorm *ööse-* substantiivi ja adverbi vahepeal: käändeparadigma on laiem kui tavalistel vaegmuutelistel substantiividel, võimalik on ühilduv täiend, kasutatakse koguni partitiivi ning võimalik on moodustada mitmust, puudub aga nominatiivi vorm, mis on üks olulisi kriteeriume noomeni määratlemisel. (Vt Alvre 1978, Velsker 2006) Sama iseloomustab ka vana kirjakeele tekste (18), ehkki nominatiiviks on vähemalt Heinrich Stahl pidanud vormi *öö* (tema sõnastikust leiame käändevormid *öö* : *ööse*). Selle sõna puhul oleme lähtunud keele varasemast arenguastmest ja teadmisest, et nimisõnade kohakäändevormide leksikaliseerumine on ajas järkjärgult toimuv protsess, mille tulemusena on tänapäeva keeles kasutusele jäänud muutumatu (leksikaliseerunud) sõnakuju.

- (18) Seh pehwlick meilt om~a paistusse wottap/ se pimmedus sest **öhsest** [S] wasto tullep .. Die Son vns jhren schein benimb :/ die Finsternüß der **Nacht** herkompt .. (Stahl 1637: 249)

Omapärased on Heinrich Stahli tekstides esinevad leksikaliseerunud alaltütleva käände vormid *hohmickult* 'hommikul' (19), *hohmselt* 'hommikul' ja *ochtult* 'õhtul' (20), millest esimene on registreeritud Häädemeeeste ja Jöelähtme murrakust ning viimane on tuntud ka mulgi murdes (VMS).

- (19) Evangeliummes lehp temma welja **hohmickult** [Adv] warra/ ninck perrast kolmandal/ kuhel/ üdixal/ ninck öheltöistkümmel tundil .. Jm Evangelio gehet er auß **am Morgen** frühe/ vnd hernacher vmb die dritte/ sechste/ neunde/ vnnd eilffte stund .. (Stahl 1641: 246)
- (20) Jsaac issi lex **ochtult** [Adv] ninck **hohmselt** [Adv] welja/ Jumjala pal-luma/ ninck sai se Rebecca Emmandax. Jsaac selber gieng **Abendes** vnd **Morgends** hinauß GOtt zu bitten/ vnd krigte die Rebecca zum Weibe. (Stahl 1641: 168)

Küsitavusi tekitavad ka nimisõnade sisekohakäänetest kujunenud (seisundi)-adverbid või käändumatud omadussõnad (21).

- (21) Tullep JEsus/ kus needt uxet **luckus** [Adv] ommat/ ninck astup nende kesckel .. (Stahl 1638: 157)

Sageli moodustavad noomenite käändevormid või neist kujunenud (**afiksaal**)-**adverbid** ühendeid tegusõnadega, nt *tähele panema*, *kihla vedama*, *maha matma*. Korpuse märgendamisel tuleb otsustada, kas tegemist on nimi- ja tegusõna ühendiga (väljendverb) või nimi- ja abimäärsõna ühendiga (ühendverb). Niisugused verbühendid sisaldavad vahel adverbe, mis peaaegu üheski muus seoses ei esine, sest algselt ühendisse kuulunud nimisõna muud käändevormid on keelest kadunud.

Definitsiooni kohaselt on väljendverbi sisuliseks tuumaks noomen, ühendverbi sisuliseks tuumaks aga verb (EKG II: 19–20). On ka niisuguseid ühendeid, mille sisuline tuum pole päris selge. Sageli ei ole võimalik leida ühest lahendust ka tänapäeva keele käsiraamatutest. Nt *kihla* on 2006. a ÕS-is nimisõna *kihl* vorm, *põhja* ja *vargile* aga eraldi märksõnad, 1976. a ÕS seevastu esitas adverbidena *kihla* ja *vargile*, kuid mitte *põhja*. Niisugust perifrastilise verbi liiki, mille sisuliseks tuumaks

oleks adverb, meie grammatikad ei esita. Korpuses oleme niisuguseid ühendeid, mille sisuliseks tuumaks on mõne noomeni erandlik või arhailine, adverbistunud vorm, käsitletud siiski väljendverbidena. Vana kirjakeele tekstide puhul on niisugune otsus seda enam õigustatud, et paljud praeguseks erandlikud vormid olid paarsada aastat tagasi veel tavalised. Nt *kihl* 'leping' on 17. sajandil olnud kasutusel ka iseseisva substantiivina, mitte ainult ühendis *kihla vedama*, vrd (22) ja (23).

(22) Mea peame meije teggema/ kui meije temma Ihu söhme/ ninck temma Werre johme/ ninck se **kichla** [S] nūnda wöttame? Was sollen wir thun/ weñ wir seinen Leib Essen/ vnd sein Blut Trincken/ vnd das **Pfand** also nehmen? (Rossihnius 1632: 28)

(23) EPs teije tehte/ eth needt/ kumbat johxmisse pehle **kihla** [S] weddawat/ needt johxwatkeick? (Stahl 1638: 335)

Vanas kirjakeeles esineb ka laenulisi verbiühendeid, mida tänapäeva keeles ei kasutata. Nt saksa keelest otsetõlkena üle võetud *kodu otsima* < *heim suchen* 'proovile panema' (24). Eesti *kodu* on nimisõna, saksa *heim* sellises ülekandelises kasutuses juba adverbistunud ning mingit sisulist seost 'kodu'-tähdendusega pole. Peame selliseid ühendeid eesti keeles laenulisteks väljendverbideks, hoolimata nende saksa vastetest.

(24) Minna se Jssand sünno Jum~al ollen üx weggiw wehane Jummal nende pehle/ kumbat münd wehastawat/ ke nedt pattud neist wannambist **Koddo** [S] **otzip** nende lapsede pehl/ Kolmande ninck neljande põlwe tagga .. Ich der HErr dein GOtt bin ein starcker Eyferiger GOtt/ vber die so mich hassen/ der die Sünde der Väter **heimsuchet** an den Kindern/ biß ins dritte vnd vierdte Glied .. (Stahl 1632: 3)

Huvitaval kombel leidub ka niisuguseid (arhailisi) noomenivorme, mis vanas kirja-keeles seostuvad ainult ühe kindla verbiga, tänapäeval aga mitmega. Selline ühend on nt *andeks andma* (sagedus 17. saj tekstides ca 450). 16. sajandi esimesest poolest pärinevas Kullamaa käsikirjas esineb ühel korral siiski ka *andeks heitma*. Põhjuseks, miks nt ühendit *andeks paluma* ei kasutatud, võib jällegi olla sõnasõnaline tõlkimine saksa keelest: *um verzeihung bitten* > *andeksandmist paluma* (25).

