

TARTU ÜLIKOOL
FILOSOOFIATEADUSKOND
EESTI JA ÜLDKEELETEADUSE INSTITUUT

Mirjam Ruutma

KAASSÕNADE *LÄBI, MÖÖDA, VASTU, ÜLE* JA *ÜMBER*
PRE- JA POSTPOSITSIONILISEST VARIEERUMISEST
EESTI MURRETES

Bakalaureusetöö

Juhendajad Kristel Uiboaed, Aki-Juhani Kyröläinen ja Maarja-Liisa Pilvik

Tartu 2014

Sisukord

Sissejuhatus	3
1. Teoreetiline taust	5
1.1. Korpuspõhine dialektoloogia	5
1.2. Kaassõnad	6
1.2.1. Kaassõnade kujunemine	7
1.2.2. Kaassõnafraasi struktuur	11
1.2.3. Ambipositsioonid <i>läbi, mööda, vastu, üle ja ümber</i>	13
2. Andmestik	16
2.1. Andmestiku kirjeldus	16
2.2. Andmestiku kodeerimine	20
3. Meetodid	26
3.1. Uurimismeetodi iseloomustus	26
3.2. Klassifitseerimispuu	26
4. Analüüs: <i>läbi, mööda, vastu, üle ja ümber</i> varieerumisest murretes	28
4.1. Pre- ja postpositsiooniline varieerumine klassifitseerimispuu põhjal	28
4.2. Tulemuste kokkuvõte	36
Kokkuvõte	39
Kirjandus	41
Pre- and postpositional variation of adpositions <i>läbi, mööda, vastu, üle and ümber</i> in Estonian dialects. Summary	44

Sissejuhatus

Käesolevas bakalaureusetöös käsitlen kahe paiknemisvõimalusega kaassõnade – *läbi*, *mööda*, *vastu*, *üle* ja *üumber* – kasutuse varieerumist pre- ja postpositsioonidena eesti murretes. Selliseid pre- ja postpositsioonidena esinevaid kaassõnu nimetan siinses töös ambipositsioonideks. Terminit ambipositsioon (*ambiposition*) on soovitanud kasutada näiteks prantsuse keeleteadlane Claude Hagège (2010: 114) ning termin on kasutusel mitmetes sama teemat käsitlevates artiklites (nt Janda jt 2014). Ingliskeelses kirjanduses on mõiste tähistamiseks esitatud veel termin *biposition* (nt Ehala 1995; Grünthal 2003; Huumo 2013). Eestikeelses kirjanduses ei ole nähtuse tähistamiseks terminit välja kujunenud.

Bakalaureusetöö eesmärk on uurida, kas kaassõnade *läbi*, *mööda*, *vastu*, *üle* ja *üumber* kasutuses pre- ja postpositsioonidena esineb geograafilist varieerumist ning selgitada, mis võib võimalikku pre- ja postpositsioonidena esinemise varieerumist mõjutada. Töö hüpoteesideks on, et kaassõnade asendil kaassõnafaasis esineb murdeti geograafilist varieerumist ning pre- ja postpositsioonina esinemise varieerumist mõjutavad uurimusse kaasatud ja kodeeritud keelesised faktorid (täpsemalt vt peatükk 2.2). Kodeeritud tegurid olen valinud varasemate uurimuste põhjal.

Siinse töö analüüsitav materjal on kogutud eesti murrete korpuse morfoloogiliselt mägendatud murdetekstidest ning seega on uurimus korpuspõhine. Kogutud materjalist koostatud andmestik sisaldab kokku 1062 uuritavate kaassõnadega lausungit. Bakalaureusetöös järgin korpuspõhise dialektoloogia põhimõtteid ning empiirilist materjali analüüsin klassifitseerimispuu (*conditional inference tree / classification tree*) abil, mida tutvustan lähemalt peatükis 3.2. Kui sageli uuritakse keelise nähtuse varieerumist vaid ühe muutuja põhjal, arvestan siinses analüüsis erinevate keelekasutust mõjutada võivate teguritega korraka. Seega sisaldab bakalaureusetöö ka katset rakendada uusi meetodeid ja tehnikaid, mis murdeuurimises pole veel laialdast kasutamist leidnud.

Töö koosneb neljast peatükist, mis omakorda jagunevad alapeatükkideks. Esimeses peatükis tutvustan töö teoreetilist tausta. Esmalt iseloomustan korpuspõhise dialektoloogia põhimõtteid ning seejärel käsitlen kaassõnu üldisemalt. Vaatlen nii kaassõnade kujunemist soome-ugri keeltes kui ka kaassõnadega seonduvaid nähtuseid. Esimese peatüki viimases osas kirjeldan täpsemalt ambipositsioone *läbi*, *mööda*, *üle*, *ümber* ja *vastu*. Töö teoreetiline osa toetub peamiselt eesti kirjakeele uurimustele, sest kaassõnu on murretes vähe uuritud. Varem on kaassõnu eesti murretes käsitlenud näiteks Harald Ploom (1955) oma deskriptiivses diplomitöös. Samuti on Pauline Palmeos (1982) kirjeldanud mõningaid murdelisi eripärasid eesti keele kaassõnasüsteemis.

Töö teises peatükis annan ülevaate analüüsis kasutatud andmestikust. Esmalt tutvustan andmestiku algmaterjali ning uurimismaterjali kogumise kriteeriumeid. Seejärel kirjeldan andmestiku kodeerimise põhimõtteid. Kolmandas peatükis annan ülevaate uurimismeetoditest, mida olen rakendanud neljandas peatükis, töö analüüsiosas. Töö viimane, neljas peatükk, sisaldabki analüüsi ambipositsioonide paiknemisest kaheksa potentsiaalselt olulise teguri põhjal, mis eelnevate uurimuste järgi võivad olla olulised pre- või postpositsioonilise asendi valikul kaassõnafraasis.

1. Teoreetiline taust

Järgnevates peatükkides annan ülevaate teoreetilisest raamistikust, millele bakalaureusetöö tugineb. Esmalt tutvustan lühidalt korpuspõhise dialektoloogia põhimõtteid. Seejärel määratlen, mis on kaassõna, annan ülevaate kaassõnasüsteemi tekkest eesti keeles ning kirjeldan tegureid, mis võivad olla seotud kaassõnade paiknemise varieerumisega teda laiendava noomenifraasi suhtes. Lisaks annan ülevaate ambipositsioonidest *läbi*, *mööda*, *vastu*, *üle* ja *ümber* eesti kirjakeeles ja võimaluse korral esitan ka varasemad tähelepanekud vaadeldava kaassõna kohta eesti murretes.

1.1. Korpuspõhine dialektoloogia

Käesolev bakalaureusetöö tugineb korpuspõhise dialektoloogia põhimõtetele, mida siinkohal täpsemalt tutvustan. Korpuspõhine dialektoloogia eristub traditsioonilisest murdeuurimisest eelkõige eesmärkide ja meetodite poolest (Szmrecsanyi, ilmumas).

Võrreldes traditsioonilise dialektoloogiaga, kus uuritakse eelkõige murrete leksikaalseid ja foneetilisi erinevusi, keskendutakse korpuspõhises dialektoloogias eeskätt murde grammatikale ja morfosüntaksile ning analüüsi kaasatakse korraga mitmeid keelenähtuseid. Seejuures on keskseks uurimisobjektiks loomulik keelekasutus, mis on traditsioonilises dialektoloogias jäänud juba materjali kogumismeetodite tõttu tagaplaanile. (Szmrecsanyi, ilmumas: 2) Traditsiooniliselt on dialektoloogilise uurimuse algmaterjaliks murdeatlased, mille andmestik põhineb reeglina küsitluskavadel. Murdeainestiku kogujad ja atlase koostajad on materjali lihtsustanud ja seetõttu pole andmestik piisavalt objektiivne ega kajasta loomulikku keelekasutust kõige paremini. Korpusandmestikud sisaldavad seevastu üldjuhul muutmata keelematerjali, spontaanset kõnet, mis võimaldab saada uuritavast keelenähtusest objektiivsema pildi. Korpusandmestik võimaldab kaasata analüüsi ka keelenähtuse sageduse ning korpuspõhist dialektoloogiat võib nimetada seega ka sageduspõhiseks. (Szmrecsanyi, Wolk 2011: 564–565)

Szmrecsanyi ja Wolki (2011: 561–566) järgi peaks korpuspõhine uurimus toetuma järgmistele põhimõtetele:

- Analüüs ei tohiks põhineda ainult ühe keelilise nähtuse uurimisel, vaid kaasatud peaks olema võimalikult palju erinevaid keelenähtuseid. Sel viisil väheneb üksikute irrelevantsete tegurite mõju analüüsile.
- Andmestiku analüüsimiseks peaks rakendama mitmemõõtmelise statistika meetodeid.
- Uurimistulemusi peaks kindlasti erinevaid tehnilisi võimalusi rakendades visualiseerima.

Siinses bakalaureusetöös uurin küll ainult kaassõnu ja teisi keelenähtusi oma analüüsi ei kaasa, kuid töö toetub muus osas korpuspõhise dialektoloogia põhimõtetele.

1.2. Kaassõnad

Kaassõnad ehk adpositsioonid on muutumatud sõnad, mis kuuluvad lauses noomeni või noomenifraasi juurde (EKG I 1995: 33). Adpositsioonide põhiline ülesanne on süntaktiline, s.t kaassõnad väljendavad erinevaid lauseliikmetevahelisi suhteid (Palmeos 1982: 7). Lisaks täidavad kaassõnad käändetunnustega sarnaseid funktsioone ja seega on neil ka morfoloogilisi ülesandeid. Käändetunnustega võrreldes annavad adpositsioonid aga edasi täpsemaid ja mitmekesisemaid tähendusvarjundeid. Kaassõnade tähendus ilmneb alles kontekstis teda laiendava noomeniga koos ning ühel ja samal kaassõnal võib seetõttu olla eri kontekstis erinevaid tähendusi ehk neil võib olla kontekstile vastavalt erinevaid tähendusfunktsioone. (EKG I 1995: 33–34)

Niisiis ei saa kaassõna esineda iseseisvalt, kaotamata oma süntaktilist või morfoloogilist funktsiooni, ja kaassõna on laiendavast noomenist eraldamatu. Koos moodustavad kaassõna ja kohustuslik laiendav noomen ehk komplement kaassõnafraasi (Karlsson 2002: 164). Kaassõnafraasis võib adpositsioon asetseada oma komplemendi ees või järel. Vastavalt sellele nimetatakse komplemendile eelnevaid kaassõnu eessõnadeks ehk prepositsioonideks ning komplemendile järgnevaid kaassõnu tagasõnadeks ehk postpositsioonideks. Lisaks leidub kaassõnu, mis esinevad nii pre- kui

ka postpositsioonidena. Seejuures võib sellise kaassõna tähendus oma paiknemise järgi kas muutuda või püsida samana. (EKG I 1995: 37) Siinses töös nimetatakse nii pre- kui ka postpositsioonidena esinevaid kaassõnu ambipositsioonideks.¹

1.2.1. Kaassõnade kujunemine

Keeles toimuvad uuendused mitmel viisil: analoogia alusel olemasoleva keelematerjali kasutamise või muutmise teel, teistest keeltest keelematerjali laenamisel või täiesti uue leiutamisel (Lass 1997: 305). Samamoodi võivad olla tekkinud ka eesti pre- ja postpositsioonid.

Kaassõnad on sõnaliigina soome-ugri keeltes hiljutine nähtus. Nad pärinevad tõenäoliselt läänemeresoome algkeele ajastust või isegi perioodist, kui eesti keel oli juba algkeelest eraldunud. Kaassõnade tüved ning struktuurid, millest kaassõnad on tekkinud, põlvnevad varasemast perioodist. Seejuures on eessõnad tekkelt hilisemad kui tagasõnad. (Palmeos 1982: 69–72)

Prepositsioonide hilise tekke põhjuseks võib olla soome-ugri keelte lausestruktuur ning prepositsioonide erinev kujunemislugu. Seda kinnitab prepositsioonide puudumine paljudes soome-ugri keeltes, näiteks ungari ja mordva keeles esinevad üksnes tagasõnad. (Ploom 1955: 11) Soome-ugri keelte tagasõnade eelistust on toodud esile juba esimestes 17.–19. sajandi võrdlevates uurimustes ja tagasõnalisusega koos esinevaid iseloomulikke jooni on võimalik taastada ka vanimas rekonstrueeritavas algkeeles (Grünthal 2003: 34).

Soomeugriline postpositsioonide eelistus ilmneb ka eesti keeles – „Eesti keele seletava sõnaraamatu” veebiversioonis tehtud otsingu põhjal on eesti keeles 266 adpositsiooni, millest 193 esinevad ainult postpositsioonidena, 48 ainult prepositsioonidena ja 25 ambipositsioonidena. Järelikult on eesti keele kaassõnadest suurem osa, ligi 73%, postpositsioonid. Prepositsioonide osakaal on 18% ning ambipositsioonide osakaal on 9%.

Postpositsioonid on tekkinud keelesisestest muutustest ning pärinevad enamjaolt nimisõnadest, millele oli liitunud käändelõpp (EKG I 1995: 38). Mõned kaassõnad peegeldavad praegugi kunagise noomeni tunnuseid, käändudes kohakäändeis (Ploom

¹ Terminit ambipositsioon (*ambiposition*) on kasutanud ka näiteks Hagège (2010: 114) ja Janda jt (2014).

