

KRISTEL UIBOAED

Verbiühendid eesti murretes



DISSERTATIONES PHILOLOGIAE ESTONICAE
UNIVERSITATIS TARTUENSIS

34



KRISTEL UIBOAED

Verbiühendid eesti murretes

Tartu Ülikooli filosoofiateaduskond, eesti ja üldkeeleteaduse instituut

Väitekirja on filosoofiadoktori kraadi saamiseks kaitsmisele suunanud Tartu Ülikooli filosoofiateaduskonna eesti ja üldkeeleteaduse instituudi nõukogu 8. oktoobril 2013.

Juhendajad: dr Liina Lindström

dr Kadri Muischnek

Oponent: dr Maria Vilkuna (Kodumaa keelte keskus, Soome)

Kaitsmine toimub 16. detsembril 2013 kell 14.15 Tartu Ülikooli senati saalis.

Doktoritöö valmimist on toetanud Euroopa Liidu Sotsiaalfond.



Euroopa Liit
Euroopa Sotsiaalfond



Eesti tuleviku heaks

ISSN 1406–1325

ISBN 978–9949–32–453–8 (trükis)

ISBN 978–9949–32–454–5 (pdf)

Autoriõigus: Kristel Uiboaed, 2013

Tartu Ülikooli Kirjastus

ww.tyk.ee

EESSÕNA

Selle töö kirjutamise ajal on minu ümber olnud palju toetavaid inimesi, kes kõik on ühel või teisel viisil väitekirja valmimisse oma panuse andnud. Siinkohal ei jõua ma nimeliselt mainida kõiki, kes seda väärivad, aga olen erakordselt tänulik oma kolleegidele, õpetajatele, sugulastele ja sõpradele, kes on mind toetanud, innustanud ja kiiretel perioodidel mõistvalt eemale hoidnud.

Minu südamlük tänu kuulub minu juhendajatele Liina Lindströmile ja Kadri Muischnekile, kelle panuse suurst selle töö valmimisse ja minu akadeemilisse arengusse on raske sõnadesse panna. Olen tänulik tingimusteta toetuse eest, mida olen tundnud doktorantuuri esimestest päevadest kuni viimaste kiirete minutiteni välja. Hindan kõrgelt oma juhendajate koostööst tekkinud positiivset ja motiveerivat õhkkonda ning kõiki mulle loodud töötingimusi ja võimalusi. Olen tänulik mulle võimaldatud vabaduse eest teha ise oma valikud ja eksimused. Olen juhendajatele tänulik erakordse kättesaadavuse, kiirete vastuste, probleemide lahendamise, suunavate mõtete, konstruktiivse tagasiside, eelkõige aga kannatlikkuse ja vastutulelikkuse eest.

Avaldan tänu kõigile oma retsensentidele, kelle parandusettepanekud andsid olulise panuse töö lõpptulemusse. Olen tänulik kõigile kolmele konstruktiivse ja sisuka tagasiside eest. Aitäh Katrin Niglasele, kes juhtis tähelepanu statistilistele probleemidele ja terminoloogilistele ebakõladele. Olen väga tänulik Heete Sahkaile, kelle terav silm märkas väiksemaidki teoreetilisi vastuolusid. Suured tänud Maria Vilkunale, kes väga detailselt tõi välja materjali kirjeldamise kriitilised kohad. Kõik töös olevad vead on loomulikult minu enda omad.

Olen abi ja toetust saanud ka kolleegidelt ja õpetajatelt. Aitäh John Nerbonnele võimaluse eest veeta viis kuud Groningeni ülikoolis ja saada osa ühest maailma tugevamaast dialektomeetriaga tegeleva uurimiskeskuse tööst. Aitäh Aki-Juhani Kyröläisele ja Pärtel Lippusele tehnilise abi ja kaasamõtlemise eest. Suured tänud Ellen Niidile töö toimetamise, vastutulelikkuse ja otsekoheuse eest.

Olen väga tänulik kõigi võimaluste eest, mis mulle väitekirja kirjutamise perioodil on enesetäiendamiseks avanenud. Aitäh Euroopa Liidu Sotsiaalfondi programmile DoRa, Sihtasutusele Archimedes, keeleteaduse, filosoofia ja semiootika doktorikoolile, Eesti Üliõpilaste Toetusfondile USAs, Tartu Ülikooli Sihtasutusele ja Eesti Keeletehnoloogia Sihtasutusele. Uurimuse valmimist on toetanud ETF-i grant 7464, Eesti keeletehnoloogia programmid EKKTT06-14, EKKTT10-74 ja EKT7, Eesti keele ja kultuurimälu programmid EKKM09-111 ja EKKM10-202. Suured tänud Eesti Teadusagentuuri projektile PUT90 toetuse eest ning võimaluse eest arendada ja rakendada selle töö tulemusena valminut edaspidi.

Olen erakordselt tänulik kõigile oma lähedastele, kes on olnud alati valmis mind kuulama ja toetama ning aidanud kontaktis püsida ülikoolivälise maailmaga. Olen tänulik kaasaelamise eest kõigile oma uutele sõpradele, kellega väitekirja kirjutamise periood on mind kokku viinud, ja vanadele sõpradele selle eest, et nad ikka veel alles on ja minu eemalolekusse mõistvalt suhtusid. Aitäh,

Hanna-Kaisa, Julianna, Kersten, Maarja-Liisa, Sigrid, Siiri, Triin, Triinu ja väga paljud teised, et te mul olemas olete. Ilma teieta oleks see periood olnud palju igavam. Olen tänulik kõigile oma toetavatele sugulastele. Kõige erilisem tänu kuulub aga minu perekonnale: ema Tiiale, isa Aivarile ja vend Janekile. Aitäh kõige eest, mis te minu heaks teinud olete.

SISUKORD

LÜHENDID	15
1 SISSEJUHATUS	17
1.1 Murdesüntaks	18
1.2 Kobardialektoloogia	19
1.3 Uurimisobjekt: verbiühendid	20
1.4 Töö ülesehitus	20
2 TÖÖ EESMÄRGID JA HÜPOTEESID	22
3 TEOREETILINE TAUST	25
3.1 Konstruksioonigrammatika	25
3.1.1 Konstruksioonipõhise lähenemise rakendamine käesolevas töös	28
3.1.2 Konstruksiooni mõiste	29
3.2 Kasutuspõhine keeleuurimine	30
3.3 Sagedus	32
3.4 Süntaktiline produktiivsus	35
3.5 (In)finiitsus	36
3.6 Eesti keele infiniitsed vormid	37
3.6.1 Infinitiivid	38
3.6.2 Partitsiibid	39
3.6.3 Konverbid	40
3.7 Abiverbistumisest ja abiverbikonstruktsioonidest	40
3.8 Dialektomeetria	41
3.9 Korpuspõhine dialektoloogia	43
3.10 Korpuspõhised dialektoloogilised uurimused	44
3.11 Eesti murdesüntaks	45
3.12 Varasemad verbiühendite uurimused	46
4 ANDMESTIK: EESTI MURRETE KORPUS	48
4.1 Murdekorpus visuaalselt ja arvudes	49
4.2 Murdekorpuse märgendus ja XML-fail	53
5 VERBIÜHENDITE TUVASTAMINE	55
6 MEETODID	59
6.1 Kvalitatiivsete ja kvantitatiivsete meetodite kombineerimine käes- olevas töös	60
6.2 Kollostruktsioonilised meetodid	61
6.2.1 Meetodi rakendamine selles töös	62
6.2.2 Kollostruktsiooniliste meetodite kriitika	65
6.3 Korrespondentsanalüüs	66
6.4 Hägus klasteranalüüs	73
6.5 Geograafilise signaali tugevus	74
6.6 Verbiühendite produktiivsuse mõõtmisest selles töös	75

6.7	Kartograafia	76
7	KVANTITATIIVSE KOBARANALÜÜSI TULEMUSED	78
7.1	Sageduste normaliseerimine	78
7.2	Korrespondentsanalüüsi tulemused	79
7.3	Hägusa klasteranalüüsi tulemused	84
7.4	Kvantitatiivse kobaranalüüsi kokkuvõte	86
8	KONSTRUKTSIOONIDE ANALÜÜSI PÕHIMÕTTED	87
8.1	Analüüsitavad verbiühendid	87
8.2	Konstruktioonide analüüsi alused	88
8.3	Konstruktioonide analüüsi probleemid ja piirangud	89
8.4	Impersonaalsuse mõistest konstruktsioonide kirjeldamisel	90
8.5	Ahelverb, perifrastiline verbivorm ja verbiühend	92
8.6	Konstruktioonide produktiivsus ida- ja läänepoolsetes murretes ...	92
9	VERBIÜHENDID EESTI MURRETES	94
9.1	<i>da</i> -infinitiivi konstruktsioonid	94
9.1.1	<i>da</i> -infinitiiv ja finiitne <i>saama</i>	95
9.1.2	<i>da</i> -infinitiiv ja finiitne <i>võima</i>	102
9.1.3	<i>da</i> -infinitiiv ja finiitne <i>tohtima</i>	105
9.1.4	<i>da</i> -infinitiiv ja finiitne <i>tulema</i>	107
9.1.5	<i>da</i> -infinitiiv ja finiitne <i>laskma</i>	110
9.1.6	<i>da</i> -infinitiiv ja finiitne <i>olema</i>	113
9.1.7	<i>da</i> -infinitiiv ja finiitne <i>tahtma</i>	119
9.1.8	<i>da</i> -infinitiiv ja finiitne <i>jõudma</i>	123
9.1.9	<i>da</i> -infinitiiv ja finiitne <i>oskama</i>	125
9.1.10	Kokkuvõte	128
9.2	<i>ma</i> -infinitiivi konstruktsioonid	134
9.2.1	<i>ma</i> -infinitiiv ja finiitne <i>hakkama</i>	135
9.2.2	<i>ma</i> -infinitiiv ja finiitne <i>minema</i>	138
9.2.3	<i>ma</i> -infinitiiv ja finiitne <i>tulema</i>	142
9.2.4	<i>ma</i> -infinitiiv ja finiitne <i>panema</i>	145
9.2.5	<i>ma</i> -infinitiiv ja finiitne <i>pidama</i>	149
9.2.6	<i>ma</i> -infinitiiv ja finiitne <i>saama</i>	153
9.2.7	<i>ma</i> -infinitiiv ja finiitne <i>olema</i>	155
9.2.8	<i>ma</i> -infinitiiv ja finiitne <i>jääma</i>	157
9.2.9	<i>ma</i> -infinitiivi inessiiv ja finiitne <i>käima</i>	160
9.2.10	<i>ma</i> -infinitiivi inessiiv ja finiitne <i>olema</i>	163
9.2.11	Kokkuvõte	167
9.3	<i>tud</i> -partitsiibi konstruktsioonid	173
9.3.1	<i>tud</i> -partitsiip ja finiitne <i>olema</i>	174
9.3.2	<i>tud</i> -partitsiip ja finiitne <i>saama</i>	178
9.3.3	Kokkuvõte	182
9.4	Konstruktioonide analüüsi kokkuvõte	185
9.4.1	Erinevused varasematest käsitlustest	185
9.4.2	Konstruktioonide kvantitatiivne kobaranalüüs	186

10 KOKKUVÕTE	189
10.1 Töö tulemused varasemate uurimuste kontekstis	189
10.2 Töö tulemuste olulisus	192
10.3 Töö kitsaskohad ja edasised uurimissuunad	194
10.4 Lõpetuseks	195
KIRJANDUS	196
SUMMARY IN ENGLISH	208
LISA 1. MURDEKORPUSE MORFOLOOGILISELT MÄRGENDATAVAD KATEGOORIAD	213
LISA 2. VERBIÜHENDITE NORMALISEERITUD SAGEDUSTE ANALÜÜSI ALGANDMED	216
LISA 3. KOLLOSTRUKTSIOONILIST ANALÜÜSI ARVESTAVA NORMALISEERITUD SAGEDUSTE ANALÜÜSI ALGANDMED	220
ELULOOKIRJELDUS	222
CURRICULUM VITAE	224

Tabelid

Tabel 1. Süntaksi-leksikoni kontiinum (kohandatud Crofti ja Cruse'i (2004: 255) järgi)	27
Tabel 2. Eesti keele infiniitsed vormid ja nende formatiivid (Viitso 2003)	38
Tabel 3. Murdekorpuse ülevaade 01.09.2010 seisuga	50
Tabel 4. Murdekorpuse morfoloogiliselt märgendatud sõnade arv murrete kaupa uurimismaterjali kogumise hetkel (seisuga 01.09.2010)	52
Tabel 5. Kahemõõtmeline sagedustabel koosvarieeruvate kolekseemide analüüsi jaoks vajalike väärtuste arvutamiseks	64
Tabel 6. Kahemõõtmeline sagedustabel. <i>da</i> -infinitiiv + <i>tahtma</i> ühend idamurdes	64
Tabel 7. Korrespondentsanalüüsi näiteandmestik	68
Tabel 8. Koos finiiitse <i>saama</i> -verbiga esinevad verbid <i>tud</i> -partitsiibis ja statistilise seose tugevuse väärtus	76
Tabel 9. <i>da</i> -infinitiivi ja <i>saama</i> -verbi konstruktsioonid ning nende esinemisjuhtude normaliseeritud sagedused. Lühendid: imp – impersonaalikonstruktsioon, püsi – püsiühend, ment – mentaalse verbi konstruktsioon, mod – modaalkonstruktsioon, pass – passiivikonstruktsioon	98
Tabel 10. <i>saama</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	102

Tabel 11.	<i>da</i> -infinitiivi ja <i>võima</i> -verbi modaalkonstruksiooni (mod) esinemisjuhtude normaliseeritud sagedused	103
Tabel 12.	<i>võima</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	104
Tabel 13.	<i>da</i> -infinitiivi ja <i>tohtima</i> -verbi modaalkonstruksiooni (mod) esinemisjuhtude normaliseeritud sagedused	105
Tabel 14.	<i>tohtima</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	107
Tabel 15.	<i>da</i> -infinitiivi ja <i>tulema</i> -verbi konstruksioonid ning nende esinemisjuhtude normaliseeritud sagedused. Lühendid: atrib – atributiivkonstruksioon, finaal – finaalkonstruksioon, mod – modaalkonstruksioon	108
Tabel 16.	<i>tulema</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	110
Tabel 17.	<i>da</i> -infinitiivi ja <i>laskma</i> -verbi konstruksioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: instrument – instrumendikonstruksioon, mod – modaalkonstruksioon	111
Tabel 18.	<i>laskma</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	113
Tabel 19.	<i>da</i> -infinitiivi ja <i>olema</i> -verbi konstruksioonid ning nende esinemisjuhtude normaliseeritud sagedused. Lühendid: lahk – lahkkonstruksioon, prag – pragmaatilise funktsiooniga konstruksioon, ment – mentaalse verbi konstruksioon, maa – vahemaad väljendav konstruksioon, mod – modaal-konstruksioon, om – omamistähendusega konstruksioon, omadus – omadust väljendav konstruksioon, tarvis_vaja – modaaladverbidega <i>tarvis</i> või <i>vaja</i> modaalkonstruksioon, üt – öeldistäite-konstruksioon	114
Tabel 20.	<i>olema</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	119
Tabel 21.	<i>da</i> -infinitiivi ja <i>tahtma</i> -verbi konstruksioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: mod – modaalkonstruksioon, proks – proksimatiivkonstruksioon, soov – soovi väljendav konstruksioon, vastupanu – vastupanu väljendav konstruksioon	120
Tabel 22.	<i>tahtma</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	123
Tabel 23.	<i>da</i> -infinitiivi ja <i>jõudma</i> -verbi suutlikkust väljendava konstruksiooni (suut) esinemisjuhtude normaliseeritud sagedused	123
Tabel 24.	<i>jõudma</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	125
Tabel 25.	<i>da</i> -infinitiivi ja <i>oskama</i> -verbi oskamist väljendava konstruksiooni (oskam) esinemisjuhtude normaliseeritud sagedused.	126

Tabel 26.	<i>oskama</i> -verbiga tugevamini tõmbuvad <i>da</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	128
Tabel 27.	Finiitse verbi ja <i>da</i> -infinitiiviga moodustatavad konstruktsioonid eesti murretes	132
Tabel 28.	<i>ma</i> -infinitiivi ja <i>hakkama</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: algus – algust väljendav konstruktsioon, püsi – püsiühend	136
Tabel 29.	<i>hakkama</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	138
Tabel 30.	<i>ma</i> -infinitiivi ja <i>minema</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: afadv – afiksaaladverbikonstruktsioon, algus – algust väljendav konstruktsioon, lok – lokatiivse tähendusega konstruktsioon, püsi – püsiühend	139
Tabel 31.	<i>minema</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	142
Tabel 32.	<i>ma</i> -infinitiivi ja <i>tulema</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: afadv – afiksaaladverbikonstruktsioon, püsi – püsiühend, lok – lokatiivse tähendusega konstruktsioon	143
Tabel 33.	<i>tulema</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	145
Tabel 34.	<i>ma</i> -infinitiivi ja <i>panema</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: algus – algust väljendav konstruktsioon, kaus – kausatiivkonstruktsioon, püsi – püsiühend	146
Tabel 35.	<i>panema</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	148
Tabel 36.	<i>ma</i> -infinitiivi ja <i>pidama</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: avert – avertiivkonstruktsioon, kvot – kvotatiivikonstruktsioon, kvot_ips – impersonaalne kvotatiivikonstruktsioon, mod – modaalkonstruktsioon, proks – proksimatiivkonstruktsioon	150
Tabel 37.	<i>pidama</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	152
Tabel 38.	<i>ma</i> -infinitiivi ja <i>saama</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: fut – tulevikukonstruktsioon, kaus – kausatiivikonstruktsioon, püsi – püsiühend, fut_ipf – imperfekti tuleviku konstruktsioon, õnnest – õnnestumistähendusega konstruktsioon	153
Tabel 39.	<i>ma</i> -infinitiivi ja <i>olema</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: mod – modaalkonstruktsioon, omadus – omadust väljendav konstruktsioon, proks – proksimatiivkonstruktsioon	156

Tabel 40.	<i>ma</i> -infinitiivi ja <i>jääma</i> -verbi algust väljendava (algus) konstruktsiooni esinemisjuhtude normaliseeritud sagedused	158
Tabel 41.	<i>jääma</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas	160
Tabel 42.	<i>ma</i> -infinitiivi inessiivi ja <i>käima</i> -verbi lokatiivse konstruktsiooni (lok) esinemisjuhtude normaliseeritud sagedused	161
Tabel 43.	<i>käima</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi inessiivi vormis verbid ida- ja läänepoolsete murrete rühmas	163
Tabel 44.	<i>ma</i> -infinitiivi inessiivi ja <i>olema</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: lok – lokatiivse tähendusega konstruktsioon, omadus – omadust väljendav konstruktsioon, progr – progressiivkonstruktsioon, proks – proksimatiivkonstruktsioon, püsi – püsiühend	164
Tabel 45.	<i>olema</i> -verbiga tugevamini tõmbuvad <i>ma</i> -infinitiivi inessiivi vormis verbid ida- ja läänepoolsete murrete rühmas	167
Tabel 46.	Finiitse verbi ja <i>ma</i> -infinitiiviga moodustatavad konstruktsioonid eesti murretes	171
Tabel 47.	<i>tud</i> -partitsiibi ja finitse <i>olema</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: imp – impersonaalikonstruktsioon, pass – passiivikonstruktsioon, possperf – possessiivne perfekt	175
Tabel 48.	<i>olema</i> -verbiga tugevamini tõmbuvad <i>tud</i> -partitsiibis verbid ida- ja läänepoolsete murrete rühmas	178
Tabel 49.	<i>tud</i> -partitsiibi ja finitse <i>saama</i> -verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: imp – impersonaalikonstruktsioon, pass – passiivikonstruktsioon, possperf – possessiivne perfekt, result – resultatiivsus-konstruktsioon	179
Tabel 50.	<i>saama</i> -verbiga tugevamini tõmbuvad <i>tud</i> -partitsiibis verbid ida- ja läänepoolsete murrete rühmas	181
Tabel 51.	Finiitse verbi ja <i>tud</i> -partitsiibiga moodustatavad konstruktsioonid eesti murretes	184

Joonised

Joonis 1.	Madal ja kõrge produktiivsus (kohandatud Crofti ja Cruse'i (2004: 309) järgi)	36
Joonis 2.	Eesti murrete korpuses kasutatav murdejaotus. Põhjaeesti murderühm: ida, kesk-, lääne- ja saarte murre; lõunaeesti murderühm: Mulgi, Seto, Tartu ja Võru murre; kirderannikurühm: rannamurre ja kirdemurre (Alutaguse)	50
Joonis 3.	Murdekorpuse tekstide päritolu. Kaardina on kasutatud Google'i Eesti kaarti < https://maps.google.com/ > [03.09.2012]	51
Joonis 4.	Murdekorpuse informantide sünniaegade jaotus	53

Joonis 5.	Murdekorpuse XML-dokumendi näide	54
Joonis 6.	Süntaksianalüsaatori väljund osalausestatud failist	56
Joonis 7.	Korrespondentsanalüüsi kaart näiteandmestiku põhjal	72
Joonis 8.	Näiteandmestiku hägusa klasteranalüüsi visuaalne tulemus	73
Joonis 9.	Korrespondentsanalüüsi kaart normaliseeritud sageduste põhjal. Lühendid: IDA – idamurre, KES – keskmurre, KIR – kirde- murre, LÄÄ – läänemurre, MUL – Mulgi murre, RAN – ranna- murre, SAA – saarte murre, SET – Seto murre, TAR – Tartu murre, VÕR – Võru murre	80
Joonis 10.	Korrespondentsanalüüsi kaart kollostruktsioonilise analüüsi tulemuste ja normaliseeritud sageduste põhjal. Lühendid: IDA – idamurre, KES – keskmurre, KIR – kirdemurre, LÄÄ – läänemurre, MUL – Mulgi murre, RAN – rannamurre, SAA – saarte murre, SET – Seto murre, TAR – Tartu murre, VÕR – Võru murre	83
Joonis 11.	Hägusa klasteranalüüsi tulemus normaliseeritud sageduste põhjal	84
Joonis 12.	Hägusa klasteranalüüsi tulemus kollostruktsioonilise analüüsi ja normaliseeritud sageduste põhjal	85
Joonis 13.	Finiitse <i>saama</i> ja <i>da</i> -infinitiivi konstruktsioonide geograafiline sagedusesitus	101
Joonis 14.	Finiitse <i>võima</i> ja <i>da</i> -infinitiivi modaalkonstruktsiooni geograafiline sagedusesitus	104
Joonis 15.	Finiitse <i>tohtima</i> ja <i>da</i> -infinitiivi modaalkonstruktsiooni geograafiline sagedusesitus	106
Joonis 16.	Finiitse <i>tulema</i> ja <i>da</i> -infinitiivi modaalkonstruktsiooni geograafiline sagedusesitus	109
Joonis 17.	Finiitse <i>laskma</i> ja <i>da</i> -infinitiivi modaalkonstruktsiooni geograafiline sagedusesitus	112
Joonis 18.	Finiitse <i>olema</i> ja <i>da</i> -infinitiivi konstruktsioonide geograafiline sagedusesitus	118
Joonis 19.	Finiitse <i>tahtma</i> ja <i>da</i> -infinitiivi soovi väljendava konstruktsiooni geograafiline sagedusesitus	122
Joonis 20.	Finiitse <i>jõudma</i> ja <i>da</i> -infinitiivi suutlikkuskonstruktsiooni geograafiline sagedusesitus	124
Joonis 21.	Finiitse <i>oskama</i> ja <i>da</i> -infinitiivi konstruktsiooni geograafiline sagedusesitus	127
Joonis 22.	<i>da</i> -infinitiivi ja finiiitse verbi ühendite sagedused (1)	129
Joonis 23.	<i>da</i> -infinitiivi ja finiiitse verbi ühendite sagedused (2)	130
Joonis 24.	<i>da</i> -infinitiivi ja finiiitse verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafik	131
Joonis 25.	Finiitse <i>hakkama</i> ja <i>ma</i> -infinitiivi alguskonstruktsiooni geograafiline sagedusesitus	137
Joonis 26.	Finiitse <i>minema</i> ja <i>ma</i> -infinitiivi konstruktsioonide geograafiline sagedusesitus	141

Joonis 27. Finiitse <i>tulema</i> ja <i>ma</i> -infinitiivi lokatiivse konstruktsiooni geograafiline sagedusesitus	144
Joonis 28. Finiitse <i>panema</i> ja <i>ma</i> -infinitiivi kausatiivkonstruktsiooni geograafiline sagedusesitus	148
Joonis 29. Finiitse <i>pidama</i> ja <i>ma</i> -infinitiivi konstruktsioonide geograafiline sagedusesitus	152
Joonis 30. Finiitse <i>jääma</i> ja <i>ma</i> -infinitiivi alguskonstruktsiooni geograafiline sagedusesitus	159
Joonis 31. Finiitse <i>käima</i> ja <i>ma</i> -infinitiivi inessiivi konstruktsiooni geograafiline sagedusesitus	162
Joonis 32. Finiitse <i>olema</i> ja <i>ma</i> -infinitiivi inessiivi konstruktsioonide geograafiline sagedusesitus	166
Joonis 33. <i>ma</i> -infinitiivi ja finiidse verbi ühendite sagedused (1)	168
Joonis 34. <i>ma</i> -infinitiivi ja finiidse verbi ühendite sagedused (2)	169
Joonis 35. <i>ma</i> -infinitiivi ja finiidse verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafik	170
Joonis 36. Finiitse <i>olema</i> ja <i>tud</i> -partitsiibi konstruktsioonide geograafiline sagedusesitus	177
Joonis 37. Finiitse <i>saama</i> ja <i>tud</i> -partitsiibi konstruktsioonide geograafiline sagedusesitus	181
Joonis 38. <i>tud</i> -partitsiibi ja finiidse verbi ühendite sagedused	182
Joonis 39. <i>tud</i> -partitsiibi ja finiidse verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafik	183
Joonis 40. Finiitse ja infiniitse verbi konstruktsioonide korrespondentsanalüüsi graafik	187

LÜHENDID

afadv	afiksaaladverbikonstruktsioon
algus	algust väljendav konstruktsioon
atrib	atributiivkonstruktsioon
avert	avertiivkonstruktsioon
CA	korrespondentsanalüüs
EKG	Erelt <i>et al.</i> 1995; Erelt <i>et al.</i> 1993
EMK	Eesti murrete korpus
finaal	finaalkonstruktsioon
FTT	Fisher-Yates'i test
fut	tulevikukonstruktsioon
fut_ipf	imperfekti tuleviku konstruktsioon
IDA	idamurre
Inf	infiniitne vorm
imp	impersonaalikonstruktsioon
instrument	instrumendikonstruktsioon
kaus	kausatiivkonstruktsioon
KG	konstruktsioonigrammatika
KES	keskmurre
KIR	kirdemurre
kvot	kvotatiivikonstruktsioon
kvot_ips	impersonaalne kvotatiivikonstruktsioon
lahk	lahkkonstruktsioon
lok	lokatiivse tähendusega konstruktsioon
LÄÄ	läänemurre
maa	vahemaad väljendav konstruktsioon
ment	mentaalse verbi konstruktsioon
mod	modaalkonstruktsioon
MUL	Mulgi murre
om	omamistähendusega konstruktsioon
omadus	omadust väljendav konstruktsioon
pass	passiivikonstruktsioon
possperf	possessiivne perfekt
progr	progressiivkonstruktsioon
proks	proksimatiivkonstruktsioon
püsi	püsiühend
RAN	rannamurre
result	resultatiivsuskonstruktsioon
SAA	saarte murre
SET	Seto murre
soov	soovi väljendav konstruktsioon
suut	suutlikkuskonstruktsioon
TAR	Tartu murre

tarvis_vaja	modaaladverbidega <i>tarvis</i> või <i>vaja</i> moodustunud modaalkonstruktsioon
vastupanu	vastupanu väljendav konstruktsioon
V _{fin}	finiitne verb
VÕR	Võru murre
õnnest	õnnestumistähendusega konstruktsioon
õt	õeldistäite-konstruktsioon

I SISSEJUHATUS

„As soon as you take dialects seriously as fully articulated linguistic systems you have demonstrated that you are interested in language structure as such and not in the construction of or adherence to prescriptive norms.” (Vogelaar & Seiler 2012: 3)

Murre on enamjaolt suuline keelevariant, mida traditsiooniliselt on uuritud rohkem leksikaalsel ja foneetilisel tasandil. Murrete lauseehituslike erijoonte kohta on siiani palju vähem teadmisi. Süntaktilist varieerumist on tunduvalt keerulisem tabada kui näiteks häälduslikke ja sõnavaralisi eripärasid. Kui keegi kasutab mingit sõna, mida kuulaja selles kontekstis kunagi enda arvates ei kasutaks, või hääldab ö-d, kus kuulaja on harjunud kuulma õ-d, siis sellised erijooned on erakordselt kergesti tabatavad ja jäävad kohe kõrva. Kui keegi aga kasutab sagedamini teistsugust sõnajärge, eelistab öelda „*jooksime*” asemel „*me jooksime*” või kasutab alati konstruktsiooni „*mul saab kohe kõik tehtud*” asemel alati „*mul on kohe kõik tehtud*”, siis seda teadlikult kuulamata ei märka. Ilmselt ei seosta keelekasutaja neid nähtusi üldse mingi võimaliku murdemõjuga, sest selline väljendusviis on kirjakeeleski täiesti tavaline. Enamasti pole sellised erinevused taandatavad vastandusele (kas keelekasutaja kasutab üht või teist väljendusviisi), vaid pigem ilmneb see eelistusmuutrites. Väitekiri uuribki ühe süntaktilise nähtuse – kahe verbi ühendite kasutuse eelistusmuutreid eesti murretes. Selliste eelistuste varieerumise uurimiseks on aga vaja tunduvalt suuremaid andmehulki kui näiteks häälikuvareerumise uurimiseks. Keeleressurside puudulikkus pole varem võimaldanud kõiki murdeid hõlmavaid kvantitatiivseid uurimusi. Nüüd, mil Eesti murrete korpus, kust pärineb sinne töö uurimismaterjal, on piisavalt suur, on sellise töö tegemine võimalikuks saanud. Mahukas elektrooniline morfoloogiliselt märgendatud andmekogu teeb uurimismaterjali kättesaadavaks ja automaattöödeldavaks.