(25) Sedda peat siña süddamest kahetzama/ Jummalalt **andix** [S] **andmisse palluma**/ ninck se Töllneri kaks üttelema .. Das soltu von Herten berewen/ GOTT den HERren **vmb Vergebung anruffen**/ vnd mit dem Zöllner sagen .. (Stahl 1638: 304)

## 6. Adjektiivide eristamine teistest sõnaliikidest

Adjektiivid on käändes, arvus ja võrdlusastmes muutuvad sõnad, mis väljendavad omadust ning võivad lauses esineda ühilduva atribuudina (EKG I: 20). Täpselt nagu substantiivivormid võivad leksikaliseeruda ja/või grammatiseeruda ka adjektiivivormid. Seega on ka adjektiividel piiriala **adpositioonide** (26) ja **adverbidega** (27).

- (26) sinna ollet Andet sahnut nende Jnnimessede **hehx** [Adp] .. du hast Gaben empfangen **für** die Menschen .. (Stahl 1641: 516)
- (27) Meije hing on errapehstnut/ kudit üx lind sest keelest ohest pühdiast/ se keel on **katki** [Adv] / ninck meije olleme wallalis. Vnser Seele ist ent-runen/ wie ein Vogel dem Strick des Vogelers/ das Strick ist **zurissen** / vnnd wir sind los. (Stahl 1641: 202)

Problemaatilised on üksikud *lt-* ja *sti-*liitelised sõnad, sest need langevad kokku kohakäänevormidega (*-sti* on vanades tekstides sageli kujul *-st*). Kõhklusi tekitavad eelkõige sellised üksused, mida tänapäeva keeles ei kasutata, samuti kollokatsioonid või raskesti tõlgendatavad laused:

- (28) JESus kostas: Minna ollen **wabbast** [Adv] **awwalickult** [Adv] rechknut se ilma ehs/ minna ollen ickas opnut kohli sees/ ninck Kircko sees .. JESus antwortet: Jch habe **frey offentlich** geredt für der Welt/ ich habe allezeit gelehrt in der Schule vnd in dem Tempel .. (Stahl 1638: 221)

Adjektiivide ja adverbide piirialale jäävad ka sõnad, mis ei seondu küll ühegi morfoloogilise kategooriaga, kuid võivad sõne tasandil käituda nii adverbina (*teeb töö valmis*) kui käändumatu omadussõnana (*valmis töö*). Verbi *olema* laiendina pole selliste sõnade puhul võimalik sõnaliiki kindlaks teha. (EKG I: 25) Mõned neist sõnadest vanas kirjakeeles veel käänduvad, vähemalt osas vormides:

- (29) Jummal sanna meile andtut ../ eth meije seperrast peame ellama/ ninck tegkema hend **walmix** [Adj]/ .. (Stahl 1637: 170)
- (30) Müñul on se ilm kawwalast sehdtntut/ walle ninck kawwala nouw kahs paljo wörckut/ ninck **sallajat** [Adj] keelet .. (Stahl 1637: 112)

Üksikute sõnade puhul on valida koguni kolme sõnaliigi vahel: adjektiiv, adpositsioon või adverb. Seegi on universaalne joon, mida võib leida eri keeltest. Seda tüüpi sõnadest esinevad vana kirjakeele korpuses eeskätt *täis* ja *väärt*, mida võiks tähenduse ja enamasti ka süntaktilise funktsiooni alusel pidada pigem adjektiivideks või muutumatuse tõttu ka adverbideks (31), (32), kuid mis koos osastavakäändelise nimisõna(fraasi)ga moodustavad siiski adpositsioonifraasilaadse ühendi (33).

- (31) Minne welja neine mah hulitzede ninck need aidwehrt möta/ ninck palle neid sisse tulla/ et minnu kodda **teus** [Adv] sahkut. (Rossihnius 1632: 122)
- (32) Ken enge needsinatzet sönnat ei ussu mitte ächk katzipeiti mötleb/ se ejolle mitte **währt** [Adj] ninck hästi walmistut .. Wer aber diesen Worten nicht gläubet/ oder zweyffelt/ der ist **vnwürdig** vnd vngeschickt .. (Rossihnius 1632: 21)
- (33) STephanus enge **teus** [Adp] usku ninck weggi/ teggi imme ninck suhret tähet raghwa sean. (Rossihnius 1632: 76)

Viimase kasutusjuhu analüüsimine eessõnafraasina pole siiski probleemitu, kuna sõna *tüis* esineb paralleelselt peaaegu samasuguses elatiivse substantiiviga fraasis, kus seda oleks ilmselt loomulikum pidada adjektiiviks, nt (34).

- (34) Christussest kuhlutap se Evangelist/ eth temma on kaswanut/ se waimo sees wegkiwax sahnut/ **töis** [Adj] tarckussest/ ninck Jummal arm temma jures olnut. Von Christo meldet det Evangelist/ das er sey gewachsen/ im Geist starck worden/ **voller** Weißheit/ vnd GOTTES gnad bey jhm gewesen. (Stahl 1641: 110)

Nii substantiividest kui ka adjektiividest võivad kujuneda **afiksaaladverbid**, mille leksikaliseerumise astet ei ole materjali vähesuse tõttu võimalik tagantjärele kindlaks teha. Näiteks tänapäevasest verbist *kõlbama* tuntud tüvi on vanas kirjakeeles kasutusel substantiivina (vrd soome *kelpo* 'hea, korralik', mis osutab, et vanem kasutus on võinud olla adjektiivne) ja vähemalt kahes grammatilises käändes (35), (36). Seni sisestatud korpusetektides ei esine see siiski kordagi atribuudina, nagu adjektiivi puhul oleks oodatav.

- (35) .. sesamma Iumala thenistussest ei olle middakit **kõlbo** [S]. (Rossihnius 1632: 113)
- (36) Minna langsin ickas sügkawambel sinna sisse/ ep olli münno ellust **kõlb** [S]/ se patt olli münd ümberandnut. Jch fiel auch immer tieffer drein/ es war kein **gutes** am Leben mein/ die Sünd hatte mich besessen. (Stahl 1641: 251)

## 7. Numeraalide eristamine teistest sõnaliikidest

Numeraalid on käändes ja arvus muutuvad sõnad, mis väljendavad asjade kvantitatiivseid tunnuseid (EKG I: 22). Numeraalidel on ühine piiriala eeskätt **pronoome-nitega**, peamiselt puudutab see arvsõnu *üks* ja *teine*. Vana kirjakeele tekstides on *üks* võrdlemisi reeglipäraselt tarvitusel saksa keele umbmäärase artikli tõlkevastena (vt ka Habicht 2001: 32–33). Numeraal *teine* on kasutusel umbmäärase asesõnana (saksa vaste enamasti *ander*), samas funktsioonis võib olla kasutusel ka *üks*.