1955: 5), näiteks deklioneeruvad väliskohakäändeis kaassõnad *järele*, *järel* ja *järelt* (Palmeos 1982: 66). Koos teise nimisõnaga moodustasid käänduvad nimisõnad possessiivse konstruktsiooni (EKG I 1995: 38), s.t teine nimisõna oli omastavas käändes. Näitena possessiivse konstruktsiooniga postpositsiooni tekkest võib tuua fraasi *maja ees*, kus *ees* on algselt olnud ruumimõistet tähistava nimisõna käändeline vorm ja *maja* seda laiendav genitiivne täiend (EKK 2007: 195). Possessiivne konstruktsioon võib põhjendada postpositsioonide rohkest soome-ugri keeltes, sest soome-ugri keelte reegleid järgides peab genitiivne täiend eelnema oma põhisõnale (Grünthal 2003: 36; Reimann 1958: 102–103). Aja jooksul abstraherusid sageli kasutatavad genitiivsele täiendile järgnevad nimisõna käänevormid tähenduslikult. Nimisõna käänevormid kaotasid oma leksikaalse sisu ja hakkasid väljendama abstraktseid suhteid lauseliikmete vahel. (EKK 2007: 195)

Prepositsioonide kujunemine pole sama läbipaistev nagu postpositsioonide tekkimine ning tõenäoliselt on prepositsioonide teke põhjustatud nii mitmetest keelesisetest muutustest kui ka keelekontaktidest (Grünthal 2005: 30). Enamjaolt pärinevad prepositsioonid adverbidest (EKG I 1995: 38; Palmeos 1982: 72). Verbi laiendanud partitiivne noomen eraldus verbist ja liitus adverbiga, mille tulemusel sai adverbist prepositsioon (Grünthal 2005: 41). Seetõttu nõuab prepositsioon tavaliselt oma komplementilt ka partitiivivormi (EKG I 1995: 37).

Ometi ei ole prepositsiooniline kaassõnafraas soome-ugri keelte lauseehitusele omane (Grünthal 2005: 29). Näiteks on peale läänemeresoome ja saami keelte kõik soome-ugri keeled SOV-sõnajärgiga ning neis keeltes esineb ainult postpositsioone. Seega on SOV-sõnajärgiga soome-ugri keeled kooskõlas keeleuniversaaliga, mille järgi SOV-keeled on peaaegu eranditult postpositsioonilised. (Grünthal 2003: 45) SVO-sõnajärgiga keeltes esinevad seevastu põhiliselt prepositsioonid. Läänemeresoome ja saami keeltes toimunud morfoloogilise erosiooni tõttu on hakatud lause moodustajaid eristama sõnajärje abil (Grünthal 2005: 29–30) ja sellega on kaasnenud ka ajaloolise SOV-sõnajärje muutumine SVO-liseks (Grünthal 2003: 45). Niisiis on võimalik, et sõnajärje muutudes on muutunud ka kaassõnafraasi järjestus prepositsiooniliseks (Grünthal 2005: 30).

Üksnes sõnajärje muutus ei seleta siiski eesti prepositsioonide teket, sest kui prepositsioonid oleksid ainult sõnajärje muutuse tagajärg, peaks neid laiendav komplement olema genitiivis nagu postpositsioonide puhul (Grünthal 2003: 35–45). Nagu juba varem täheldatud, nõuavad prepositsioonid aga oma komplementilt enamjaolt partitiivi (EKG I 1995: 37). Nagu ilmnes „Eesti keele seletava sõnaraamatu” veebiversioonis tehtud otsingus, on eesti kaassõnad enamasti postpositsioonid, kuigi SVO-keeled peaksid eelistama prepositsioonilist kaassõnafraasi.

Teiste keelte mõju kaassõnasüsteemile ei saa samuti olla ainus põhjus prepositsioonide tekkele, sest kaassõnad pole võõrmõjutustele sama vastuvõtlikud nagu muud sõnaliigid (Lass 1997: 190). Näiteks on läti keel mõjutanud küll liivi keele leksikat, foneetikat ja süntaksit, kuid läti keelest on liivi keelde laenatud vaid üksikuid prepositsioone (Grünthal 2003: 53). Kui keelekontaktidel oleks tugev mõju kaassõnafraasi ülesehitusele, tõuseks tõenäoliselt lisaks laenatud kaassõnafraasi sõnajärjele esile ka kaassõnade laenamine sõnavarasse. Kuna prepositsioonid on oletatavasti tekkinud hiljem kui postpositsioonid (ehk pärast eesti keele eraldumist läänemeresoome algkeelest) ja prepositsioone esineb siiski ka teistes läänemeresoome keeltes, on eeldused prepositsioonide tekkeks olnud olemas juba algkeele oma struktuuris. (Grünthal 2005: 37–38)

Niisiis on prepositsioonid kujunenud mitme teguri tulemusena, milleks on muutused adverbide grammatilistes suhetes ning sõnajärjes ja, kuigi harvem, laenud teistest keeltest.

Pole tavaline, et keeles esineks samaaegselt nii pre- kui ka postpositsioone. Veelgi haruldasem on ühe ja sama kaassõna kasutamine mõlemas positsioonis. Sellele, miks mõned adpositsioonid on kujunenud ambipositsioonideks ja eiravad kaassõnafraaside süntaktilist järjepidevust, ei ole ilmset seletust. Eesti ambipositsioonide puhul võib täheldada süntagmatilist sobitumist kaassõnafraasi komponentide järjestuse järgi: postpositsioonidena esinedes nõuavad nad genitiivi ja prepositsioonidena esinedes partitiivi. Näiteks on prepositsioonina *vastu* komplement partitiivis (*vastu ust*) ja postpositsioonina esinedes on *vastu* komplement genitiivis (*ukse vastu*). Ometi on ambipositsioone, mis ei järgi seda reeglit. Näiteks esineb kaassõna *läbi* komplement nii pre- kui ka postpositsioonina genitiivis: *ukse läbi* ja *läbi ukse*. Niisiis võib

kaassõnafraasile omane morfosüntaktiline struktuur olla tekkinud ka hilisema kohanemisega komponentide järjekorra järgi – esmalt on olnud nii pre- kui ka postpositsiooniga komplemendi käändeks genitiiv ning partitiivne vorm on komplemendile kujunenud alles hiljem. Seega on võimalik, et adpositsioonid olid esialgu ambipositsioonilised ja omandasid alles hiljem praeguste adpositsioonidele iseloomuliku morfosüntaktilise struktuuri, nõudes komplemendilt teatud kääned. (Grünthal 2003: 67, 84)

Ehala (1995: 328–338) järgi võib adpositsioonidel olla keeles kaks paiknemise võimalust vaid sel juhul, kui pre- ja postpositsioonina nõutakse komplemendilt erinevat kääned, muidu tekiks adpositsiooni kahe võimaliku paiknemisega liigset mitmetähenduslikkust. Niisiis võib prepositsioonide partiivivormis komplement ja postpositsioonide genitiivivormis komplement seletada, miks eesti keeles on mõne kaassõna puhul võimalik kasutada nii pre- kui ka postpositsioonilist paiknemist. Ometi selgub siinse töö peatükis 1.2.3, et eesti ambipositsioonid seda reeglit alati ei järgi. See kinnitab taas Grünthali (2003: 84) esitatud oletust, et kui adpositsioonid tekkisid, oli kaassõnafraasi sõnajärg vaba ja adpositsioonidega kaasnevate komplementide käänded kujunesid keeles hiljem.

Eesti keele ambipositsioonide asendis on täheldatud prepositsioonide eelistamise kasvu (Ehala 1994: 177; Hint 1990: 1401). Seda on seletatud prepositsioonirohke vene keele mõjuga eesti keelele (Hint 1990). Sarnast keelekontaktide tagajärjel suurenevat prepositsioonide eelistamist on toimunud ka põhjasaami murretes, kus kontakt rootsi ja norra keelega seostub prepositsioonide eelistamisega ning kontakt soome keelega jällegi postpositsioonide eelistamisega (Janda jt 2014). Kuid nagu juba eelnevalt märgitud, pole adpositsioonide süsteem kuigi aldis võõrmõjutustele. Näiteks oli eesti keel enne vene keele mõju alla sattumist pikalt kontaktis saksa keelega, ent ometi ei toimunud saksa keele mõjul kaassõnade kasutuses muutuseid (Ehala 1994: 177–178). Niisiis peab lisaks võõrmõjutustele olema muutuste ajendiks ka muid keelesiseseid tegureid. Mõnigaid neist teguritest kirjeldan täpsemalt peatükis 1.2.2.

Üldjoontes võib järeldada, et adpositsioonid järgivad seda sõnajärge, mida on järginud ka lekseemid, millest nad on tekkinud (Hagège 2010: 110). Kaassõnade kujunemisprotsess järgib tüüpilisi grammatiseerimise tunnuseid: sõnavorm omandab

muutumatu vormi, siirdub avatumast klassist rohkem suletud klassi, kinnistub süntaktiliselt ja ei saa esineda iseseisvalt ilma noomenita (Jürine, Habicht 2013: 740). Seejuures võivad kaassõnad olla oma kujunemiselt eri etappides. Lehmann (1985: 303–318) esitab oma grammatisatsioonimudelis viit adpositsioonide arenemisjärku: suhet väljendav substantiiv > sekundaarne adpositsioon > primaarne adpositsioon > aglutinatiivne afiks > fusiivne afiks. Nii-öelda sekundaarsete kaassõnade seos lähtevormiga on läbipaistev ning selline kaassõna on alles kujunemisfaasis. Primaarsete kaassõnade seos lähtesõnaga pole enam selge. Lähtesõna on foneetiliselt kulunud ning on keeles juba juurdunud. (Habicht, Penjam 2007: 52–54)

Mõned postpositsioonid on jõudnud selles grammatisatsiooni protsessis viimastele etappidele ja on muutunud käändelõppudeks. Näiteks on sel moel tekkinud eesti kaasaütlev kääne. (Palmeos 1982: 70) Niisiis on kaassõnade süsteem pidevas muutuses.

1.2.2. Kaassõnafraasi struktuur

Adpositsioone võib paiknemise järgi jagada kolme rühma (Grünthal 2005: 31):

- 1) komplementidele eelnevad adpositsioonid ehk prepositsioonid,
- 2) komplementidele järgnevad adpositsioonid ehk postpositsioonid,
- 3) mõlema positsiooni võimalusega ambipositsioonid.

Selles peatükis keskendun kolmandale rühmale ning esitan varasemate uurimuste põhjal, mis võib mõjutada valikut pre- ja postpositsioonide vahel, kui mõlemad valikud on võimalikud.

Pre- ja postpositsioonifraasi keskne erinevus on nende morfosüntaktilises struktuuris. Tavaliselt on eesti keeles prepositsioonifraas oma morfosüntaktilise ülesehituse poolest partitiivse rektsiooniga [Prep+N[PART]] ja postpositsioonifraas genitiivse rektsiooniga [N[GEN]+Post]. (EKG I 1995: 37–38)

Tegelikus keelekasutuses ei ole jagunemine rektsiooni järgi sama ilmne. Komplement võib olla lisaks genitiivile ja partitiivile ka muudes käänetes ja mõned ambipositsioonid võivad paikneda emmas-kummas positsioonis, muutmata oma komplementi käänat. Kui ambipositsiooni puhul pre- või postpositsiooniline

paiknemine komplemendi käänet ei muuda, lisab genitiivne rektsioon postpositsiooniline sõnajärje eelistust ja partitiivne rektsioon prepositsioonilise sõnajärje eelistust. (Grünthal 2005: 32–33)

Samuti on kaassõnafraasi struktuur seotud adpositsiooni tähendusfunktsiooniga (Palmeos 1982). Varasemates uurimustes on leitud, et näiteks soome keeles on aega tähistades valdav prepositsioonide kasutus ja ruumisuhteid tähistades on tavalisem postpositsioonide kasutus (Grünthal 2005: 33; Huumo 2013: 320; Janda jt 2014: 92). Isegi kui ambipositsioon võib esineda aja tähenduses mõlemas positsioonis, on prepositsiooniline sõnajärg loomulikum valik (Huumo 2013: 320).

Kaassõna paiknemist võib mõjutada analoogia alusel ka teine sünonüümne kaassõna. Näiteks on *läbi* kasutuses toimunud *kaudu* analoogial nihe, mille tulemusena on *läbi* prepositsiooniline kasutus sagenenud. *Kaudu* esineb alati tagasõnana, kuid täidab seejuures kahte eri tähendusfunktsiooni – vahendi tähendust (*kohtu kaudu*) ja ruumilist tähendust (*metsa kaudu*). Kuna postpositsioonina esinev *kaudu* kannab kahte tähendust, on ta sünonüümiks nii prepositsioonina ruumisuhteid väljendavale *läbi*-le (*läbi metsa*) kui ka postpositsioonina esinevale vahendi tähendusfunktsiooni täitvale *läbi*-le (*kohtu läbi*). See on aga tekitanud *läbi* kasutuses ebastabiilsust ning eessõnalist *läbi* on hakatud tarvitama ka vahendi tähendusfunktsiooniga *kaudu* asemel. Nõnda on ees- ja tagasõnaline *läbi* muutunud sünonüümseteks (*kohtu läbi* ~ *läbi kohtu*), kuigi algselt väljendas *läbi* eri positsioonides eri tähendust. (Ehala 1994: 184–192)

Lisaks määrab ambipositsiooni paiknemise ka komplemendi struktuur. Postpositsiooni ja komplemendi vahel ei saa olla ühtegi teist sõna, kuid prepositsiooni ja komplemendi vahel võib olla noomeni määruseid, pronoomeneid ja muid atribuute. (Palmeos 1982: 3) Seega, mida keerukam² on komplemendi struktuur, seda tavalisem on prepositsiooni kasutus (Huumo 2013: 320).