Miks murdeid uurida ja mida võiks see anda juurde meie teadmistele keelest? Erinevad autorid (Auer 2004: 69; Berthele 2004: 118 ; Kortmann 2004: 8; Matras 2004: 278) on rõhutanud, et murdematerjali analüüsimine võib aidata tuvastada grammatilisi nähtusi, mida on kas väga harva või pole üldse mitte kirjeldatud senistes uuritava keele tüpoloogilistes, sugulaskeelte või areaalsetes uurimustes. Seiler (2004: 368) rõhutab, et murded pakuvad andmeid ja materjali, mida tüpoloogilistes uurimustes pole sageli arvesse võetud, seega annavad murded meile keele kohta lihtsalt rohkem ja mitmekülgsemat informatsiooni. Kuna eesti murdeid pole süntaktilisest perspektiivist varem uuritud, võimaldab töö tuua esile laiemat tausta murrete seostest kirjakeelega ja aitab seeläbi vastata küsimustele, millised kirjakeele konstruktsioonid on erinevatele murretele iseloomulikud, millistest murretest on need pärit, kas ja kuidas ilmnevad murdejooned keelekasutaja lauseehituses jne.

Uurimus on uudne mitmest aspektist. Töö suurim panus on kindlasti metodoloogiline, nimelt on esmakordselt välja töötatud meetodite ja vahendite komplekt, mis loob kindla aluse ükskõik missuguse murdekorpuse põhjal tehtava süntaktiliste töö tegemiseks tulevikus. Rakendan töös mitmemõõtmelise

statistika meetodeid, mida siiani pole eesti murrete uurimisel (ega ka muu eesti (all)keele varieerumise uurimisel) üldse rakendatud, kuid mis on selles valdkonnas üldkasutatavamad (vt nt Grieve 2009; Keune *et al.* 2005; Spruit 2008; Szmrecsanyi 2013). Uudne on kollestruktsiooniliste meetodite rakendamine murdeandmestikul, mida autorile teadaolevalt pole siiani murrete uurimisel tehtud. Kasutan uusi tehnilisi võimalusi oma andmestiku ja töö tulemuste visualiseerimiseks ning väljatöötatud kartograafilised võtted võimaldavad tulevikus visualiseerida igasugust eesti murretest tehtavat tööd.

Töö teoreetiline panus on dialektoloogia sidumine tänapäevase keeleteooriaga – konstruktsioonigrammatika ja kasutuspõhise keeleuurimise kirjeldusvahendite rakendamisega – mida on siiani dialektoloogilistes töodes tehtud vähe. Käesolev töö esindab murdeuurimise uut suunda, mis püüab näidata, et murrete uurimine ei pea olema omaette, muust keeleteadusest eraldiseisev distsipliin, vaid kõik tänapäevased teoreetilised keelekirjeldusvõimalused on ühendatavad regionaalselt varieeruvate keelevariantide kirjeldamiseks. Samuti annab töö panuse uude arenevasse murdesüntaksi valdkonda ja eesti keele kui valdkonna seisukohast uue keele informatsiooni lisamisega uurimisaslasse.

1.1 Murdesüntaks

Traditsiooniline dialektoloogia on oma uurimisobjekti reeglina piiranud häälikulise, leksikaalse ja morfoloogilise varieerumisega, olles vähem huvitatud süntaktilisest varieerumisest. Viimase paarikümne aasta jooksul on tõusnud huvi ka süntaktiliste nähtuste varieerumise vastu ning praeguseks on süntaksist saamas väga aktiivne ja elav uurimisala, mis pretendeerib täiesti omaette valdkonnaks üldnimetusega murdesüntaks (vt ka Fried 2010: 9; Macaulay 2010: 67; Schutter 2010: 74). Auer (2004: 71–72) väidab, et murdesüntaks on oma iseloomult suulise keele süntaks, ent selles on jooni, mida suulises standardkeeles ei leidu ning mida seetõttu ei saa taandada ainult suulise registri eripäradele. Auer rõhutab, et murdesüntaks on vaid osa suulise keele süntaksist, mille oluline komponent on geograafiline mõõde.

Võrreldes häälikulise muutusega, on süntaktiline muutus vaevumärgatav (Labov 1991: 277), kuid näiteks Szmrecsanyi (ilmumas (a): 13) rõhutab, et pole mitte mingisugust põhjust eeldada, et süntaktilised nähtused geograafiliselt ei varieeru. Kortmann (2010: 842) omakorda lisab, et süntaktilistel nähtustel on laiem piirkondlik levik ja need on harvem piiratud väiksemate aladega ja seetõttu on süntaksi uurimiseks vaja tunduvalt suuremaid andmehulki kui need, mida saab kasutada fonoloogilisest ja leksikaalsest varieerumisest ülevaate saamiseks.

Viimastel aastatel on kas juba läbi viidud või on käimas mitmeid murdesüntaksialaseid projekte. Olulisim neist on aastatel 2005–2012 kestnud üle-euroopaline projekt Edisyn¹, mille üks eesmärkidest oli uurida Euroopa murde-

¹ Infot projekti kohta leiab aadressilt: http://www.dialectsyntax.org/wiki/About_Edisyn

süntaksi olukorda ning luua ühtne uurijate ning ressursside võrgustik. Projekti tulemusena on valminud internetikeskkond, kust leiab infot pea kõigi Euroopas lõpetatud või käimasolevate murdesüntaksi projektide kohta ja viiteid vabalt kättesaadavatele elektroonilistele murdesüntaksiressurssidele. Projekti üks olulisematest praktilistest väljunditest on erinevatele olemasolevatele murde-korpustele ühtse kasutajaliidese loomine (Kunst & Wesseling 2011). Projekti on kaasatud ka Eesti murrete korpus.

Edukamad on siiani olnud kindlasti hollandi murdeid hõlmav projekt SAND (Barbiers & Bennis 2007; Barbiers *et al.* 2007) ja Skandinaavia murdeid hõlmav ScanDiaSyn (Vangsnes 2007). Eesti keele sugulaskeeltest on murdesüntaksi projekt FinDiaSyn varem läbi viidud Soomes (<http://www.kotus.fi/index.phtml?s=2524>). Edisyn'i kodulehelt leiab erinevate käimasolevate ja lõpetatud murdesüntaksi projektide loetelu, kirjeldused ning ülevaated nende olulisematest tulemustest.

2013. aastast käivitus Liina Lindströmi juhtimisel ka eesti murdesüntaksi projekt EstDiaSyn, mis keskendub süntaktiliste nähtuste uurimisele eesti murretes. Projekti üks olulisemaid eesmärke on välja töötada metodoloogia, mis on hiljem rakendatav ükskõik missuguste (all)keelte varieerumise uurimisel. Doktoriväitekiri on eeltöö sellele järgnevad kolm aastat kestvale projektile. Siinses töös väljatöötatud rakendused ja metodoloogiline raamistik on EstDiaSyni sisendiks.

I.2 Kobardialektoloogia

Töö mõistmise seisukohalt pean vajalikuks selgitada kohe sissejuhatuses kobardialektoloogia (*aggregate dialectology*) mõistet, mida pole siiani eesti keeleteaduses kasutatud. Mõnevõrra põhjalikumalt ja kontekstiga seotult tuleb sellest juttu peatükis 3. Kobardialektoloogia on murdeuurimisvaldkond, kus ei tegeleta ühe nähtuse uurimisega, nagu see enamasti traditsiooniliselt murdeuurimises on olnud, vaid vaadeldakse koos paljusid erinevaid keelenähtusi ja nende kõigi varieerumist korraga ning järeldused tehakse geograafilise varieerumise kohta paljude nähtuste koosvarieerumise põhjal (Nerbonne 2009). Kobardialektoloogia on vastupidine lähenemisviis traditsioonilisele murdeuurimisele, kus vaadeldakse vaid ühte keelenähtust ning selle varieerumist, nn „liblikakogumise lähenemine” (*butterfly collecting mentality*) (Filppula *et al.* 2005: vii). Kobardialektoloogia näol pole tegemist siiski mingi selgelt välja-töötatud raamistikuga, samuti pole olemas ühemõttelist piiri, kui palju nähtusi peaks usaldusväärse tulemuse saamiseks uurimusse kaasama. Üldine põhimõte on, et kaasata tuleb võimalikult palju eri tunnuseid. Süntaktilise kobardialektoloogia headeks näideteks võib pidada Grieve'i (2009) ja Spruiti (2008) doktoritöid ning Szmrecsanyi (2013) värselt ilmunud monograafiat. Tavaliselt keskenduvad kobardialektoloogilised tööd siiski mingi kindla keeleteaduse valdkonna uurimisele, nagu näiteks eelnevalt mainitud süntaksimonograafiad.

Käesolev uurimus ei ole traditsioonilises mõttes kobardialektoloogiline töö, kuna uurimise fookuses on vaid üks süntaktiline nähtus – kahe verbi ühendid. Siiski pole ma keskendunud vaid üksikute vormide või väljavalitud verbide abil moodustatud konstruktsioonidele, vaid vaatlen koos kõiki korpusest tuvastatud infiniitse vormi ja finiiitse verbi ühendeid. Seega võib siinset tööd pidada kobardialektoloogiliseks, kuna rakendan vastavaid põhimõtteid ja metoodikat.

1.3 Uurimisobjekt: verbiühendid

Töö uurimisobjekt on kahe verbi, finiiitse ja infiniitse vormi ühendid, mille kaudu kirjeldan ühe süntaktilise nähtuse varieerumist eesti murretes. Sellise valiku eesmärk oli esiteks kirjeldada nähtust, mis on ühelt poolt piisavalt kõrge esinemissagedusega, teiselt poolt eesti keelele omane ja tavaline; teiseks kirjeldada keelenähtust, mis eesti murretes teadaolevalt suurel määral varieerub. Kuigi põhjalikumad verbiühendite uurimused murretes siiani puuduvad, lubas murdekorpusega töötamise kogemus eeldada, et just selliste verbiühendite kasutus on murrete lõikes väga erinev.

Nii nagu hiljem ka töö tulemused kinnitavad, pole vaadeldavate verbiühendite näol tegemist homogeense ja selgelt süntaksiga piirneva keelenähtusega. Finiiitse ja infiniitse verbi ühendid võivad küll moodustada selgelt süntaktilisi konstruktsioone, aga ka näiteks leksikaalsemaid üksusi erinevate püsiühendite kujul. See omakorda sobib töö teoreetiliseks raamistikuks oleva konstruktsioonigrammatikaga, mille põhieeldus on, et süntaks, morfoloogia ja leksika pole üksteisest eraldiseisvad valdkonnad, vaid keelenähtused paiknevad nende kolme kontinuumil. Teiselt poolt sobituvad uuritavad verbiühendid töö teise põhilise eesmärgiga rakendada kobardialektoloogilisi uurimismeetodeid ja -põhimõtteid. Kuna kobardialektoloogia koondab erinevaid keelenähtusi ja teeb üldistusi nende põhjal, siis seda ülesannet sobib verbiühendite uurimine suurepäraselt täitma, sest oma olemuselt on see üsna heterogeenne ning võimaldab teha üldistusi erinevate nähtuste koondanalüüsi põhjal.

1.4 Töö ülesehitus

See osa annab ülevaate töö struktuurist. Töö teine peatükk kirjeldab põhjalikumalt uurimistöö eesmärgi ja hüpoteesi. Kolmandas peatükis tuuakse esile töö teoreetilised lähtekohad ja nende rakendamine siinses töös. Teoreetiline peatükk jaguneb alapeatükkideks, käsitledes ka sageduse temaatikat keeleteaduses, ning tutvustab kasutuspõhise keeleuurimise põhimõtteid. Samas peatükis annan lühikese ülevaate infiniitsetest verbivormidest ning ka varasematest eesti murrete süntaksi uurimustest. Põhjalikumalt käsitlen samas peatükis ka korpuspõhist dialektoloogiat. Neljas peatükk tutvustab Eesti murrete korpust, kust on pärit uurimismaterjal ja viies peatükk annab ülevaate poolautomaatse verbiühendite tuvastamise algoritmist. Kuues peatükk tutvustab konstrueeritud näidete põhjal kvantitatiivses analüüsis kasutatavaid meetodeid. Väitekirja

seitsmes peatükk esitab kvantitatiivse analüüsi tulemused. Kaheksas peatükk tutvustab konstruktsioonide analüüsi põhimõtteid ja läbiviimise protseduuri. Konstruktsioonide analüüs üheksandas peatükis moodustab töö kõige mahukama osa, kus esitan nii kvalitatiivse iseloomuga konstruktsioonide analüüsi kui ka kvantitatiivse analüüsi põhjal tuvastatud verbikonstruktsioonid eesti murretes ning esitan nende liigituse. Töö viimases, kümnendas peatükis toon kokkuvõtvalt välja töö tulemused koos nende uurimuse kitsaskohtade ja võimalike suundadega edaspidises uurimistöös.

2 TÖÖ EESMÄRGID JA HÜPOTEESID

„Keele grammatilise struktuuri selgitamisel tuleb lähtuda keele enese materjalist, mitte aga lasta ennast kõrvale kallutada valmisskeemidest ja teistes keeltes eksisteerivaist süsteemidest.” (Rätsep 1957: 79)

Väitekirja eesmärk on kirjeldada uuritava nähtuse varieerumist eelarvamustevabalt, ainult uurimismaterjali põhjal, seadmata eelnevalt materjalile suuri teoreetilisi piiranguid. Teoreetilist raamistikku kasutan vaid kirjeldusvahendina ning töös esitatud konstruktsioonide liigitus ei ole kõikehõlmav ja ainuvõimalik.

Varem on esitatud üleskutseid, et dialektoloogia peaks end palju rohkem siduma teiste lingvistikavaldkondadega ning rakendama vastavaid meetodeid. Funktsionaalne ja kasutus põhine raamistik pakub näiteks Berthele (2004: 121–122) arvates varieerumise uurimiseks suurepäraseid võimalusi. Olen selliste väidetega igati nõus ning püüan oma töös vastavaid põhimõtteid järgida.

Tööl on mitu suuremat eesmärki. Töö olulisim keeleteaduslik eesmärk on saada esmast teavet eesti murrete süntaktilise varieerumise kohta ühe süntaktilise nähtuse – finiitse ja infiniitse verbi ühendite – analüüsi kaudu. Juba murdekorpuse tekste lugedes võib märgata, et selliste verbiühendite murretevaheline varieerumine on märkimisväärne. Annan töös põhjaliku ülevaate murretes esinevate verbiühendite kohta ja vaatlen nende varieerumist kõigis kümnes eesti murdes². Olen seadnud eesmärgiks tuvastada, kas eesti murded grupeeruvad $V_{fin}+Inf$ ühendite põhjal samamoodi nagu siiani esitatud murdejaotustes, mis on enamjaolt tehtud foneetika, morfoloogia ja leksika põhjal ja üldjuhul süntaksit arvesse ei võta. Siiani pole olemas ühtegi eesti murdeliigendust, mis kaasaks süntaktilise mõõtme.

Selles töös püüan tuua teineteisele lähemale üldkeeleteadust ja dialektoloogiat ning näitan, et murdeid ei peaks uurima teistsuguste meetoditega kui need, mida kasutatakse ükskõik missuguse keelevariandi või -nähtuse uurimiseks. Minu seisukoht on, et dialektoloogia ei peaks olema eraldiseisev uurimisvaldkond, vaid mis tahes keeleteaduslikud teooriad ja meetodid on rakendatavad ka murrete uurimisel, aga ka vastupidi: usun, et dialektoloogia pakub omalt poolt palju üldkeeleteadusteooriatele ning on suureks abiks oma mitmekülgse ja uudse materjaliga nii teooriate kui ka keeleuurimise meetodite arendamisel.

Töö kolmas suurem eesmärk on ühtse aluse ja tehnilise baasi väljatöötamine, mis võimaldaks tulevikus eesti murrete korpuse põhjal uurida ükskõik missugust süntaktilist nähtust. Uurimismaterjali kogumise käigus olen välja töötanud tehnilised vahendid, mis lihtsate modifikatsioonidega on rakendatavad uurimismaterjali kogumiseks ka muude nähtuste uurimisel. Töö rakendab automaatset süntaksianalüüsi lausepiiride määramiseks ning mõõdab selle kvaliteeti. Kuigi süntaktilise automaatanalüüsi korrektsuse hindamine on omaette huvitav ja suur teema, pole ma sellele siinses töös aja- ja mahupiirangute

² Siinses töös järgitavat murdejaotust on kirjeldatud peatükis 4.1.

tõttu põhjalikumalt keskendunud, kuid töö loob kõik eeldused sellise analüüsi tegemiseks tulevikus. Uurimuse tulemusena on olemas põhjalik ülevaade murdekorpuses sisalduvast ja uurimismaterjaliga seotud probleemidest. Kõik see võimaldab tulevikus saada usaldusväärsemat uurimismaterjali ja samuti saavad üsna selgeks probleemid, millele korpuse arendamisel tulevikus tähelepanu pöörama peaks. Töö üks tehnilisi väljundeid on visualiseerimisraamistiku väljatöötamine, mis on tulevikus rakendatav samuti ükskõik missuguse murdenähtuse uurimisel, kirjeldamisel ja uurimistulemuste visualiseerimisel.

Töö neljas oluline eesmärk on metodoloogiline. Murdesüntaks on väheuuritud valdkond ja eriti korpuspõhised murdesüntaksiuurimused on ikka veel üsna harvad. Siiani on selline töö enamjaolt takerdunud keeleressursside kättesaadavuse taha. Korpuste loomisega ja andmemahutude suurenemisega on hakanud ilmuma süntaksi alalt ka korpusdialektoloogilisi töid (Grieve 2009; Hernández 2010; Schulz 2010; Szmrecsanyi 2013). Valdonna nooruse tõttu on vastavad meetodid alles arenemisjärgus. Siinses töös rakendan üht tänapäevast korpuslingvistilist konstruktsioonigrammatikal põhinevat meetodit – kollostruktsioonilist analüüsi – ning keele varieerumise uurimisel laiemalt kasutatud mitmemõõtmelise statistika meetodeid – korrespondentsanalüüsi ja hägusat klasteranalüüsi. Kuigi mainitud meetodid pole statistikas uued, on neid siiani keele-, aga eriti just murrete uurimisel rakendatud vähe.

Töö hüpoteesid ja olulisemad eesmärgid jagunevad keeleteaduslikeks ja metodoloogilisteks.

- Töö põhiline hüpotees on, et finiiitse ja infiniitse verbi ühendid varieeruvad murrete lõikes suurel määral. Eeldan, et uuritava süntaktilise nähtuse põhjal kujunev murdejaotus erineb traditsioonilisest, mis süntaktilisi nähtusi ei arvesta. Oletan eelneva murdetekstidega töötamise kogemuse põhjal, et ida- ja läänepoolsete murrete vahelised erinevused verbiühendite kasutuses on tunduvalt suuremad kui erinevused põhja- ja lõunaeesi murrete vahel, mida traditsiooniliselt on olulisemaks peetud.
- Eelmisest hüpoteesist tulenevalt eeldan, et läänepoolsed murded kasutavad rohkem kahest verbist koosnevaid ühendeid kui idapoolsed murded. Seda lubab oletada senine põhjalik töö murdetekstidega ja eelnevad uurimistulemused (Uiboed 2010a; Uiboed *et al.* 2013).
- Väidan, et murdeuurimine ei vaja omaette murdeuurimise teooriat ja meetodeid. Püüan näidata, et varieerumist murdeis on võimalik kirjeldada tänapäevases keeleteoreetilises raamistikus ning dialektoloogiat ei peaks muust (keele)teadusest teoreetiliselt ja metodoloogiliselt eraldama. Näitan, et konstruktsioonigrammatika sobib süntaktilise varieerumise kirjeldusvahendiks.
- Metodoloogiliselt eeldan, et laialt kasutatavad kvantitatiivse korpuslingvistika meetodid sobivad suurepärast ka korpusdialektoloogiliseks tööks ning nendest on palju abi konstruktsioonilise varieerumise tuvastamisel.
- Keeleteaduses on palju vaidlusi tekitanud keelenähtuste sageduse olulisus keelesüsteemis (Arppe *et al.* 2010; Bybee 2010; Gries *et al.* 2005).

Varasemad tööd on näidanud, et esinemissagedus kirjeldab üsna hästi uuritavat keelenähtust ja ka ainult sageduse põhjal on võimalik tuvastada keeles eksisteerivaid konstruktsioone (Evert & Krenn 2001; Uibo 2010a). Olen seisukohal, et sagedus on keelenähtuste uurimisel ja keelenähtuste kirjeldamisel oluline ning (koos)esinemissageduse põhjal on võimalik tuvastada nii varieerumismustreid kui ka keelekonstruktsioone.

Niisiis on töö eesmärkideks siduda tänapäevaseid üldkeeleteaduslikke teooriaid murdeuurimisega, rakendada uusi kvantitatiivseid korpuslingvistika meetodeid dialektoloogilise ainestiku peal, saada esmane ülevaade süntaktilisest varieerumisest eesti murretes, töötada välja metodoloogia eesti murrete korpuse süntaktiliste nähtuste uurimiseks ja uurimistöö tulemuste esitamiseks ning tuvastada murretes esinevad finiiitse ja infiniitse verbi konstruktsioonid ning kirjeldada nende varieerumise mustreid.

3 TEOREETILINE TAUST

See peatükk kirjeldab uurimuse teoreetilist tausta. Esmalt annan ülevaate konstruktsioonigrammatikast ja kasutuspõhisest keeleuurimisest, mis on töö keeleteoreetiliseks aluseks. Esitan lühidalt varasemate uurimuste taustal süntaktilise produktiivsuse ja sageduse mõisted, nii nagu neid selles töös rakendan, jätkan infiniitsuse ja infiniitsete vormide ülevaatega. Seejärel selgitan abiverbi mõistet ja selle seost analüüsimaterjaliga. Jätkan dialektomeetria ja korpuspõhise dialektoloogia raamistiku kirjeldamisega ning esitan lühidalt varasemad nii eesti murrete süntaktilised uurimused kui ka eesti keele verbiühendite uurimused.