Sõnade *üks* ja *teine* numeraalsete ja pronominaalsete kasutuste eristamine on enamasti suhteliselt lihtne, vrd (37) ja (38). Esineb siiski ka kontekste, kus arv- ja asesõna määratluse vahel otsustamine osutub keerukaks (39). Sellisel juhul eelistame vahetegemisel numbriga asendamise testi: kui probleemset üksust on võimalik numbriga väljendada, siis on tegemist arvsõnalise kasutusega, kui mitte, siis asesõnalisega.

- (37) .. Ei kuhle temma sinno mitte/ sihs wötta wehl **ütte** [Num] ächk katte hennesse mannu/ .. (Rossihnius 1632: 24)
- (38) Ninck Jummal keskis neile unnes/ eth nemmat mitte piddit jelle Herodesse jure minnema/ ninck lexit jelle lebbi **öhe** [Pron] **töise** [Pron] te omma mah sisse. Vnd Gott befahl jhnen im traum/ das sie sich nicht solten wider zu Herodes lencken/ vnd zogen durch **einen andern** Weg wider in jhr Land. (Stahl 1638: 21)

- (39) Sihs lasck Iumal ütte süwwa unne langeda se Inimesse pähle/ ninck tema maggas/ ninck wöttis **ütte** [Num] temma Külgluhst/ ninck panni se asseme kindi Liha kahn. Da ließ GOtt einen tieffen Schloff fallen auff den Menschen/ vnd er entschlieff/ vnd nam seiner Rieben **eine**/ vnd schloß die stete zu mit Fleisch. (Rossihnius 1632: 39)

**Adverbideks** peetakse arvsõnatüvelisi järjekorda märkivaid vorme *esiteks, teiseks, kolmandaks, neljandaks* jne. Vanas kirjakeeles on samas tähenduses kasutusel muidki vorme, mida oleme samuti adverbideks märgendanud, nt *esimelt* 'esiteks' ja *kolmandel* 'kolmandaks'.

- (40) **Essimelt** [Adv] A. R. piddat teje löhitelt opma ninck meles piddama/ mea sÿß v"che eike Palwe iures piddab ollema .. (Müller 1603: 10)
- (41) **Kolmandel** [Adv] ninck Wimatel/ Nüit on v"x Maddo v"x Kurrÿ ninck wihane Loius .. (Müller 1601: 5)

## 8. Konjunktsioonide eristamine teistest sõnaliikidest

Konjunktsioonid on muutumatud sõnad, mille ülesandeks on siduda lause moodustajaid (EKG I: 40). Enamik eesti keele konjunktsioone on kas kujunenud **adverbidest** või laenatud teistest keeltest (Alvre 1987). Mitmed konjunktsioonid (*ehk, või, aga, nagu*) saavad tänapäevalgi esineda lisaks ka modaaladverbina.

Nt *aga* (17. sajandi tekstides veel koos possessiivsufiksi rudimendga -s, s.t kujul *agkas*) paikneb sageli selle fraasi ees, mille kohta vastandus käib, olles nii pigem rõhumäärsõna kui lause moodustajaid siduv sõna (42).

- (42) Keick mea igganes on sihn ilma pehl/ olckut höbbe/ kuld/ echk raha/ rickus ninck ajalick willi/ se kestop **agkas** [MdAd] öhe ürrikesse aja/ ninck ep awwitap mitte sex önnistussex. Alles was ist auff dieser Welt/ es sey Silber/ Gold oder Gelt/ Reichthumb vnd zeitlich Gut/ das wäret **nur** ein\_kleine zeit/ vnd hilfft doch nicht zur Seligkeit. (Stahl 1637: 97)

Tänapäevases erinev on ka **proadverbide** *kuna, kus* ja konjunktsiooni *kui* kasutus. Et konjunktsioonid on sageli kujunenud adverbide grammatiseerumise tulemusena (proadverb *kui* ebamäärase tähendusega üksus on abstraktse grammatilise funktsiooni arenemiseks soodne alus), siis võib sellist arengut näha ka siinse sõnarühma puhul. Vanades tekstides leidub näiteid tänapäevases vähem grammatiseerunud funktsioonide kohta, nt *kui* esineb sageli tähenduses 'kuidas' (vrd saksa *wie* 'kuidas; kui'), sel juhul saab ta korpuses märgendi *proadverb* (43). *kuna* ja *kus* on samuti enamasti proadverbid, v.a kui need väljendavad selgelt tingimust (44), mitte aega ega kohta (45), (46). Kindlat piiri määr- ja sidesõnalise kasutuse vahele siiski tõmmata ei saa ja täielikku ühtsust me nende sõnade puhul ei taotlegi.

- (43) **Kui** [PrAd] seisap sehl Kirjotut? **Wie** stehet da Geschrieben? (Stahl 1632: 7)

- (44) **Kuß** [Konj] ux aßi üx kort leppit ut on see ep pehe uheste kaiwatut sama (16. saj, Liivimaa talurahva õigus)
- (45) Omma willitzusse kahs/ peat sinna rahwul olla/ sest **kus** [PrAd] se hedda keicke suhremb/ tahap Jummal sünd issi tröstima/ Mit dem Elende dein/ soltu zu frieden seyn/ denn **wenn** die Noth am grösten/ wil dich Gott selber trösten. (Stahl 1637: 125)
- (46) Se v" x Herris kumb meddi Issanda Christuße kaas rÿsti pæle podi/ saab se Rÿsti pæl oma Ello parranduße pole kutzututh/ **kus** [PrAd] tæma sen Surma oma Silmade kaas negki. (Müller 1601: 38)

**Interjektsiooni** ja konjunktsiooni piirialal paikneb 17. sajandi esimese poole tekstides kasutatud partikkel *ja(a)*, mis tänapäeva eesti keeles esineb rinnastava konjunktsioonina. Vana kirjakeele tekstides on see sageli funktsioneerinud veel modaalsõnaku või interjektsioonina, mille kasutamise eeskujuks on saksa keele partikkel *ja*. Tegemist on juhtumiga, kus laenuisest partiklist areneb sidesõna: grammatiseerumisprotsessis kaotab üksus emfaatilisust ning hakkab vähehaaval täitma neutraalset rinnastava sidesõna funktsiooni. Selle kohta pakuvad 17. sajandi esimese poole tekstid ilmeka näiteid, kus on võimalik näha funktsionaalset varieeruvust ja nn üleminekukontekste (48), (49), samal ajal kui üksus on kasutusel ka algses partiklilises funktsioonis (47).