Niisiis võib kaassõna paiknemine sõltuda prepositsioonifraasi ja postpositsioonifraasi morfosüntaktilisest ülesehitusest, kaassõna tähendusfunktsioonidest, komplemendi raskusest, analoogiast ja peatükis 1.2.1 esitatud keeleajaloolisest taustast.

² Inglise keeles *heaviness*, mis võib tähendada fraasi pikkust sõnadena, silpidena või häälikutena ning fraasi grammatilist keerukust (Huumo 2013).

1.2.3. Ambipositsioonid *läbi, mööda, vastu, üle ja ümber*

Töös uuritavad kaassõnad *läbi, mööda, vastu, üle, ja ümber* võivad esineda nii pre- kui ka postpositsioonidena. Olenevalt paiknemisest võib nende ambipositsioonide tähendusfunktsioon muutuda ning seejuures võivad need ambipositsioonid nõuda oma komplemendilt üht ja sama käänat või erinevaid käändeid (EKG I 1995: 37).

Järgnevalt tutvustan viie uuritava kaassõna erinevaid tähendusfunktsioone ja komplementide võimalikke käändeid. Palmeose (1982) ja Ploomi (1955) esitatud kaassõnade tähendused on kokku sobitatud EKG tähendusfunktsioonidega. Kui murretes on kirjakeelest erinevusi, olen välja toonud ka murrete eripärad.

Läbi on funktsioonirohke kaassõna. Oma laiendavalt noomenilt nõuab ta enamasti genitiivi, kuid harvem ka nominatiivi. Nominatiivse reksiooniga esinedes tähistab *läbi* üksnes ajasuhteid (*öö läbi*) ning esineb ainult postpositsioonina. Kõikidel muudel juhtudel esineb *läbi* komplement alati genitiivis. Genitiivse komplemendiga ajakaassõnana on *läbi* prepositsioon (*läbi aastate*). Lisaks võib *läbi* prepositsioonina väljendada ka viisi-seisundi suhteid (*ütleb läbi hammaste*) või kohasuhteid (*läbi värava*). Viisi-seisundi tähendust ja kohasuhteid tähistades esineb *läbi* ka postpositsioonina (*sai sõja läbi hukka, värava läbi*). Postpositsioonina väljendab *läbi* ka vahendi (*käe läbi*), abinõu (*õnne läbi*) ja põhjuse (*ülekohtu läbi*) tähendusfunktsioone. (Palmeos 1982: 25–39, 60–61; Ploom 1955: 152–155)

Kaassõna *mööda* tarvitatakse ainult partitiivse komplemendiga. Seejuures kohakaassõna funktsiooni täites on *mööda* kasutamine prepositsioonina või postpositsioonina paralleelne ning selle tähendusfunktsioon ei muutu oluliselt (*mööda teed ~ teed mööda*). Harvem tähistab *mööda* prepositsioonina sama tähendust, mis kaassõna *vastu* (*andis mööda kõrvu ~ andis vastu kõrvu*). Postpositsioonina väljendab *mööda* sama tähendust nagu kaassõna *järgi* genitiivse komplemendiga (*enese heaksarvamist mööda ~ enese heaksarvamise järgi*). (Palmeos 1982: 26, 36, 53–54) Lõunaeeesti murretes kasutatakse ambipositsiooni *mööda* tunduvalt vähem kui põhjaeeesti murretes, sest *mööda* funktsiooni täidab kaassõna *pidi*, mis erinevalt kaassõnast *mööda* esineb üksnes postpositsioonina (Ploom 1955: 155–158).

Vastu kannab postpositsioonina rohkem tähendusi kui prepositsioonina. Seejuures on postpositsioonina esinedes kaassõna *vastu* laiendav noomen genitiivis.

Postpositsioonina väljendab *vastu* koha, aja, viisi, suhtumise, suhte ja asenduse tähendusi (*ukse vastu, õhtu vastu, silm silma vastu, minu vastu ollakse kuri, ei tunne tema vastu huvi, vahetab kinga teise vastu*). Prepositsioonina nõuab *vastu* oma laiendavalt noomenilt partitiivi. Sel juhul esineb *vastu* kohta, aega ja viisi väljendava kaassõnana (*vastu akent, vastu ööd, vastu tahtmist*). (Palmeos 1982: 23–47, 58–59; Ploom 1955: 186–193)

Üle nõuab oma komplemendilt üksnes genitiivi. Prepositsioonina esineb *üle* kohta, aega, viisi-seisundi ja kvantiteeti väljendava kaassõna tähenduses (*üle tee, üle kesköö, üle jõu, üle kahe kilo*). Postpositsioonina esineb *üle* harva ja võib täita koha ja kvantiteedi tähendusfunktsiooni (*maja üle ripuvad oksad, viie kilo üle*). Prepositsioonina ei väljenda *üle* põhjuslikke suhteid, kuid postpositsioonina esineb *üle* ka põhjuskaassõnana, olles sünonüümiks kaassõnale *pärast* (*selle üle hea meel ~ selle pärast hea meel*). (Palmeos 1982: 24–43, 50)

Ümber võib ruumilisi suhteid väljendades esineda mõlemas positsioonis, muutmata oluliselt oma tähendust (*ümber aia ~ aia ümber*). Ligikaudsust tähistades võib aga *ümber* esineda üksnes postpositsioonina (*kahekümne kilo ümber*). *Ümber* nõuab mõlema tähendusfunktsiooni puhul oma laiendavalt noomenilt genitiivi. (Palmeos 1982: 24, 42, 50; Ploom 1955: 200–204)

Tabel 1 võtab Palmeose (1982) põhjal kokku, mis on morfosüntaktilise ülesehituse järgi kaassõnade *läbi, mööda, vastu, üle* ja *ümber* tähendusfunktsioonid.

Tabel 1. Tähen­dus­funktsioonid (Palmeose (1982) põhjal)

	[Prep + N[GEN]]	[Prep + N[PART]]	[N[GEN] + Postp]	[N[PART] + Postp]	[N[NOM] + Postp]
läbi	aeg viis seisund koht	-	vahend viis seisund koht põhjus	-	aeg
mööda	-	koht	-	koht viis	-
vastu	-	koht aeg viis	koht aeg viis suhtumine suhe asendus		-
üle	koht aeg kvantiteet viis seisund	-	koht põhjus kvantiteet	-	-
üumber	koht	-	koht ligikaudsus	-	-

Kokkuvõttes on ambipositsioonid *läbi*, *mööda*, *vastu*, *üle* ja *üumber* polüseemsed ehk neil kõigil on mitu tähendust. Siiski võiks neid ambipositsioone liigitada ka hoopis homonüümseteks ehk sama vormi ja kõlaga, kuid eri tähendusega lekseemideks. Kognitiivses keeleteaduses aga oletatakse, et kuigi ühel ja samal kaassõnal võib olla mitu tähendust, on sellel siiski üks algne tähendus, millest teised tähendused on tuletatud. (Veismann 2011: 9–10) Sellest oletusest lähtun ka selles uurimuses, ehkki kognitiivse keeleteaduse teoreetilised seisukohad on siinses töös jäänud tagaplaanile.

Kuigi üldiselt võib märkida, et prepositsioonid nõuavad oma komplemendilt partitiivi ning postpositsioonid genitiivi, siis siinses viiest ambipositsioonist kõik seda reeglit ei järgi. Prepositsioonina nõuavad partitiivset komplementi vaid *mööda* ja *vastu*. *Mööda* nõuab ka postpositsioonina partitiivset konstruktsiooni. Lisaks on erandlik kaassõna *läbi*, mis nõuab postpositsioonina ajasuhteid tähistades oma komplemendilt nominatiivi.

2. Andmestik

Selles peatükis kirjeldan töö analüüsis kasutatud andmestikku. Annan ülevaate algmaterjalist ja kriteeriumitest, mille alusel on uurimismaterjal kogutud. Lisaks esitan andmestiku kodeerimise põhimõtted.

2.1. Andmestiku kirjeldus

Käesoleva töö empiiriline materjal on pärit eesti murrete korpuse morfoloogiliselt märgendatud murdetekstidest. Tartu Ülikooli ja Eesti Keele Instituudi koostöös valminud murdekorpuse sisaldab autentset keelematerjali kõigilt Eesti murdealadelt. Korpuse koostamise põhieesmärkideks on teha murdematerjal elektrooniliselt kättesaadavaks ja võimaldada suure andmehulga põhjal murrete kõrvutav uurimine. (EMK) Tabel 2 esitab murdekorpuse morfoloogiliselt märgendatud sõnade arvu murrete kaupa andmestiku kogumise hetkel seisuga september 2013.

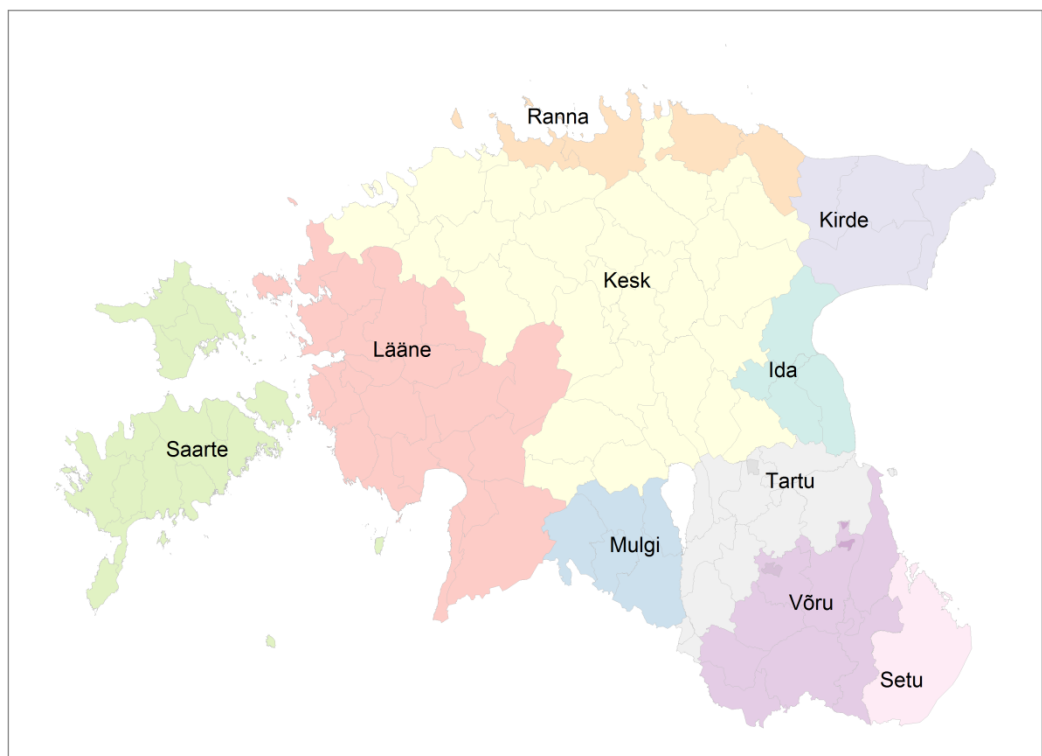
Tabel 2. Murdekorpuse morfoloogiliselt märgendatud sõnade arv murreti

	Sõnu
Ida	45280
Kesk	130086
Kirde	52637
Lääne	154400
Mulgi	63516
Ranna	54910
Saarte	158407
Setu	39175
Tartu	65591
Võru	66301
Kokku	830303

Eesti murrete korpus koosneb suuremas osas 1960.–1970. aastatel salvestatud murdelindistustest ja neil põhinevatest murdetekstidest – lihtsustatud transkriptsioonis

tekstidest, morfoloogiliselt märgendatud tekstidest ning süntaktiliselt analüüsitud tekstidest. Lisaks on korpusesse kogutud andmeid keelejuhtide, lindistuste ja litereeringute kohta. (EMK)

Korpuse murdeliigendus põhineb õpiku „Eesti murded ja kohanimed” esitatud murdejaotusel. Kuigi õpiku murdejaotuses on Setu murrakut vaadeldud Võru ühe murrakurühmana (Pajusalu jt 2009: 184), eraldatakse murdekorpuses Setu omaette murdeks. Nõnda on korpuses esitatud 10 murdeala: põhjaeesti murderühma moodustavad ida-, kesk-, lääne- ja saarte murre, lõunaeesti murderühma kuuluvad Mulgi, Võru, Tartu ja Setu murre ja kirderanniku murderühm jaguneb kirdemurdeks ja rannamurdeks. (EMK) Joonis 1 esitab murdealad kaardil.³



Joonis 1. Eesti murdealad

³ Murdeandmestiku visualiseerimiseks kaartidel olen kasutanud tehnikaid, mille on esitanud Kyröläinen ja Uiboaed (valmimas).