3.1 Konstruktsioonigrammatika

Käesoleva töö teoreetiliseks raamistikuks on konstruktsioonigrammatika. Konstruktsioonigrammatika (edaspidi KG) on mitte-komponentne keeleteooria, mida on nimetatud ka kognitiivse lingvistika (Langacker 1987) lähenemiseks süntaksile (Croft & Cruse 2004: 257). Konstruktsioonigrammatika ei ole homogeenne keeleteooria, vaid koosneb mitmetest eri suundumustest. Ühine on kõigile see, et keele väikseimaks alusosaks peetakse konstruktsiooni, s.t süntaktiliste, prosoodiliste, pragmaatiliste, semantiliste, tekstuaalsete jne tunnuste konventsionaliseerunud gruppe, mis korduvad vormi ja tähenduse kombinatsioonidena. Vorm konstruktsioonigrammatikas võib viidata nii süntaktiliste, morfoloogiliste kui ka prosoodiliste tunnuste kombinatsioonile ning tähendust ja/või funktsiooni käsitletakse väga laialt, hõlmates leksikaalse semantika, sündmuse struktuuri, diateesi, pragmaatika ja diskursuse struktuuri. (Fried ilmusas: 1) Niisiis käsitleb konstruktsioonigrammatika vormi ja tähendust teineteisest lahutamatuena.

KG ei erine põhimõtteliselt ühestki teisest keeleteooriast, püüdes leida parimat viisi keele struktuuri, tähenduse ja kasutuse suhete kirjeldamiseks. KG lähtub järgnevatest eeldustest: inimesed toetuvad erinevate tähenduste edasiandmisel suhteliselt keerulistele tähenduse-vormi paaridele (konstruktsioonidele); lingvistilised väljendid peegeldavad keelekasutuse ja keele materjali vahelisi suhteid; konstruktsioonid moodustavad üksteisega kattuvaid võrgustikke, mis on omavahel seotud sarnaste omadustega. (Fried & Östman 2004: 14)

Käesolev töö ei võta aluseks ühtegi konkreetset konstruktsioonigrammatilist lähenemist, vaid kombineerib konstruktsioonigrammatika kirjeldusvahendeid. Lähtun kõige üldisemast konstruktsioonigrammatilisest eeldusest, et keel koosneb konstruktsioonidest. Uuritav nähtus, finiiitse ja infiniitse verbi vorm koos oma tähendusega moodustab konstruktsiooni, ilma et ühe või teise verbi esinemine selles konstruktsioonis olulisem oleks. Kaks verbi moodustavad koos konstruktsiooni, mille tähendust ja funktsiooni ei ole võimalik tuletada kumagi verbivormi eraldiseisvast tähendusest või vormist. Samuti võib sellistes

konstruktsioonides esineda muid obligatoorseid ja/või fakultatiivseid elemente. Järgnev osa annab lühikese ülevaate KG olulisematest suundadest ja üldisematest ühistest põhimõtetest, kuid pole kindlasti kõikehõlmav.

KG on ühekihiline kitsenduspõhine süsteem, milles süntaktiline ja semantiline informatsioon on esitatud ühes tunnustestruktuuri skeemis (Kay & Fillmore 1999). Ühekihilisuse all mõeldakse siin seda, et KG raamistik ei esita mingeid ümberkirjutamisreegleid ning et pole olemas keelekirjeldussüsteemi eritasandilisi osi, vaid kõiki keeletasandeid (nt semantiline ja süntaktiline) kirjeldatakse ühtselt, ühes kihis. Konstruktsioonigrammatika on välja kasvanud käändegrammatikast (Fillmore 1968) ning on tihedalt seotud freimisemantikaga (Fillmore 1982). Esimesena hakkasid konstruktsioonipõhise lähenemisega tegelema Fillmore ja tema õpilased Californias Berkley ülikoolis. See KG suund (Fillmore 1988; Fillmore 1989; Fillmore *et al.* 1988; Kay & Fillmore 1999) on Crofti ja Cruse'i (2004: 266) hinnangul kõige lähemal formalistlikele keeleteooriatele ja keskendub konstruktsioonide pärilusseoste ja nende omavaheliste suhete detailsele uurimisele. Goldberg (2006: 215) peab seda teooriat teistest KG suundadest kõige vähem kasutuspõhiseks.

Crofti (2001) radikaalne konstruktsioonigrammatika on KG grammatika kasutuspõhise suuna esindaja, mille oluline lähtekoht on keeletüpoloogia; suundumus lähtub põhimõttest, et varieerumine on keele omadus, millest ei saa keele uurimisel mööda vaadata ja mida uurija peab kirjeldama (Croft 2001: 7). Radikaalne KG eitab universaalsete semantiliste ja süntaktiliste kategooriate olemasolu (Croft & Cruse 2004: 283) ja eeldab, et keelt saab kirjeldada vaid konstruktsioonide ning konstruktsioonis sisalduvate suhete ja kategooriate abil. Näiteks pole olemas universaalset subjekti kategooriat, vaid intransitiivses konstruktsioonis on intransitiivse konstruktsiooni subjekt (Croft 2001: 7–62). Ehk siis lingvistilised kategooriad pole ainult keelepõhised, vaid keele- ja konstruktsioonipõhised.

Kognitiivse KG tuntum esindaja on Adele Goldberg (1995; 2006), kes on uurinud erinevaid argumentstruktuuri konstruktsioone ja neis esinevate verbide semantikat. Kognitiivse KG suuremaks erinevuseks teistest KG lähenemistest peetakse keskendumist konstruktsioonide seletamisel keelekasutaja kognitiivsele kogemusele; konstruktsioonide esitamist ja kirjeldamist püütakse võimalikult vähe formaliseerida (Boas 2013: 248–249).

Kõige üldisemad KG pooltargumendid on, et mudel on ökonoomne, sest see ei tekita erinevaid grammatikakomponente, mille sidumine hilisemas analüüsis ja kirjelduses võib muutuda probleemiks; mudel ei omista mitte ühelegi grammatikakomponendile eristaatust ega määra keelenähtusi ebaregulaarseteks või erandlikeks. Seega on KG-l potentsiaali moodustada ühtne keelekirjeldussüsteem, mis suudab kirjeldada kogu grammatilist süsteemi; KG tundub olevat kooskõlas sellega, mida me teame inimese kognitiivsetest kogemustest ja käitumisest. (Fried & Östman 2004: 18)

Fillmore (1989: 19–20) kummutab KG-le sageli tehtavad etteheited, et tegeletakse n-ö perifeersete ja keeruliste keelenähtustega (nt idioomid), kuid ignoreeritakse lihtsamini kirjeldatavaid nähtusi. Fillmore'i järgi see tänapäeval

kindlasti enam nii ei ole ja KG moodustab grammatika kõige üldisemate ja spetsiifilisemate nähtuste vahelise kontiinumi ning KG eesmärk on kirjeldada kõiki selle kontiinumi osi ühtses raamistikus.

KG seisukoht on, et süntaksi ja leksikoni vahelist piiri ei eksisteeri või seda on väga raske määrata. Semantiline struktuur on tihedalt lõimunud grammatiliste konstruktsioonidega. Semantiline motivatsioon tingib sageli teatava grammatilise konstruktsiooni kasutamise (Fillmore 1989: 34). Tabel 1 esitab Crofti ja Cruse'i järgi süntaksi-leksikoni kontiinumi, mida olen kohandanud eestikeelse näitematerjali põhjal.

Tabel 1. Süntaksi-leksikoni kontiinum (kohandatud Crofti ja Cruse'i (2004: 255) järgi).

Konstruktsiooni tüüp	Traditsiooniline nimetus	Näited
Kompleksne ja skemaatiline	Süntaks	[modaalverb + <i>ma</i> -infinitiiv]
Kompleksne koos täistähendusliku verbiga	Subkategorisatsiooni printsiip	[OBJ <i>huvitama</i> SUBJ]
Kompleksne ja täistähenduslik	Idioom	kärbseid pähe ajama
Kompleksne, kuid seotud	Morfoloogia	[nimisõna+d], [verb+ajakatekooria]
Atomaarne ja skemaatiline	Süntaktiline kategooria	[adverb], [nimisõna]
Atomaarne ja täistähenduslik	Sõna/leksikon	[ilusti], [keel]

Käesoleva töö kontekstis on olulised konstruktsioonigrammatika kirjeldusvõimalused, mis kaasavad keelevälised tegurid (regionaalsus, sotsiolingvistilised faktorid jmt), pidades neid konstruktsiooni kirjeldamise seisukohalt relevantseteks, kuid samas ei eelda raamistik iga võimaliku detaili esitamist, kui puudub vastav info (Fried 2008). KG kirjeldusvõimalusi eesti keele näitel on põhjalikult analüüsinud Sakhai (2011).

Matras (2004: 282, 299) toob välja, et keele kirjeldamisel kombineeritakse praktikas sageli kahte strateegiat: vormilt funktsioonile ja funktsioonilt vormile, ja ta väidab, et näiteks tüpologia uurimisvahendite kaasamine on funktsiooni uurimise kaasamine murde-uurimisse. Seda analoogiat võiks käesoleva töö autori arvates kasutada konstruktsioonigrammatika ja dialektoloogia ühendamisel.

3.1.1 Konstruktsioonipõhise lähenemise rakendamine käesolevas töös

Väitekirja eesmärk ei ole KG teoreetiliste aspektide analüüs, vaid rakendada teooriat kirjeldusvahendina. See on antud töö materjali silmas pidades sobivaim tegutsemisviis, kuna võimaldab käsitleda nii morfoloogilisi, süntaktilisi kui ka semantilisi muutujaid samas raamistikus ühtse süsteemina. Konstruktsioonipõhine lähenemine on olnud aluseks ka kahele üsna mahukale eestikeelsele uurimusele (Penjam 2008; Sahlkai 2011), kus on samuti käsitletud siinse töö uurimisobjektiga sarnaseid konstruktsioone kirjakeeles. Sahlkai (2011: 28–29) on esitlenud KG lähenemisvõimalust abiverbikonstruktsioonide kirjeldamisel ning osutanud sellise lähenemisviisi eelistele teiste teoreetiliste raamistike ees. Samas toob Sahlkai (2011: 32) välja ühe olulise KG tahu, mis on eriti kriitiline käesoleva töö kontekstis, nimelt suudab KG toime tulla uute, senisesse süsteemi mittesobivate näidete ilmnemisega, võimaldades olemasolevate kategooriate süsteemi muutmist. See teoreetiline aspekt on murdematerjali käsitlemisel eriti oluline, kuna sisuliselt analüüsin kümnet eri keelt, ning pole põhjust eeldada, et kõik materjalis esile tulevad kategooriad on kõigis murretes olemas ja/või käituvad kõigis murretes sarnaselt. Seetõttu on teatav teoreetiline pindlikkus murrete kirjeldamisel hädavajalik.

Käesoleva töö materjali analüüsi on suuresti toetanud Rätsepa (1978) monograafia eesti keele lihtlausetüüpidest, mis on üsna sarnane tänapäevase konstruktsioonipõhise lähenemisega. Siiski eristab mainitud tööd KG-st üks oluline aspekt, nimelt pole lausetüüpide rühmitamisel arvesse võetud semantikat. Teise siinse töö materjali käsitlemist toetava konstruktsioonigrammatika vaimus kirjutatud eesti süntaksikäsitluse autor Tauli (1980) keskendub rohkem ka tähenduste analüüsimisele. Eesti kirjakeele verbikonstruktsioone käsitlevas KG-põhises monograafias (Penjam 2008) on konstruktsioonide liigitamisel võetud aluseks konstruktsiooni moodustavate elementide süntaktiline funktsioon, mis põhineb eesti keele akadeemilisel grammatikal (Erelt *et al.* 1993). Selline tegutsemisviis oleks käesoleva töö kontekstis hetkel veel teostamatu. Süntaktilistest suhtest lähtuvas analüüsil tuleks tõmmata selge piir teatud määral grammatiseerunud ühendite vahele, kus nii põhiverb kui infiniitne verbivorm kuuluvad predikaati (ahelverbid, perifrastilised verbid), ning verbide vahele, kus infiniitse verbi kasutus on tingitud finiiitse verbi valentsist (tomib süntaktiliselt subjekti või objektina) (vt ka Erelt 2013: 51–55). Käesoleva töö valguses on selline lähenemine problemaatiline, kuna puudub info erinevate verbide grammatiseerumisastme kohta murretes. Grammatiseerumise ühe indikaatorina nähakse selles töös sagedust (sagedasemad verbiühendid on suurema tõenäosusega teatud määral grammatiseerunud), aga ka verbide grammatiliste kasutusfunktsioonide lisandumist. Kuna grammatiseerumine pole diskreetne vaid skalaarne nähtus, siis selline liigitamine on problemaatiline juba ka ühe keele uurimise kontekstis.

Mainitud probleemidest tulenevalt on siinses töös püütud vältida selgete piiride tõmbamist erinevate lingvistiliste kategooriata ja tasandite vahele,

millest tulenevalt võib rakendatavat lähenemisviisi pidada Rätsepa, Tauli ja Penjami lähenemiste kombineerimiseks. Sarnaselt Rätsepa ja Tauliga võtan materjali analüüsimisel aluseks vormilise kriteeriumi ning lauseliikmelisust konstruktsioonide tähenduse määramisel ei arvesta. Süntaktiline funktsioon osutub aeg-ajalt oluliseks vormipaaride tähenduste vaatlemise etapis ning teatud konstruktsioonide puhul võib ilmneda, et mingit tähendust kandvas konstruktsioonis võib teatud positsioonis esineda vaid kindla süntaktilise funktsiooniga element.

Autori arvates on vormipõhine lähenemine siinse töö kontekstis õigustatud vähemalt kahel erineval põhjusel. Kuna töö uurimismaterjal on äärmiselt heterogeenne ja ühtsete kirjeldusvõimaluste leidmise seisukohalt eriti problemaatiline, siis on konstruktsioonide liigitamiseks süntaktilisest tunduvalt selgem alus vormiline kriteerium. Süntaktilise kriteeriumi rakendamise teeks keeruliseks kui mitte võimatuks murrete süntaktiliste teadmiste puudulikkus. Teiseks kuna töö kvalitatiivsema iseloomuga konstruktsioonide analüüsi osale eelneb mahukas kvantitatiivne osa, kus materjali kogumise ja töötlemise seisukohalt on keskne morfoloogiline kriteerium, siis oleks raske põhjendada mõne vormipaari materjalist väljajätmist puhtalt süntaktilise funktsiooni tõttu. Kuna tegemist on täielikult kasutuspõhise tööga, siis konstruktsioone kirjeldatakse ainult olemasoleva materjali põhjal. Kõigi vormipaaride hoidmine analüüsis ja nende kirjeldamine võimaldab heterogeenset materjali paremini kontrolli all hoida ja analüüsitulemused on vähem mõjutatud analüüsija suvast.

3.1.2 Konstruktsiooni mõiste

Grammatika on terviklike mustrite ehk konstruktsioonide kogum ja igal sellisel mustril on oma funktsioon tähenduslike lausungite produtseerimisel (Fillmore 1989: 18). KG raamistikus on keele põhiline ehitusosa konstruktsioon. Grammatilise konstruktsiooni all mõeldakse mis tahes süntaktilist mustrit, mis esindab üht või mitut keeles konventsionaliseerunud funktsiooni (Fillmore 1988: 36). Fried ja Östman (2004: 18) esitavad omapoolselt sarnase definitsiooni, mille kohaselt konstruktsioon on abstraktne, mingit suuremat elementide hulka esindav üksus, lingvistilise struktuuri konventsionaaliseerunud skemaatiline muster, mis võimaldab moodustada korrektseid lingvistilisi väljendeid. Suhtlusesituatsioonis esile tulevad konkreetseid väljendid (laused, fraasid, tekstid) ei ole aga konstruktsioonid, vaid konstruktsiooni esindajad, esiletuleku juhud, ilmnemised jne ehk konstruktid. Seega on konstruktsioon mingi lingvistilise väljendusvahendi abstraktne skeem ning konstrukst selle skeemi rakendumine ja ilmnemine loomulikus suhtlusesituatsioonis. Fried (ilmumas) lisab, et alguses regulaarsena näiv, läbipaistev süntaktiline struktuur võib sisaldada ennustamatuid kitsendusi, mis ei tulene ainult süntaksist ja kuna konstruktsioonid on n-ö märgid, pole nad mitte kunagi kompositsioonilised. Ka semantiliselt läbipaistmatutel väljenditel (idioomidel) võib olla süntaktiliselt regulaarne ja produktiivne struktuur (Fillmore *et al.* 1988).

See töö käsitleb konstruktsioone, mille moodustavad finiiitne ja infiniitne verb. Sellised ühendid võivad kanda erinevaid tähendusi, s.t moodustada erinevaid konstruktsioone. Finiitse modaalverbi *saama* abil võib moodustada *da*-infinitiivis verbiga muu hulgas näiteks modaalkonstruktsioone (nt *sain lõpuks koju minna*), impersonaalkonstruktsioone (nt *lapsena sai palju ringi joosta*), passiivikonstruktsioone (nt *saime ema käest riielda*) või mentaalseid konstruktsioone (nt *saime lõpuks linna näha*). Kõigi nende konstruktsioonide ilmnemised keelest (nagu esitatud näited) on konstruktid. Seega on konstruktsioon mingi vormi ja tähenduse abstraktsioon ning selle abstraktsiooni konkreetsed esinemisjuhud keeles on konstruktid. Esitatud näidetest võib näha, et konstruktsioonidevahelised piirid pole sugugi nii selged ja erinevad tähendused võivad ilmneda samas konstruktsioonis. Seda tuleb loomulikus keeles sageli ette ning siinses töös käsitletud konstruktsioonide puhul saab harva väga rangeid konstruktsioonidevahelisi piire tõmmata, pigem on tegemist kontiinumiga, kus ühes konstruktsioonis võib üks tähendus olla esildumam kui teine.

Konstruktsiooni kirjeldus võib sisaldada Friedi ja Östmani (2004: 19–20) järgi järgmist informatsiooni.

- Morfosüntaktiline informatsioon (moodustajate struktuursed suhted, konstruktsiooni elementide järjekord, morfoloogiline informatsioon).
- Prosoodiline ja foneetiline informatsioon.
- Info tähenduse või funktsiooni kohta (nt agent, patsient, koht).
- Kontekstuaalne informatsioon.

Konstruktsioonigrammatika kasutab konstruktsioonide kirjeldamisel atribuudi-väärtuste kastdiagramme (Kay & Fillmore 1999; Fillmore 1988), mida siinses töös hetkel veel üsna puudulike murdesüntaktiliste teadmiste tõttu ei rakendata.

3.2 Kasutuspõhine keeleuurimine

„... the fact that such Usage-Based approaches to language change have refrained from engaging in large-scale dialectological research is perhaps even more surprising than the neglect from the generative side.” (Vogelaar & Seiler 2012: 11)

Kasutuspõhise keeleuurimise mõiste võttis esmakordselt kasutusele Langacker (1987) ning see lähenemisviis on leidnud keeleteaduses viimasel ajal palju rakendust (tuntumad tööd Barlow & Kemmer 2000; Bybee 2007; 2010). Kasutuspõhine keeleuurimine lähtub põhimõttest, et keeleteadlase esmane uurimisobjekt peaks olema keel, mida keelekasutajad ise loovad ja mõistavad. Keelekasutuse uurimine on parim tõendusmaterjal, mille põhjalt teha järeldusi keelesüsteemi kohta. Ideaalne kasutuspõhine analüüs toetub keelekorpusetele kui tegeliku keelekasutuse avaldumiskogule. (Kemmer & Barlow 2000: xv).

Kasutuspõhise keeleuurimise olulisuse rõhutamine murdeuurimises tundub mõnevõrra üleliigsena, kuna dialektoloogia on läbi aegade olnudki kasutuspõhine keeleuurimine selle kõige konkreetsemas tähenduses. Dialektoloogias on kasutuspõhine lähenemine olnud iseenesest mõistetav, dialektoloogid on alati

toetunud tegelikule keelekasutusele selle loomulikus kontekstis. Kasutuspõhise lähenemise suurele potentsiaalile murdeuurimisse integreerimisel on osutatud varemgi (vt nt Vogelaer & Seiler 2012: 11). Kasutuspõhisus on siinse töö autori arvates ainumõeldav viis keele kirjeldamiseks ja keelesüsteemi mõistmiseks. Kuna töö üks peamistest eesmärkidest on dialektoloogia sidumine tänapäevaste üldkeeleteaduslike suundade ja kirjeldusvahenditega, siis järgnevalt annan lühikese ülevaate kasutuspõhise mudeli põhijoontest tänapäeva keeleuurimise kontekstis.

Kasutuspõhise mudeli idee kohaselt määrab loomulik keelekasutus grammatiliste struktuuride ja nende esituse kõneleja lingvistilises tajusüsteemis. Kemmer ja Barlow (2000) loetlevad kasutuspõhise mudeli olulisemad omadused:

- rõhutatakse kasutussageduse olulisust, kuna sagedasematel keelenähtustel on omadus laieneda, need muutuvad teatavaks kognitiivseks rutiiniks;
- keele mõistmine ja töötlemine on lingvistilises süsteemis kesksel kohal;
- keele mõistmise ja omandamise kirjeldamine on olulised;
- keel tekib ja kujuneb kasutuses, mitte ei ole kindlate üksustena salvestatud kõneleja ajju;
- keele uurimisel ja kirjeldamisel on oluline toetumine tegelikule keelekasutusmaterjalile, tegelik keelekasutus on parim tõestusmaterjal keelesüsteemi kirjeldamise ja keelesüsteemi kohta käivate väidete tõestamisel;
- keelekasutus, sünkrooniline varieerumine ja diakrooniline muutumine on tugevasti seotud;
- keelesüsteem ja mittekeeleline kognitiivsüsteem on omavahel seotud;
- keeleväline kontekst on keelekasutuses oluline.

Järgnevalt toon esitatud põhimõtetest välja selle töö seisukohalt olulisemad ja kirjeldan lühidalt, kuidas need väited suhestuvad siinse tööga. Selle töö puhul on kesksel kohal mõiste sagedus. Töö konstruktsioonide analüüsi osa (ptk 9) käsitleb ainult sagedasi verbiühendeid, mis omakorda võivad moodustada erinevaid konstruktsioone. Töö autor on veendunud, et sagedusel on keelekasutussüsteemis oluline osa ja see on üks paljudest faktoritest, mis annab keele kohta informatsiooni. Väidan, et sageduslik geograafiline varieerumine on tähenduslik ning see kirjeldab vaadeldavat keelt, antud juhul murret (täpsemalt peatükis 7). Kui mingi ühend on teatud murdes eriti sage ning teistes murretes sageduse poolest marginaalne, siis võib väita, et selline konstruktsiooniline (konstruktiline) omadus on selle murde eripära. Tõenäoliselt on sama tähenduse edasiandmiseks teistes murretes kasutusel kas mingi muu konstruktsioon või grammatiline/süntaktiline võimalus. Samuti on tähenduslik, kui palju erinevaid konstruktsioone teatud ühendi abil on võimalik moodustada. Materjali kvalitatiivses analüüsis osas võib näha, et see on murede üsnagi erinev.

Selle töö aluseks on materjal, mis on tekkinud ja ilmneb reaalses keelekasutuses. Uurimismaterjal sisaldab nähtusi, mida kirjakeeles ei pruugi olla ning mille sagedus võib ka eri murretes olla marginaalne. Siiski on see reaalne keelekasutus ning kasutuspõhise eelduse kohaselt on igal kõnelejal olnud

vaadeldava väljendi loomiseks motivatsioon, mis tõenäoliselt ilmnes keelekasutussituatsioonis. Sellised harvad, kuid enamasti huvitavad näited võivad olla heaks tõestusmaterjaliks ka kirjakeele vastavate nähtuste kirjeldamisel. Käesolevas töös ei kirjeldata siiski väga harvade nähtuste potentsiaalset kasutusmotivatsiooni.

Olulisim ülalpool väljatoodud punktidest on tegeliku keelekasutusmaterjali kirjeldamine. Töö on täielikult kasutuspõhine ning pürib keelt kirjeldama ainult materjali põhjal. Eesmärgiks on kirjeldada materjali ilma sellele kitsamaid teoreetilisi piiranguid loomata. Kuigi seda pole võimalik täielikult vältida, on püütud piirduda minimaalsega, mis materjali kirjeldamiseks vajalik.

Siinne töö kirjeldab murdeid küll vaid sünkroonilisest vaatepunktist, kuid sellest hoolimata on töö autor veendunud, et keelekasutus, sünkrooniline varieerumine ja diakrooniline muutumine on tihedalt seotud. Keelesüsteem ise loob piirangud ja võimalused teatud muutumistendentside tekkimiseks, mille analüüsimist kahjuks sinse töö materjal ei võimalda. Regionaalset varieerumist võib teatud määral käsitleda keele diakroonilise varieerumisena, sest eeldatavasti on erinevad murded keelemuutumise erinevates etappides. Murdeid mõjutavad ka kontaktkeeled, mis geograafilisest asukohast ja ajaloolistest põhjustest võivad olla erinevad ja erineva intensiivsusestmega.

Keeleväline kontekst on kahtlemata keelekasutuses oluline, selles töös ilmneb see analüüsimaterjalist, mis on traditsiooniline murdeintervjuu. Selline kindla kasutussituatsiooniga piiratud žanr loob kindlasti eeldused ühtede keelenähtuste ilmnemiseks ning teiste mitteilmnemiseks.

Üks põhjustest, miks lähtun konstruktsioonigrammatikast, on konstruktsioonigrammatika valdavalt kasutuspõhine lähenemisviis keeleuurimisele. Barlow ja Kemmer (2000: xv) nendivad, et tegeliku keelekasutuse andmestiku uurimine on olnud tüüpiline keeleteaduses perifeerseteks peetavates valdkondades, nagu foneetika, sotsiolingvistika ja ajalooline keeleteadus, mille mõju keeleteaduslike teooriate kujunemisele on olnud siiani väike. Sellesse loetellu võib kindlasti lisada dialektoloogia. Kemmer ja Barlow (2000: xv) märgivad samas artiklis tabavalt, et süntaksi on pikka aega peetud keelesüsteemi keskseks osaks, kuid pikka aega on süntaksialased uurimused toetunud ainult konstrueeritud näidete analüüsile ja hindamisele.

3.3 Sagedus

„... frequency effects are everywhere.” (Bod *et al.* 2003: 10)

Keelenähtuste sageduse roll on tänapäeva keeleteaduses küllaltki erinevaid arvamusi tekitav (vt põhjalikumat diskussiooni nt Arppe *et al.* 2010) ja viimasel ajal diskuteeritakse selle üle järjest rohkem. Üha enam ollakse veendunud, et sõnade, konstruktsioonide jne kasutussagedusel on oluline osa keele nähtuste ilmnemisel ja muutumisel (nt Divjak & Gries 2012; Gries & Divjak 2012). Eriti

rõhutavad sageduse olulisust grammatilise süsteemi kujunemisel kasutus-põhised keelekäsitlused (Barlow & Kemmer 2000; Bybee 2010; Bybee & Hopper 2001; Croft & Cruse 2004), samas kui mõned teised teooriad sagedusele keelekirjeldamisel ja -kujunemisel erilist tähtsust ei omista (eriti generatiivsed teooriad (nt Chomsky 1965)).