- (47) **Ia** [Int] / minna ussu/ minna olle ütZ Pattane inimene. **Ia**/ Ich gläubs/ ich bin ein Sünder. (Rossihnius 1632: 25)
- (48) Allandas hend issi/ ninck kuhlis omma Jssa/ emmis se surma/ **ja** [Konj] se surma sest ristist. Er nidriget sich selbst/ vnd ward gehorsam biß zum Tode/ **Ja** zum Tode am Creutz. (Stahl 1638: 59)
- (49) Sÿn oppep meit Chrus/ et tæma oma Apostlit/ **ia** [Konj] keick truwit Kircko Ibandat ninck Oriat/ eb mitte suhre auwo/ echk hëe Peiwade siße on kutznut. (Müller 1601: 38)

## 9. Adverbide eristamine teistest sõnaliikidest

Adverbid on tüüpjuhul muutumatud sõnad, mis esinevad lauses adverbiaali funktsioonis ja iseloomustavad tunnuse tunnust (EKG I: 23). Probleemne piiriala on adverbidel adpositsioonidega, sest vana kirjakeele tekstides esineb tänapäevasest oluliselt rohkem **adpositsioone**. Kui praeguses eesti keeles öeldakse *kellelegi vastu panema*, siis vanas kirjakeeles *kellegi vastu panema*, seda isegi tõlkelistes tekstides, kus saksa eeskujuga *jemandem* (daativ) *widerstehen* on selgelt olemas (50). Võiks oletada, et neis lausetes on siiski tegu ühendverbidega, mille rektsioon on tänapäevasest erinev, kuid sellisele tõlgendusele räägib vastu adverbiaali võimalik paiknemiskoht lauses (51).

- (50) Peame semperrast ussinat ollema/ ninck walwama/ ninck se **kurrati wasto** [Adp] **pannema** uscko sees. Sollen des halben nüchtern seyn/ vnd Wachen/ vnnd **dem Teuffel widerstehen** im Glauben. (Stahl 1641: 26)

- (51) Sihs peame meije pattude sees mitte magkama/ erranis se kurrati wasto taplema/ ninck **kindlast tem~a wasto** [Adp] **seisma** uscko sees. Als sollen wir in Sünden nicht schlaffen/ sondern wieder den Teuffel streiten/ vnd **jhm fest widerstehen** im Glauben. (Stahl 1641: 291)

Iga väljendi puhul pole piisavalt näiteid, selgitamaks adverbiaali võimalikku asukohta. Eelnevale tuginedes märgendame sellistel juhtudel abisõnad alati kaassõnadeks (52).

- (52) SE tennapehwase Evangeliumme Oppetusset sahwat **meije ette** [Adp] **kandtut** se kescko/ Vscko/ ninck Patri siddes. DJe Lehren des heutigen Evangelij werden **vns fürgetragen** im Gesetz/ Glauben/ vnd Vater Vnser. (Stahl 1641: 19)

Seda tüüpi fraaside reanalüüsil on arvatavasti toimunud seoste ümbertõlgendus nii, et varasemat kohaadverbi on hakatud kasutama kaassõnafraasis, kus seesama kohasuhe väljendub nüüd noomeniga seoses (vt ka Habicht 2000).

Nii adverbi kui ka **pronoomenina** võib esineda *palju*, mille kasutus erineb tänapäevasest. Kui tänapäeval on kasutusel fraas *paljud inimesed* (fraasi põhjaks on nimisõna ja sellega ühilduv täiend on sõnaliigiliselt asesõna) või *palju inimesi* (fraasi põhjaks on määrsõnaline kvantor, mille laiend on mitmuse osastavas käändes) või *palju tegemist* (mille laiend on ainsuse osastavas), siis vanas kirjakeeles esineb lisaks nimetatutele ka varianti nagu *palju inimesed* (53).

- (53) Sedda ewat teh needt Valschit Oppiat mitte/ neil omat **paliu** [Adv] **Lehet** ilma Wiliata/ se on/ neil eb olle mitte se eike Iumala Sana .. (Müller 1601: 37)

## 10. Adverbiliikide omavaheline eristamine

Vana kirjakeele tekstides leiduvate adverbide määratlemisel tekib põhilisi probleeme ühelt poolt vaba **adverbi** ja koos verbiga püsiühendit moodustava **afik-saaladverbi** eristamisel ning teiselt poolt selliste arhailiste adverbide määratlemisel, mis esinevad tänapäeva keele seisukohalt ootamatutes lauseümbrustes, nagu *üle* näites (54), või mis on tänapäevaks keelest kadunud, nagu arhailised 3. isiku possessiivsufiksi rudimenti sisaldavad adverbid *ihulis* ja *hingelis* 'ihulikult ja hingeliselt' (vt Mägiste 2000: 33) näites (55). Näites (54) on sõna *üle* tõlgendatud (määra)adverbiks, mis laiendab teist adverbi. Tänapäeval oleks sellises lauseümbruses kasutusel pigem augmentatiivne adverb *üli*-.

- (54) .. temma röhmustap se ilm **ülle** [Adv] **paljo**/ Er tröst die Welt **mit massen**/ (Stahl 1637: 122)
- (55) .. üchtekit pehw/ üchtekit tund/ üchtekit silmapilgkminne lehp möhda/ kummal Jummal sünnul **iholis** [Adv] ninck **hingelis** [Adv] mitte hehdt teep. ... kein Tag/ kein Stund/ kein Augenblick vergehet/ da dir GOTT **an Leib** vnd **Seel** nicht gutes thuet. (Stahl 1641: 280)

Kui üksus on küsimustatav ja väljendab tekstis mingit adverbile omast tähendust, on seegi loetud täistähenduslikuks adverbiks, nt liitsõna *niipea* 'kohe'<sup>4</sup> (56) kollokatsioonis *niipea kui*, kuigi leidub ka kontekste, kus seesama üksus on määratletav proadverbina (57).

(56) **Ni pea** [Adv] kudt JEsus ristitut olli/ on temma sest Pöhist Waimust körbe sisse sahtetut .. **Als bald** JEsus getaufft war/ ist er vom Heiligen Geist in die Wüsten geführet .. (Stahl 1641: 299)

(57) Ninck eth se Jumjala komb/ eth temma mitte **ni pea** [PrAd] tullep/ kudt meije tahame .. Vnd weil Gottes weise/ das er nicht **so bald** kompt/ als wir wollen .. (Stahl 1638: 121)

**Afiksaaladverbe** eristab vabadest adverbidest see, et need kuuluvad püsiühendite, nn ühendverbide koosseisu. Ühendverb omakorda on tänapäeval defineeritud kui püsiühend, millesse lisaks verbile kuulub ka koha-, seisundi- või perfektiivsusmäärsõna. Suur osa sellistest ühenditest on eesti keeles tõlkelaenulised (vt Hasselblatt 1990), kuid tänapäevani produktiivsed. Nii näiteks võimaldavad ühendverbide komponendiks olevad afiksaaladverbid morfoloogilise aspektikategooria puudumise tingimustes muuhulgas väljendada aspektisuhteid, nt *välja maksma*, *ära ajama*, *minema jooksuma*, *läbi vaatama* jpt.