Uurimismaterjali olen kogunud murdekorpuse morfoloogiliselt märgendatud tekstidest ning olen kaasanud lausungid, mis vastavad järgmistele põhimõtetele: lausung sisaldab adpositsiooni, mis võib esineda nii pre- kui ka postpositsioonina; nende adpositsioonide esinemissagedus murdekorpuses on nii pre- kui ka postpositsioonina piisavalt suur kvantitatiivseks analüüsiks; nende adpositsioonide esinemissagedus nii pre- kui ka postpositsioonina on ühtlane. Kuigi eesti keeles on mitmeid ambipositsioone, vastas murdekorpuses nendele kriteeriumitele vaid viis: kaassõnad *läbi, mööda, vastu, üle* ja *üumber*.

Analüüsitav materjal hõlmab kõiki murdekorpuses esinenud kaassõnafraase, mis sisaldasid uuritavaid kaassõnu, ehk arvestatud on kogu korpuse üldkogumiga. Seega on siinses töös kasutatud kõikset uuringut, mis vastandub vaid ühte osa üldkogumist käsitlevale valikuuringule (Tooding 2007: 122–124). Kuna aga murdekorpus on käsitsi märgendatud, on märgendusvigade esinemine paraku vältimatu. Näiteks on kaassõnadeks märgendatud tihti ka adverbe ja afiksaaladverbe. Seetõttu olen andmestikku kogunud lausungid üle kontrollinud ja eemaldanud juhtumid, kus kaassõna asemel on tegemist (afiksaal)adverbiga. Lisaks olen välja jätnud lausungid, kus kaassõnafraasi tähendust polnud konteksti põhjal võimalik tuvastada.

Tabel 3 esitab uuritavate ambipositsioonide absoluutesinemissagedused andmestikus.

Tabel 3. *Läbi, mööda, vastu, üle* ja *üumber* esinemine andmestikus

	Post	Prep	Kokku
Ida	10	41	51
Kesk	50	132	182
Kirde	21	31	52
Lääne	106	150	256
Mulgi	41	41	82
Ranna	24	61	85
Saarte	73	88	161
Setu	7	34	41
Tartu	24	61	85
Võru	24	43	67

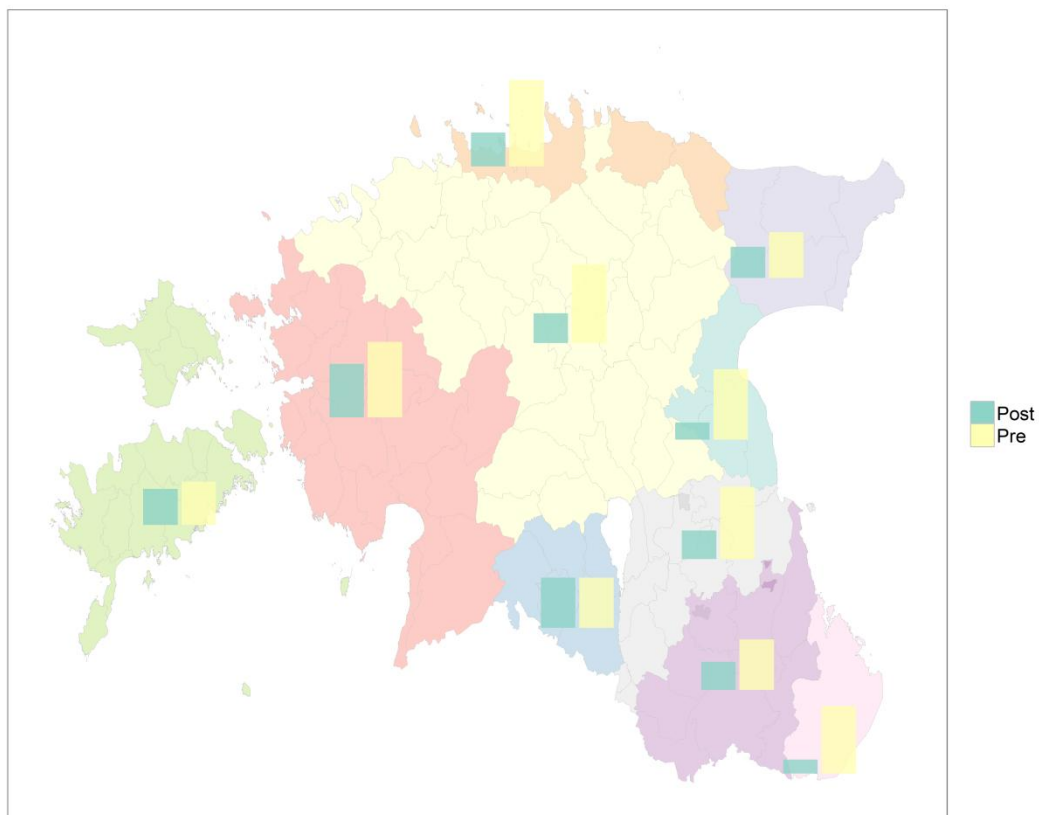
Märgendatud sõnade arv pole korpuses murdeti ühtlane ja seetõttu ei ole ambipositsioonide absoluutsagedused omavahel võrreldavad. Tabel 4 esitab uuritavate ambipositsioonide normaliseeritud sagedused murdekorpuses. Normaliseerimine on tehtud 100 000 sõna põhjal, mis tähendab, et tabelis on esitatud uuritavate ambipositsioonide oletatav arv, kui iga murde materjalimaht oleks 100 000 märgendatud sõna.

Tabel 4. *Läbi, mööda, vastu, üle ja ümber* normaliseeritud sagedused

	Post	Pre	Pre %
Setu	18	87	83
Ida	22	91	80
Kesk	38	101	73
Tartu	37	93	72
Ranna	44	111	72
Võru	36	65	64
Kirde	40	59	60
Lääne	69	97	59
Saarte	46	56	55
Mulgi	65	65	50

Tabelis 4 selgub, et prepositsioone esineb kõikides murdealades rohkem kui postpositsioone ning ida- ja Setu murre on teistest murretest suurema prepositsioonide osakaaluga (üle 80%). See võib olla seotud vene keele mõjuga, kuid nagu selgus tabelis 3 (vt lk 18), on andmestikus ida- ja Setu murdest kõige vähem kaassõnade esinemisi. Seetõttu võib üksik kaassõna, mis esineb rohkem prepositsioonina (näiteks *üle*), ja mida esineb nendes murdetekstides juhuslikult rohkem, mõjutada normaliseeritud sageduste tulemusi. Joonis 2 illustreerib tabelis 4 esitatud normaliseeritud pre- ja postpositsioonide esinemissageduste geograafilist jaotust.⁴

⁴ Visualiseerimiseks olen taas kasutanud tehnikaid, mille on esitanud Kyröläinen ja Uiboed (valmimas).



Joonis 2. Normaliseeritud sagedused geograafiliselt

2.2. Andmestiku kodeerimine

Kodeeritavad tunnused on valitud suuresti varasemate uurimuste põhjal, mis on esitanud kaassõnade paiknemist ja kasutust mõjutavaid tegureid. Analüüsi uuritavaks tunnuseks ehk sõltuvaks tunnuseks on niisiis adpositsioonide asendi varieerumine pre- ja postpositsioonina. Sõltuva tunnuse väärtuste varieerumist mõjutavaid asjaolusid püüan selgitada seletavate tunnuste abil ehk siinse andmestiku ülejäänud kodeeritud tunnuste abil. Andmestikus on kokku 9 tunnust ja analüüsitud on 1062 lausungit.⁵

Järgnevalt annan ülevaate kõikidest kodeeritud tunnustest, nende tasemetest ning kodeerimise põhimõtetest.

⁵ Kodeerisin lisaks veel viit tunnust, mida analüüsis ei kasutanud. Nende tunnuste tasemetes kordumiste arv andmestikus oli väike ja seetõttu kvantitatiivseks analüüsiks ebapiisav. Näiteks kodeerisin kirjakeelestatud kujul *ma*-infinitivis verbivormi, eristasin verbi paradigmaatilisi vorme, esitasin ambipositsiooni komplemendi nimetavas käändes kirjakeelestatud kujul, kodeerisin komplemendi pikkust märkides ning liigitasin kaassõnade tähendusfunktsioone EKG põhjal.

Murre (MURRE). Tähistab murdeala, kust analüüsitav lause pärit on. Tasemeteks on Eesti murrete korpuses esitatud 10 murdeala: idamurre (*ida*), keskmurre (*kes*), kirdemurre (*kir*), läänemurre (*laa*), Mulgi murre (*mul*), rannamurre (*ran*), saarte murre (*saa*), Setu murre (*set*), Tartu murre (*tar*) ja Võru murre (*vor*). Tunnuse kodeerimise eesmärk on tuvastada võimalikku geograafilist varieerumist, kas ja kuidas uuritavate kaassõnade kasutus pre- või postpositsioonidena võib murdeti erineda.

Adpositsiooni asend (ADPOSITSIOON). Binaarne tunnus, mis märgib, kas analüüsitavas lauses esineb kaassõna prepositsioonina (kodeering *Pre*) (1) või postpositsioonina (kodeering *Post*) (2).

- (1) (IDA) `kar'juss `ikke läheb **mööda** teed⁶
- (2) (KES) *tema ei val'ind et nüüd lähäb tied mööda*

Selles töös on ADPOSITSIOON sõltuv ehk uuritav tunnus, mille varieerumist uuritakse teiste kodeeritud tunnuste abil.

Vaadeldav ambipositsioon (AMBIPOS). Tähistab kaassõna, milleks on *läbi* (3), *mööda* (4), *vastu* (5), *üle* (6) ja *üumber* (7). Andmestikus on tähistused vastavalt *labi*, *mooda*, *vastu*, *yle* ja *yumber*. Peatükk 1.2.3 kirjeldas nende ambipositsioonide paiknemisvõimalusi täpsemalt.

- (3) (IDA) *aga sialt (.) **läbi** metsa onn (.) enne minna nattukke*
- (4) (MUL) *sõss `lasken `nurme = **müüdü** (.) obesege*
- (5) (RAN) *siis *laine *viskas minut kolinal ***vasta** *paatt`i*
- (6) (TAR) *ma `üt'li = t (.) ärge **üle** aija `sisse tulgõ*
- (7) (SAA) *ja sis (õnn)= et et üks (.) kahe+ `gümne inimese **üumber***

⁶ Näited on esitatud lihtsutatud transkriptsioonis (vt Lindström 2009: 7–9).

Komplemendi arv (KOMPL_ARV). Tegemist on binaarse tunnusega, mis tähistab, kas kaassõnafraasi komplement on ainsuses või mitmuses. Niisiis tähistasin tunnuse kahte taset vastavalt *sg* ehk singular (8) ja *pl* ehk pluural (9).

- (8) (RAN) *paistab tore suur *tiilis+kivi+maja siit säält üle **jue** säält*
- (9) (LAA) *`võetti seält siis ja `pandi `sinna **käär+buude** `ümmer*

Grammatilist arvu on peetud pluurali markeerituse tõttu oluliseks grammatilise konstruktsiooni varieerumise mõjutajaks (Klavan 2012: 75–76). Ka peatükis 1.2.2 tõin välja, et mida keerukam on komplemendi struktuur, seda tavalisem peaks olema prepositsiooni kasutus (Huumo 2013: 320). Selle alusel olen ka siinses analüüsis seda tunnust kodeerinud.

Komplemendi kääne (KOMPL_KAANE). Kaassõna paiknemist lauses võib mõjutada ka kaassõna juurde kuuluva fraasi kääne, mida olen täpsemalt kirjeldanud peatükis 1.2.2. Siinses töös uuritud kaassõnade juurde kuuluvad noomenid esinevad ainult nominatiivis (10), genitiivis (11) või partitiivis (12). Kodeeritud tasemed on vastavalt *nom*, *gen* või *par*.

- (10) (SAA) *talvõd läbi tass'si `einu puid*
- (11) (VOR) *ja ol'liq = `pandnu siss üle **niidü** minemä*
- (12) (KIR) *mei *läksima **jääd** *möödä*

Adpositsiooni asend verbi suhtes (SONAJARG). Tunnus tähistab kaassõnafraasi paiknemist öeldisverbi suhtes. Kui kaassõnafraas järgneb verbile, on kodeeringuks *parem* (13) ja kui eelneb verbile, on kodeeringuks *vasak* (14). Verbi puudumist tähistab *ellips*.

- (13) (VOR) *jaq = tuu lammass olnu säält `vasta `perve maahh*
- (14) (KIR) *kesk raut tied = **müöd** mina *kõnnin*

Grammatilise konstruktsiooni varieerumist käsitlevates uurimustes on oletatud, et keerukamad konstruktsioonid paiknevad lause lõpus (nt Klavan 2012: 80–84). Võib seega eeldada, et selline kaassõnafaasi konstruktsioon, mis on murdele loomulikum, eelneb verbile, ning verbile järgneva kaassõnafaasiga rõhutatakse edasi antavat teavet mitte nii tavapärase konstruktsiooniga.