Sagedust ei peeta oluliseks ainult kasutuspõhistes teooriates, vaid ka keele varieerumise seletamisel. Bod jt (2003: 3) rõhutavad, et teadmised varieerumise kohta peavad sisaldama infot sageduse kohta ning nende väitel ei saa sagedust muust lingvistilisest informatsioonist keele varieerumise kirjeldamisel kuidagi vähemoluliseks pidada. Sarnast seisukohta toetab Szmrecsanyi (ilmumas (c)), kes lisab, et sagedus peegeldab oluliselt täpsemalt keelelist reaalsust kui dialektoloogias traditsiooniliselt kasutatav atlasandmestik. Kuigi selline sageduspõhine lähenemine ei võimalda uurida harvaesinevaid keelenähtusi ja automaatselt kogutud sagedusandmestikud ei sisalda alati kaugeltki mitte ainult n-õ puhast informatsiooni, ei pea Szmrecsanyi (ilmumas (c)) neid fakte kobaruuringutes probleemiks, sest esiteks pole harvaesinevad keelenähtused siin huvi keskpunktis, teiseks on sagedusmüra osa lingvistilisest reaalsusest, mis peabki analüüsis kajastuma. Eesti murrete lauseehituse uurimisel on enamasti keskendutud eripärastele ja suhteliselt harva esinevatele üksustele; sagedaste ja n-õ tavaliste nähtuste varieerumisest on siiani info lünklik. Esimest lähenemisviisi on Szmrecsanyi (ilmumas (b)) nimetanud n-õ vana kooli dialektoloogiaks, mille põhiline moto oli „mida murdelisem ja ebastandardsem keelenähtus, seda parem”, tänapäevaseks murdeuurimiseks peab ta korpuspõhist ja grammatikale keskenduvat dialektoloogiat.

Kasutuspõhistes teooriates eristatakse enamasti tüübisagedust (*type frequency*) ja sõnesagedust (*token frequency*). Sõnesagedus vastab vormi, sõna, konstruktsiooni kasutusjuhtude arvule, tüübisagedus osutab skemaatilisele nähtuse (konstruktsioonimall, kindel vormitunnus) kasutusulatusse. Sõnesagedus peegeldab kindla üksuse juurdumisastet (*entrenchment*) keeles; tüübisagedust arvatakse viitavat malli produktiivsusele. (Bybee 2007: 9; Croft & Cruse 2004: 309) Konstruktsiooni produktiivsuse määramisel on oluline, kui paljudele üksustele mingi muster või konstruktsioon rakendub. Mida rohkem leksikaalseid üksusi konstruktsiooni teatud positsioonis võib ilmuda, seda tõenäolisem on, et see konstruktsioon võib laieneda uutele sõnadele või vormidele ehk on produktiivsem. Kuna tüübisagedus võib varieeruda ühest kuni väga kõrge sageduseni, võib konstruktsioonide produktiivsuse aste olla väga erinev. (Bybee 2007: 275)

Bybee (2007: 10–22) on omistanud sõnesagedusele säilitava ja taandava mõju. Säilitav mõju Bybee järgi sõltub asjaolust, et kordus tugevdab keeleüksuste esinemisvormi mälus ja teeb need kergemini ligipääsetavaks. Ligipääsetavus tähendab siin seda, et inimesed tunnevad ära tunduvalt kiiremini kõrge sagedusega keeleüksusi kui madalama sagedusega sõnu, häälikuid, vorme jne. Sagedaste keeleüksuste juurdumine mälus seletab, miks sagedasemad keelenähtused muutuvad aeglasemalt ja pole nii vastuvõtlikud analoogiapõhistele muutustele, samas kui harvem esinevad keelenähtused on muutustele

vastuvõtlikumad. Sõnesageduse säilitava mõju teine väljendusvorm on sagedaste vormide toimimine analoogiamustritena, mille järgi uusi keelenähtusi mugandatakse. Taandaval mõjul on Bybee järgi foneetilised, süntaktilised ja semantilised mõõtmed. On üldtuntud fakt, et sagedasti kasutatavad sõnad, fraasid jne kalduvad häälikuliselt lühenema. Lühenemine võib toimuda ka häälikulisest muutusest komplekssemal tasandil (nt süntaktilisel). Bybee nendib säilitava ja taandava mõju näilist vastuolu: ühelt poolt väidetakse, et sagedus mõjutab keeleüksuste ümberkujunemist, ning teisalt, et sagedus muudab keeleüksuse teatavatele muutustele vastuvõtlikumaks. Selle vastuolu lükkab Bybee aga samas ümber. Keelemuutus tähendab väga paljusid eri liiki muutusi. Keeleüksuse juurdu mine mälus teeb keerulised konstruktsioonid vastupidavamaks analoogiamuutustele ja ümbervormumisele. Mingite sõnade, lausete jne pidev kordamine teeb need tema järgi vastuvõtlikumaks foneetilistele ja semantilistele muutustele. (2007: 10–22; 270)

Sõne- ja tüübisageduse mõisted on selles töös kesksel kohal. Analüüsimaterjal on kogutud tüübisageduse põhjal³, s.t kogutud on poolskemaatilisi ühendeid, mis esinevad samas osalauses (infiniitse verbivormi kategooria + finiiitne verb). Sõnesagedus on aga mingi konkreetse ühendi ilm nemine, näiteks mitu korda esineb konstrukt *minema pidama* (*pean minema*, *pidi minema*, *peame minema* jne), ehk siis konkreetne finiiitne verb koos teise konkreetse infiniitse verbiga. Sõnasagedus on selles töös sünonüümne konstrukti sagedusega⁴ ja sellele keskendun rohkem peatükis 9, kus sõnesagedust rakendatakse konstruktsioonide produktiivsuse mõõtmisel. Mida rohkem erinevaid infiniitseid verbe vaadeldavas konstruktsioonis võib teatud finiiitverbiga koos esineda, seda kõrgem on selle konstruktsiooni produktiivsus. Süntaktilisest produktiivsusest üldiselt tuleb juttu järgmises peatükis ning produktiivsusest vaadeldakse iga konstruktsiooni puhul eraldi töö konstruktsioonide analüüsi peatükis.

Kuigi tüübi- ja sõnesageduse vahelist suhet on peetud heaks produktiivsuse indikaatoriks, siis otseselt pole seda käesolevas töös mõõdetud. Tähelepanu keskpunktis on siin sõnesagedus. Tüübisagedust vaadatakse kaudselt produktiivsuse arvutamisel (täpsemalt selle kohta vt peatükke 6.6 ja 9). Ehk siis mida madalam on mingi ühendi statistilise seose tugevuse väärtus, seda kõrgemaks võib pidada ühendi tüübisagedust, s.t konkreetne finiiitne verb kombineerub suurema arvu erinevate verbidega, mis esinevad infiniitses positsioonis.

Kindlasti ei saa väita, et sagedus ise ongi selgitus. Szendrői (ilmumas (c)) illustreerib inglise keele genitiivkonstruktsioonide näitel, kuidas sagedus ei pruugi viidata mitte ainult muutuvale grammatikale, vaid selle võib tingida ka tekstiteema, -liik jne (Biber 1995; Biber & Conrad 2009). See on korpuslingvistikas tuntud probleem, millest sageli mööda vaadatakse. Käesolev töö on sünkrooniline uurimus murretest ja seetõttu saan sageduse muutusi selgitada

³ Materjali kogumise protsessi on kirjeldatud peatükis 5.

⁴ Käesolevas töös kasutan mõistet *sõnesagedus*, mis on teatav korpuslingvistiline konventsioon. Mõtlen sellega siiski konstruktisagedust, kui pole mainitud teisiti.

siin geograafia ehk eri murrete kaudu. Korpuse tekstid ja temaatika on suhteliselt homogeensed (täpsemalt vt ptk 4).

3.4 Süntaktiline produktiivsus

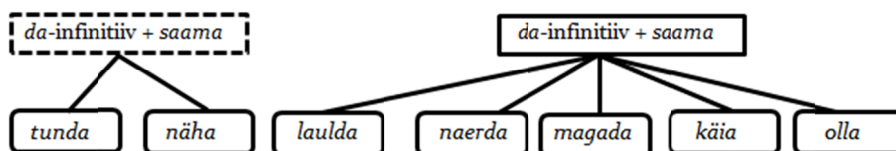
Mälus säiliv keel koosneb suurematest erineva suurusega üksustest (sõnadest fraaside ja osalausesteni), mis on erineva juurdumisastmega, erineva produktiivsuse ja seostega teiste keeleüksuste suhtes (Bybee 2007: 281). Morfoloogiline produktiivsus on siiani leidnud laialdast käsitlemist ning selle mõõtmiseks on arendatud palju arvutuslikke meetodeid (Baayen 2009; Baayen 1994; Bauer 2001; Keune *et al.* 2005). Süntaktiline produktiivsus on viimasel ajal küll hakanud rohkem tähelepanu saama, kuid selle täpne definitsioon on siiani ebaselge. Barðdal (2008) on esitanud põhjaliku ülevaate sellest, mida keeleteaduslikus kirjanduses on produktiivsuse all käsitletud, ning on leidnud 19 erinevat lähenemist (2008: 10–11). See kinnitab, et tegemist pole kaugeltki mitte üheselt mõistetava nähtusega, mille suhtes kõik keeleteadlased oleksid samal meelel.

Kasutuspõhine raamistik defineerib produktiivsuse astmelise nähtusena, sest tüübisagedus on astmeline, seega kasutuspõhine mudel prognoosib, et produktiivsus võib ulatuselt varieeruda (Croft & Cruse 2004: 297). Konstruktsiooni produktiivsust nähakse tüüpiliselt sõne- ja tüübisageduse suhtena, nii on selle esitanud ka Croft ja Cruse (2004: 307). Barðal (2008: 29–33) defineerib süntaktilise produktiivsuse selle kaudu, kuidas keelekasutaja on võimeline genereerima lauseid, mida ta varem kuulnud pole, ning kui suures ulatuses uued verbid võivad vaadeldavas süntaktilises konstruktsioonis ilmned.

Käesolevas töös käsitlen produktiivsust just selles teises mõttes ehk pean produktiivseks finiiitse verbi ja infiniitse vormi ühendit, kus infiniitses positsioonis esineb rohkem erinevaid verbe. Seda mõõdan küll kaudselt kahe üksuse statistilise seose abil, seades eelduseks, et kui infiniitne verb pole finiiitse verbiga väga tugevalt seotud, on see ühend vähem kinnistunud ning võimaldab kombineeruda rohkema arvu erinevate infiniitsete verbidega. Täpsemalt tuleb produktiivsuse mõõtmisest juttu peatükis 6.6.

Töö autor on veendunud, et produktiivsus on kindlasti astmeline ning seetõttu pole võimalik tõmmata selget piiri produktiivsete ja ebaproductiivsete konstruktsioonide vahele, vaid võib rääkida ainult vähem või rohkem produktiivsetest konstruktsioonidest. Mida püsiühendilaadsem konstruktsioon on, seda vähem produktiivne ta on. Näiteks konstrukt *hakkama saama* tähendusega 'toime tulema' on selgelt ebaproductiivne, kuna tegemist on kinnistunud ühendiga ühes konkreetsetes tähenduses. Joonis 1 illustreerib kõrge ja madala sõnesagedusega keeleüksuse esinemist, kus ümarate nurkadega kastid tähistavad konkreetset kasutusjuhtu (konstrukti) ning kandiliste nurkadega kastid konstruktsiooni. Punktiiriga riskülik tähistab madalama juurdumisastmega keeleüksust. Vasakpoolne, vähemproduktiivne konstruktsioon tähistab tajuverbi-konstruktsiooni, mis moodustub *da*-infinitiivi ja finiiitse *saama*-verbiga ning see

konstruktsioon on vähem produktiivne, kuna infiniitse verbi positsioonis saab esineda vaid teatud hulk kindla tähendusega verbe. Parempoolne konstruktsioon joonisel tähistab aga kõrge produktiivsusestmega modaalse verbi konstruktsiooni, kus infiniitse verbina võib põhimõtteliselt esineda mis tahes verb. Nendest konstruktsioonidest tuleb detailsemalt juttu peatükis 9 ning joonis on siin ainult illustreeriva funktsiooniga produktiivsuse mõiste selgitamiseks.



Joonis 1. Madal ja kõrge produktiivsus (kohandatud Croft ja Cruse'i (2004: 309) järgi).

3.5 (In)finiitsus

Infiniitsus ei ole tüpoloogiliselt üheselt mõistetav nähtus. Traditsiooniliselt jagatakse verbivormid finiiitseteks ja infiniitseteks ning nende eristamise traditsioon pärineb klassikalise kreeka ja ladina keele kirjeldamisest ning põhineb verbi-morfoloogial. Finiiitsetes vormid markeerivad reeglina aega, kõneviisi, aspekti, isikut ja arvu ning nad võivad esineda lauses või osalauses üksinda predikaadina, samas kui infiniitsetes vormid sellis el kujul iseseisvana esineda ei saa. (Ylikoski 2003) Infiniitsetes vormid, mis traditsiooniliselt jagatakse nelja põhitüüpi: infiniitivid, partitsiibid, konverbid ja nominalisatsioon, ei saa tavaliselt lauses olla iseseisvaks öeldiseks, vaid kannavad lauses mingit muud funktsiooni.

Koptjevskaja-Tammi järgi (1999) on infiniitsetel vormidel maailma keeltes palju erinevaid funktsioone, mille võib üldiselt jagada kahte gruppi. Ühelt poolt sekundaartarindi funktsioon pikemas lauses ning teiselt poolt võivad infiniitvormid grammatiseeruda ja muutuda perifrastilise verbi üheks osaks. Sellised verbid võivad väljendada aega, aspekti, kõneviisi, eitust jne. Infiniitsete vormide muutumist analüütiliste verbivormide osaks võib täheldada ka siinse töö fookuses olevate verbiühendite puhul.

Finiitsust on käsitletud nii diskreetse nähtusena (formaalsed keeleteooriad) kui ka skalaarse nähtusena (funktsionaalsed keeleteooriad). Funktsionaal-tüpoloogiliste uurimuste tulemused näitavad, et finiiitsust ei saa defineerida üksikute universaalsete morfoloogiliste tunnuste või parameetrite põhjal, vaid pigem parameetrite rühmade abil. (Cristofaro 2007; Bisang 2007)

Skalaarsena nähakse sageli ka infiniitsete konstruktsioone. Konstruktsioon ei saa olla finiiitne või infiniitne, vaid võib olla rohkem või vähem finiiitne, olenevalt sellest, kui palju prototüüpilisest pealause mustrist kõrvalekalduvaid jooni ta sisaldab. Erinevate keelte jaoks tuleb finiiitsuse määramisel kasutada erinevaid kriteeriume. Pole mingit põhjust arvata, et eri keeltes on mingi infiniitsete

vormi puhul tegemist täpselt sama kategooriaga. Finiitsust tuleks vaadata keeltevahelise tendentsina, kus teatud parameetrid üksteisega korreleeruvad. (Nikolaeva 2010) See väide on kooskõlas ideega, mille järgi grammatilised kategooriad on konstruktsiooni- ja keelepõhised ning grammatilised universaalid avalduvad vormi ja tähenduse ühenduspõhimõtetes (näiteks Croft 2001). Sellised väited on kooskõlas nii siinse töö põhimõtete kui ka valitud teoreetilise raamistikuga. Esiteks sobib see hästi konstruktsioonigrammatika eeldustega ning teisalt pole mingit alust eeldada, et kõigis murretes käituvad infiniitsed vormid täpselt samamoodi ning on keele arengu seisukohalt samas etapis. Siinne töö lähtub konstruktsioonikandidaatide ekstraheerimisel küll ainult vormikriteeriumist, kuid edasises analüüsis ei eeldata, et kõik infiniitsed vormid kõigis murretes samu funktsioone täidavad.

Tüpoloogiliselt on püütud erinevaid infiniitsete vorme sisaldavaid konstruktsioone liigitada. Käesolevas töös käsitletavad infiniitset verbi sisaldavad konstruktsioonid kuuluvad rühma, mille Maas (2004) on infiniitsete konstruktsioonide klassifitseerimisel nimetanud abiverbikonstruktsioonideks, mis koosnevad finiiitset verbivormist ja lause põhitähendust kandavast infiniitset vormist. Analooget liigitust järgib ka Anderson (2006), kes nimetab sarnase struktuuriga konstruktsioone abiverbikeskseteks abiverbikonstruktsioonideks. Siinne töö keskendub just sellistele infiniitset vormi sisaldavatele konstruktsioonidele. Täpsemalt tuleb mõistest abiverbikonstruktsioon juttu peatükis 3.7.

3.6 Eesti keele infiniitsed vormid

Soome-ugri keeltele, k.a eesti keelele on tüüpiline infiniitsete vormide rohkus. Akadeemilise eesti keele grammatika järgi (Erelt *et al.* 1993) on eesti keeles järgmised infiniitsed vormid: 1) substantiivile lähedased tegevusnimed (*da-*, *ma-*, *vat*-infinitiivid). *ma*-infinitiivile võib liituda veel käändelõppe; 2) adjektiividele lähedased oleviku ja mineviku partitsiibid; 3) adverbile lähedane gerundiiv. Üldiselt kasutatakse infiniitsete vorme keerulisemate perifrastiliste konstruktsioonide moodustamiseks, mis seejärel samm-sammult grammatiseeruvad, integreeruvad verbi paradigmasse ja võivad muutuda sünteetilisteks vormideks (Haspelmath 1994). Tabel 2 esitab eesti keeles esinevate infiniitsete vormide loendi ja nende formatiivid ning järgnev alapeatükk annab lühikese ülevaate nende vormidest ja funktsioonidest eesti kirjakeeles. Murretes infiniitsete vormide süntaktiliste funktsioonide tervikkäsitus puudub.

Tabel 2. Eesti keele infiniitsed vormid ja nende formatiivid (Viitso 2003).

INFINIITSED VORMID	PERSONAAL	IMPERSONAAL
Partitsiibid		
Olevik	<i>v</i>	<i>dav tav</i>
Minevik	<i>nud</i>	<i>dud tud</i>
Supiin		
Illatiiv	<i>ma</i>	<i>dama tama</i>
Inessiiv	<i>mas</i>	
Elatiiv	<i>mast</i>	
Translatiiv	<i>maks</i>	
Abessiiv	<i>mata</i>	
Infinitiiv	<i>da a ta</i>	
Gerundiiv	<i>des es tes</i>	

3.6.1 Infinitiivid

Eesti keeles on kaks infinitiivi: *ma-* ja *da-*infinitiiv, esimest nimetatakse vahel ka supiiniks⁵. Laialt levinud seisukoht on, et infinitiiv on markeerimata verbivorm, mis peale verbitüve tähenduse mingit muud tähendust ei kanna, kuid Haspelmathi (1989) järgi on paljude Euroopa keelte infinitiivid arenenud otsustatavusest. Ylikoski (2003) defineerib infinitiivid veelgi abstraktsemal tasandil, väites, et infinitiivid on infiniitsed vormid, mis funktsioneerivad komplementidena. Seotud laienditena võivad infinitiivid väljendada erinevaid modaalseid tähendusi, kuid infintiivide semantilise funktsiooni määrab siiski temaga koos esinev põhiverb. (Ylikoski 2003)

***ma*-infinitiiv ehk supiin** on keeleajalooliselt supiini illatiivivorm, mis võib lauses esineda sihtkoha- või rektiooniadverbiaalina, adverbiaaltribuudina, perifrastilise verbivormi osana ja erijuhtudel ka iseseisva predikaadina (Erelt *et al.* 1995: 65). Kohamäärusena väljendab *ma*-infinitiiv sihtkohta ning laiendab verbe, mis väljendavad liikumist (*tulema, minema, hüppama, heitma*), tähendust 'tahtma väga minna' (*tükkima, kibeleva*), võimalikkust või piisavust (*jõudma, mahtuma, jätkuma*) ning kausatiiv-direktaalsust (*panema, istuma, ulatama*), kausatiivsust (*kutsuma, lubama, palkama*) ja sobivust (*sobima, kõlbma*). Sõltuvusmäärusena laiendab *ma*-infinitiiv lisaks eelnevatele mittelokatiivseid kausatiivverbe (*sundima, õhutama*), mõningaid modaalpredikaate (*nõusutuma, suuteline olema*), tegevuse või tegija omadusi väljendavaid adjektiive ja substantiive (*aeglane, helde, meister*). (Erelt *et al.* 1993: 252–255) *ma*-tegevusnimi moodustab perifrastilise verbi faasiverbidega, verbidega, mis väljendavad protsessi või seisundi kauseerimist. *minema* ja *tulema* võivad esineda perfektiivse afiksaaladverbina koos siirdumist ja siirmist väljendavate verbidega. Koos verbiga *saama* võib *ma*-infinitiiv moodustada perifrastilise tuleviku vormi. Verbiga *pidama* moodustab *ma*-infinitiiv modaalset või kvotatiivi väljendava konstruktsiooni. (Erelt *et al.* 1993: 19–20; 258–259)

⁵ Edaspidi kasutan supiini ja *ma*-infinitiivi mõisteid sünonüümsetena.

mas-vorm ehk *ma*-infinitiivi inessiiv väljendab tavaliselt mingi muu tegevuse suhtes samaaegselt toimuvat tegevust. *mas*-vorm esineb lauses koha- või reksiooniadverbiaalina, adverbiaaltribuudina ning moodustab koos *olema*-verbiga aspektilise tähendusega konstruktsiooni. (Erelt *et al.* 1995: 65)

mast-vorm ehk *ma*-infinitiivi elatiiv väljendab tavaliselt mingile teisele tegevusele eelnenud tegevust. Vorm võib lauses esineda koha- või reksiooniadverbiaalina, adverbiaaltribuudina või koos *lakkama*-verbiga tegevuse lõppemist väljendavas konstruktsioonis. (Erelt *et al.* 1995: 66)

maks-vorm ehk *ma*-infinitiivi translatiiv esineb lauses vaba otstarbemäärusena, kuid viimasel ajal on vormi kasutus levimas väljapoole vaba adverbiaali funktsiooni (Erelt *et al.* 1993: 258).

mata-vorm ehk *ma*-infinitiivi abessiiv väljendab sooritamata tegevust. Üksikute verbidega võib ta esineda reksiooniadverbiaalina nende deverbaalidega adverbiaaltribuudina. Enamasti on *mata*-vorm kas viisi või mingi muu määrus, adjektiivtribuut, predikatiiv või seisundimäärus. (Erelt *et al.* 1995: 66)

da-infinitiiv väljendab tegevust ilma spetsiifiliste grammatiliste tähendusteta (Erelt *et al.* 1995: 64). *da*-infinitiivil on eesti keeles palju süntaktilisi funktsioone (Erelt 1984), olles süntaktiliselt kõige mitmekesisem infiniitvorm. *da*-infinitiiv võib lauses esineda subjektina, objektina, adverbiaalina, predikatiivina, atribuudina, perifrastilise verbivormi osana ja iseseisva öeldisena. (Erelt *et al.* 1995: 64)

vat-infinitiiv on päritolult oleviku partitsiibi partitiivivorm, kuid tänapäeval käitub infiniitvina. *vat*-vorm võib lauses esineda reksiooniadverbiaalina ja sihitisena, väljendades tajumistegevuse objekti. (Erelt *et al.* 1995: 65)

3.6.2 Partitsiibid

Partitsiipidel on nii verbidele kui nimisõnadele omaseid morfosüntaktilisi jooni (Ylikoski 2003) ja neid võib defineerida verbaalsete adjektiividenä, see tähendab et tegemist on sõnadega, mis käituvad adjektiividenä, kuid on tuletatud verbidest (Haspelmath 1994).

nud- ja tud-partitsiibid on vastavalt isikulise ja umbisikulise tegumoe mineviku partitsiibid, mis võivad lauses esineda predikaadi, atribuudi, predikatiivi ja seisundimäärusena. Predikaadi funktsioonis esinevad mineviku partitsiibid koos verbi *olema* finitise vormiga või eitussõnaga, moodustades liitaegu või impersonaalse konstruktsiooni. (Erelt *et al.* 1995: 67–68)

v- ja tav-partitsiibid on vastavalt isikulise ja umbisikulise tegumoe olevikupartitsiibid. Nende partitsiipide süntaktilised funktsioonid kattuvad täielikult adjektiivide süntaktiliste funktsioonidega, kuid nad on siiski säilitanud verbile omased laiendid. Oleviku partitsiibid võivad lauses esineda atribuudi, predikatiivi või seisundimäärusena. (Erelt *et al.* 1995: 66–67)

3.6.3 Konverbid

Konverbid on infiniitsed verbivormid, mida peetakse funktsioonilt analoogseteks adverbidega. **Gerundiiv** ehk *des*-vorm väljendab lause põhitegevusega samaaegselt kaasnevat tegevust. Gerundiiv võib lauses esineda adverbiaali või adverbiaalatribuudina ning võib moodustada analüütilisi verbivorme, millel on perfekti tähendus. (Erelt *et al.* 1995: 69)

3.7 Abiverbistumisest ja abiverbikonstruktsioonidest

„The moment a verb is given an infinitive complement, that verb starts down the road of auxiliariness.” (Bolinger 1980: 297)

Suurem osa selles töös käsitletavate verbühendite finiiitses positsioonis esinevatest verbidest on kas omandanud abiverbifunktsioone või on neil selliseks arenguks olemas selge potentsiaal. Abiverbi mõiste suhtes ollakse väga erinevatel seisukohtadel ja erinevad autorid käsitlevad selle mõiste all üsna laia lingvistiliste nähtuste kogumit. Heine (1993: 22–24) esitab oma abiverbe käsitlevas monograafias selleteemalise kirjanduse põhjal üle 20 erineva abiverbi omaduse, millest mõnedki on omavahel vastuolus. Abiverbi kategooria määramise raskusele osutab ka Anderson (2006) oma põhjalikus tüpoloogilises abiverbikonstruktsioone käsitlevas monograafias. Kuteva (2001: 5, 13) märgib, et abiverbikonstruktsioonide ning muude analüütiliste verbide vahele pole võimalik selget piiri tõmmata ning erinevad verbid võivad olla väga erineva abiverbistumisastmega. Bolinger (1980: 297) väidab, et kohe, kui verb seob endaga infinitiivse komplemendi, algab sellega alistava verbi abiverbiks kujunemise teekond, mis ei pruugigi kaugemale kulgeda, kuid selline verb võib kujuneda ka täiesti omaette abiverbikategooria esindajaks. Siinses töös vaadeldavates ühendites võib teatud infinitiive käsitleda kui verbi valentsist tulenevaid grammatilisi objekte, subjekte jne, kuid kõigil neilgi juhtudel on finiiitsel verbil siiski teatav potentsiaal abiverbistuda. Grammatiseerumise potentsiaali lisab kindlasti vaadeldavate nähtuste sagedus; töö käsitleb põhjalikumalt vaid sagedasemaid ühendeid. Hopper ja Traugott (2003: 126–130) peavad keelenähtuse kasutussagedust nende grammatilisemaks muutumist soodustavaks teguriks.