Vana kirjakeele materjali puhul tekib probleeme ennekõike laenuliste verbiühenditega, mis võivad olla tekstides sagedased ja käibida püsiühenditena (sageli küll laenulistena), kuid mis tänapäevase definitiooni järgi siiski afiksaaladverbi määratluse alla ei mahu. Üheks selliseks näiteks on saksa *wieder* vastena ühendite koosseisus esinev *jälle*. Kuna tegemist on ajaadverbiga, siis tänapäevase definitiooni kohaselt ei moodustaks see ühendverbe. Samal ajal on selle esikomponendiga ühendid, eriti *jälle tulema* (sks *wiederkommen*), *jälle tooma* (sks *wiederholen*; *wiederkehren*), *jälle andma* (sks *wiedergeben*), *jälle otsima* (sks *wiedersuchen*) jt vanades tekstides sagedased ja *jälle* verbi tõkelise esikomponendina produktiivne, mistõttu on sellised ühendid praegu märgendatud püsiühenditeks ja ühendi esikomponent määratud afiksaaladverbiks (58), (59).

(58) Se kaks temma meid teep hendesse sarnasex/ ninck **jelles** [AfAd] **sah-tap** Jumjala rickusse pohle. .. damit er vns jhm machet gleich/ vnd **wiederbringt** zu GOTTes Reich. (Stahl 1641: 67)

(59) Ep woinut sesamma suggu/ ommast wehjest **jelle** [AfAd] **jetma**/ echk temma kül sagedast kiusati .. Nicht [..]üglich war dieselbig Art/ auß eygnen Kräfteffen **lassen**: Wiewol es oft versuchet ward .. (Stahl 1637: 79)

Vana kirjakeele tekstides leidub ka seda tüüpi keerukaid juhtumeid, kus ühendverbil võib olla määrsõnaline laiend, mille määratlemine tekitab küsimuse, kas tegemist on ühendverbi püsiühendilise komponendi või vaba adverbiga (60). Sellistel juhtudel on modifitseeriv adverb loetud vabaks adverbiliseks laiendiks.

(60) Agkas temma ep tehþ sedda mitte om~a armo perrast/ ninck on kannatalick meite kaks/ temma **laotap** ommat kehdt **laijalist** [Adv] **welja** [AfAd] keicke piddose pehwa/ öhe kangke usckmatta rahwa wasto ..

<sup>4</sup> Ehkki korpuse märgendamisel määratakse esmalt liitsõna või ühendi osadele eraldi sõnaliik, on võimalik märgendada ka liitse üksuse sõnaliik. Seda ongi siinse näite puhul silmas peetud.



Aber er thut solches nicht/ von wegen seiner Barmhertzigkeit/ vnd hat Gedult mit vns/ er **recket** seine Hände **auß** den gantzen Tag zu einem vngehorsamen Volck .. (Stahl 1641: 222)

Modaaladverbide puhul tuleb eeldada, et üksus on läbinud leksikaliseerumis- ja grammatiseerumisprotsessi, muutudes vaadeldava hetke seisukohalt diskursigrammatika vahendiks. Tegemist on pikaajalise muutumisprotsessiga, mille tulemusena varasemad noomeni- või verbiparadigmaste kuulunud vormid iseseisuvad tähenduselt ning hakkavad toimima hoiakuid ja hinnanguid väljendavate fraasi- või lauselaienditena. See astmeline protsess on jälgitav ka vana kirjakeele tekstide põhjal, kuigi sageli on probleemiks otsustada üksuste arengustaadiumi või grammatiseerumistasme üle, s.t määratleda, millal saab üksust juba modaaladverbiks lugeda. Probleemiks on see, et üksuse grammatiseerumisega kaasneb sageli aastakümneid kestev varieerumise periood, mil varasem paradigmaatiline ja uus diskursigrammatiline kasutus figureerivad kõrvuti ning modaaladverbilist kasutust võib olla raske eristada. 17. sajandi kirjakeele modaaladverbide kujunemise ja kasutuse kohta on lähemalt kirjutanud Külli Habicht (2009).

Sellise eristamise raskust näitavad mõned vana kirjakeele muutumatud sõnad, mille funktsiooni üle otsustamisel tuleb toetuda kontekstile. Üheks probleematiliseks üksuseks on *koggonis* 'koguni; kogunisti', mis võib olla kasutusel nii määrämäärsõnaliselt kui ka modaalses tähenduses. Esimesel juhul on tegemist täistähendusliku adverbiga, millest hilisem diskursigrammatiline funktsioon lähtub. Täistähendusliku ja modaalse funktsiooni eristamiseks on vaja analüüsida konteksti, s.t üksuse asukohta lauses ja seost teiste üksustega, ning teha küsimustamise ja väljajätmise test. Kui uuritav üksus on täistähenduslikuna küsimustatav, nt *kuidas, mil määral?*, ei ole ilma tähenduskaota lausest väljajäetav ega paikne lause- või fraasialgulisel positsioonil, on tegemist täistähendusliku adverbilise kasutusega (61), vastasel juhul tuleb kõne alla üksuse modaaladverbiks pidamine (62). Samasuguseid teste saab kasutada ka teiste modaaladverbi ja adverbi piirimaile jäävate juhtude korral.

(61) Sihs sawhat sünnul üx meeheh olla needt offrit sest öigkusses/ needt tulle offrit/ ninck **koggonis** [Adv] offrit/ sihs sawhat nemmat wassikat sünnu Altari pehle offrime. Denn werden dir gefallen die Opfer der Gerechtigkeit/ die Brandopffer vnd **gantzen** Opfer: Denn wird man Farren auff deinen Altar opffern. (Stahl 1638: 9)

(62) Wihendel: nemmat oppetawat WaheAsja Oppedusse **koggonis** [MdAd] mitte öikest. Zum fünften: si handeln der Lere von den Mittel-Dingen **gantz und gahr** zu wider. (Blume 1662: 117)

Probleeme võivad tekitada ka juhtumid, mil tänapäeva keeles modaalseks arenenud üksus esineb vanas kirjakeeles veel valdavalt täistähenduslikuna, nagu muutumatu sõna *iseäranis*, mis võib 17. sajandi tekstides esineda nii viisimäärsõnana tähenduses 'eraldi' (63) kui ka grammatiseerunud modaalpartiklina (64).