Verbi aeg (AEG). Tunnus märgib verbi aega, eristades olevikku (*olevik*) (15) ja mis tahes minevikuvormi (*minevik*) (16). Verbi puudumist tähistab *ellips*.

(15) (SET) *kattõgese pois'iq lääväq läbi mõtsa*

(16) (SAA) *see kahe nädali `ümber läks `ikka levaga ära*

Peatükis 1.2.1 selgus, et pre- ja postpositsioonid on kujunenud eri moel ning et tekkelt on prepositsioonid verbiga kaudselt seotud. Niisiis on võimalik, et verbil on mõju ka ambipositsioonide paiknemisele ning seega võib minevikuvormi markeeritus mõjutada kaassõnafaasi sõnajärge sarnaselt nagu pluurali markeerituski.

Komplemendi elusus (ELUSUS). Eesmärgiks on liigitada kaassõna juurde kuuluv noomenifraas vastavalt oma referendile elusaks või elutuks. Kodeering põhineb Garretsoni (2004) kolmeastmelisel elususe jaotusel:

ÜLEMINE (<i>TOP</i>)	KESKMINE (<i>MIDDLE</i>)	ALUMINE (<i>BOTTOM</i>)
<i>inimene</i> (<i>human</i>)	> <i>elusolend</i> (<i>animal</i>)	> <i>konkreetne elutu</i> (<i>concrete inanimate</i>)
	> <i>organisatsioon</i> (<i>organization</i>)	> <i>mittekonkreetne elutu</i> (<i>non-concrete inanimate</i>)
		> <i>koht</i> (<i>place</i>)
		> <i>aeg</i> (<i>time</i>)

Oma andmestikus kasutan Garretsoni (2004) esitatud hierarhiat lihtsustatult. Ühendasin kategooriad *inimene*, *elusolend* ja *organisatsioon* ning tähistasin neid

kodeeringuga *elus* (17). Hierarhia viimase astme kategooriaid, mis sisult väljendavad kõik elutust, vaatlesin kooslusena *elutu* (18). Niisiis koosneb tunnus ELUSUS kahest tasemest: *elus* ja *elutu*.

(17) (LAA) *pärast saadetti meit liinile `jälle `sakslase vastu*

(18) (RAN) *selle Aksi = pialt tuli ära läks üle *Prangli juo*

Elututele objektidele viitavad noomenid kuuluvad taseme *elutu* alla. Taseme *elus* alla kuuluvad kõik elusolenditele viitavad noomenid nagu pärisnimed (*Anna*), üldnimed (*kolmkümmend lehma*) ja asesõnad (*nemad*). Lisaks kuuluvad taseme *elus* alla ka referendid, mis tähistavad organisatsioone, institutsioone ja muid kollektiivseid entiteete. Organisatsioonilisus põhineb kriteeriumil, et referendil on oma rühmaidentiteet ja et see koosneb inimestest (Garretson 2004).

Elusust kodeerides on oluline meeles pidada, et vaadeldakse mitte konkreetset noomenit, vaid entiteeti, millele sõnakuju viitab (Garretson 2004).

Elususe tunnust on kasutanud näiteks Jane Klavan (2012) oma doktoriväitekirjas, mis käsitleb eesti keele alalütleva käände ja kaassõna *peal* paralleelset kasutust kohasuhete väljendamisel kirjakeeles. Klavan (2012: 74–75) on ka pikemalt kirjeldanud, kuidas elusust on peetud oluliseks semantiliseks tunnuseks teisteski grammatilist sünonüümiat käsitlevates uurimustes. Kuigi käesolev töö ei tegele lähisünonüümiat uurimisega, olen siinses töös kaasanud elususe kui kaassõnade kasutust potentsiaalselt mõjutava teguri. Elusus võiks mõjutada kaassõna paiknemist näiteks analoogia alusel: kui mõne teise adpositsiooniga eelistatakse elusa referendiga rohkem pre- või postpositsioonilist kasutust, siis selle analoogial eelistatakse sama referendiga, aga teise adpositsiooniga, samuti rohkem vastavalt pre- või postpositsioonide kasutust.

Adpositsiooni väljendatav suhe (ADP_SUHE). Töös uuritavatel kaassõnadel on vastavalt oma asendile lauses sageli erinevaid tähendused. Seega on kaassõna tähendus oluline mõjutaja kaassõna paiknemisel pre- või postpositsioonina. Peatükid 1.2.2 ja 1.2.3 kirjeldasid täpsemalt, kuidas eri tähendusfunktsioonid mõjutavad ambipositsiooni paiknemist.

Andmestiku kodeerimisel eristasin järgnevad suhted: *abstraktne* (19), *aeg* (20), *mõõde* (21) ja *ruum* (22). Tasemed on vastavalt: *abs*, *aeg*, *moode* ja *ruum*.

(19) (LAA) *aga`raadjo läbi alatti räägib*

(20) (KES) *neli+gümmend meest tegivad talve läbi metsa+tüed*

(21) (SET) *üle kolmõ+saa`ektur'i tr-`saijõ siss*

(22) (MUL) *üits(.)`seantte imelik läits müödä mäge*

Kodeering põhineb Ehala (1994: 184) uurimuses kasutatud adpositsioonide semantilisel liigitusel, kus kaassõnad on jagatud samuti tähenduselt ruumiliseks, ajaliseks, mõõtmeliseks ja abstraktseks. Kategooriad kattuvad EKG-s (1995: 34–36) esitatud kaassõnade semantilise jaotusega, kuid ei ole sama üksikasjalikud. Nõnda vastavad tähenduskategooriale *abstraktne* EKG-s *põhjускаassõnad*, *viisi- ja seisundikaassõnad*, *vahendit või abinõu väljendavad kaassõnad*, *suhtumist või suhet väljendavad kaassõnad* ning *asendust väljendavad kaassõnad*. Kategooria *aeg* kattub *ajakaassõnade* kategooriaga ning kategooria *mõõt* langeb kokku kategooriatega *kvantiteeti (hulka, mõõtu, määra) väljendavad kaassõnad* ja *ligikaudsust väljendavad kaassõnad*. Kaassõnade tähenduskategooria *ruum* vaste on EKG-s *kohakaassõnad*.

Siinses andmestikus polnud põhjendatud eristada kaassõnade tähendusi sama täpselt nagu EKG-s, sest paljud EKG esitatud tähendused esineksid andmestikus vaid üksikjuhtudel. Näiteks oleks andmestiku 1062 kaassõnast asendust väljendavaid kaassõnu vaid 6. Nii võrd haruldase esinemise põhjal pole võimalik selle funktsiooni kohta aga järeldusi teha. Samuti kuna andmestik on väga varieeruv, on siin kaassõnade tähendusi EKG-s esitatud detailsusastmega raske (kui mitte võimatu) määrata.

3. Meetodid

Järgmistes alapeatükkides kirjeldan bakalaureusetöös rakendatud uurimismeetodeid ning tutvustan töö materjali analüüsis kasutatud klassifitseerimispuud.

3.1. Uurimismeetodi iseloomustus

Käesolev bakalaureusetöö järgib korpuspõhise dialektoloogia põhimõtteid, mida olen täpsemalt tutvustanud peatükis 1.1. Uurimismaterjal on pärit murdekorpusest, mis kajastab loomulikku spontaanset keelekasutust. Empiirilist materjali analüüsin klassifitseerimismeetodeid rakendades ning oma uurimistulemuste visualiseerimiseks olen kasutanud olemasolevaid tehnilisi võimalusi.

Empiirilise materjali analüüsi olen teinud statistikaprogrammiga R (R Core Team 2013), rakendades klassifitseerimispuud (*conditional inference tree / classification tree*). Analüüs on läbi viidud R-i paketiga *party* (Hothorn jt 2006). Andmestikus olen kodeerinud erinevaid tunnuseid ja seega on töö analüüs mitmefaktoriline, mis tähendab, et arvestatud on erinevate keelekasutust potentsiaalselt mõjutavate teguritega (Glynn 2010: 17–18).

3.2. Klassifitseerimispuu

Valikut kahe võimaliku keelevormi vahel võivad mõjutada mitmed tunnused (nii koos kui eraldi). Selleks, et tuua selgust, mis tunnused on ühe või teise valiku tegemisel olulised ja mil määral, on keelematerjali analüüsides vajalik kasutada statistilisi mudeleid. (Janda jt 2012: 1) Klassifitseerimispuu esitabki, kuidas erinevad seletavad tunnused koos või eraldi mõjutavad uuritava ehk sõltuva tunnuse esinemist (Tagliamonte, Baayen 2012: 135).

Klassifitseerimispuu põhineb korduval ositamisel (*recursive partitioning*), kus sarnaste sõltuva tunnuse väärtustega nähtused jagatakse rühmadeks ehk sarnased väärtused hargnevad nn sõlmkohtadest (*node*). Siinses töös on sõltuvaks tunnuseks

ADPOSITSIION ehk pre- ja postpositsioonide varieerumine ning klassifitseerimispuu hargnemiskohad näitavad, milline seletav tunnus ennustab järgmiseks pre- ja postpositsioonide paiknemist andmestiku põhjal kõige paremini. Selle tunnuse põhjal jaguneb andmestik igas sõlmkohas kaheks omavahel kõige selgemini eristuvaks rühmaks. Sellel põhimõttel ositatakse kõik andmestikus esinevad tunnused, kuni edasine jagamine pole põhjendatud ja/või võimalik (kas statistilise ebaolulise tõttu või andmepunktide lõppemisel). Igal sõlmkoha hargnemisel ositatakse seletavat tunnust, millel on järgmiseks tugevaim seos sõltuva tunnusega. Iga sõlmkoha juures on märgitud p -väärtus, mis on sõltuva ja seletava tunnuse seose tugevuse üheks indikaatoriks. Mida väiksem on p -väärtus, seda tugevam on seos tunnuste vahel. Andmestiku ositamine peatub, kui tunnustevahelisi seoseid ei leidu või kui hargnemiseks pole piisavalt materjali (minimaalne maht on 7 andmepunkti). Ühte ja sama tunnust võib klassifitseerimispuus ositada mitu korda. Statistiliselt ebaolulised tunnused eemaldatakse ja analüüsipuu need ei kajastu. Sel viisil jaguneb kogu andmestik väikesteks rühmadeks, mis on väärtuste poolest suhteliselt homogeenid. Viimase hargnemise juures on kuvatud tulemi väärtuste jagunemine ehk kuvatakse sõltuva tunnuse kahe taseme andmepunktide jagunemise proportsioonid (ehk selles analüüsis, mis osakaal on prepositsioonidel ja postpositsioonidel). Tähistega $n=$ on märgitud, mitu andmepunkti tulemi sisaldab. (Janda jt 2012; Tagliamonte, Baayen 2012; Strobl jt 2009)

Kokkuvõtvalt annab klassifitseerimispuu ülevaate, kuidas andmed andmestikus esinevad ning võimaldab selgitada, mis tunnused on uuritava nähtuse esitamisel statistiliselt olulised (Janda jt 2012: 20).

4. Analüüs: läbi, mööda, vastu, üle ja ümber varieerumisest murretes

Käesolevas peatükis vaatlen joonisel 3 esitatud (vt lk 35) klassifitseerimispuu abil, kuidas ja kas kaassõnade *läbi, mööda, vastu, üle ja ümber* kasutus pre- ja postpositsioonidena varieerub eesti murretes. Analüüsis olen kasutanud peatükis 2 tutvustatud andmestikku, mis sisaldas 9 tunnust ja 1062 ambipositsioone *läbi, mööda, vastu, üle ja ümber* sisaldavat lausungit. Klassifitseerimispuu hargnemisi seletan nende sõlmkohtade hargnemise järjekorras ning seda järjekorda kuvavad iga hargnemissõlme ning tulemi kohal olevad numbrid (näiteks tunnus AMBIPOS on number 1, sellest järgmine ADP_SUHE number 2 ja seejärel on märgistusega *node* 3 esitatud nende tunnuste tulem⁷). Seletamisel kasutan tunnuse tasemete täispikki nimetusi ja iga tulemi kohta toon ühe andmestikus esinenud murdenäite. Näideteks olen valinud lausungid tulemi väärtuste järgi – kui kaassõna esinemissagedus on prepositsioonina suurem, olen toonud näiteks prepositsioonilise ülesehitusega lausungi.

Klassifitseerimispuu analüüsis on uuritavaks tunnuseks adpositsiooni paiknemine kaassõnafraasis pre- või postpositsioonina (ADPOSITSIION). Analüüsis olen kasutanud 8 potentsiaalselt paiknemist mõjutavat tunnust: murre (MURRE); ambipositsioonid *läbi, mööda, vastu, üle ja ümber* (AMBIPOS); komplemendi arv (KOMPL_ARV); komplemendi kääne (KOMPL_KAANE); adpositsiooni asend verbi suhtes (SONAJARG); verbi aeg (AEG); komplemendi elusus (ELUSUS) ja adpositsiooni tähendusfunktsioon (ADP_SUHE). Tunnuste kodeerimise põhimõtteid olen seletanud täpsemalt peatükis 2.2. Klassifitseerimispuu korduva ositamise põhimõtteid olen esitanud peatükis 3.2.