Anderson (2006: 4–5) käsitleb abiverbi kui üksust, mis asetseb leksikaalse funktsiooniga verbi ja afiksi kontinumil ning on vähemalt mingil määral semantiliselt pleekinud ja grammatiseerunud, väljendamiseks ühte või mitut verbiga seotud grammatilist kategooriat, tüüpiliselt aspekti, modaalsust, aega, polaarsust või tegumoodi. Abiverb võib seega olla keeleüksus, mis kombineerudes mõne leksikaalse verbiga, moodustab samas osalauses verbifraasi, ning millel on rohkem või vähem defineeritav grammatiline funktsioon. Abiverbikonstruktsiooni määratleb Anderson (2006: 7) kui ühes osalauses esinevat struktuuri, mis koosneb vähemalt ühest leksikaalsest verbist, mis annab konstruktsioonile leksikaalse sisu, ning abiverbist, mis lisab konstruktsioonile

kas grammatilise või funktsionaalse sisu. Anderson (2006) jagab maailma keeltes esinevad abiverbikonstruktsioonid nelja suuremasse rühma, millest eesti keele puhul on olulisim nn abiverbikeskne rühm (*AUX-headed constructions*). Selles rühmas on Andersoni (2006: 15) järgi abiverbide tekkimise allikaks just verbi ja selle komplemendi paarid ning kaheosalauselised ühendid, kus teine osalause talitleb esimeses osalauses esineva verbi komplemendina. Siinses töös käsitletavaid kõiki ühendeid iseloomustavad just need jooned.

Käesoleva töö fookuses olevatest verbiühenditest on paljud juba omandanud abiverbikonstruktsiooni jooni ja käsitletavad finiitsed verbid on, kui mitte enam, siis abiverbikandidaadid. Eesti kirjakeeles jagatakse finiitse ja infiniitse verbi ühendid grammatiseerimise astme verbi liitvormideks, perifrastilisteks verbivormideks, ahelverbideks ja öeldis+komplement-ühenditeks, samas mõeldakse, et nende vahelised piirid ei ole väga selged (Erelt 2013: 51–55). Siinse töö käsitluses pole nende eristamise poole püritud just seetõttu, et rühmade vahelised piirid on hägusad ning eristamiskriteeriumid ebaselged. Keerukust lisab geograafilise mõõtme kaasamine, sest abiverbistumise protsess võib eri murretes olla erinevas etapis, ent iga murde kohta ei ole piisavalt informatsiooni, et selle üle otsustada. Verbiühenditest räägitakse järgnevalt kui (potentsiaalsetest) abiverbikonstruktsioonidest ning abiverbi käsitus on kõige lähedasem Andersoni (2006) omale, kus kahe verbi ühendeid käsitletakse suuremal või vähemal määral grammatiseerunud abiverbikonstruktsioonidena. Paljudel finiitsetel verbidel on teatavate infiniitsete vormidega koos esinedes abiverbidele iseloomulikke jooni, mõned neist on rohkem abiverbistunud, teised vähem. Vaadeldavates ühendites on võimatu tõmmata piiri selgelt abiverbistunud finiitsete verbide ja mitteabiverbistunud verbide vahele, pigem saab neid vaadelda abiverbi kontiinumina.

3.8 Dialektomeetria

Siinne töö on oma lähenemisviisilt ja kasutatavate meetodite poolest dialektomeetiline töö. Tänapäeval mõeldakse dialektomeetria all mis tahes dialektoloogilise andmestiku statistilist tötlust ja analüüsi (Chambers & Trudgill 1998: 137). Suured korpuspõhised murdeuurimused ei olnud pikka aega keeleressursside piiratuse tõttu võimalikud. Viimastel aastakümnetel kättesaadavaks muutunud murdekorpused on murdeainese statistilise analüüsi ja automaattötluse võimalikuks teinud. Murdekorpused on tänapäeval siiski suhteliselt vähe, kuid neid tekib järjest juurde ja ka nende mahud suurenevad. Murdekorpuste loomine on pikaajaline ning palju inimressurssi ja aega nõudev töö, kuna palju on tekstide morfoloogilist märgendamist, litereerimist, ülekantavust jne. Viimastel aastatel on hakanud tähelepanu keskmesse tõusma ka mittestandardse keele, s.h murrete automaattöötlus ja selleks püütakse järjest rohkem arendada erinevaid automaatrakendusi.

Traditsiooniliselt on dialektomeetria aga kitsam valdkond ning selle rajajaks peetakse Séguyd, kes oli „Gascogne'i keele ja etnograafia atlase” („*Atlas*

linguistique et ethnographique de la Gascogne” I–VI) tööde juht. Enda väljatöötatud dialektomeetrilisi meetodeid (Séguy 1971; 1973) rakendas ta küll ainult 1973. aastal ilmunud atlase viimases köites (Grieve 2009: 68). Séguy võrdles kõrvuti asetsevate piirkondade keelelisi sarnasusi ja erinevusi. Selleks kaasas ta oma analüüsi leksikaalsed, häälduslikud, foneetilised/fonoloogilised, morfoloogilised ja süntaktilised nähtused ning mõotis iga tüübipaari (mitte iga esinemisjuhtumi) vahelised kaugused vastavalt sellele, kas mingi tüüp esines naaberpiirkonnas või mitte. Andmestikus rohkem esinenud nähtused said suurema kaalu, millega Séguy elimineeris suured sageduserinevused. (Chambers & Trudgill 1998: 138; Grieve 2009: 69; Heeringa 2004: 14; Spruit 2008: 19) Selline ongi kobardialektoloogia ideaalis, kus analüüsi kaasatakse suur hulk eri keeletasandite nähtusi ning mõõdetakse nende omavahelisi kaugusi. Käesolev töö erineb eespool mainitud traditsioonilisest lähenemisest, keskendudes ainult ühe keeletasandi (süntaksi) ühele nähtusele (verbiühendid), kaasates analüüsi siiski kõik tuvastatud eri vormide tüübid. Séguy ei olnud traditsiooniliste analüütiliste meetoditega rahul ning oli veendunud, et isoglosside kaardistamise teed minek pole õige, kuna see eeldab absoluutsete murdepiiride olemasolu, samas kui tema enda uurimistulemused näitasid, et keele varieerumine moodustab geograafilise kontinuumi (Chambers & Trudgill 1998: 138; Grieve 2009: 69).

Tänapäevase dialektomeetria, Salzburgi dialektomeetria koolkonna rajaja ja Séguy’ põhimõtete ja osaliselt ka tehnika järgija on Goebel (Grieve 2009: 69–70). Goebli panus dialektomeetria arengusse on olnud märkimisväärne (vt nt Goebel 1982; 2006; Goebel, *et al.* 1984), tema metodoloogilisteks uuendusteks teiste seas on olnud näiteks klasteranalüüsi kasutuselevõtmine dialektomeetrilises töös ning mitmete visualiseerimisvõimaluste arendamine (Spruit 2008: 20).

Omaette koolkonnaks võib pidada ka Hollandis, eriti aga Groningeni ülikoolis tegutsevaid dialektolooge. Koolkonnale panid aluse John Nerbonne ja Wilbert Heeringa 1990. aastate alguses ja põhiliselt on nad tegelenud häälduslike ja fonoloogiliste nähtuste varieerumise mõõtmisega (Grieve 2009: 72). Groningenis on välja arendatud elektrooniline keskkond, mida on võimalik kasutada kõigi keeletasandite varieerumisnähtuste uurimisel – programm Gabmap (Nerbonne *et al.* 2011). Programmi olen klasteranalüüsi tegemiseks rakendanud ka siinses töös. Lisaks eespool mainituile tegelevad dialektomeetriaga laialdaselt teadlased üle maailma, näiteks Kretschmar USA-s, Cichocki Kanadas, Ogura Jaapanis jne (Chambers & Trudgill 1998: 140).

Eestis on dialektomeetrilisi selle kitsamas tähenduses, aga ka laiemalt statistilisi meetodeid rakendavaid töid tehtud vähe. Üks esimesi dialektomeetria rakendajaid Eestis oli Murumets, kes on mõõtnud eri kihelkondade sõnavaralisi suhteid (Murumets 1982; Murumets 1983; 1984) ning samuti on sõnavaralistele suhetele keskendunud dialektomeetrilises võtmes Pajusalu ja Krikmann (2000). Statistilisi meetodeid laiemalt erinevate verbiühendite uurimisel on rakendanud käesoleva töö autor (Uiboed 2010a; Uiboed *et al.* 2013).

3.9 Korpuspõhine dialektoloogia

Kuna traditsiooniliselt kasutab dialektomeetria algandmetena digitaliseeritud atlasandmestikku, siis peetakse eraldi valdkondadeks korpuspõhist dialektomeetria ja dialektoloogiat (Szmrecsanyi & Wolk 2011: 562). Korpuspõhine dialektoloogia on üsna uus uurimisvaldkond ja selleteemalisi töid on hakanud ilmuma alles viimastel aastatel.

Korpuse põhjal tehtav dialektoloogiline uurimus võib teenida erinevaid eesmärgi: eesmärgiks võivad olla funktsionaal-tüpoloogiline ja varieerumise analüüs, keele tõenäosuslike mudelite arendamine, keeleajalooliste teadmiste täiendamine või mõne keeleteooria arendamine/ümberlõkkamine (Anderwald & Szmrecsanyi 2008: 1131). Korpuspõhine dialektoloogia üritab sageli vastata küsimustele, kas ja kuidas mingite keelenähtuste põhjal on võimalik eri keelevariante klassifitseerida. Tavaliselt lähenetakse sellele küsimusele kahel viisil: kas teoreetiliselt kindlas raamistikus, vastavalt olemasolevatele uurimustele ja keeleuurija oma teadmistele või empiirilisel, kus klassifitseerimiskriteeriumid ilmnevad korpusest mingi arvutusliku algoritmi alusel, ilma et arvestataks kindlat teoreetilist raamistikku, eelnevaid uurimusi või isiklike teadmisi. (Moisl 2009)

Szmrecsanyi ja Wolk (2011) esitavad kolm olulist punkti, mis peaksid olema korpuspõhise murdeuurimuse lahutamatud osad.

- Korpuspõhine murdeuurimine peaks hõlgama senise, ühe nähtuse põhise lähenemise ning analüüsi tuleks kaasata võimalikult palju erinevaid lingvistilisi muutujaid, uurida tuleks n-ö metsa, mitte puid.
- Andmestikku tuleks analüüsida tänapäevaste mitmemõõtmelise statistika meetoditega (nt peakomponentanalüüs, mitmemõõtmeline skaleerimine, klasteranalüüs, korrespondentsanalüüs).
- Tulemuste tõlgendamisel tuleks kasutada erinevaid visualiseerimisvõimalusi (nt kaardid, graafikud, dendrogrammid).

Käesolev töö järgib üldjoontes sellist uurimistöö ülesehitust. Pisut erinev on siiski see, et töö ei keskendu paljude eri keelenähtuste uurimisele, vaid ainult ühele neist – kahe verbi ühenditele – kuid kaasab siiski kõik tuvastatavad ühendid, kitsendamata uurimust mõnele konkreetsele verbile või infiniitsele vormile. Peatükis 9 ilmneb, et sellised verbiühendid oma iseloomult küllaltki heterogeensed ning seega on töö oma iseloomult ikkagi kobardialektoloogiline.

Võrreldes murdeatlastega annab korpusandmestik lingvistiliselt realistlikuma pildi. Korpusandmestikul on atlasandmestiku ees kaks olulisemat eelist. Esiteks esitab atlasandmestik ainult mingi nähtuse ilmnemise või mitteilmnemise, selles esinev info on liialt üldistatud ja redutseeritud ning seetõttu võib korpusandmestikuga võrreldes olla ebatäpsem. Korpuspõhist dialektomeetria võib nimetada ka sageduspõhiseks dialektomeetriaks, kuna see pakub keelenähtuste kohta arvulist informatsiooni. Teine atlasandmestiku suurem puudus on see, et seal sisalduv info pole loomulik keelekasutus, info on metalingvistiline ja põhineb atlase koostaja teadmistel. Atlasandmestik põhineb tüüpiliselt küsimustikel ning on analüüsi käigus vähemalt kaks korda redutseeritud (välitööde läbiviija ja

atlase koostaja poolt). Tekstikorpus seevastu pakub otsest, reaalsest keelekasutusest esile tulnud materjali ning seetõttu võib aidata jõuda lingvistiliselt usaldusväärsemate tulemusteni. (Szmrecsanyi ilmumas (c))

Korpuslingvistika üldtuntud puudus on, et see ei võimalda uurida harvaseinevaid keelenähtusi, kuid teisalt võib küsida, kas keelenähtused, mis ei kajastu suuremahulises tekstikogus, peaksid kobaruuringutes üldse kaasatud olema. Geograafiliselt ebaolulised keelenähtused ning sagedusmüra on osa lingvistilisest reaalsusest ning on raske õigustada nende analüüsist eemaldama. Võrreldes korpuspõhiste sagedusanalüüsidega, ülehindab atlasandmestik keelenähtuste kirjeldamisel geograafilist mõõdet. (Szmrecsanyi ilmumas (c)) Korpusandmestiku kasutamise eeliseid atlasandmestiku ees demonstreerib hästi ka eesti murrete *tarvis/vaja*-konstruktsioonide pilootuuring (Lindström & Uihoaed 2013), mis näitas ilmekalt, kuidas korpuspõhine sagedusandmestik, võrreldes traditsioonilise atlasandmestikuga, võimaldab saada varieerumisest mitmekülgsema ja detailsema pildi.

Siiani on murdekorpuste põhjal enim uuritud häälikulist varieerumist. Samuti on mõnevõrra uuritud sõnavara. Vaieldamatult on kõige vähem murdesüntaktilisi uurimusi. Põhjalikumad korpuspõhised murdesüntaksi uurimused käsitlevad briti-inglise keele murdeid (Szmrecsanyi 2013) ja hollandi murdeid (Spruit 2008).

3.10 Korpuspõhised dialektoloogilised uurimused

Nii nagu eelpool juba korduvalt rõhutatud, on korpuspõhine dialektoloogia noor valdkond ning seetõttu on veel ka vähe selleteemalisi uurimusi. Järgnevalt annan lühikese ülevaate maailma korpuspõhistest dialektoloogilistest töödest ning keskendun ainult käesoleva töö seisukohalt relevantsematele ehk süntaksiteemalistele ja/või statistilisi meetodeid rakendanud töödele.

Kõige põhjalikumalt on korpuste põhjal uurinud briti-inglise keele murrete süntaktilist varieerumist Szmrecsanyi (nt 2012; 2013). Oma 2013. aastal ilmunud põhjalikus monograafias annab ta väga põhjaliku ülevaate briti-inglise keele murrete süntaktilisest varieerumisest, kaasates analüüsi üle kümne süntaktilise nähtuse. Uuritavad tunnused ekstraheerib ta suurest tekstikorpusest automaatselt. Samuti esitab ta samas töös väga täpse metodoloogia samalaadse töö läbiviimiseks. Analüüsis rakendab ta mitmemõõtmelise statistika meetodit (peakomponentanalüüsi) ning kasutab erinevaid visualiseerimisvõimalusi andmete ja tulemuste esitamiseks ning tõlgendamiseks. Oma töös käsitleb ta ka probleeme, mis sellise kvantitatiivse korpuse põhjal tehtava dialektoloogilise töö läbiviimisel ette tulevad ja pakub välja lahendusi nendest probleemidest ülesaamiseks.

Grieve (2009) on suure tekstikorpuse põhjal uurinud ameerikainglise kirjakeele süntaktilist varieerumist. Töö on tähelepanuväärne, kuna tüüpiline murdekorpus sisaldab reeglina suulist kõnet intervjuude kujul. Grieve kogus oma korpuse aga elektrooniliselt üle USA ilmuvate ajalehtede lugejakirjadest

ning üritas selle materjali põhjal tuvastada süntaktiliste nähtuste geograafilisi erijooni. Autor rakendab taas mitmemõõtmelise statistika meetodeid, aga ka geostatistikast tuntud meetodeid nagu Morani *I*-koefitsient ning ruumiline autokorrelatsioon. Grieve'i tulemused on küll osaliselt kattuvad varem esitatud murdejaotustega, kuid üks huvitavamaid tulemusi näitas, et esineb süntaktilisi jooni, mille distributsioon korreleerub hoopis kultuuriipiiridega.

Spruit (2008) on uurinud hollandi murrete süntaktilist varieerumist digiteeritud atlasandmestiku põhjal. Spruit uurib oma töös komplekselt erinevaid süntaktilisi nähtusi ning näitab, et ka süntaktiliste nähtuste varieerumist mõjutab tõepoolest geograafia. Lisaks mõõdab ta seoseid keele eri tasandite vahel (nt süntaks-leksika, hääldus-süntaks jne). Töös on peatähelepanu kvantitatiivsetel meetoditel ja kartograafial.

Eespool on mainitud ainult mõned näited töödest, mis uurivad murdesüntaksi korpusandmestiku põhjal ja/või rakendavad samas väljaarendatud kvantitatiivseid meetodeid. Viimastel aastatel on selliste tööde ilmumine siiski sagenenud (vt näiteks ka Hernández 2010; Schulz 2010). Eesti murrete uurimisel on korpusi ja statistilisi meetodeid rakendanud näiteks käesoleva töö autor (Uiboed 2010a; 2010b; Uiboed *et al.* 2013) ning Lindström jt (2009).

3.1.1 Eesti murdesüntaks

Eesti murrete süntaksi pole siiani eriti süstemaatiliselt uuritud. Senised vähesed morfosüntaktilised uurimused keskenduvad enamjaolt eripärasematele ja rohkem varieeruvatele nähtustele. Järgnev osa annab põgusa ülevaate seni tehtud eesti murdesüntaksi alasest töödest. Käesoleva töö seisukohalt olulisematest uurimustest tuleb põhjalikumalt juttu peatükis 9.

Kõiki eesti murdeid hõlmavaid töid on siiani tehtud vähem. Põhjalikumalt on uuritud kaudse kõneviisi kasutust eesti murretes, mida väljendatakse nii morfoloogiliselt kui ka erinevate analüütiliste verbikonstruktsioonidega (Kask 1984). Neetar on uurinud aluse ja öeldise ühildumist (Neetar 1965a; 1965b; 1964; 1963) ning leiab, et aluse ja öeldise ühildamatust esineb rohkem põhjaeesti murretes ning lõunaeesti murretes on see haruldasem nähtus. Lühidalt on Neetar käsitlenud *eks*-lauseid (1966) ja määrsõnalist täiendit (1970). Viimase puhul leiab ta, et määrsõna kasutamine omadussõna funktsioonis on levinum saarte murdes, vähem kirde-eesti rannikumurdes ja lõunaeesti murretes, mis osutab ranniku- ja lõunaeesti murrete mõningasele sarnasusele. Käsikirjalisena on Tartu ülikooli lõputöödena olemas eituse-kategooria käsitlusi (Vitsberg 1958; Klaus 2009). Käesoleva töö autor on uurinud ühendverbide kasutust eesti murretes (Uiboed 2010a; 2010b) ning põgusalt vaadanud verbikonstruktsioonide varieerumist (Uiboed *et al.* 2013). Siinse töö temaatikale kõige lähemal on Potsepsi (1956) *da*-infinitiivi käsitlev diplomitöö, kus lisaks vormilisele varieeruvusele on põgusalt vaadeldud *da*-infinitiivi sisaldava liitöeldise varieerumist kõigis eesti murretes. Põhjalikumalt on murretes käsitletud isikukategooria kasutust (Lindström *et al.* 2009) ja isikuühildumisest on tehtud

üks magistritöö (Kalmus 2009). Kontrastiivselt läti keelega on tehtud tähelepanekuid mineviku partitsiibi kasutuse ja funktsioonide kohta eesti murretes (Muižniece *et al.* 1999). Ka kõik kõited sarjas „Eesti murded” sisaldavad vähemal või rohkemal määral tähelepanekuid käsitletavate murrete süntaktiliste nähtuste kohta (Tanning 1961; Must 1965; Keem 1970; Univere 1996; Must 1995; Keem & Käsi 2002; Lonn & Niit 2002; Juhkam & Sepp 2000). Lühikese ülevaate varasematest murrete morfosüntaktilistest uurimustest esitab õpik „Eesti murded ja kohanimed” (Pajusalu *et al.* 2009: 44), kus autorid taas kord möönavad, et sellest lingvistikavaldkonnast eesti murretes on uurimusi napilt (2009: 25).

Rohkem on tehtud üksikuurimusi väiksemate murdealade kohta. Lindström (1997) on uurinud sõnajärge Võru murdes. Objekti käitumist Nõo murrakus on vaadelnud Keem (1961) ning põhiliselt Risti murrakus ja selle naaberaladel Juhkam (1983). Samas artiklis on Juhkam teinud üksikuid tähelepanekuid uuritava nähtuse kohta muudel murdealadel, osutades võõr mõjudele objekti-käänete valikul: idamurdes on tugev vene mõju ja läänes suur rootsi mõju. Juhkam (2009) on teinud ülevaate refleksiivverbi konstruksioonist läänemurdes ning oma magistritöös (1998) annab ta muu hulgas ülevaate mõnedest grammatilistest nähtustest, mis on tingitud eestirootsi murdekontaktidest. Harju-Madise murraku erinevaid süntaktilisi nähtusi käsitleb sama autor oma värskes põhjalikus monograafias (Juhkam 2012). Karksi murrakut ja selle erinevaid süntaktilisi nähtusi on käsitlenud Tanning (2004). K. Pajusalu (1996; 1989) on põhjalikult uurinud Karksi murraku verbimorfoloogiat ning sellega seoses esitanud tähelepanekuid ka süntaktiliste nähtuste kohta. Karksi murrakus on *nud*-partitsiip väga produktiivne, liitudes ka afiksaaladverbidele (Pajusalu 1989). Eraldi on uuritud saarte murde eitust (Koit 1963) ning kitsamal alal, Kihnu murrakus (Sang 1974; Sang 1975). Mets (2010) on käsitlenud Võru murde näitel suhtlusvõrgustikke ning nendega seoses teinud tähelepanekuid ka süntaktiliste nähtuste kohta.

Eespool esitatud tööde loetelu pole kindlasti ammendav ja kõikehõlmav, kuid tuleb möönda, et eesti murdesüntaksi on uuritud suhteliselt ebasüsteemaatselt ning üldjuhul on tegemist üksikuurimustega ja/või vaadeldud on üksikuid kitsamaid piirkondi. Vähem on kõiki murdeid käsitlevaid töid. Käesolev uurimus püüab laiendada teadmisi ühe süntaktilise nähtuse varieerumisest kõigi eesti murdealade lõikes.

3.12 Varasemad verbiühendite uurimused

Käesoleva töö uurimisobjekt – verbiühendite varieerumine eesti murretes – on murretes väheuuritud teema. Puudub informatsioon võimalike finitiitse ja infiniitse verbi ühendite esinemiste kohta eri murretes, rääkimata nende esinemis-sagedustest ja morfosüntaktilistest omadustest. Potseps (1956) on uurinud *da*-infinitiivi kasutust eesti murretes ja sellega seoses vaadelnud mõningaid *da*-infinitiiviga moodustunud liitõeldise variante. Neetar (1988) on avaldanud

artikli, kus ta vaatleb mõnd finiiitse ja infiniitse verbikonstruktsiooni eesti murretes. Univere (1966) on võrrelnud põhjaeestilise *hakkama* ja lõunaeestilise *nakkama* tähenduserinevusi, millega seoses on ta analüüsinud ka mitmeid nende verbidega esinevaid verbikonstruktsioone. Tähelepanekuid erinevate tarindite kohta on ilmunud sarja „Eesti mured“ iga köite sissejuhatavas peatükis (Tanning 1961; Must 1965; Keem 1970; Univere 1996; Must 1995; Keem & Käsi 2002; Lonn & Niit 2002; Juhkam & Sepp 2000). Verbiühendite varieerumisest on lühikese ülevaate esitanud Uibo et al. (2013).

Kirjakeeles on uurimisobjekt leidnud tunduvalt laialdasemat käsitlemist, ilmunud on mitu doktoritööd. Trigel (2003) on uurinud eesti keele tuumverbe ning nende grammatilisi funktsioone erinevates verbiühendites kognitiivse lingvistika perspektiivist. Penjam (2008) on korpuse põhjal uurinud erinevaid *ma-* ja *da-*infinitiivi konstruktsioone ja töö raamistikuks on samuti konstruktsioonigrammatika. Trigel ja Habicht (2012) on teinud põhjaliku diakroonilise uurimuse verbist *saama* ja sellega moodustunud verbikonstruktsioonidest. Kõne all olevate verbiühendite kohta on ilmunud palju tähelepanekuid mitmete grammatiliste ja süntaktiliste nähtuste kirjeldamisel (nt Erelt & Metslang 2009; Erelt 2009; Habicht et al. 2010; Kasik 2001; Kehayov & Torn-Leesik 2009; Lindström & Trigel 2010; Lindström & Trigel 2007; Metslang 2006; Metslang 1993a; Mägiste 1936; Muischnek & Sahkai 2010; Penjam 2005 jne).

Üksikasjalikumalt tuleb erinevatest verbikonstruktsioonidest ja nende käsitlusest nii eesti kirjakeeles kui ka murretes juttu töö teises, tulemusi kirjeldavas osas iga ühendi juures üksikasjalikumalt.