(63) Sedda agkas keick teep sesamma ainus waim/ ninck jagkap öhel igkalickul **issiehranis**/ [Adv] kudit temma tahap. Disz aber alles wircket derselbige

einige Geist/ vnd theilet einem jglichem **seines zu**/ nach dem er wil.  
(Stahl 1638: 112)

- (64) Eth meil nühd't aigk on/ sihs techkem hehdt öhel igkal meehel/ **issieh-ranis** [MdAd] agkas neil öigkel üsckolissel. Als wir denn nu zeit haben/ so lasset vns gutes thun/ an jederman/ **Allermeist** aber an den Glaubens genossen. (Stahl 1638: 123)

Sarnane on ka sõnade *ühtlasi*, *ikka* ja *väga* määratlemise probleem. Needki võivad olla kasutusel nii täistähendusliku adverbi kui ka modaalpartiklina. Nt muutumatu sõna *ickas* 'ikka' võib vana kirjakeele tekstides esineda kord täistähendusliku ajaadverbina (65), ent ka täistähenduseta modaalpartiklina, mille ainus funktsioon on kõneleja seisukoha rõhutamine (66).

- (65) Christussel olekut **ickas** [Adv] auw/ ke sehl tüllis Jummal nimmel/ immelickul kombel/ meije wainlaset keick üchlisse/ weggiwast on üllewoidnut/ ninck omma rickusse sissewotnut .. Christo sey **allzeit** Preiß/ der da kam in Gottes Nahmen mit wunderlicher weiß vnser Feinde all zusammen/ gewaltig hat vberwunden/ vnd sein Reich eingenommen.. (Stahl 1637: 51)

- (66) Paulus agkas sai **ickas** [MdAd] wegkiwambax/ ninck üllewoijas needt Judalisset .. Paulus aber ward **je** mehr kräftiger/ vnd treib die Jüden ein .. (Stahl 1638: 168)

**Proadverbid** moodustavad suletud sõnaklassi, mida iseloomustab üldine ja ebamäärane tähendus erinevalt adverbidest, mis on konkreetse tähendusega. Proadverbid on asendussõnad, mis on olulised tekstisiseste viitesuhete vormistamisel, aidates vältida kordusi. Tüüpiliste proadverbide määratlemisel probleeme ei teki, sest tegemist on sõnaklassiga, mille koosseis on võimalik loendada kirjeldada, nt kohasuhteid väljendavad *sinna*, *seal*; ajasuhteid väljendavad *millal*, *seni*; viisi väljendavad *nii*, *nõnda*; põhjussuhteid väljendavad *seepärast*, *mistõttu* jne. Vanas kirjakeeles esineb ka arhailisi proadverbe, mille tuvastamine ei pruugi nii lihtne olla. Selline on nt *situ(kord)* 'nii mitu korda' (67).

- (67) Sest mitto kord üx Jnnimenne hend pöhrap nende Laußjade ninck Tarckade pohle/ ninck neist middakit küssip/ **sitto** [PrAd] **kord** piddap tem~a hohrajachto nende kahs .. Denn so oft ein Mensch sich Lev. 19. 31. wendet zu den Warsagern vnd Zeichendeutern/ vnd von jhnen etwas forschet/ **so oft** huret er jhnen nach/ .. (Stahl 1649: 618)

Proadverbidel on probleemne piiriala **modaaladverbidega**, s.t sõnaliiki määratledes tekib küsimus, kas üksus on kasutusel üldises tähenduses asendussuhte väljendajana või kannab see lisaks ka mõnd modaalset, selgitavat-rõhutavat lisatähendust (68). Selliseid probleeme tekitavad nt *minkepärast* ja *se(m)pärast*, mis enamasti on märgendatud põhjust väljendavateks proadverbideks (69), küsipartikli funktsioonis aga modaaladverbideks (70).

- (68) **Minckperrast** [MdAd] se Leerpois eb olle mitte v"lle tæma Meistri/ echk v"x Sullane v"lle tæma Issanda. (Müller 4: 8)

- (69) **Semperrast** [PrAd] eike ninda/ kudit needt Merre Lainedt mæßawat/  
kudit se Tuhl weikasti puhub/ eike ninda on kaas meddÿ Südda/ kudit  
meÿe Iumala pæle motleme. (Müller 7: 14)
- (70) Sÿn tahax kekit v"ttelda: Kuÿ nuit Armas Simeon/ **minckperrast** [MdAd]  
eb karta sina hend mitte sen Surma eddest .. ? (Müller 35: 5)

Proadverbide ja modaaladverbide piiriala on sageli kokkuleppeline, sest üks ja seesama üksus võib olla tekstis kasutusel nii täistähendusliku adverbi asendajana kui funktsioneerida ka näiteks küsilause eristajana, toimides sellisel juhul pigem grammatikavahendina, mille sõnaliigiliseks määratluseks on modaaladverb.

## 11. Kokkuvõte

Artiklis käsitleti mõningaid keskseid sõnaliikide määratlemisega seotud probleeme eesti vana kirjakeele korpuse märgendamisel. Põhitähelepanu all olid sõnaliikide omavahel kattuvad piirialad, kuhu kuuluvate üksuste süntaktilise funktsiooni ja muutumiseviisi üle otsustamine ei pruugi isegi konteksti põhjal lihtne olla. Vana kirjakeele spetsiifilised sõnaliigiprobleemid seostuvad ennekõike vormide arhailisuse ja laenulisusega. Diakrooniline vaatenurk võimaldas välja tuua ka kirjakeele arengus huvitavaid grammatiseerumis- ja leksikaliseerumisteid – paljud vaadeldud üksused on aja jooksul läbi teinud funktsioonimuutusi. Kattuvate piirialade käsitus võimaldas iseloomustada ka mitmeid üksuste sõnaliigilise taaskasutamise juhtumeid.

Ülevaate eesmärgiks oli tutvustada vana kirjakeele korpuse sõnaliigimärgendussüsteemi ühelt poolt sõnaliikide tüpoloogia ja teiselt poolt eesti keele seniste sõnaliigikäsitluste taustal ning põhjendada korpuse kasutaja jaoks seniseid märgendamisotsuseid.