4.1. Pre- ja postpositsiooniline varieerumine klassifitseerimispuu põhjal

Klassifitseerimispuu (joonis 3, lk 35) esimesel ositamisel osutub, et sõltuva tunnusega on tugevaim seos vaadeldavatel ambipositsioonidel (AMBIPOS) ning esimesel hargnemisel eraldub uuritavatest ambipositsioonidest kaassõna *üle*. Järgmisel ositamisel

⁷ Sõlmkohtade numbrid osutavad ainult hargnemise järjekorrale, mitte vastava hargnemiskoha olulisusele.

jaguneb kaassõna *üle* adpositsiooni tähendusfunktsiooni (ADP_SUHE) järgi kaheks: esimese rühmana eralduvad *aja-*, *mõõtme-* ja *ruumikaassõnad* ning teise rühmana *abstraktsust* väljendavad kaassõnad. Klassifitseerimispuu esimeses tulemis (*node* 3) selgub, et ambipositsioon *üle* esineb siinse töö andmestikus *abstraktsust* tähistades 70%-lise osakaaluga postpositsioonina (23).

(23) (MUL) *nõu pidäme ikki `jälle selle üle*

See tähendab, et *abstraktsust* väljendades on tavalisem, et kaassõna *üle* järgneb oma komplemendile. 326 esinemissagedusega *aja*, *mõõtme* ja *ruumi* tähendusfunktsioone tähistades esineb kaassõna *üle* andmestikus jällegi peaaegu alati prepositsioonina (*node* 4) (24).

(24) (SAA) *see kodus küll `ikke üle `kümne `küindra `päavas*

Niisiis ei tule siinse analüüsi andmestikus esile ambipositsiooni *üle* kasutusel pre- ja postpositsioonina geograafilist varieerumist. Peatükis 1.2.3 selgus, et ambipositsioon *üle* esinebki vähemalt kirjakeeles postpositsioonina harva, kuigi selle kaassõna genitiivne rektsioon (ülekaalukalt tüüpiline ka käesolevas andmestikus) peaks keele struktuuralseid tendentse järgides just nimelt suurendama postpositsioonilise sõnajärje eelistamist. Seda ometi siinse töö andmestiku põhjal väita ei saa.

Esimese hargnemise teine rühm sisaldab kaassõnu *läbi*, *mööda*, *vastu* ja *ümber*. See rühm jaguneb järgmisel ositusel tähendusfunktsioonide järgi kaheks: esiteks eraldub *abstraktne*, *aeg* ja *mööde* ning teisena eraldub *ruum*. Niisiis hargnes puu esimene ositus nii rühmade *üle* kui ka *läbi*, *mööda*, *vastu* ja *ümber* puhul esmalt tähendusfunktsioonide järgi. See pole üllatav, sest ka peatükis 1.2 selgus, et kaassõnade süntaktiline järjekord mõjutab kaassõna tähendust.

Järgmiseks (ositus nr 6) jaguneb klassifitseerimispuu taas vaadeldud ambipositsioonide kaupa (välja arvatud *üle*, mis eraldus teistest kaassõnadest juba esimesel ositusel ja *mööda*, mis esineb andmestikus eranditult ruumisuhteid tähistades). Kaassõnad *vastu* ja *ümber* käituvad sarnasemalt ning hargnevad lõplikeks tulemiteks

tähendusfunktsioonide järgi. *Abstraktsust* tähistades (*node* 8) esinevad need kaks kaassõna andmestikus ainult postpositsioonidena (25). Tähistades *aega* ja *möödet* (26) (*node* 9) on *vastu* ja *üumber* (30) samuti proportsionaalselt rohkem postpositsioonid (umbes 90%). Esinemissagedus on seejuures esimeses tulemis 57 ja teises 26 andmepunkti.

(25) (LAA) *rohi onn (...) ikke külmattuse `vastu*

(26) (IDA) *raha+balk tul'i koa ikke nii kahe+`kümne rubla `üumber*

Kaassõna *vastu* on ka töö teoreetilise osa peatükis 1.2.3 tutvustatud varasemate uurimuste põhjal postpositsioonina tähendusterohkem kui prepositsioonina. Lisaks esines *vastu* andmestikus *abstraktsuse* ja *aja* funktsiooni tähistades enamasti genitiivse komplemendiga ning tabelis 1 (vt lk 15) ka ilmnis, et genitiiviga esineb *vastu* vaid postpositsioonina. Seega ei ole *vastu* paiknemisel murretes *abstraktseid*, *ajalisi* ja *möötmelisi* suhteid tähistades kirjakeelest erinevusi, ning käesoleva töö andmestikus geograafilist varieerumist ei esine. Kaassõna *üumber* tähistab siinses jagunemises vaid *möötime* tähendusfunktsiooni ning selles funktsioonis esineb *üumber* ka kirjakeeles vaid postpositsioonina.

Klassifitseerimispuu kuuenda jagunemise teiseks rühmaks on kaassõna *läbi*, mis hargneb komplemendi käände (KOMPL_KAANE) järgi. Kui *läbi* komplement on genitiivis või partitiivis, on postpositsioonina esinemise osakaal 60% (*node* 11) (27). Nominatiivise komplemendiga *abstraktseid*, *ajalisi* ja *möötmelisi* suhteid väljendades esineb *läbi* 19 andmepunktiga ainult postpositsioonina (*node* 12) (28). Samuti ilmnis ka peatükis 1.2.3 esitatud tabelis 1, et nominatiivse komplemendiga *aega* tähistades esineb *läbi* postpositsioonina. Murretes ei näi siin olevat olulisi erinevusi kirjakeelest, mida kirjeldasin täpsemalt peatükis 1.2.3 varasemate uurimuste taustal. Samuti ei esine siinses analüüsis kasutatud andmestikus ka murretevahelisi erinevusi. Ometi saab erandina tuua välja kaassõna *läbi*, mis esineb prepositsioonina saarte murdes koos partitiivse komplemendiga (29). Partitiivset rektsiooni pole kaassõnaga *läbi* varasemates uurimustes täheldatud. Kuna seejuures esineb *läbi* alati koos nimisõnaga *öö*, võib oletada, et tegemist on kinnistunud väljendiga.

- (27) (KIR) *ga säääl *käisid (...) *päiva läbi tie tüöd*
 (28) (TAR) *nii kõik tal'v läbi egä päev üits'*
 (29) (SAA) *`öhtu viisime ää ja `öhtu tulime läbi üüd tagasi `jälle*

Viiendas hargnemiskohas (ADP_SUHE) eralduvad teistest tähendusfunktsioonidest *ruumilisi* suhteid väljendavad ambipositsioonid, kus *läbi, mööda, vastu* ja *ümber* paiknemisel kaassõnafaasis ilmnevad ka statistiliselt olulised geograafilised erinevused (ositus nr 13). *Ruumiliste* suhete eraldumine on ootuspärane, sest suurem osa siinses analüüsis kasutatud andmestiku kaassõnu väljendavadki ruumisuhteid. *Ruumi* väljendavate kaassõnade rohkus andmestikus võib olla tingitud murdeintervjuude iseloomust ja temaatikast. Ruumisuhete väljendamisel eralduvad esimesena ida-, kesk-, lääne- ja Setu murre ning teise rühma moodustavad kirde-, Mulgi, ranna-, saarte, Tartu ja Võru murded. Esmalt vaatlen esimest moodustunud murderühma.

Klassifitseerimispuus eraldunud ida-, kesk-, lääne- ja Setu murdes tulevad esile erinevused kaassõnade *läbi, mööda, vastu* ja *ümber* paiknemisel. Kaassõnad *läbi* ja *vastu*, mille komplemendi kääne võib varieeruda genitiivi ja partitiivi vahel, jagunevadki komplemendi käände järgi genitiivi ja partitiivi (KOMPL_KAANE). *Läbi* (30) esineb genitiivse komplemendiga ruumisuhteid tähistades ida-, kesk-, lääne-, ja Setu murdes enamjaolt prepositsioonina (osakaal umbes 95%) (*node* 17), kuigi teisi suhteid tähistades esines *läbi* rohkem postpositsioonina.

- (30) (SET) *kattõgese pois'iq lääväq läbi mõtsa*

Kaassõna *vastu* esineb nendes murretes genitiivse komplemendiga ruumisuhteid tähistades aga 90%-lise osakaaluga postpositsioonina (31) (*node* 18). Partitiivse komplemendiga esineb analüüsitud materjali põhjal *vastu* ida-, kesk-, lääne- ja Setu murdes ainult prepositsioonina (32) (*node* 19).

- (31) (LAA) *pane = see nägu nii sippelga+pesa `vastu*
 (32) (KES) *läks öhö poisil `vastu `körvu*

Need tulemused langevad taas kokku tabelis 1 esitatud andmetega ja partitiivse reksiooniga esineb kaassõna *vastu* prepositsioonina. Siinses andmestikus esinevad partitiivsed kasutused peamiselt eranditult *ruumi* suhteid tähistades. Seejuures peab märkima, et andmestikus ei esine *läbi* komplement *ruumi* tähistades ida-, keks-, lääne- ja Setu murdes kordagi partitiivis.

Ida-, kesk-, lääne- ja Setu murdes on kaassõnade *läbi* ja *vastu* paiknemist mõjutavad tegurid erinevad kui kaassõnade *mööda* ja *üumber* paiknemist mõjutavad tegurid. Viimaste puhul mõjutab asendi varieerumist pre- ja postpositsioonina adpositsiooni asend verbi suhtes (SONAJARG). *Ruumi* suhet tähistades püsib *mööda* ja *üumber* komplement nii pre- kui ka postpositsioonina samas käändes. Kui kaassõna paikneb verbist vasakul või puudub lausungist täielikult, esinevad kaassõnad *mööda* ja *üumber* nendes murretes natuke alla 60%-lise osakaaluga postpositsioonina (33) (*node* 21). Paiknedes verbist paremal, esinevad need kaks kaassõna sama tähendust tähistades ja samades murretes prepositsioonina umbes 70%-l esinemisjuhtudest (34) (*node* 22).

(33) (LAA) *aga siin maea`üumber onn meil ea maa*

(34) (IDA) *`kar`juss`ikke läheb mööda teed*

Klassifitseerimispuu tunnuse MURRE ositusel (ositus nr 13) eralduvad teise murderühmana kirde-, Mulgi, ranna-, saarte-, Tartu ja Võru murre. Esmalt hargneb rühm uuritavate ambipositsioonide järgi. Kaassõna *mööda* eristub teistest, ning esineb ligikaudu 20%-lise osakaaluga prepositsioonina (35) (*node* 24).

(35) (KIR) *mei *kierasimma siis jõe jääd *mööda*

Järgmisena hargnevad kaassõnad *läbi*, *vastu* ja *üumber* komplemendi käände järgi partitiivse komplemendiga ja genitiivse komplemendiga ambipositsioonide rühmadeks. Partisiivse komplemendiga on *vastu* paiknemise varieerumisel kaassõnafraasis murdelisi erinevusi (kaassõnad *läbi* ja *üumber* selles osituses partitiivse komplemendiga ei esine). Tartu murdes esineb kaassõna *vastu* partitiivse komplemendiga nii pre- kui ka

postpositsioonina (36 ja 37) suhteliselt ühtlaselt, kuid tuleb märkida, et kogu tulemis on andmepunktide arv kõigest 14 (*node 27*).

(36) (TAR) *jaa (...) läits `ümbre = ja (.) mina `vasta puud = ja `vasta vangerd*

(37) (TAR) *siss tynõ miiss lei siin (.) tynõ siin `penkki `vasta noid*

Kirde-, Mulgi, ranna-, saarte ja Võru murdes mõjutab nende ambipositsioonide paiknemist verbi aeg. Kui verb on olevikus või kui verb lausungist puudub, esinevad ambipositsioonid *läbi*, *vastu* ja *üumber* ruumisuhteid tähistades andmesikus vaid 8 korda ning prepositsioonide osakaal on ligikaudu 75% (38) (*node 29*).

(38) (SAA) *me lähme `vastu lainet*

Minevikuvormiga esineb kaassõna *vastu* nendes murretes 34 korda ning kaassõnafraasis paikneb ta ainult prepositsioonina (39) (*node 30*).

(39) (RAN) **purjet *läksit *vasta vett*

Selliseid tunnuse AEG järgi osituvaid ambipositsioone oli aga andmestikus vähe – kokku esines neid vaid 42 korral. Ka on tunnuse MURRE (ositus nr 26), mille kahes lõpptulemis tunnus AEG esineb, *p*-väärtus klassifitseerimispuu kõrgeim: $p = 0,035$. Ometi ei ületa see statistiliselt oluliseks peetava *p*-väärtuse ülempiiri, milleks on 0,05, mis tähistab, et tulemus on 5%-lise tõenäosusega juhuslik.

Lisaks eristub ambipositsioonide *läbi*, *vastu* ja *üumber* paiknemine kaassõnafraasis kirde-, Mulgi, ranna-, saarte, Tartu ja Võru murdes ruumilisust tähistades genitiivse komplemendi järgi. Kui komplement on genitiivis, esinevad kaassõnad *läbi* ja *üumber* 112-st esinemiskorrast prepositsioonidena 60%-lise osakaaluga (40) (*node 32*).