4 ANDMESTIK: EESTI MURRETE KORPUS

Eesti murrete korpus on elektrooniline andmebaas, mis sisaldab autentseid murdetekste kõigilt Eesti murdealadelt. Korpust täiendatakse jooksvalt, kuid tööd alustati 1998. aastal. Korpuse loomise ja arendamise põhieesmärkideks on teha murdematerjal uurijatele elektrooniliselt kättesaadavaks ning luua võimalused eesti murrete kontrastiivseks uurimiseks võimalikult suure andmehulga põhjal. Korpus valmib koostöös Tartu Ülikooli ja Eesti Keele Instituudi vahel. Eesti murrete korpus (edaspidi EMK) koosneb:

- murdelindistustest;
- foneetilises transkriptsioonis murdetekstidest;
- lihtsustatud transkriptsioonis murdetekstidest;
- morfoloogiliselt märgendatud tekstidest;
- keelejuhtide ja lindistuste metaandmetest;
- alates 2009. aastast on tehtud algust murdetekstide süntaktilise automaatanalüüsiga ning praeguseks sisaldab korpus ka üksikuid süntaktiliselt analüüsitud tekste, kuid hetkel puudub interaktiivne süntaktiline otsinguvõimalus. (EMK 2011)

Korpuse ühe suurema ja olulisema osa moodustavad soome-ugri foneetilises transkriptsioonis murdetekstid. Transkribeerida on üritatud võimalikult täpselt: lisatud on partiklid, üneemid, parandused, kordused. Kõiki neid korpuse tekstides transkribeeritud nähtusi ei ole traditsioonilises murdeuurimises oluliseks peetud. Lisaks keelejuhi tekstidele on transkribeeritud intervjuuerija tekst. (Lindström & Pajusalu 2003)

Foneetilises transkriptsioonis olevad tekstid on üle viidud lihtsustatud transkriptsiooni ning on esitatud *txt*-laiendiga failides. Lihtsustatud transkriptsioonis tekstid on aluseks morfoloogilisele märgendamisele, töödeldavad muude programmidega ning nende põhjal saab teha arvutianalüüse. (Lindström 2009)

Siinse töö aluseks on morfoloogiliselt märgendatud tekstid ja märgendamiseks on kasutatud abiprogrammi Mark (Goba 2006). Morfoloogiliselt märgendatud tekstid on XML-formaadis ning märgendatud tekstid on vabalt kättesaadavad ja viidud ühtse otsingumootori alla. Praegu on kasutusel kaks otsingumootorit, millest üks pärib otsingu morfoloogiliselt märgendatud korpusest ning teine võimaldab otsida lihtteksti lihtsustatud transkriptsioonis tekstidest ilma morfoloogilise märgenduseta. Andmebaasi kasutamisel on vajalik teada märgendamisel kasutatud lühendeid, märgendatud kategooriaid jmt. EMK on vabalt kasutatav ja kättesaadav Interneti-aadressil <<http://www.murre.ut.ee/murdekorpus/>>. Eesti murrete korpus on ka osa rahvusvahelisest murdesüntaksi projektist Edisyn ning korpus on ligipääsetav ka nende kodulehelt: <http://www.dialectsyntax.org/wiki/Network_by_country>.

EMK-s sisalduvad tekstid on traditsioonilised murdelindistused, kus küsituleja intervjuuerib keelejuhti temale tuttavale territooriumil enamasti kodus või koduõues. Lindistatud on vanemaid, umbes 70–90aastaseid, inimesi, peamiselt

naisi (Lindström 2001). Murdelindistused on sisult suhteliselt sarnased ning teemad puudutavad enamjaolt erinevaid töid, kombeid, kooliskäimist ja muud elu-olu. Üksikud tekstid on muinasjutud või jutustused. Temaatika võib varieeruda geograafilisest asukohast tingitud aktuaalsemate tööde lõikes. Näiteks on rannikualadel tunduvalt rohkem juttu kalapüügist ja merega seotud teemadest.

Murdekorpuse tekstid on lindistatud vahemikus 1957–1986. Viis teksti on stuudiolindistused varasemast perioodist, aastast 1938. Sellised tekstid moodustavad siinse töö koguandmestikust vaid 2% ja on pärit erinevatelt murdealadelt. Tekstide temaatika ja žanr on sarnane hilisemate lindistustega. Nendel põhjustel pole neid, erinevas keskkonnas lindistatud tekste materjalist välja arvatud.

Murdekorpuse tekstid ei ole päris loomulik dialoog, vaid seda võib pidada omaette spetsiifiliseks žanriks. Tegemist on pigem institutsionaalse vestlusega, sest vestlevad omavahel võõrad inimesed. Tekstid on narratiivsed ja sisaldavad pikki monoloogilisi lõike. Need asjaolud kajastuvad ilmselt ükskõik millise keelelise analüüsi tulemustes, mille aluseks murdekorpus on. Lindström (2001) on võrrelnud murdekorpuse dialooge ja argivestlusi ning välja toonud nende vahelised olulisemad erinevused. Näiteks on suured erinevuse isikuviietamises ja seetõttu ka isikuvormide kasutuses, ajavormide kasutuses, kuna enamasti räägitakse minevikusündmustest jne. Samuti on piiratud tekstide temaatika. Seega on kõigi nende asjaolude mõju uurimismaterjalis ja -tulemustes paratamatu. Murdekorpuse tekstid on siiski temaatika ja suhtlussituatsiooni poolest omavahel sarnased ning seetõttu ka omavahel võrreldavad. Kindlasti pole aga keelelise analüüsi alusmaterjal representatiivne ning seda tuleb tulemuste tõlgendamisel arvesse võtta. Siinse töö tulemustes avaldub see enamasti üsna selgelt verbikasutuses, s.t võrreldes tavalise argivestlusega esineb ebaproportsionaalselt palju verbe, mis on seotud konkreetsete (majapidamis)töödega, kooliskäimisega jmt.

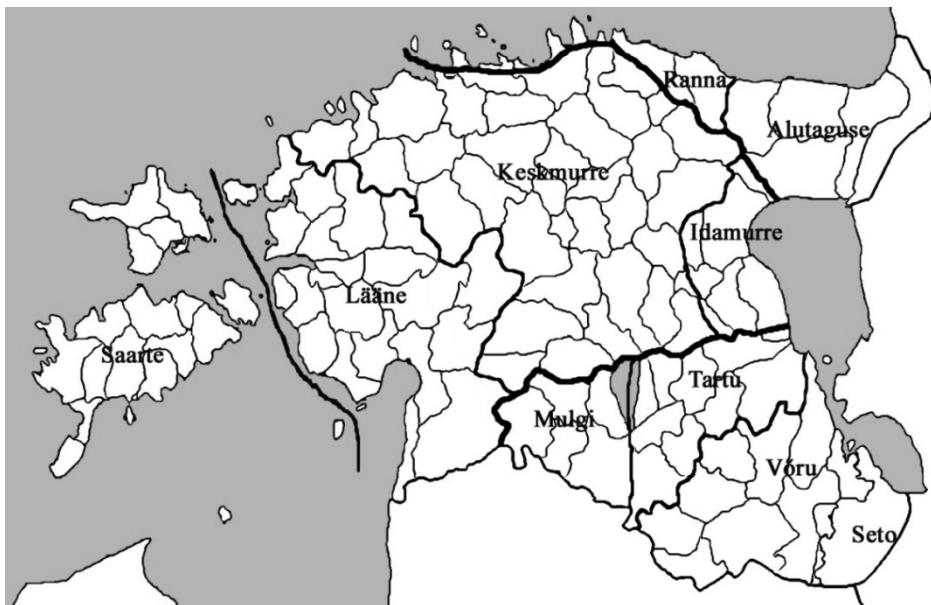
4.1 Murdekorpus visuaalselt ja arvudes

Käesoleva töö uurimismaterjali kogumise hetkel sisaldas murdekorpus 613 124 morfoloogiliselt märgendatud tekstisõna. Tabel 3 esitab mõned murdekorpuse arvulised andmed. 613 124 morfoloogiliselt märgendatud sõna jagunevad 204 erineva teksti vahel. Keelejuhte oli kokku 190, kellest 55 ehk umbes kolmandik olid mehed ning ülejäänud kaks kolmandikku keelejuhtidest on naised; naiste osakaal on selgelt ülekaalus. Uuritavaid murdealasid on korpuses kokku 10, tekste on korpuses 53 erinevast kihelkonnast ning 154 erinevast külast.

Tabel 3. Murdekorpuse ülevaade 01.09.2010 seisuga.

Sõnu	613 124
Tekste	204
Keelejuhte kokku	190
Keelejuhte (naised)	135
Keelejuhte (mehed)	55
Murdealasid	10
Kihelkondi	53
Külasid	154

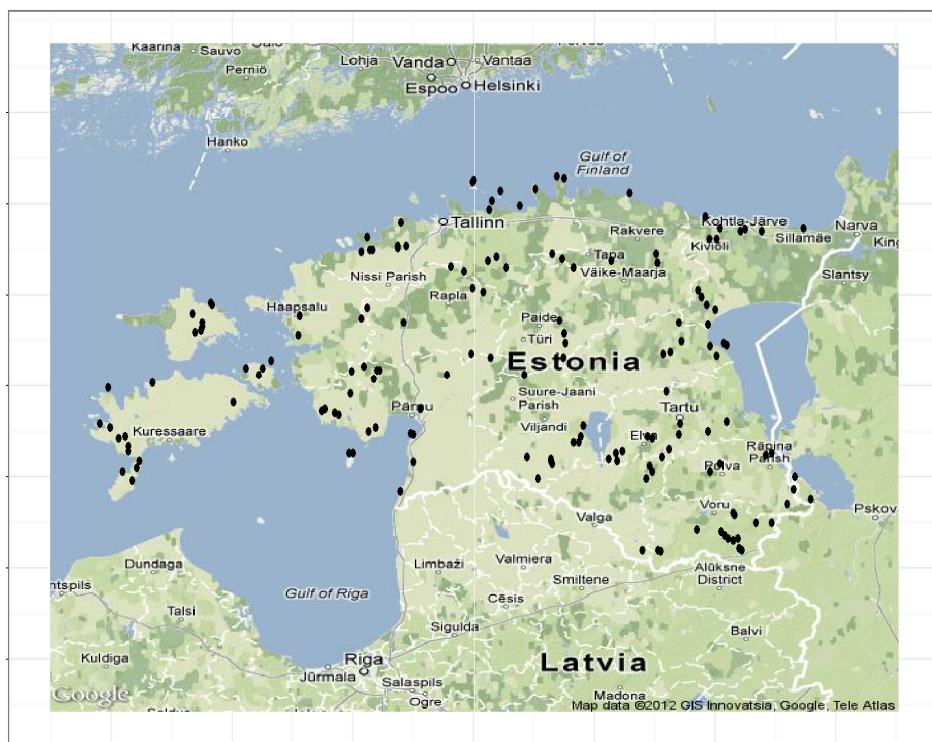
Murdekorpuses on tekstid jaotatud vastavalt EMK murdeliigendusele, mis järgib traditsioonilist murdejaotust, eristades 10 murdala: põhjaeesti rühma kuuluvad ida-, kesk-, lääne- ja saarte murre, lõunaeesti rühma moodustavad Mulgi, Võru, Tartu ja Seto murre ning eraldi kirderanniku murderühma moodustavad kirde- (ehk Alutaguse) ja rannamurre. Joonis 2 esitab loetletud murdealad kaardil.



Joonis 2. Eesti murrete korpuses kasutatav murdejaotus. Põhjaeesti murderühm: ida-, kesk-, lääne- ja saarte murre; lõunaeesti murderühm: Mulgi, Seto, Tartu ja Võru murre; kirderannikurühm: rannamurre ja kirdemurre (Alutaguse).

Murdekorpuse murdejaotus erineb pisut traditsioonilistest murdejaotustest. Nimelt loetakse tavaliselt samaks murdeks Võru ja Seto murded. Nii murde-

korpuses ja selle eeskujul on siinses töös neid kahte eristatud ja edaspidi käsitletlen neid eri murretena. Murdeõpiku autorid (Pajusalu *et al.* 2009) mõõnavad, et Võru ja Seto murde eristamine on eesti dialektoloogias problemaatiline, kuid peavad murdealasid siiski kultuuriliselt ja ajalooliselt erinevateks ja käsitlevad mainitud alasid eraldi murrakurühmana. Võru ja Seto murret on eraldiseisvate murretena käsitletud varasemateski töödes (Lindström *et al.* 2009: 92–93), et pöörata rohkem tähelepanu omapärasele kultuurilisele alale. Käesoleva töö demonstreerib hiljem, kuidas selline eristamine on vägagi õigustatud. Joonis 3 esitab murdekorpuses tekstide päritolu külade kaupa. Iga punkt kaardil tähistab ühte teksti. Kõik punktid on esitatud külade koordinaatide põhjal, kuid üksikutel juhtudel polnud kohta võimalik külanime järgi määrata. Sellistel juhtudel on punktid esitatud kihelkonna keskpunkti põhjal. Nagu kaardilt näha, on murdekorpuse tekstide päritolu geograafiliselt suhteliselt ühtlaselt jaotunud. Ükski piirkond ei ole väga tugevalt teiste suhtes ülesindatud, küll aga leidub alaesindatumaid piirkondi, näiteks saarte, lääne- ja Mulgi murde alal.



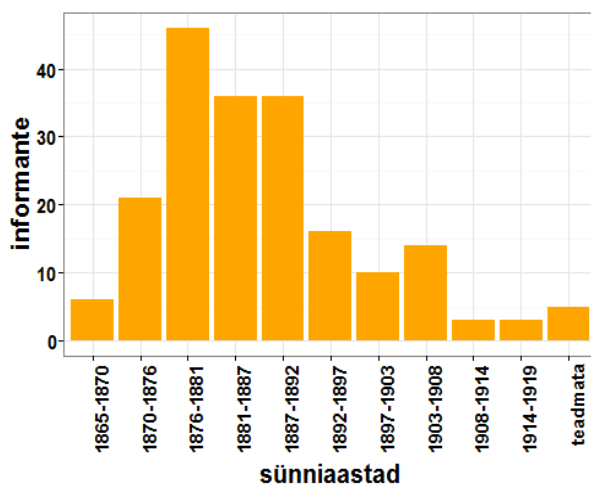
Joonis 3. Murdekorpuse tekstide päritolu. Kaardina on kasutatud Google'i Eesti kaarti < <https://maps.google.com/> [03.09.2012].

Murdekorpuse põhjalt tehtava kvantitatiivse uurimistöö teeb keeruliseks murdealade erinev esindatus korpuses, s.t et erinevatelt murdealadelt on korpuses erinev arv sõnu. Metodoloogiliselt on selline probleem lahendatav ning sellest tuleb täpsemalt juttu peatükis 7.1. Tabel 4 esitab morfoloogiliselt märgendatud sõnade arvu igas murdes, nagu see oli materjali kogumise hetkel. Tabelitest võib näha, et kõige paremini on esindatud saarte, kesk- ja läänemurre ning kõige vähem on morfoloogiliselt märgendatud sõnu Seto murdes. Sõnade arvu erinevused võivad murrete lõikes olla vägagi suured ning see kõik seab lisapiirangud kvantitatiivsele andmetötlusele. Sellised probleemid on väga tavalised nii dialektoloogiliste kui diakrooniliste uurimuste aluseks olevates keeleandmestikes: materjal on väike, ent selle suurendamine ei ole võimalik; võrreldakse väga erineva suurusega (all)korpusi. Tänapäevased meetodid ja tehnikad võimaldavad selle probleemiga, küll teatud piirangutega, toime tulla.

Tabel 4. Murdekorpuse morfoloogiliselt märgendatud sõnade arv murrete kaupa uurimismaterjali kogumise hetkel (seisuga 01.09.2010).

murre	sõnade arv
saarte	124 215
lääne	109 833
kesk	99 127
Võru	53 825
ida	43 965
Tartu	43 215
ranna	42 257
kirde	40 667
Mulgi	34 416
Seto	21 604
kokku	613 124

Korpuse tekstid on pärit 196-lt erinevalt keelejuhilt, kellest umbes üks kolmandik on mehed, seega on korpuse sooline tasakaal tugevalt kaldu naiste suunas. Kõik keelejuhid on sündinud vahemikus 1865–1919, nii et noorima ja vanima keelejuhi sünniaastate vahe on 54 aastat. Viie keelejuhi sünniaasta on teadmata. Joonis 4 esitab murdekorpuse keelejuhtide sünniaastade jaotuse viie aasta kaupa. Kõige rohkem on korpuses esindatud vahemikus 1876–1881 sündinud keelejuhid ning vahemikus 1876–1892 sündinud keelejuhid moodustavad 60% kõigist keelejuhtidest. Seega on korpuse keel on sünkrooniliseks uurimiseks täiesti sobiv.



Joonis 4. Murdekorpuse informantide sünniaegade jaotus.

4.2 Murdekorpuse märgendus ja XML-fail

Murdekorpuse on morfoloogiliselt märgendatud, s.t igale informandi sõnale tekstis on lisatud selle sõnaliigi ja muu morfoloogiline informatsioon. Seda tööd tehakse käsitsi abiprogrammiga Mark (Goba 2006). Kõik morfoloogiliselt märgendatud tekstid on XML-formaadis. XML on lihtsustatult öeldes struktureeritud tekstiformaat, mis võimaldab tekste lihtsalt kasutada andmebaasides ja otsingumootorites ning samuti on sellise tekstiga lihtne teha automaatanalüüsi. Igas morfoloogiliselt märgendatud tekst sisaldab lisaks intervjuueerija ja intervjuueeritava tekstile ning selle morfoloogilisele märgendusel ka järgmiseid metaandmeid iga faili kohta eraldi: kihelkond, küla, keelejuht, lindistusaeg, lindistajad, kommentaarid (lindistust häirivad asjaolud, kõrvalised hääled) jmt. Murdekorpuse tekstides märgendatavad sõnaklassid ja muud märgendatavad morfoloogilised kategooriad on esitatud lisa 1. Joonis 5 esitab murdekorpuse morfoloogiliselt märgendatud XML-faili näite. Struktuuriliselt näeb fail välja nii, nagu see on esitatud joonisel, kuigi failist on esitatud vaid fragment. Fail algab märkusega, et tegemist on XML-failiga, defineeritakse kodeering (*encoding*), viidatakse dokumendi tüübi failile DTD. DTD-s on kirjeldatud murdekorpuse XML-faili formaat ning selle põhjal järgitakse, et dokument oleks vormistatud õigesti, s.t kõik märgendid, nende järjekord jne peab olema esitatud DTD-s kirjeldatud viisil, vastasel juhul edastatakse veateade. Järgneb info <record>-märgendite vahel, kus esitatakse vastava faili lindistuse kihelkond ja küla. Järgneb kommentaari rida <com>, mis sisaldab muud olulist metainfot lindistuse ja keelejuhi kohta. Erinevad kõnevoorud algavad <u who=...> märgendiga ning keelejuht on alati tähistatud lühendiga KJ ning intervjuueerija oma ees- ja perekonnanime esitähedega, antud juhul on see Mari Must (MM).

<mark> märgendite vahel on iga märgendusüksus, <sne> tähistab tekstisõna ning <msn> märksõna sellele tekstisõnale kirjakeelestatud kujul. Eraldi märgendite vahel on veel morfoloogiline info <mrf ...>. Sellised failid nagu on esitatud joonisel 5 olid käesoleva töö algmaterjaliks, mille põhjal tuvastati verbiühendid (sellest vaata täpsemalt peatükis 5). XML-i kohta täpsemalt esitab hea sissejuhatuse Tidwell (2002) ja Wilcock (2009) annab mitmekülgse ülevaate XML-st lingvistilise märgendamise kontekstis.