Osa vana kirjakeele korpuse sõnaliikide märgendamisel tekkivatest küsimustest on universaalsed (näiteks substantiivide ja adjektiivide või verbi partitsiivvormide ja adjektiivide eristamise keerukus), mõned ebamäärasused on eesti keelele üldomased (näiteks kaassõnade ja nimisõnade vahelise piiri hägusus). Lisaks niisugustele tavapärastele probleemidele tekib vana kirjakeele puhul veel hulk spetsiifilisi küsimusi, mis on sageli seotud lauseliikmena toimiva fraasi ebahariliku morfoloogilise vormistusega (nt verbi mineviku partitsiibi vormide ja *nu-/tu*-liiteliste substantiivide eristamine). Märgendusraskusi põhjustab ka grammatiseeruvatele keelenditele ühese määratluse andmise vajadus, näiteks 17.–19. sajandil interjektioonist konjunktsiooniks arenenud abisõna *ja* puhul.

Sõnaliigi määratlemise tänapäevaste arusaamade kohaselt võib raskeks muuta see, et vanemad tekstid on tõkelised ning lähtetekstide üksuste funktsioonid võivad üle tulla ka sihtkeelde, ilma et eesti keeles oleks toimunud vastavasuunalist loomulikku grammatiseerumist või leksikaliseerumist. Sellist võõr mõju sisaldavad näited on seotud adverbiliikide eristamisega või ka muutuvate ja muutumatute sõnaliikide piirialadega. Sageli võimaldab aga just saksakeelne rööptekst paremini mõista, missuguses funktsioonis on teksti kirjapanija mõnd üksust kasutanud.

Enim määratlemisprobleeme põhjustasid adverbid, millel on kokkupuuteala nii enamiku muutumatute sõnaliikidega kui ka substantiivide, verbide ja

adjektiividega, mille vormide leksikaliseerumise tulemusena tekib keelde uusi adverb. Et adverb on iseseisvate muutumatute täistähenduslike sõnade hulgas ainus, lisaks ka avatud sõnaliik, on mõistetav, et sellesse liiki lisandub eri tüüpi ainet, millest osa on üleminekustaadiumis mitme sõnaliigi vahel.

Artiklis tutvustatud sõnaliigisüsteemi on kasutatud vana kirjakeele korpuse märgendamisel alates 2005. aastast. Süsteemi rakendamist hõlbustab märgendaja elektrooniline töövahend, mis võimaldab iga üksuse puhul näha varasemaid märgendusotsuseid. Märgendatud materjali analüüs võimaldab praktilise töö käigus selgunud vähese eraldustäpsusega kohti ka hiljem täpsustada, niisiis on tegemist paindliku süsteemiga.

### Lühendid

Adj	adjektiiv
Adp	adpositsioon
Adv	adverb
AfAd	afiksaaladverb
Int	interjektsioon
Konj	konjunktsioon
MdAd	modaaladverb
Num	numeraal
PrAd	proadverb
Pron	pronoomen
Pts	partitsiip
S	substantiiv
V	verb

### Viidatud kirjandus

- Alvre, Paul 1978. Erandlik adverbitüüp. – Keel ja Kirjandus, 3, 146–154.
- Alvre, Paul 1987. Soome-ugri keelte ajalooline grammatika I. Sõnaliigid. Arvu- ja käändekategooria. Tartu: Tartu Riiklik Ülikool, soome-ugri keelte kateeder.
- Anward, Jan 2000. A dynamic model of part-of-speech differentiation. – Petra M. Vogel, Bernard Comrie (Eds.). Approaches to the Typology of Word Classes. Berlin, New York: Mouton de Gruyter, 3–39.
- EKG I = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1995. Eesti keele grammatika I. Morfoloogia. Sõnamoodustus. Eesti Teaduste Akadeemia Eesti Keele Instituut. Tallinn.
- EKG II = Erelt, Mati; Kasik, Reet; Metslang, Helle; Rajandi, Henno; Ross, Kristiina; Saari, Henn; Tael, Kaja; Vare, Silvi 1993. Eesti keele grammatika II. Süntaks. Lisa: kiri. Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut. Tallinn.
- EKK 2007 = Erelt, Mati; Erelt, Tiiu; Ross, Kristiina 2007. Eesti keele käsiraamat. Kolmas, täiendatud trükk. Tallinn: Eesti Keele Sihtasutus.
- Erelt, Mati 1980. Some notes on the classification of words and the transposition of adjectives in Estonian. – Grammar and Semantics. Academy of Sciences of the Estonian SSR Institute of Language and Literature. Tallinn: Valgus, 50–72.
- Habicht, Külli 2000. Grammaticalization of adpositions in Old Literary Estonian. – Mati Erelt (Ed.). Estonian: Typological Studies IV. Tartu Ülikooli eesti keele õppetooli toimetised, 14. Tartu: Tartu Ülikool, 19–58.