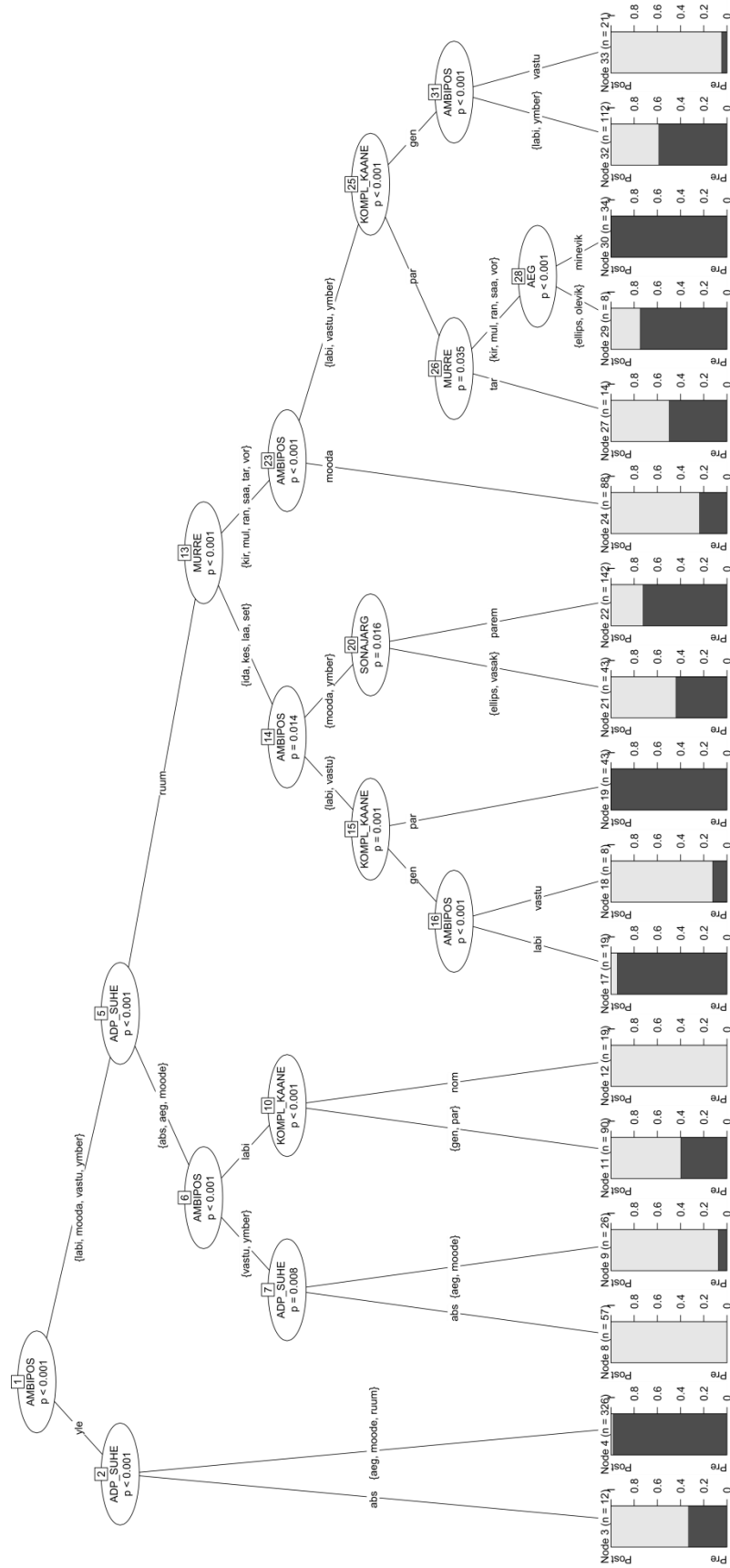
(40) (RAN) *paned siit *üumber *paela *solmed *pääle*

Kaassõna *vastu* aga esineb genitiivse komplemendiga ruumilisust tähistades kirde-, Mulgi, ranna-, saarte, Tartu ja Võru murdes enamasti postpositsioonina (41) (*node* 33).

(41) (VOR) *sõss tõne ol' sinnä küün'i `saina `vasta*

Klassifitseerimispuu kõikide jagunemiste *p*-väärtused on statistiliselt olulised ja klassifitseerimispuus statistiliselt mitteolulisi hargnemisi ei esitatagi. Vaadeldud kaheksast tunnusest osustusi komplemendi arv (KOMPL_ARV) ja komplemendi elusus (ELUSUS) analüüsis ebaolulisteks. Seega mõjutab klassifitseerimispuu järgi ambipositsioonide paiknemist murretes kuus andmestikus kodeeritud tunnust: murre (MURRE); ambipositsioonid *läbi*, *mööda*, *vastu*, *üle* ja *üumber* (AMBIPOS); komplemendi kääne (KOMPL_KAANE); adpositsiooni asend verbi suhtes (SONAJARG); verbi aeg (AEG) ja adpositsiooni tähendusfunktsioon (ADP_SUHE).

Kogu andmestikust moodustasid prepositsioonid 64% ja postpositsioonid 36%. Kui uuritavate kaassõnade paiknemise üle otsustatakse seega lihtsalt andmestiku proportsioonide põhjal ning valitakse iga kord suurema osakaaluga paiknemine (andmestikus prepositsioonina), tehtaks õige valik 64%-l juhtudest. Kuna aga klassifitseerimispuu täpsus on 83%, selgub, et kodeeritud tunnused lisavad ambipositsioonide paiknemise ennustamisele informatsiooni tervelt 19% (83% - 64%). Niisiis on meetodi rakendamine õigustatud ja edukas.



Joonis 3. Klassifitseerimispuu pre- ja postpositioonide varieerumisest

4.2. Tulemuste kokkuvõte

Klassifitseerimispuu analüüsi põhjal osutusid andmestikus olulisimaks pre- ja postpositsioonilise paiknemise mõjutajateks töös vaadeldavad ambipositsioonid. Jagunemine toimus niisiis põhiliselt kaassõnade *läbi*, *mööda*, *vastu*, *üle* ja *ümber* järgi, mis tähendab, et vaadeldud kaassõnade positsiooni varieerumist kaassõnafaasis mõjutab eelkõige konkreetne kaassõna ise. Selle analüüsi põhjal on kaassõnade paiknemise üheks olulisemaks mõjutajaks veel adpositsiooni väljendatav suhe, selle tähendusfunktsioon. Kaassõnade *läbi*, *mööda*, *vastu* ja *ümber* puhul vastanduvad *abstraktse*, *aja* ja *mõõtme* tähendusega ning *ruumi* tähendusega kaassõnad. Ambipositsioon *üle* eristub ülejäänud kaassõnadest ning selle kaassõna kasutuses vastanduvad väljendatav *abstraktne* suhe ning *aja*, *mõõtme* ja *ruumi* suhted. Tähendusfunktsiooni olulisus kaassõna paiknemisel ei ole üllatav, sest nagu selgus peatükis 1.2.3, on vaadeldud kaassõnadel tihti vastavalt paiknemisele eri tähendusi.

Analüüsis ei selgu, et ambipositsioonide paiknemisel oleks märkimisväärselt olulisi murretevahelisi erinevusi. Klassifitseerimispuus eristuvad küll *ruumi* suhet väljendavate ambipositsioonide *läbi*, *mööda*, *vastu* ja *ümber* puhul kaks murderühma: esimese rühmana ida-, kesk-, lääne- ja Setu murre ning teise rühmana kirde-, Mulgi, ranna-, saarte, Tartu ja Võru murre. *Ruumi* tähendusfunktsiooni ositus ei kajasta siiski pre- ja postpositsioonilist paiknemist ainult murrete põhjal – põhiliselt kajastab *ruumi* tähendusfunktsiooni ositus iga kaassõna pre- ja postpositsioonilise paiknemise varieerumist eraldi. Huvitav on, et olulisi erinevusi ei teki traditsiooniliste põhja- ja lõunaestli murderühmade vahel, pigem on murretevahelised erinevused ambipositsioonide kasutuses tagaplaanil ja mõningane murretevaheline varieerumine ei ole selgelt seotud kindlalte murderühmadega. Selle põhjal võib järeldada, et ambipositsioonide paiknemine ja kasutus on pigem seotud keelesiseste teguritega ja mitte niivõrd geograafilistega.

Kuna analüüsis osutus olulisimaks kaassõnade pre- või postpositsioonidena esinemist seletavaks tunnuseks ambipositsioon (AMBIPOS), vaatlen selle tunnuse tasemete kaupa, mis on klassifitseerimispuu järgi iga kaassõna sagedasem positsioon kaassõnafaasis ning mis tegurid paiknemist mõjutavad.

Tähistades *abstraktset*, *ajalist* ja *mõõtme* suhet, ei tule kaassõna *läbi* kasutuses esile statistiliselt olulisi murretevahelisi erinevusi. Neid tähendusi väljendades esineb *läbi* genitiivse komplemendiga rohkem ning nominatiivse komplemendiga ainult postpositsioonina. *Ruumi* suhteid tähistades esineb *läbi* ida-, kesk-, lääne ja Setu murdes genitiivse rektsiooniga ning on peamiselt prepositsioon. Kirde, Mulgi, ranna-, saarte, Tartu ja Võru murdes laiendab kaassõna *läbi* samuti genitiivne komplement ning ka neil juhtudel esineb see kaassõna pigem prepositsioonina.

Kaassõna *mööda* väljendab selles andmestikus vaid *ruumi* suhteid. Kusjuures kui *mööda* paikneb öeldisverbist vasakul (või kui öeldisverb lausungist puudub), on see kaassõna ida-, kesk- ja läänemurdes rohkem postpositsiooniline. Öeldisverbist paremal paiknedes on *mööda* nendes murretes rohkem prepositsiooniline. Andmestikus ei esinenud seda kaassõna kordagi Setu murdes, mis oli ka aimatav, sest lõunaeesti murretes kasutatakse *mööda* asemel rohkem postpositsiooni *pidi*. Kirde, Mulgi, ranna-, saarte ja Võru murdes on *mööda* enamasti postpositsioon.

Vastu esineb tähendusfunktsioonis *abstraktne*, *aeg* ja *mõõde* põhiliselt postpositsioonina. *Ruumi* suhteid tähistades on kaassõna *vastu* ida-, kesk-, lääne- ja Setu murdes kas genitiivse või partitiivse komplemendiga. Genitiivse rektsiooniga esineb *vastu* harva ja on seejuures sagedamalt postpositsioon. Partitiivse rektsiooniga esineb *vastu* rohkem prepositsioonina. Kirde-, Mulgi, ranna-, saarte Tartu ja Võru murdes on partitiivse komplemendiga *vastu* samuti partitiivse komplemendiga rohkem prepositsioon ning genitiivse komplemendiga on *vastu* samades murretes ainult postpositsioon.

Kaassõna *üle* paiknemisel pre- ja postpositsioonina ei esine selle töö andmestikus olulisi murretevahelisi erinevusi. *Abstraktseid* suhteid tähistades esineb *üle* andmestikus harva, vaid 12 korda, ning seejuures esineb *üle* rohkem postpositsioonina. Teisi suhteid tähistades on kaassõna *üle* pea alati prepositsioon. Seejuures võib tuua välja, et selle kaassõna puhul ei eraldu teistest mitte *ruumi* väljendavad suhted nagu teistel kaassõnadel, vaid *abstraktsed* suhted.

Ümber tähistab klassifitseerimispuu järgi nii *abstraktseid*, *aja* kui ka *mõõtme* suhteid, esinedes enamasti postpositsioonina. Väljendades *ruumi* suhteid, on *ümber* nii ida-,

kesk, lääne- ja Setu murdes kui ka kirde-, Mulgi, ranna-, saarte, Tartu ja Võru murdes enamjaolt prepositsioon. *Ümber* esineb peaaegu alati genitiivse komplemendiga.

Andmestikus ilmnenud geograafilisel ühtlusel võib olla erinevaid põhjuseid. Mõningaid neist kirjeldasin töö teoreetilises osas. Näiteks selgus peatükis 1.2, et kaassõnade ülesanne on väljendada lauseliikmete vaheliseid suhteid ning süntaktiline järjekord mõjutab kaassõna tähendust. See leidis kinnitust ka siinse analüüsi põhjal. Lisaks on kaassõna positsioon seotud kaassõnafraasi morfosüntaktilise struktuuriga: soomeugrilise genitiivse konstruktsiooniga on loomulikum postpositsioonide kasutus ning partitiivse rektsiooniga esinevad rohkem prepositsioonid. See kehtib aga rohkem nende adpositsioonide puhul (mitte küll eranditeta), mis ei või paikneda emmas-kummas positsioonis. Ambipositsioonid seda reeglit sama järjekindlalt ei järgi, millele osutas ka kvantitatiivne analüüs.