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE record SYSTEM "morfo.dtd">
3 <record khk="KUU" kla="Salmistu">
4   <com> Rannikumurre, Kuusalu, Salmistu küla. Jakob Lindström (s
      1892). EMH 2117, KKI MT 350. Lindist. M. Must, E. Ross, A. Kask
      1973. Lit. M. Must?, L. Lindström jaan. 1999 </com>
5   <u who="MM">kust need `saad'i (...)</u>
6   <u who="KJ"><mark><sne>nied</sne><msn>need</msn><mrf slk="ProS">pl n
      </mrf></mark><mark><sne>sai</sne><msn>saama</msn><mrf slk="V">ps
      ind ipf sg 3</mrf></mark><mark><sne>*ostettud</sne><msn>ostma
      </msn><mrf slk="V">tud</mrf></mark>(...) <mark><sne>li-</sne><mrf slk
      ="X"/></mark><mark><sne>*linnast</sne><msn>linn</msn><mrf slk="S">
      sg el</mrf></mark><mark><sne>puest</sne><msn>pood</msn><mrf slk="S">
      sg el</mrf></mark>(...)</u>
7   <u who="MM">jaa (...)</u>
8 </record>

```

Joonis 5. Murdekorpuse XML-dokumendi näide.

5 VERBIÜHENDITE TUVASTAMINE

Iga korpuslingvistiline töö algab automaatse või poolautomaatse materjali kogumisega. Järgnev peatükk kirjeldab käesolevas töös kasutatud uurimismaterjali automaatset kogumisviisi. Uurimismaterjal pärineb peatükis 4 kirjeldatud eesti murrete korpusest seisuga 01.09.2010.

Murdekorpused on küll morfoloogiliselt märgendatud, kuid süntaktiline märgendus suuremal osal korpustekstidest puudub. Siiski on võimalik olemasolevat morfoloogilist märgendust maksimaalselt ära kasutades tuvastada ka süntaktiliselt huvipakkuvaid nähtusi. Alustasin murdetekstide automaatse osalausestamisega selleks otstarbeks kohandatud eesti keele süntaksianalüsaatoriga (Lindström & Müürisep 2009; Müürisep 2000). Kasutasin analüsaatorit ainult osalause piiride märgendamiseks, kuna murrete süntaktiline automaatanalüüs on ülimalt keeruline ja tulemus eeldatavasti eri murrete lõikes ebaühtlane. Seetõttu ei saa süntaktiliste rollide määramist automaatsel teel pidada piisavalt usaldusväärseks algmaterjaliks. Kuigi ka automaatne osalausestamine pole ideaalne, täidab analüsaator seda ülesannet ühtlaselt üsna hästi kõigi murrete peal. Lindströmi ja Müürisepa (2009) uurimistulemused näitavad, et valesti tuvastab analüsaator 0,4% osalausepiiridest. Joonis 6 esitab fragmendi automaatselt osalausestatud murdekorpuse failist. Süntaksianalüsaatori sisendiks on murdekorpuse morfoloogiliselt märgendatud XML-fail (täpsemalt vt ptk 4.2). Analüsaator lisab automaatselt igale sõnale tema oletatava lauseliikme märgendi⁶ ja (oletatava) osalause algust tähistava märgendi (CLB või CLB-C). Väljundis esitatakse kõige pealt tekstisõna (nii nagu see jooksvas tekstis esineb), järgneb selle tekstisõna algvormi tüvi ning tekstisõna morfoloogiline ja süntaktiline informatsioon lühendite ja märgendite kujul. Morfoloogiline info on sama, mis käsitsi märgendatud murdekorpuse failides (vt ka lisa 1), süntaktilise info ja osalause piirid lisab süntaksianalüsaator. Esitatud näitefailis algab lause partikliga *e* ning see on ühtlasi ka osalause algus (CLB-C), järgmine osalause algab proadverbiga *siis* ning näite viimane, kolmas osalause sidesõnaga *ja*. Siinses töös kasutatakse vaid osalausepiiride märgendeid ja morfoloogilist informatsiooni; automaatselt analüüsitud süntaktilist infot ei rakendata.

⁶ Lühendite selgitused leiab veebiaadressilt <http://math.ut.ee/~kaili/papers/syntax.html> [viimati kontrollitud 05.10.2013].

e
 e+0 // _B_ #cap // **CLB-C @B
 tahad
 taht+0 // _V_ main ps indic pres sg ps2 #cap #FinV #Part-P #InfP // @+FMV
 ruudulest
 ruuduline+0 // _A_ pos sg part #cap // @ADVL
 teha
 tege+0 // _V_ main inf #cap #NGP-P // @OBJ
 suure
 suur+0 // _A_ pos sg gen #cap // @AN>
 lappidest
 lapp+0 // _S_ com pl el #cap // @ADVL
 siiss
 siis+0 // _D_ #cap // **CLB @ADVL
 tulid
 tule+0 // _V_ main ps indic impf pl ps3 #cap #FinV #Intr // @+FMV
 värvilehed
 värviline+0 // _A_ pos pl nom #cap // @AN>;
 lõngad
 lõng+0 // _S_ com pl nom #cap // @SUBJ
 ja
 ja+0 // _J_ coord #cap // **CLB-C @J
 jälle
 jälle+0 // _D_ // @ADVL
 ko-
 !!! tüvi+0 // _S_ com #S? #- #cap // @REP
 kujutti
 kudu+0 // _V_ main impv indic impf #cap #FinV #NGP-P // @+FMV
 ära
 ära+0 // _D_ affix #cap // @ADVL

Joonis 6. Süntaksianalüsaatori väljund osalausestatud failist.

Peale osalause piiride seadmist on võimalik vaadelda materjali ainult osalauseste kaupa. Käesoleva töö eelduseks on, et kaks verbi moodustavad konstruktsiooni ainult siis, kui nad esinevad koos samas osalauses ning ka vastupidi, kui kaks teatud vormitunnustega verbi esinevad samas osalauses, moodustavad nad tõenäoliselt konstruktsiooni. Moodustasin automaatselt kõik võimalikud kahe verbi paarid ühe osalause sees ja seda kõigis murretes eraldi.

Eesti keele sõnajärg on väga varieeruv. Finiitverb esineb enamasti teisel kohal pealause ja viimasel kohal kõrvallause, kuid on samal ajal siiski väga tundlik infostruktuuri suhtes (Lindström 2005: 174; Tael 1988). Konstruktsiooni moodustavad sõnad võivad paikneda lauses erinevas järjekorras (näiteks (1), (2)), kuid konstruktsiooni tähendus sellest ei muutu. Antud juhul on tegemist finiiitse verbiga *käima* ning supiini inessiivi vormis verbidega *jooma* ja *tegema*. Kahes esimeses näites on infiniitse ja finiitse verbi järjekord lausingis erinev. Esimesel juhul on infiniitne verbivorm enne finiiitset (*joomas käivad*) ja teisel juhul vastupidi (*käisimi jooman*), kuid tähenduselt on tegemist siiski samade vormide ja nagu peatükis 9 selgub ka samade konstruktsioonidega. Verbiühendi

eri osad ei pruugi ilmtingimata teineteise kõrval esineda (näiteks (2) ja (3))⁷, seega ei piisanud lihtsalt kahe verbi järjendite eraldamisest, vaid tuli arvestada asjaoluga, et kahe verbi vahele võib jääda erinev hulk sõnu.

- (1) et säält allikast **joomas käiväd** (LÄÄ)
- (2) **käisimi** Pikkäsillal ja Liival ja **jooman** (TAR)
- (3) vahõst **kääsime** siäl e Vana Kurasõl oksõ kua **tegemäs** (SAA)

Analüüsi kaasasin ainult finiiitse verbi tema algvormi ehk lemma kujul ning infiniitse verbi morfoloogilise kategooria. Seega uuritavate ühendite ning nende sageduste korpusest ekstraheerimise algoritm on järgmine.

- 1) Moodustasin automaatselt kõik võimalikud kahe verbi paarid ühe osalause sees.
- 2) Järjestasin ühendid ümber finiiitverbi järgi, et näidetes (1) ja (2) esitatud juhtudel *joomas käiväd* ja *käisimi jooman* loetaks samaks ühendiks.
- 3) Ekstraheerisin ainult verbipaarid, mis koosnesid finiiitsest ja infiniitsest komponendist, s.t analüüsimatest jäid välja ühendid, mis koosnesid näiteks kahest finiiitsest verbist (*lähen toon piima*) ehk võisid moodustada potentsiaalse sariverbi.
- 4) Eemaldas finiiitse komponendi lemma ja kaasasin ainult morfoloogilise kategooria märgendi ning seega jäid alles ainult finiiitse verbi lemma ja infiniitse verbi morfoloogilise kategooria paarid, nagu nad on esitatud näites (4).

Et kontrollida, kui usaldusväärne selline suhteliselt lihtsakoeline konstruktsioonikandidaatide tuvastamisprotsess on, tuli leida algoritmi täpsus. Selleks moodustasin testkorpuse, mis koosnes igast murdest juhuslikult valitud 500 sõnast. Konstruktsioonide tuvastamise täpsus oli 80%, kuid peale madala esinemissagedusega $V_{fin}+Inf$ paaride eemaldamist (<3) tõusis täpsus 92%-ni. Kuigi sellist materjali väljaarvamist võib alati küsimuse alla seada, leidsin, et antud juhul on madala esinemissagedusega paaride eemaldamine analüüsist õigustatud mitmel põhjusel. Esiteks konstruktsioonide tuvastamisprotsessi täpsus tõusis peale seda märgatavalt, mis näitab, et esialgu oli tulemustes palju juhuslikkust. See asjaolu on nii varieeruva ja keerulise materjali automaattöötlusel paratamatu. Analüsaator pidi samade reeglite põhjal hakkama saama siiski kõigi kümne murde, sisuliselt kümne erineva keele, analüüsimisega. Kuna suuline kõne üldiselt on väga varieeruv ja seal esineb palju juhuslikkust (parandused, ümbersõnastused jmt), siis on paratamatu, et see juhuslikkus kajastub ka kogutud andmetes. Madala esinemissagedusega üksuste eemaldamine on sarnases olukorras üsna levinud praktika. Iseenesest mõistetavalt kaasneb sellega ka huvitava ja relevantse materjali kadu, kuid analüüsimatest usaldusväärsuse

⁷ Siin ja edaspidi, kõik murdekorpuse näitelauseid on normaliseeritud, s.t muudetud võimalikult kirjakeelestatud kujule. Eemaldatud on transkriptsioonimärgid ning muudetud on korpuses kasutatavat ortograafiat.

ja homogeensuse huvides jätkan selles töös materjaliga, kus kõik vähem kui kolm korda esinenud ühendid on materjalist välja arvatud.

Lõplik analüüsitav materjal näeb välja nagu see on esitatud idamurde näites (4) ja selline loend on moodustatud iga murde kohta eraldi.

(4)	<i>ma</i>	hakkama
	<i>tud</i>	olema
	<i>ma</i>	panema
	<i>ma</i>	hakkama
	<i>ma</i>	pidama

Esimeses veerus kursiivkirjas on infiniitse verbi morfoloogiline kategooria ja teises veerus temaga samas osalauses esinenud finiitverbi lemma. Esitatud näites oleks võimalikud konstruktsioonide konstruktid *hakkasid nutma, oli palgatud, pannakse kuivama, hakkas liikuma, peab sööma*. Seega finiitse verbi sagedused on koondatud kõik tema lemma alla ning verbi muid grammatilisi kategooriaid (aeg, isik, tegumood jne) eristatud ei ole.

Kuigi põhjaeestilist *hakkama*- ja lõunaeestilist *nakkama*-verbi peetakse etimoloogiliselt eri verbideks (Uibo 2010; Univere 1966), olen materjali analüüsimiseks ühtlustanud verbi *hakkama* esinemiskujud korpuses. Olen muutnud lõunaeestilise *nakkama*-verbi kirjakeelseks. Esiteks ei tundu verb vaadeldavates konstruktsioonides kandvat mingit erinevat semantilist sisu ning teiseks olulisemaks põhjuseks on asjaolu, et korpuses on see sõna märgendatud väga ebaühtlaselt. Mõnikord on lõunaeestiline *nakkama* saanud märksõnaks kirja-keelse *hakkama* ning mõnikord on jäetud algvorm lõunaeestiliseks. Materjali homogeensuse huvides olen selle ebaregulaarsuse ühtlustanud, muutes kõik verbi esinemiskujud kirjakeelseteks.

Sellise ekstraheerimisalgoritmi üks suuremaid puudusi on kolme verbi ühendite käsitlemine. Kui osalauses esineb kolme verbi ühendeid, siis algoritmi põhjal genereeritakse ühendeid liiga palju ning materjali satub kandidaate, mis sellisel kujul tegelikult keeles ei esine. Seda illustreerib näide (5), kus rakendatud algoritm genereeriks ühendid *võis hakata* ja *võis tegema*, millest viimane eesti keeles konstruktsioonina ei esine.

(5)	<i>võis tööd hakata tegema</i>
-----	--------------------------------

Probleemi ei tekita kaks infiniitset verbi (nt *hakata tegema*), sest meetod võimaldab automaatselt välja võtta ainult finiidist ja infiniidist koosnevad ühendid. Kuna ülalkirjeldatud probleem ei olnud siiski suur ja sellised ühendid moodustavad materjalist väikse osa ning vaatamata valedele ühenditele tuvas-tatakse ka õiged, on seda probleemi materjali kogumisel ignoreeritud. Selline teguviis on põhjendatav mitme asjaoluga. Korpuslingvistilises töös ja keele automaattöötlusel on selline müra tekkimine paratamatus ning sellega tuleb lihtsalt arvestada. Samuti kasutan edasises analüüsis kollostruktsioonilisi meetodeid ja vaatlen üleüldist ühendite sagedust, mis võimaldavad selliseid problemaatilisi juhtumeid materjalist enamasti eemaldada ja/või analüüsis marginaliseerida.

6 MEETODID

„In the long run, linguistic theorising without a firm grip on statistical methods will lead corpus linguistics into a dead-end street.” (Mukherjee 2007: 141)

Dialektoloogid on eelkõige huvitatud sellest, kuidas uuritavad keelenähtused geograafiliselt paiknevad ning grupeeruvad. Sageli rakendatakse dialektoloogias meetodeid, mis kuuluvad mitmemõõtmelise statistika meetodite perekonda (Chambers & Trudgill 1998: 137). Mitmemõõtmeliste meetodite kasutamist on Bisang (2004: 16) põhjendanud väitega, et kaardid on staatilised ja kahemõõtmelised, samal ajal kui keele varieerumine on dünaamiline ning mitmemõõtmeline (lisaks geograafilise mõõtmele on olemas sotsiaalne, sooline, hariduslik jne mõõde), seetõttu tuleb rakendada erinevaid meetodeid, mis võimaldavad võtta arvesse keerulisemaid keelekasutussituatsioone. Siinne töö rakendab mitmemõõtmelise statistika meetoditest korrespondentsanalüüsi, mida kirjeldan peatükis 6.3.

Kõige vanemad murdeliigendused põhinesid täielikult keeleuurija intuitsioonil ja üleüldistel teadmistel murdejoonte kohta. Levinud on murrete klassifitseerimine isoglosside põhjal, mida on murretevaheliste erinevuste üle otsustamisel peetud subjektiivseks meetodiks; keelealad moodustavad enamjaolt kontiinumi ja murdepiiride range tõmbamine on küsitav tegutsemisviis. (Heeringa 2004: 10–11)

Chambers ja Trudgill (1998: 135–136) väidavad, et murretevahelised erinevused võivad olla nii kvalitatiivsed kui kvantitatiivsed. Autorid näevad kvalitatiivseid erinevusi absoluutsetena, s.t mingi nähtus esineb murdes A, kuid ei esine murdes B. Kvantitatiivsed erinevused on Chambersi ja Trudgilli järgi suhtelised ning fookuses pole mingi nähtuse (mitte)ilmnemine, vaid see, kui sageli mingit nähtust murdealal esineb. Siinse töö eeldused on sellise liigituse puhul kindlasti kvantitatiivsed, s.t et ma ei eelda mingi nähtuse esinemist ainult ühes murdes, välistades sama nähtuse esinemise teises. Mainitud käsitluse põhjal võib süntaktilisi erinevusi üleüldiselt pidada pigem kvantitatiivseteks, kuna selle lingvistilise tasandi puhul on eriti keeruline jõuda absoluutsete järeldusteni. Spruiti (2008: 19) järgi käsitleb kvalitatiivne keeleuurimine reeglina detailselt väikest hulka nähtusi ning tavaliselt on aluseks väiksemahuline, kuid fokuseeritud ja kompaktne andmehulk, kvantitatiivne keeleuurimine seevastu analüüsib suurt hulka nähtusi, kasutab suuri andmehulki, kuid vaatab keelenähtusi väiksema lingvistilise detailsusega. Siinse töö analüüsi osas püüan neid kahte poolt siduda, kuid domineerib siiski kvantitatiivne põhimõte, mille järgi vaatlen uurimisobjekti väiksema lingvistilise detailsusega.

Kvantitatiivsete erinevuste tuvastamine nõuab kvantitatiivseid meetodeid, mis on keele varieerumise uurimisel äärmiselt olulised. Dialektoloogia ja sotsiolingvistika kasutavad paljuski sarnaseid meetodeid, kuid on olemas meetodeid, mida eriti eelistatakse rakendada dialektoloogilises uurimistöös. Kvantitatiivsed meetodid mitte ainult ei võimalda töödelda suuremat andmehulka, vaid on ka objektiivsemad. (Chambers & Trudgill 1998: 135–136)

Sarnaselt sotsiolingvistilistele andmetikele on murdeandmetikes tavaliselt palju probleeme, mis seavad statistiliste meetodite kasutamisele piirangud. Mendoza-Denton jt (2003: 106) toovad välja kolm sellist probleemset kohta:

- geograafiliselt ebaühtlaselt jaotunud keelejuhid,
- sageduste suured erinevused,
- palju lünki (n-õ null-lahtrite probleem) andmestikus.

Kõigi need probleemid on ilmsed ka siinse töö uurimismaterjalis. Kvantitatiivsete tulemuste täiendamiseks olen seetõttu rakendanud kvalitatiivsema iseloomuga konstruktsioonide analüüsi (täpsemalt peatükis 9).

6.1 Kvalitatiivsete ja kvantitatiivsete meetodite kombineerimine käesolevas töös

Siinne töö lähtub eeldusest, et parim viis nii heterogeense materjali uurimiseks on kvalitatiivsete ja kvantitatiivsete meetodite kombineerimine. Olen nõus Szmrecsanyi ja Kortmanni (2009) väitega, et kvantitatiivne analüüs ei suuda kunagi asendada põhjalikku kvalitatiivset analüüsi, kuid kindlasti suudavad kvantitatiivsed andmed oluliselt täiendada kvalitatiivseid uurimistulemusi. Manning (2003: 295) arutleb süntaktilise varieerumisandmestiku kogumise probleemide üle ja leiab, et harvem esinevate süntaktiliste nähtuste uurimiseks, tuleb toetuda intuitsioonile või koondada analüüsi andmestik, mis on pärit erinevatelt keelejuhtidelt ja eri ajaperioodidest. Sellega osutab ta nähtavasti statistiliselt sobiliku valimi koostamise probleemile. Samas artiklis möönabki Manning (2003: 295), et isegi kui otsustada kas introspektsiooni või andmete koondamise kasuks, ei pruugi see lahendada lünkliku andmestiku probleemi. Selles töös olen otsustanud minna teiste teed: koondada kõigi murrete kogu olemasolev andmestik ning pole seda eraldi kategoriseerinud näiteks informantide vanuse, soo, kihelkonna jmt järgi.

Mukherjee (2007) rõhutab, et statistilisi uurimistulemusi tuleb alati hoolikalt tõlgendada ja analüüsida kvalitatiivselt. Käesolev töö rakendab lisaks tänapäevastele kvantitatiivsetele korpuslingvistika meetoditele traditsioonilisi kvalitatiivseid meetodeid konstruktsioonide väljaselgitamiseks ning nende tähenduste tuvastamiseks. Kvantitatiivsed meetodid võimaldavad küll tuvastada ja analüüsida vormi põhjal potentsiaalseid konstruktsioone, kuid kvalitatiivsed meetodid kindlasti täiendavad tulemust ja võimaldavad detailsemalt analüüsida konstruktsioonide tähendusi. Kombineerin meetodeid mitmel põhjusel. Esiteks on materjal liialt problemaatiline ja kasutatavad kvantitatiivsed meetodid siiani murrete uurimisel vähe rakendust leidnud, et nende põhjal võiks teha põhjapanevaid lingvistilisi järeldusi. Teiseks üritan kvalitatiivsete meetodite kaasamisega saada infot kvantitatiivsete meetodite usaldusväärsuse kohta. Kolmandaks nõustun Mukherjee (2007) ja leian, et oluline on vaadelda ka keelenähtusi, mis pole kas statistiliselt olulised või on madala esinemissagedusega, sest sellised nähtused võivad osutada oluliseks lingvistiliselt.

6.2 Kollostruktsioonilised meetodid

Kollostruktsioonilised meetodid⁸ on kvantitatiivse korpuslingvistika meetodite perekond, mis uurib keelekonstruktsioonide ja nendes esinevate sõnade või grammatiliste kategooriate vahelisi seoseid (Stefanowitsch & Gries 2003; Gries & Stefanowitsch 2004). Meetod järgib konstruktsioonigrammatika (eriti Goldberg 1995; Lakoff 1987) terminoloogiat ja põhimõtteid, mille kohaselt keele põhiüksuseks on konstruktsioon kui tähenduse ja vormi ühisus (Stefanowitsch & Gries 2003: 211). Kollostruktsiooniline analüüs uurib alati ühte konstruktsiooni ning vaatleb, millised sõnad on konstruktsiooni teatud positsioonides konstruktsiooniga eriti tugevalt seotud (Stefanowitsch & Gries 2003: 214–215). Kollostruktsioonilised meetodid on oma põhimõttelt sarnased rohkem tuntud kollokatsioonide tuvastamismeetoditega (Evert 2004; Evert 2008; Pecina 2008), mis mõõdavad korpuses kahe potentsiaalselt püsiühendit moodustava sõna vahelist statistilise seose tugevust.

Kollostruktsioone on võimalik analüüsida kolmel erineval viisil. **Koleksee-mide analüüsi** (Stefanowitsch & Gries 2003) kasutatakse teatud konstruktsiooni ning selle konstruktsiooni kindlas positsioonis esinevate sõnade (või vormide) uurimiseks. **Kolekseem** on sõna, mis on tugevalt seotud mingi kindla konstruktsiooniga, s.t on selle konstruktsiooni kolekseem. Sellise meetodiga saab uurida näiteks kausatiivset konstruktsiooni, mis moodustatakse *ma*-tegevusnime ja finiiitse *panema*-verbi abil. Selle uurimiseks peaks korpusest leidma kõik sellekujulised potentsiaalsed kausatiivkonstruktsioonid, misjärel on võimalik teha vajalikud arvutused. Meetodi rakendamise tulemusena näeksime, millised *ma*-tegevusnimes verbid oleksid kausatiivses konstruktsioonis finiiitse *panema*-verbiga eriti tugevalt seotud. Võib ilmned, et selline kausatiivkonstruktsioon eelistab *ma*-infinitiivi positsioonis ainult kindlaid (või kindlasse tähendusrühma kuuluvaid) verbe. Võimalik oleks uurida ka vastupidist: millised verbid oleksid *panema*-kausatiivkonstruktsioonis statistiliselt eriti mitte-eelistatud. Kui analüüsist ilmneks, et kausatiivkonstruktsioonis on finiiitse *panema*-verbiga eriti tugevalt seotud supiinivormis verb *kiirustama*, siis *kiirustama* on kausatiivse *panema*-konstruktsiooni kolekseem.

Eristavate kolekseemide analüüsi (Gries & Stefanowitsch 2004) eesmärk on uurida kahte või enamat konstruktsiooni, mis on semantiliselt ja/või funktsionaalselt sarnased. Gries ja Stefanowitsch (2004) analüüsivad inglise keeles palju uuritud ditranstiivset konstruktsiooni ja eessõnaga *to* moodustatava daativi konstruktsiooni ning vaatavad, millised ülekandmistähendusega verbid eelistavad statistiliselt ühte või teist konstruktsiooni. Eesti keele näitena meetodi rakendamiseks võib tuua analüütilised kausatiivkonstruktsioonid, mida on uurinud Kasik (2001). Selliseid konstruktsioone on võimalik moodustada näiteks finiiitsete verbide *ajama* ja *panema* ning supiinivormis verbi abil. Kasik on oma artiklis esitanud konstruktsioonide erinevusi, kuid seda oleks võimalik suure tekstikorpuse põhjal vaadelda veelgi suurema materjali põhjal. Selleks

⁸ Kollostruktsioon on sulam sõnadest *kollokatsioon* ja *konstruktsioon*.

tuleks korpusest leida kõik potentsiaalsed *panema*- ja *ajama*-kausatiivid ja vaadata milliste supiinivormis verbidega need koos esinevad. Eristavate kolekseemide analüüs võimaldaks uurida, millised supiinivormis verbid või verbirühmad eelistavad ühte või teist finiiitset verbi (*ajama* vs. *panema*) ning millised on eri kausatiivkonstruktsioonide semantilised erinevused, kui need üldse peaks ilmnenema. Peale statistilise analüüsi läbiviimist saab uurija ise tulemuste põhjal otsustada, kas tegemist on semantiliselt ja/või funktsionaalselt sarnaste või erinevate konstruktsioonidega.

Koosvarieeruvate kolekseemide analüüs (Stefanowitsch & Gries 2005) uurib sama konstruktsiooni erinevates positsioonides esinevate sõnade omavahelisi suhteid. Kasutades eelmist näidet, võib vaadata, millised sõnad esinevad enamasti koos kausatiivses *panema*-konstruktsioonis. Konstruktsioon esineb kujul PÕHJUSTAJA: PANEMA_{finiiitne} MÕJUTATAV: TEGEVUS_{supiin}. Koosvarieeruvate kolekseemide analüüs võimaldab näiteks uurida, millised põhjustaja ja mõjutatava positsioonis olevad sõnad on statistiliselt rohkem seotud sellekujulise kausatiivkonstruktsiooniga. Erinevalt kolekseemide analüüsist vaadeldakse siin konstruktsiooni kahte erinevat positsiooni ja nendes positsioonides esinevate sõnade omavahelisi suhteid.

6.2.1 Meetodi rakendamine selles töös

Käesolev töö rakendab kollostruktsioonilisi meetodeid natuke erineval viisil, kui seda traditsiooniliselt tehakse. Suurim erinevus on, et statistilist tugevust ei mõõdeta ainult mõne kitsamalt eeldefineeritud konstruktsiooni jaoks, vaid analüüsi eesmärk on laiem: eeldusel, et finiiitne ja infiniitne verb moodustavad samas osalauses esinedes konstruktsioone, mõõdetakse statistilist tugevust ainult finiiitse verbi ja infiniitse verbi **vormi** (infinitiiv, partitsiip jne) vahel, s.t ei arvestata finiiitse verbi leksikaalset sisu. Niisiis erinevalt traditsioonilistest kollostruktsioonilistest meetoditest, kus uuritakse sõnade ja konstruktsioonide omavahelist suhet, rakendan selles töös meetodit skemaatilisemal tasandil, uurides infiniitse verbi morfoloogilise kategooria ja sellega koosineva finiiitse verbi omavahelisi suhteid, kuna paljud finiiitsed verbid, mis esinevad koos teatud infiniitsete vormidega, on tugevalt grammatiseerunud. Seega keskendun infiniitse verbi morfoloogilisele kategooriale ning temaga koosinevale finiiitse verbi lemmale, mõõtes nende kahe üksuse vahelist statistilist tugevust eesti murrete korpuse tekstides. Analüüsi eesmärk on jõuda selgusele, millised finiiitsed verbid on statistiliselt tugevamalt seotud erinevate infiniitsete vormidega. Kui teatud ühend on statistiliselt piisavalt tugevasti seotud, võib eeldada, et tegemist on konstruktsiooniga. Analüüsi järgnevad sammud peaksid põhjalikumalt välja selgitama, kas vaadeldavad ühendid peidavad enda taga tõepoolest konstruktsioone või on tegemist lihtsalt statistilise müraga. Eelmises alapeatükis kirjeldatud kolmest erinevast kollostruktsioonilisest meetodist rakendan siin koosvarieeruvate kolekseemide analüüsi, vaadeldes ühe potentsiaalse konstruktsiooni.

siooni kahte erinevat positsiooni (finiitset ja infiniitset) ja nendes positsioonides esinevate üksuste omavahelisi suhteid.