- Habicht, Külli 2001. Lexicalization and grammaticalization – opposite phenomena? – Mati Ereht (Ed.). *Estonian: Typological Studies V*. Tartu Ülikooli eesti keele õppetooli toimetised, 18. Tartu: Tartu Ülikool, 26–76.
- Habicht, Külli 2002. Pilguheit peaaegu nelja sajandi tagusele sõnatuletusele. Heinrich Stahli omapärased tuletised. – *Keel ja Kirjandus*, 4, 246–261.
- Habicht, Külli 2009. Heinrich Stahli eripärased modaaladverbid. – *Keel ja Kirjandus*, 8–9, 608–619.
- Habicht, Külli; Penjam, Pille 2007. Kaassõna keeleuurija ja -kasutaja käsutuses. – *Emakeele Seltsi aastaraamat*, 52 (2006). Tallinn: Eesti Teaduste Akadeemia Emakeele Selts, 51–68.
- Hasselblatt, Cornelius 1990. Das estnische Partikelverb als Lehnübersetzung aus dem Deutschen. *Veröffentlichungen der Societas Uralo-Altaica*, 31. Wiesbaden: Otto Harrassowitz.
- Hengeveld, Kees 1992. *Non-Verbal Predication: Theory, Typology, Diachrony*. Functional Grammar Series, 15. Berlin: Mouton de Gruyter.
- Hennoste, Tiit 2002. Suulise kõne uurimine ja sõnaliigi probleemid. – Renate Pajusalu, Ilona Tragel, Tiit Hennoste, Haldur Õim (Toim.). *Teoreetiline keeleteadus Eestis*. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised, 4. Tartu: Tartu Ülikool, 56–73.
- Hermann, Karl August 1884. *Eesti keele Grammatik*. Koolide ja iseõppimise tarwis. Tartu. Trükitud Wilhelm Just'i juures.
- Jänes, Henno 1942. *Eesti keele grammatika gümnaasiumile I. Hääliku- ja vormiõpetus: I klass*. Tallinn: Eesti Kirjastus.
- Kaalep, Heiki-Jaan; Muischnek, Kadri; Rääbis, Andriela; Habicht, Külli 2000. Kas tegelik tekst allub eesti keele morfoloogilistele kirjeldustele? – *Keel ja Kirjandus*, 9, 623–633.
- Karelson, Rudolf 1972. Märkmeid kaassõna piirimaolt. – *Emakeele Seltsi aastaraamat*, 18, 1972. Tallinn, 71–86.
- Karelson, Rudolf 1985. Probleme der Beziehungen zwischen Prä- bzw. Postposition und Adverb im Estnischen. – Heikki Leskinen (Toim.). *Ostseefinnische Untersuchungen*. *Studia Fennica*, 28. Helsinki: Suomalaisen Kirjallisuuden Seura, 41–47.
- Kasik, Reet 1996. *Eesti keele sõnatuletus*. Tartu: Tartu Ülikool.
- KKSM 1977 = Kuidas käsitada sõnaliigi mõistet? (Diskussiooni materjale.) – *Keel ja Kirjandus*, 9, 520–537.
- Lindström, Liina; Bakhoff, Liisi; Kalvik, Mari-Liis; Klaus, Anneliis; Läänemets, Rutt; Mets, Mari; Niit, Ellen; Pajusalu, Karl; Teras, Pire; Uiboed, Kristel; Veismann, Ann; Velsker, Eva 2006. Sõnaliigituse küsimusi eesti murrete korpuse põhjal. – Ellen Niit (Toim.). *Keele ehe*. Tartu Ülikooli eesti keele õppetooli toimetised, 30. Tartu: Tartu Ülikool, 154–167.
- Mägiste, Julius 2000. Possessiivsufiksrite rudimentidest eestis, eriti eesti vana kirjakeele (1520–1739) adverbides jm partiklites. – Julius Mägiste 100. Tartu Ülikooli eesti keele õppetooli toimetised, 15. Tartu: Tartu Ülikool, 17–154.
- Pajunen, Anneli 1998. Pääsanaluokkien eriytymättömyydestä uralilaiskielissä. – Anneli Pajunen (Toim.). *Kieliopillistumisesta, analogiasta ja typologiasta*. *Suomi*, 185. Helsinki: Suomalaisen Kirjallisuuden Seura, 59–109.
- Sepper, Maria-Maren 2006. Indirektaal eesti 19. sajandi lõpu ja 20. sajandi aja- ja ilukirjanduskeeles. *Magistritöö*. (Käsikiri Tallinna Ülikooli eesti keele õppetoolis.)
- Vare, Silvi 2004. Põgusalt ühest leksika ja süntaksi piirinähtusest. – *Keel ja Kirjandus*, 12, 915–922.
- Vare, Silvi 2005. Eesti keele verbimoodustus: desubstantiivne konversioon. – *Emakeele Seltsi aastaraamat*, 50 (2004). Tallinn: Eesti Teaduste Akadeemia Emakeele Selts, 39–67.
- Vare, Silvi 2006. Adjektiivide substantivatsioonist ühe tähendusrühma näitel. – Ellen Niit (Toim.). *Keele ehe*. Tartu Ülikooli eesti keele õppetooli toimetised, 30. Tartu: Tartu Ülikool, 205–222.

Velsker, Eva 2006. Ööd tähistavad ajaväljendid eesti murretes, ühis- ja kirjakeeles. – Emakeele Seltsi aastaraamat, 51 (2005). Tallinn: Eesti Teaduste Akadeemia Emakeele Selts, 184–207.

VMS = Väike murdesõnastik. Toimetanud Valdek Pall, koostanud A. Haak, E. Juhkam, M. Kallasmaa, A. Kask, V. Lonn, E. Niit, P. Norvik, V. Oja, P. Raud, E. Ross, A. Sepp, J. Simm, J. Viikberg. <http://www.eki.ee/dict/vms/> (20.09.2010).

### **Võrgumaterjalid**

ESTMORF. <http://www.eki.ee/keeletehnoloogia/projektid/estmorf/> (28.09.2010).

Tartu Ülikooli vana kirjakeele korpus. <http://www.murre.ut.ee/vakkur/korpused/> (25.09.2010).

**Külli Habichti** (Tartu Ülikool) teaduslikeks huvialadeks on eesti vanema kirjakeele morfosüntaks ja sõnavara.  
[kulli.habicht@ut.ee](mailto:kulli.habicht@ut.ee)

**Külli Prillopi** (Tartu Ülikool) põhilised uurimisvaldkonnad on eesti vana kirjakeele morfofonoloogilised nähtused ja korpuslingvistika.  
[kulli.prillop@ut.ee](mailto:kulli.prillop@ut.ee)

**Pille Penjam** (Tartu Ülikool) tegeleb oma uurimistöös peamiselt eesti vanema ja tänapäeva kirjakeele morfosüntaktiliste nähtuste uurimisega.  
[pille.penjam@ut.ee](mailto:pille.penjam@ut.ee)

# **PARTS OF SPEECH AS A FUNCTIONAL AND LINGUISTIC PROBLEM: ANNOTATION OF PARTS OF SPEECH IN THE CORPUS OF OLD WRITTEN ESTONIAN**

**Külli Habicht, Pille Penjam, Külli Prillop**

University of Tartu

The article provides an overview of parts of speech annotation adjusted for the corpus of Old Written Estonian, against the background of parts of speech typology on the one hand and the previous treatments of parts of speech on the other. The corpus user is introduced to the principles of annotation and previous treatments of parts of speech. The core of the paper is constituted by practical issues of annotation concerning the boundaries of different parts of speech. Examples are drawn from the University of Tartu Corpus of Old Literary Estonian (<http://www.murre.ut.ee/vakkur/korpused>).

The system of 12 parts of speech used in the existent grammatical description has been extended for the morphosyntactic annotation of texts by adding a few parts of speech relevant for the analysis of old texts. The categories of *article*, *noun (name)*, *participle* and *(uninflected) modifier* have been added. While deciding on the number of parts of speech to be distinguished, the possibilities of semiautomatic annotation have also been taken into account.

The annotation of parts of speech makes it possible to specify the distinctive nature of the textual usage of words. Parts of speech also form an important basis for syntactic analysis. Issues related to parts of speech in Old Literary Estonian are mainly associated with the frequent use of archaic or borrowed word forms. The diachronic view used in the article also enables emphasizing interesting grammaticalization and lexicalization paths in the history of the literary language – a number of the units discussed in the article have undergone changes in function over time. An approach that recognizes overlapping borders permits the characterization of several instances of parts-of-speech-related recycling of units.

If on the one hand the paper introduces a particular instance of linguistic application, on the other it aims to raise more general issues of distinguishing parts of speech in order to enliven the discussion on the topic.

**Keywords:** morphosyntactical tagging, corpus linguistics, old literary Estonian, Estonian