Kokkuvõte

Käesolevas bakalaureusetöös uurisin ambipositsioonide *läbi, mööda, vastu, üle* ja *ümber* paiknemise varieerumist pre- ja postpositsioonidena eesti murretes. Töö eesmärk oli selgitada, kas nende ambipositsioonide paiknemise puhul esineb geograafilist varieerumist, ning analüüsida, mis keelesisesed tegurid võivad kaassõna kasutust pre- või postpositsioonina mõjutada. Hüpoteesiks oli, et kaassõna paiknemine on murdeti erinev ning sõltub kodeeritud keelesisestest nähtustest.

Hüpoteeside kontrollimiseks viisin läbi korpuspõhise dialektoloogia põhimõtetele tuginedes kvantitatiivse analüüsi (ptk 1.1 ja 3.1). Uurimismaterjal pärines eesti murrete korpuse morfoloogiliselt märgendatud murdetekstidest ning andmestik koosnes 1062-st ambipositsioone *läbi, mööda, vastu, üle* ja *ümber* sisaldavast lausungist. Kodeerisin kaheksa kaassõnade kasutust potentsiaalselt mõjutavat tunnust, mille valisin välja varasemate uurimuste põhjal.

Andmete analüüsil rakendasin klassifitseerimispuud, mille abil tuvastasin, kuidas andmed andmestikus jagunevad ja mis kodeeritud kaheksast tunnustest mõjutavad uuritavat pre- ja postpositsioonide vaheldumist kõige rohkem. Klassifitseerimispuu põhjal osutusid kodeeritud kaheksast tunnusest analüüsis statistiliselt oluliseks kuus: murre (MURRE); ambipositsioonid *läbi, mööda, vastu, üle* ja *ümber* (AMBIPOS); komplemendi kääne (KOMPL_KAANE); adpositsiooni asend verbi suhtes (SONAJARG); verbi aeg (AEG) ja adpositsiooni tähendusfunktsioon (ADP_SUHE). Komplemendi arv (KOMPL_ARV) ja elusus (ELUSUS) ei osunud kvantitatiivses analüüsis statistiliselt olulisteks teguriteks.

Kõige enam kaassõnade paiknemist mõjutavateks tunnusteks osutusid vaadeldav ambipositsioon (ehk kaassõnad *läbi, mööda, vastu, üle* ja *ümber*) ning ambipositsiooni väljendatav tähendus (*abstraktne, aeg, mõõde* ja *ruum*). Tulemustes ei selgunud, et ambipositsioonide paiknemisel oleks märkimisväärset geograafilist varieerumist. Kaassõnad *läbi, mööda, vastu* ja *ümber* moodustasid vaid *ruumi* suhteid tähistades kaks murderühma (ida-, kesk-, lääne-, ja Setu; kirde-, Mulgi, ranna-, saarte, Tartu ja Võru).

Klassifitseerimispuu lõpptulemid kajastasid pigem iga vaadeldud kaassõna pre- ja postpositsioonilise paiknemise varieerumist eraldi. Sellest võib järeldada, et adpositsioonide paiknemine kaassõnafraasis pole niivõrd seotud geograafiliste teguritega kuivõrd keelesisestega.

Keele ambipositsioonilisusele ei ole ilmset seletust ning teemat on eesti keeles siiani vähe uuritud. Enamjaolt on varasemad uurimused kirjakeelt analüüsivad ning keskenduvad vaid üksikutele pre- või postpositsioonidena esinevatele kaassõnadele. Samuti on siiani veel vähe korpuspõhiseid kvantitatiivseid kaassõnateemalisi uurimusi. Käesolev bakalaureusetöö täidab eelkõige seda lünka ning pakub empiirilistele meetoditele tuginedes uudset teavet adpositsioonide kasutuse kohta eesti murretes. Põhjalikumate järelduste tegemiseks peaks bakalaureusetöös käsitletud ambipositsioonide paiknemise varieerumist pre- ja postpositsioonidena kindlasti üksikasjalikumalt edasi uurima, suurendades korpuse mahtu ja kaasates analüüsi veel tunnuseid, mis kaassõnade positsiooni varieerumist mõjutada võiksid.

Kirjandus

Eesti keele seletav sõnaraamat. <http://www.eki.ee/dict/ekss/>. Vaadatud 10.04.2014

EMK = Eesti murrete korpus. <http://www.murre.ut.ee/murdekorpus/>. Vaadatud 24.03.2014

Ehala, Martin 1994. Russian Influence and the Change in Progress in the Estonian Adpositional System. – *Linguistica Uralica*, 3, 177–199.

Ehala, Martin 1995. Explaining the Bipositional Head-Complement Order in Adpositional Systems. In: *Proceedings of LP'94: Item Order in Natural Languages*, Prague, August 16-18, 1994, *Proceedings of the Conference*. Ed. by Bohumil Palek, František Dane, Osamu Fujimura, Jiri V. Neustupný. Prague: Charles University Press, 325–339.

EKK = *Mati Ereht, Tiiu Ereht ja Kristiina Ross* 2007. Eesti keele käsiraamat. Kolmas, täiendatud trükk. Tallinn: Eesti Keele Sihtasutus.

EKG I = *Mati Ereht, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael ja Silvi Vare* 1995. Eesti keele grammatika I. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.

Garretson, Gregory 2004. Coding Practices Used in the Project Optimal Typology of Determiner Phrases. Unpublished manuscript. Boston University, Boston, MA. <http://npcorpus.bu.edu/html/documentation/index.html>. Vaadatud 6.03.2014

Glynn, Dylan 2010. *Corpus-Driven Cognitive Semantics. Introduction to the field. Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*. New York: De Gruyter Mouton.

Grünthal, Riho 2003. *Finnic Adpositions and Cases in Change*. – *Suomalais-Ugrilaisen Seuran Toimituksia* 244. Helsinki: Société Finno-Ougrienne.

Grünthal, Riho 2005. Miksi itämerensuomessa on prepositioita? – *Virittäjä*, 1, 28–51.

Habicht, Külli, Pille Penjam 2007. Kaassõna keeleuurija ja -kasutaja käsituses. *Emakeele Seltsi Aastaraamat* 52 (2006). Tallinn: Emakeele Seltsi aastaraamat, 51–56.

- Hagège, Claude* 2010. *Adpositions*. Oxford: Oxford University Press.
- Hint, Mati* 1990. Vene keele mõjud eesti keelele. – *Akadeemia*, 2, 1383–1404.
- Hothorn, Torsten, Kurt Hornik, Achim Zeileis* 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. – *Journal of Computational and Graphical Statistics*, vol. 15, No 3, 651–674.
- Huumo, Tuomas* 2013. Many Ways of Moving Along a Path: What Distinguishes Prepositional and Postpositional Uses of Finnish Path Adpositions?. – *Lingua*, 133, 319–335.
- Janda, Laura A., Lene Antonsen, Berit Anne Bals Baal* 2014. A Radial Category Profiling Analysis of North Sámi Ambipositions. – *High Desert Linguistics Society Proceedings*, vol. 10, 91–102.
- Janda, Laura A., Tore Nessel, Stephen Dickey, Anna Endresen, Anastasia Makarova, R. Harald Baayen* 2012. Making Choices in Slavic: Pros and Cons of Statistical Methods for Rival Forms. <http://www.sfs.uni-tuebingen.de/~hbaayen/publications/JandaEtAL2012.pdf>. Vaadatud 05.03.2014
- Jürine, Anni, Külli Habicht* 2013. Kaassõnade tsükliline areng. – *Keel ja Kirjandus*, 10, 736–750.
- Karlsson, Fred* 2002. *Üldkeeleteadus*. Tartu: Eesti Keele Sihtasutus.
- Klavan, Jane* 2012. *Evidence in Linguistics: Corpus-Linguistics and Experimental Methods for Studing Grammatical Synonymy*. *Dissertationes linguisticae Universitatis Tartunensis* 15. Tartu: University of Tartu Press.
- Kyröläinen, Aki-Juhani, Kristel Uibo* (valmimas). *Visualization Techniques for Spatial Linguistic Data Using ggplot2 and ggmap*.
- Lass, Roger* 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
- Lehmann, Christian* 1985. Grammaticalization: Synchronic Variation and Diachronic Change. – *Lingua E Stile*, 20, 303–318; http://www.christianlehmann.eu/publ/syn_dia.pdf. Vaadatud 23.04.2014
- Lindström, Liina* 2009. Ülevaade eesti murrete korpusest. Käsikiri. <http://www.murre.ut.ee/assets/EMK.PDF>. Vaadatud 20.05.2014

- Pajusalu, Karl, Tiit Hennoste, Ellen Niit, Peeter Päll ja Jüri Viikber* 2009. Eesti murded ja kohanimed. Tallinn: Eesti Keele Sihtasutus.
- Palmeos, Pauline* 1982. Eesti keele grammatika. II osa. Neljas vihik. Kaassõna. Teine Trükk. Tartu: Tartu Riiklik Ülikool.
- Ploom, Harald* 1955. Kaassõnad eesti murdeis. Diplomitöö: Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- R Core Team* 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Reimann, Maire* 1958. Kaassõnad kaasaegses eesti keeles. Diplomitöö: Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- Strobl, Carolin, James Malley, Gerhard Tutz* 2009. An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. – Technical Report Number 55.
- Szmrecsanyi, Benedikt* (ilmumas). Methods and Objectives in Contemporary Dialectology. – *Современные Методы В Диалектологии. Ареал Северных, Северо-Западных Русских И Белорусских Говоров*. Ed. Ilja A. Seržant & Björn Wiemer.
- Szmrecsanyi, Benedikt, Cristoph Wolk* 2011. Holistic Corpus-Based Dialectology. – *Revista Brasileira de Linguística Aplicada*, vol. 2, No 2, 561–592.
- Tagliamonte, Sali A., R. Harald Baayen* 2012. Models, Forests, and Trees of York English: Was/Were Variation as a Case Study for Statistical Practice. – *Language Variation and Change* 24. Cambridge: Cambridge University Press, 135–178.
- Tooding, Liina-Mai* 2007. Andmete analüüs ja tõlgendamine sotsiaalteadustes. Tartu: Tartu Ülikooli Kirjastus.
- Veismann, Ann* 2011. Kaassõnade uurimisest. – *Oma Keel*, 1, 5–11.

Pre- and postpositional variation of adpositions *läbi*, *mööda*, *vastu*, *üle* and *üumber* in Estonian dialects. Summary

In case of certain adpositions, the use of both prepositional and postpositional order in adpositional phrase is possible (Grünthal 2003; Hagège 2010). These adpositions that can appear in both positions are termed *ambipositions* (e.g. Hagège 2010; Janda et al. 2014). This thesis focuses on the variation of the ambipositions *läbi* ‘through’, *mööda* ‘along’, *vastu* ‘against’, *üle* ‘over’ and *üumber* ‘around’ in Estonian dialects. The primary aim of the study is to investigate whether regional variation occurs in pre- and postpositional usage and to discover which language-internal factors are associated with the ambipositional word order. Two main hypotheses are tested in this study: the positions of ambipositions vary in Estonian dialects; and the pre- and postpositional variation is correlated with factors coded in this study.

Following the principles of corpus-based dialectology (Szmrecsanyi (to appear); Szmrecsanyi & Wolk 2011), a quantitative analysis was carried out in order to test the hypotheses. The empirical data was collected from the morphologically tagged texts of the Estonian Dialect Corpus. The collected dataset consists of 1062 phrases that contain ambipositions *läbi*, *mööda*, *vastu*, *üle* and *üumber*. Although linguistic alternations are typically examined based on merely one variable, the current study analyses several variables simultaneously. Eight variables were chosen to be coded because previous research proposes that these factors potentially affect the pre- and postpositional usage of ambipositions.

The quantitative analysis was conducted by utilising recursive partitioning tree model (*conditional inference tree / classification tree*). Quantitative analysis revealed that six of the eight coded variables were statistically significant: dialect (MURRE); ambipositions *läbi*, *mööda*, *vastu*, *üle* and *üumber* (AMBIPOS); case of the complement (KOMPL_KAANE); position of the adposition in relation to the verb (SONAJARG); tempus (AEG); and ambiposition’s expression of meaning (ADP_SUHE). Thus, the pre- and

postpositional variation does correlate with multiple factors as was presupposed in the hypothesis.

The variables with the strongest influence on the pre- and postpositional order in adpositional phrase were the observed ambipositions (*läbi, mööda, vastu, üle* and *üumber*) and the expression of meaning (*abstract, temporal, measure* and *spatial*). In contrast, the contribution of regional variation to the classification of ambipositions was smaller than hypothesized based on the statistical analysis. Hence, this analysis does not support the hypothesis that differences in usage of pre- and postpositional adpositions is primarily driven by geographical variation. It must be noted, however, that expressing spatial relations, four of the ambipositions (*läbi, mööda, vastu* and *üumber*) divided into two larger dialect groups (Eastern, Mid, Western and Setu; and North-Eastern, Mulgi, Coastal, Insular, Tartu and Võru dialects).

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina MIRJAM RUUTMA

(*autori nimi*)

(sünnikuupäev: 10.11.1990)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
„Kaassõnade läbi, mööda, vastu, üle ja ümber pre- ja postpositsioonilisest
varieerumisest eesti murretes”,
(*lõputöö pealkiri*)

mille juhendajad on Kristel Uiboaed, Aki-Juhani Kyröläinen ja Maarja-Liisa Pilvik,
(*juhendaja nimi*)

- 1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 28.05.2014