Eelnevast järeldub, et meetodi rakendusviis on väga sarnane traditsiooniliste kollokatsioonide tuvastamise meetoditega, mille käigus mõõdetakse kahe keeleüksuse statistilist seost. Oma iseloomult on siinse töö rakendusviis aga siiski kollostruktsiooniline. Esiteks eeldan, et kaks vaadeldavat komponenti moodustavad potentsiaalse konstruktsiooni, seega ei mõõdetata statistilist seost kõigi potentsiaalselt konstruktsiooni moodustavate üksuste vahel. Teiseks on korpusest välja otsitud ainult huvi keskmes olevad vormipaarid, mis moodustavad kollostruktsioonilise analüüsi andmestiku. Kuigi eeldan, et finiiitne verb ja infiniitne vorm moodustavad samas osalauses koos esinedes konstruktsiooni, tuvastan kollostruktsiooniliste meetoditega üldisemalt siiski konstruktsioonide kandidaate. Analüüsi lähtekohaks on küll eeldus, et uurimisobjektiks olevad infiniitne vorm ja finiiitne verb moodustavad konstruktsiooni, ent kogu analüüs on siiski mitmetasandiline. Esiteks tuvastatakse (potentsiaalsed) konstruktsioonid, mida statistiliste tulemuste põhjal võib juba konstruktsiooniks pidada. Kvantitatiivne analüüs keskendub nende (statistiliste) konstruktsioonide varieerumisele. Peatükis 9 jätkatakse nende statistiliste konstruktsioonide detailsema analüüsiga, et selgitada välja, kas kvantitatiivse analüüsi käigus leitud konstruktsioonikandidaadid toimivad ka tegelikult konstruktsiooni(de)na ning mis (eelkõige grammatilisi) tähendusi nad väljendavad.

Otsustamaks, milliseid vaadeldavatest paaridest saab kollostruktsioonilise analüüsi tulemusena pidada konstruktsioonideks ning mitte juhuslikeks koosesinemisteks, rakendasin vabavaralist Coll.analysis.3 programmi (Gries 2007), mis on välja töötatud just kollostruktsioonilise analüüsi läbiviimiseks. Programm on hästi dokumenteeritud ja rakendustarkvaraga R (R Development CoreTeam 2011) kergesti kasutatav. Programm võimaldab korpusandmestiku põhjal mõõta infiniitse verbi morfoloogilise kategooria ja temaga koosesineva finiiitse verbi lemma vahelist statistilist tugevust, võttes arvesse mõlema üksuse eraldi- ja koosesinemissagedused andmestikus. Arvutused on tehtud kõigi kümne murde jaoks eraldi, s.t sisuliselt töötan kümne eri korpusega.

Põhimõtteliselt võib kollostruktsioonilise analüüsi jaoks valida ükskõik missuguse sõnadevahelise seose tugevuse mõõtmise statistiku (põhjalikku ülevaadet eri võimalustest vt nt Evert 2004; Pecina 2008; Wiechmann 2008). Kollostruktsioonilises analüüsis on levinumad Fisheri täpne test⁹ (*Fisher-Yates exact test*), šansisuhte mõõdik (*odd's ratio*), ja log-tõepära (*log-likelihood*) mõõdik (Stefanowitsch & Gries 2003; Gries 2007), kuid mõned testid on näidanud ka minimaalse tundlikkuse mõõdiku (*Minimum Sensitivity measure*) paremust teiste ees (Wiechmann 2008).

Enne murdematerjali analüüsiks sobivaima väljavalmist katsetasin kolme eri mõõdikut: šansisuhet, Fisheri täpset testi ja minimaalse tundlikkuse mõõdikut.

⁹ Kuigi ISI (Rahvusvaheline Statistikainstituut) soovib kasutada terminit 'Fisher-Yatesi test', siis sellest töös kasutan terminit 'Fisheri täpne test', mida kasutatakse enamasti eesti-keelses statistikaalases kirjanduses.

Lõpuks valisin Fisheri täpse testi (edaspidi FTT), kuna see töötas suhteliselt hästi kõigi kümne murde peal ja seda peetakse eriti sobivaks keeleandmestikuga töötamisel, sest statistik ei sea erilisi piiranguid valimi suurusele ega eelda normaaljaotust (Pedersen 1996; Pedersen & Bruce 1996), mida keeleteaduslike andmetike puhul esineb harva. Samuti on seda kõige rohkem rakendatud kollostruktsioonilise analüüsi läbiviimiseks.

FTT väärtused arvutatakse kahemõõtmelise sagedustabeli väärtuste põhjal, mis saadakse korpusandmestikust. Tabel 5 illustreerib, kuidas vajalikud väärtused FTT arvutamiseks korpusandmestikust saadakse. Vajalikud ühendid, mille sagedusi analüüsitakse, on leitud viisil, mida kirjeldas peatükk 5.

Tabel 5. Kahemõõtmeline sagedustabel koosvarieeruvate kolekseemide analüüsi jaoks vajalike väärtuste arvutamiseks.

	V_{fin} Y positsioonis 2	Kõik ülejäänud V_{fin} positsioonis 2
inf-vorm X positsioonis 1	Sagedus (X pos1 + Y pos 2)	Sagedus (X pos1 + $\neg$ Y pos2)
kõik ülejäänud inf-vormid positsioonis 1	Sagedus ($\neg$ X pos1 + Y pos2)	Sagedus ($\neg$ X pos1 + $\neg$ Y pos2)

Tabel 6 esitab kahemõõtmelise sagedustabeli idamurdes esineva verbiühendi *da*-infinitiivi ja finiiitse *tahtma* jaoks. Ühend esineb idamurdes samas osalauses viis korda. Kuus korda esineb finiiitne *tahtma* koos mõne muu infiniitse vormiga kui *da*-infinitiiv. Samas osalauses esineb *da*-infinitiiv koos mõne muu finiiitse verbiga (v.a *tahtma*) 47 korda. Idamurdes on 1096 finiiitse ja infiniitse verbi ühendit, mis ei sisalda ei finiiitset verbi *tahtma* ega ka *da*-infinitiivi. Kokku on idamurde alamkorpuses 1107 V_{fin} +Inf ühendit.

Tabel 6. Kahemõõtmeline sagedustabel. *da*-infinitiiv + *tahtma* ühend idamurdes.

	V_{fin} <i>tahtma</i>	Kõik ülejäänud V_{fin}	ridade summad
<i>da</i> -infinitiiv	5	47	52
kõik ülejäänud inf vormid	6	1049	1055
veergude summad	11	1096	N=1107

Sõnadevahelise seose tugevuse mõõtmise statistikud võrdlevad kahe uuritava üksuse koosesinemissagedust sagedusega, mida võiks oodata (nn. oodatav sagedus), kui need kaks elementi esineksid koos juhuslikult. Näiteks tabeli 6 eeskujul *da*-infinitiiv esineb 52 korda 1107 osalauses ja tema esinemise oodatav sagedus on vastavalt 52/1107 (umbes 5%). Finiiitse *tahtma*-verbi oodatav esinemissagedus on 11/1107 (umbes 1%). Kui need kaks elementi oleksid statistiliselt sõltumatud, võiks nende oodatav koosesinemissagedus olla $0.05 \times 0.01 = 0.0005$.

FTT arvutatakse järgmise valemi järgi:

$$FTT = \sum_{k=O_{11}}^{\min\{R_1, V_1\}} \frac{\binom{V_1}{k} \cdot \left(\frac{V_2}{R_1 - k}\right)}{\frac{N}{R_1}}$$

kus V tähistab tabeli veerge ja R ridu ning valemi väärtused saadakse kahe-mõõtmelisest sagedustabelist. Mida kõrgem on kahe uuritava üksuse FTT väärtus, seda suurema tõenäosusega saab ühendit pidada uurimismaterjali põhjal keeles esinevaks konstruktsiooniks, mitte kahe üksuste juhuslikuks koos-esinemiseks korpusel. Käesolevas töös kasutan 0,05 usaldusnivood, s.t et kõik ühendid, mille FTT kriitiline väärtus jääb lubatud piiridesse, on loetud konstruktsioonideks. Sellisel juhul võib väita, et uuritav ühend on piisava statistilise kindlusega keelekonstruktsioon.

Töös ei võrdle ma FTT väärtusi murrete ja konstruktsioonide lõikes. Esiteks pole FTT kasutades seda võimalik teha, kuna kümme uuritavat murret on korpusel erineva sõnade arvuga ning järelikult mõjutab erinev valimimaht ka FTT-i väärtusi. See ei võimalda FTT-i väärtusi murrete lõikes võrrelda¹⁰. Teiseks pole selle töö eesmärgiks jõuda lõplike järeldusteni ainult kollo-struktsioonilisi meetodeid kasutades. Kuna meetod sobib hästi automaatseks ja statistiliseks konstruktsioonide (kandidaatide) tuvastamiseks, rakendan selles töös ainult meetodi neid võimalusi.

6.2.2 Kollostruktsiooniliste meetodite kriitika

Kollostruktsioonilisele analüüsile heidetakse sageli ette, et arvesse ei võeta uuritavate elementide sagedusi kogu korpusel, vaid sagedusi vaadeldakse ainult uuritavate konstruktsioonide kontekstis. Keeleandmestikes, eriti aga diakroo-nilistes ja dialektoloogilistes andmestikes, võivad uuritavate nähtuste jaotused eri alamkorpusel olla väga erinevad, mis muudab kollostruktsiooniliste uurimistulemuste võrdlemise mõttetuks (Hilpert 2006: 250). Viimane probleem on käesolevas töös lahendatud nii, et analüüsid on tehtud kõigi kümne murde jaoks eraldi ning sagedusi on vaadeldud ainult ühe alamkorpusel põhjal. Nagu ka eelnevalt mainitud, ei püüa ma siinses töös kollostruktsioonilise analüüsi tulemusi omavahel murrete lõikes võrrelda, vaid kasutan meetodit iga murde jaoks eraldi ainult konstruktsioonide (kandidaatide) tuvastamiseks.

Probleeme tekitavad ka statistiliste meetodite üldised tööpõhimõtted, mille kohaselt eeldatakse sõnade esinemise juhuslikkust (vt eelmist alapeatükki), kuid keel ei ole juhuslik, keelekasutaja teeb iga lingvistilise valiku kindlas kontekstis kindlatel põhjustel (Bybee 2010: 97). See probleem on korpuslingvistikas üldtuntud ja seda tuleb tulemuste tõlgendamisel arvesse võtta. Samuti puudutab

¹⁰ On olemas siiski mõõdikuid, mis normaliseerivad lõplikud väärtused, nii et neid on või-malik rakendada eri suurusega valimite põhjal, nt šansisuhte mõõdik (Gries 2006). Kuna aga selle töö eesmärgiks pole mõõdikute väärtuste võrdlemine ja FTT osutus selgelt paremaks mõõdikuks, kasutan selles töös FTT-i.

see probleem rohkem normaaljaotustele tuginevaid meetodeid, käesolevas töös kasutatav FTT ei eelda uuritavate tunnuste normaaljaotust.

Kollostruktsiooniliste meetodite probleemne aspekt on kindlasti madala esinemissagedusega üksustega konstruktsioonide analüüs, kuna need saavad tavaliselt madala statistiku väärtuse ja võivad niimoodi analüüsist üldse välja jääda (Bybee 2010: 101). Seda tuleb tulemuste tõlgendamisel arvestada. Käesolev töö keskendub keelenähtustele, mis enamjaolt on suhteliselt kõrge esinemissagedusega ning infiniitse vormi ja finiiitverbi ühendite näol pole eesti keeles ega murretes tegemist haruldase nähtusega. Samuti koosnevad sellised ühendid eeldatavasti pigem sagedasti esinevatest vormidest ning sõnadest. Lisaks analüüsin töö 9. peatükis detailsemalt kõiki ühendeid nii kollostruktsioonilises kui lihtsageduste analüüsis (vt ptk 7) tuvastatud ühendeid. See kõik võimaldab pöörata tähelepanu ka madalama esinemissagedusega üksustele.

On erinevaid seisukohti, kas lihtne sõnesagedus on parem lingvistiline indikaator kui mingi keerulisema arvutuse põhjal saadud tulemus. Nt Gries jt (2005: 665) paljude seas on veendunud, et lihtsagedustega opereerimine võib viia valede lingvistiliste järeldusteni, samas kui Bybee jagab teistsugust seisukohta (Bybee 2010: 98; Bybee & Eddington 2006) ning lihtsageduse häid tulemusi kollokatsioonide tuvastamise meetoditega võrreldes on näidanud teisedki uurimused (nt Evert & Krenn 2001). Mingil määral on kindlasti õigus mõlemal poolel ja käesoleva töö üks eesmärkidest ongi mõlema poole väidete testimine, sest analüüsid viin läbi nii kollostruktsiooniliste meetodite kui ka lihtsageduste põhjal (täpsemalt vt ptk 7). Samuti ei rakenda käesolev töö kollostruktsioonilisi meetodeid traditsioonilisel kujul ning seega pole eesmärk teha nende tulemuste põhjal semantilisi järeldusi, milleks kollostruktsiooniliste meetodite arendamise esmane eesmärk (Stefanowitsch & Gries 2003). Rakendan meetodit vaid konstruktsioonide tuvastamiseks. Need asjaolud võimaldavad mõnes mõtte meetodile sageli esitatavast kriitikast mööda vaadata, kuid samas on kasutusele võetud n-õ ettevaatusabinõud ja meetodit ei kasutata lõplike järelduste tegemiseks ega konstruktsioonide semantilise analüüsi läbiviimiseks. Neid aspekte vaatlen siinses töös kvalitatiivsemalt peatükis 9.

Statistilisest perspektiivist on kollostruktsiooniliste meetodite puhul kahtluse alla seatud p-väärtuse kasutamine mõõdikuna (Baayen 2011), kuid see on laiem statistikateooria probleem ning puudutab kõiki kollokatsioonide tuvastamiseks kasutatavaid meetodeid. Käesolev töö lähtub valdkonnas eelnevalt tehtud tööst ning see statistiline probleem jääb põhjalikuma uurimistöö teemaks tulevikus.

6.3 Korrespondentsanalüüs

Korrespondentsanalüüs on mitmemõõtmelise statistika uuriv-kirjeldav meetod tabelikujulise andmestiku analüüsimiseks ja graafiliseks esitamiseks (Greenacre 2007: 1). Meetod on sarnane peakomponent- ja faktoranalüüsiga (Baayen 2008: 129). Paljude mitmemõõtmelise statistika meetodite põhimõte on sarnane: iga riskülikukujulise arvuliste väärtustega tabeli põhjal on võimalik mõõta kaugust

tabeli eri elementide (ridade ja veergude) vahel, seega on võimalik arvutada kaugusi tabeli ridade ja veergude vahel ning analüüsi peaülesanne on tulemust visualiseerida nii selgelt ja vähese infokaoga kui võimalik. Selliste ülesannete lahendamist võimaldavaid meetodeid on kahte liiki: 1) klasterdamis- ja klassifitseerimismeetodid (siinses töös kasutan hägusat klasteranalüüsi, mida kirjeldan täpsemalt järgmises peatükis) ning 2) peatelgede meetodid (*principal axes methods*), mis põhinevad enamasti lineaaralgebral (nende hulgas ka korrespondentsanalüüs), mõlemat liiki meetodeid saab rakendada samasugusel andmematriksil ja meetodid on üksteist täiendava iseloomuga. (Lebart *et al.* 1998: 46) Seetõttu rakendan sellest töös nii klasteranalüüsi kui ka korrespondentsanalüüsi.

Mitmemõõtmelise statistika meetodid on arenenud erinevates valdkondades ja koolkondades ning seetõttu võib terminoloogia kirjanduses varieeruda, samuti võib sarnastel meetoditel olla väikseid tehnilisi erinevusi. Korrespondentsanalüüsi (*correspondence analyses*) asemel esineb ingliskeelses kirjanduses ka termineid *dual scaling* ja *canonical correlation analyses*. Meetoditel on siiski rohkem sarnasusi kui erinevusi. (Chambers & Trudgill 1998: 137) Korrespondentsanalüüsile (edaspidi kasutan lühendit meetodi ingliskeelsest nimetusest CA) alusepanijaks peetakse Benzécrit 1960. aastate Prantsusmaal (Benzécrit 1973; Lebart *et al.* 1998: 47). Meetodit on pidevalt unustatud ja taasavastatud. Korrespondentsanalüüsist on erinevaid versioone ja käesolev töö rakendab lihtsat korrespondentsanalüüsi.

Meetodi eesmärgiks on vähendada andmematriksi mõõtmelisust ja visualiseerida seda kahemõõtmelisena (Nenadic & Greenacre 2007). CA sisendiks on tavaline sagedustabel, mille veergudes ja ridades paiknevate tunnuste vahelisi seoseid soovitakse uurida (käesolevas töös murded ja verbiühendite sagedused). Meetodi eeliseks on objektiivsus, kuna statistikaprogrammis sisestatakse n-ö toores sagedusandmestik seda enne kuidagi lingvistiliselt analüüsimata, sisuliselt teeb programm tuhandeid võrdlusi, püüdes nendes seast tuvastada kõige olulisemaid (Chambers & Trudgill 1998: 105).

Järgnevalt kirjeldan korrespondentsanalüüsi meetodit ja selle terminoloogiat oma andmestikust kunstlikult moodustatud näiteandmestiku põhjal¹¹, mille esitab tabel 7. Veergudes on uuritavad ühendid: infinitiivivormi formatiiv koos samas osalauses esineva finiitse verbiga. Vasakult paremale on uuritavad ühendid *ma*-infinitiiv + *pidama*, *tud*-partitsiip + *saama*, *ma*-infinitiiv + *hakkama*, *da*-infinitiiv + *olema*, *ma*-infintiiv + *minema* ning *ma*-infinitiivi inessiiv + *käima*. Ridades on vastavate ühendite sagedused kõigis murretes. Kasutan meetodi selgitamiseks väikest andmestikku, kuid peamine eelis on just võimalus analüüsida suuri sagedustabeleid, millest palja silmaga mingit struktuuri tuvastada on sageli võimatu. Tabelit võib tõlgendada kui kümmet punkti (murded) kuue mõõtmelises (verbiühendid) ruumis.

¹¹ Kõigi selgituste puhul olen järginud analoogiat, mis on esitatud (Greenacre 2007; Lebart *et al.* 1998 ptk 3; Abdi & Williams 2010).

Tabel 7. Korrespondentsanalüüsi näiteandmestik.

	ma_pidama	tud_saama	ma_hakkama	inf_olema	ma_minema	mas_käima	Σ
RANNA	96	181	173	19	52	26	547
IDA	84	28	124	20	54	33	343
SAARTE	176	192	215	16	40	44	683
KESK	167	212	165	25	51	23	643
MULGI	162	21	34	5	21	14	257
KIRDE	184	21	134	18	35	27	419
SETO	65	11	108	9	114	0	307
TARTU	61	16	82	7	51	4	221
LÄÄNE	153	186	140	26	43	34	582
VÕRU	80	19	108	5	50	7	269
Σ	1228	887	1283	150	511	212	

Sellest andmestikust moodustub maatriks X , millel on $I = 10$ rida ja $J = 6$ veergu, ning see on võrdne:

X	$=$	96	181	173	19	52	26
		84	28	124	20	54	33
		176	192	215	16	40	44
		167	212	165	25	51	23
		162	21	34	5	21	14
		184	21	134	18	35	27
		65	11	108	9	114	0
		61	16	82	7	51	4
		153	186	140	26	43	34
		80	19	108	5	50	7

X on CA sisendiks, kus read tähistavad murdeid ning veerud uuritavat verbiühendit ning rea ja veeru lõikekohas on verbiühendi esinemissagedus vastavas murdes. Olulisem on aga, kui palju mingis murdes uuritavat ühendit suhteliselt kasutatakse, selleks arvutatakse ridade **profiilid** (suhtelised sagedused), jagades rea kogusumma (kõigi ühes murdes kasutatavate ühendite summa) iga ühendi esinemissagedusega vastavas murdes. Niimoodi saadakse ridade profiilide maatriks R , mis on võrdne:

R	$=$	0,176	0,331	0,316	0,035	0,095	0,048
		0,245	0,082	0,362	0,058	0,157	0,096
		0,258	0,281	0,315	0,023	0,059	0,064
		0,260	0,330	0,257	0,039	0,079	0,036
		0,630	0,082	0,132	0,019	0,082	0,054
		0,439	0,050	0,320	0,043	0,084	0,064
		0,212	0,036	0,352	0,029	0,371	0,000
		0,276	0,072	0,371	0,032	0,231	0,018
		0,263	0,320	0,241	0,045	0,074	0,058
		0,297	0,071	0,401	0,019	0,186	0,026

Rannamurdes on näiteandmestikus kokku 547 ühendit ja *ma*-infinitiiv + *pidama* esineb seal 96 korda, seega saadakse profiilimaatriksi esimene väärtus tehtega $96/547 = 0,176$. Samamoodi tehakse arvutused iga lahtri jaoks ja nii saadakse iga murde (rea) profiil. Maatriksi R esimene rida on rannamurde profiil üle kõigi andmestikus olevate verbiühendite. Keskmise ridade profiili vektor (CA terminoloogias tsentroid) on $[0,2875 \ 0,2077 \ 0,3004 \ 0,0351 \ 0,1196 \ 0,0496]$, mis saadakse iga verbiühendi sageduse jagamisel kogusageduste rea kogusummaga. Keskmise profiil näitab verbiühendite suhtelise kasutussageduse keskmist, s.t kui kõik murded oleksid sarnased, oleks *ma*-infinitiivi ja *pidama* ühendi

suhteline sagedus näiteandmestikus 0,2875, *tud*-partitsiibi ja *saama* suhteline sagedus 0,2077 jne. Uurimaks murretevahelisi erinevusi, tuleks vaadelda iga murde häälbimist n-ö keskmise murde suhtes ehk häälbemaatriksit **Y**:

$$\mathbf{Y} = \begin{bmatrix} -0,112 & 0,123 & 0,016 & 0,000 & -0,025 & -0,002 \\ -0,043 & -0,126 & 0,061 & 0,023 & 0,038 & 0,047 \\ -0,030 & 0,073 & 0,014 & -0,012 & -0,061 & 0,015 \\ -0,028 & 0,122 & -0,044 & 0,004 & -0,040 & -0,014 \\ 0,343 & -0,126 & -0,168 & -0,016 & -0,038 & 0,005 \\ 0,152 & -0,158 & 0,019 & 0,008 & -0,036 & 0,015 \\ -0,076 & -0,172 & 0,051 & -0,006 & 0,252 & -0,050 \\ -0,012 & -0,135 & 0,071 & -0,003 & 0,111 & -0,032 \\ -0,025 & 0,112 & -0,060 & 0,010 & -0,046 & 0,009 \\ 0,010 & -0,137 & 0,101 & -0,017 & 0,066 & -0,024 \end{bmatrix}$$

CA-s omistatakse igale reale kaal (inglise keeles on eelistatud termin *mass*) (Greenacre 2007: 21). Iga rea kaal näitab selle rea tähtsust andmestikus. Ridade kaalud arvutatakse nii, et *i*-nda rea väärtuste kogusumma jagatakse kõigi ridade väärtuste kogusummaga ning rea massi vektor on seega $\mathbf{m} = [0,128 \ 0,080 \ 0,160 \ 0,151 \ 0,06 \ 0,098 \ 0,072 \ 0,052 \ 0,136 \ 0,063]$, kaalu väärtused vastavalt ranna-, ida-, saarte, kesk-, Mulgi, kirde-, Seto, Tartu, lääne-, ja Võru murde jaoks.

Profiilide vahelisi kaugusi mõõdetakse CA-s χ^2 (hii-ruut) kauguse abil, mis arvutatakse reaprofiilide väärtustest järgmise valemiga:

$$d_{i,i'}^2 = \sum_j^J w_j (r_{i,j} - r_{i',j})^2$$

Profiilide varieerumist tsentroidi suhtes näitab inerts (*inertia*), mis on sarnane dispersiooniga. Ridade inerts oma tsentroidi suhtes on ridade kauguste (tsentroidist) ruutude kaalutud summa. Kõrge inerts väärtus näitab ridade profiilide suurt häälbimist keskmisest ja väike inerts vähesemat häälbimist keskmisest. Kõrge inerts tähendab suurt ridade ja veergude vahelist seost ning näiteandmestiku põhjal tähendaks see seda, et verbiühendi kasutussagedus sõltub murdest ehk verbiühendite kasutus ja nende kasutussagedus on murretega seotud.

Kõik eelpool kirjeldatud arvutused tehakse analoogselt ka veergude jaoks. CA analüüsi käigus tehaksegi kaks kaarti eraldi ridade ja veergude jaoks ning tulemused asetatakse piltlikult öeldes ühel kaardil teineteise peale. Seetõttu pole ridade ja veergude punktide vahelised kaugused kaardil üksüheselt tõlgendatavad, kuid kaart on siiski nende kauguste võimalikult täpne üldistus, mis annab ühele punktikogumikule teise suhtes ruumilise tõlgenduse.

Kaks olulist mõistet CA-s on veel **panus** (*contribution*) ja **korrelatsioon** (*correlation*), mis esitatakse ruutväärtustena. Panus näitab, kui suur on vaadeldava punkti tähtsus vaadeldaval teljel suhtes kõigi teiste punktidega. Ruudus korrelatsioon näitab jälle, kui hästi antud telg elemendi positsiooni analüüsis seletab.

Joonis 7 esitab näiteandmestiku põhjal tehtud kahemõõtmelise kaardi, s.t esitatud mitmemõõtmeline andmestik on taandatud kahemõõtmeliseks. Teljed tähistavad kahte esimest kõige olulisemat mõõdet ja paiknevad nende keskmiste profiilide kohal. Punktid (murded ja verbiühendid) paiknevad ruumiliselt keskmise profiili suhtes. Mida lähemal on punkt keskmisele, seda sarnasem ta keskmisega on ja vastupidi. Kaart esitab kaks kõige olulisemat mõõdet: horisontaalne seletab üle 57% andmestiku variatiivsusest ja vertikaalne üle 31%, s.t et ainult kaks esimest mõõdet seletavad kokku juba üle 88% variatiivsusest ning ülejäänud mõõtmekokku 12% ringis. Joonise keskel kaks tumedamat joont, telge tähistavad mõlema mõõtme tsentroide.

Esimese, horisontaalse mõõtme põhjal tekib kaks selgemat gruppi. Esimese neist moodustavad lääne-, saarte, kesk- ja ranna murre ning teise Mulgi, Seto, Tartu, Võru, ida- ja kirdemurre, kusjuures teise grupi sisene varieeruvus on natuke suurem ja teistest erinevam näib siin olevat Seto murre. Teise, vertikaalse mõõtme põhjal eristuvad esiteks Mulgi ja kirdemurre ning teiseks kõik ülejäänud murded, kuid selle mõõtme põhjal pole varieeruvus nii suur. Kahe mõõtme lugemist lihtsustab ehk see kui vaadata mõlemat mõõdet kui venimist vastavalt kas horisontaalselt või vertikaalselt. Esimese mõõtme põhjal peaks punkte kaardil n-õ venitama horisontaalelt ja teise mõõtme põhjal vertikaalselt.

Jooniselt võib näha, et kasutatud fiktiivse andmestiku põhjal on *tud-*partitsiibi ja *saama* ning *mas-*vormi ja *käima* ühendi kasutus iseloomulikum läänepoolsetele murretele. *ma-*infinitiivi ja *pidama* ühend on iseloomulikum kirdemurdele, supiini ja finiiitse *minema* ühend aga Seto murdele. *da-*infinitiivi ja *olema* ühend paikneb üsna keskel, mis lubab oletada selle ühtlaselt sarnast kasutust kõigis murretes. Oluline on silmas pidada, et kaart ei esita sagedusandmestikku, vaid analüüsitava tabeli väärtuste vahelisi seoseid.

Näide korrespondentsanalüüsi graafikust



Joonis 7. Korrespondentsanalüüsi kaart näiteandmestiku põhjal.

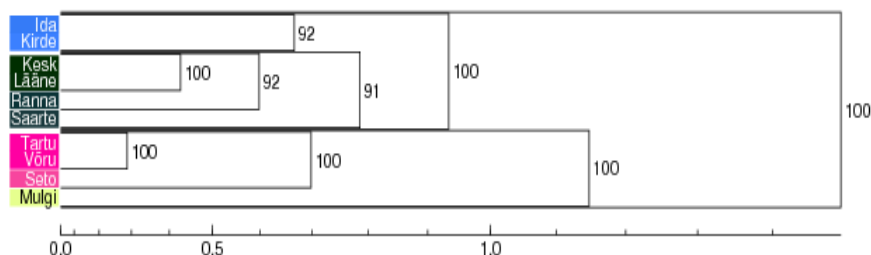
6.4 Hägus klasteranalüüs

Klasterdamismeetoditest rakendan selles töös hägusat klasteranalüüsi. Nii nagu eelmises peatükis juba mainitud, kuuluvad klasteranalüüs ja korrespondentsanalüüs erinevatesse klassifitseerimismeetodite rühma, mis on teineteist täiendava iseloomuga. Kuna siinses töös rakendan rohkem korrespondentsanalüüsi ning klasteranalüüsi kasutangi ainult täiendava meetodina, kontrollimaks kas analüüsitulemused on sarnased korrespondentsanalüüsi tulemustega, ei ole järgnevalt klasteranalüüsi nii detailselt kirjeldatud kui eelnevalt korrespondentsanalüüsi.

Klasteranalüüs on katusermin suurele hulgale meetoditele, mille ühine joon on püüe tuvastada andmestikust uuritavate tunnuste põhjal väiksemaid gruppe, mis üksteisest eristuvad (Everitt & Hothorn 2011: 165; Everitt *et al.* 2011). Tulemused esitatakse reeglina puukujulise dendrogrammina. Mõõtmete vähendamismeetodid (nagu eelnevalt kirjeldatud korrespondentsanalüüsi) ja klasterdamismeetodid (klasteranalüüs) on teineteist täiendava iseloomuga ja mõlemat on võimalik rakendada täpselt samal andmestikul (Lebart *et al.* 1998: 46). Klasteranalüüs on dialektoloogias soositud meetod ja leidnud palju rakendamist (Goebel 1982; Goebel *et al.* 1984; Heeringa 2004; Nerbonne *et al.* 2008). Töös tehtud klasteranalüüs on läbi viidud vabavaralise programmiga Gabmap (Nerbonne *et al.* 2011), kuna R ei paku hetkel veel meetodi visualiseerimiseks nii häid võimalusi ja neid võib kohati olla keerulisem tõlgendada.

Selles töös kasutan hägusa klasteranalüüsi meetodit. Hägusa klasteranalüüsi eelis tavaliste klasteranalüüsi meetodite ees on selle meetodi parem sobivus keeleandmestikule. Kui lihtsamad meetodid eeldavad mingi objekti kuulumist või mittekuulumist mingisse rühma, siis hägusa klasteranalüüsi käigus antakse igale objektile pigem mingi tugevuse väärtus, mis näitab, kui tugevalt või nõrgalt see element vaadeldava rühmaga seotud on. Hägusa klasteranalüüsi rakendamisel tuleb uurijal endal ette anda rühmade arv, mille kohta ta andmestiku põhjal analüüsi saada soovib. (Everitt *et al.* 2011: 242) Algoritmi põhjal moodustatakse korduvalt etteantud arv gruppe ning igal iteratsioonisammul lisatakse andmestikku juhuslikku andmemüra (Nerbonne *et al.* 2008).

Joonis 8 esitab dendrogrammi, mis on tehtud eelmise peatüki tabelis 7 esitatud näiteandmestiku põhjal. Programmile on eelnevalt tingimuseks seatud kahe suurema grupi moodustamine.



Joonis 8. Näiteandmestiku hägusa klasteranalüüsi visuaalne tulemus.

Numbrid iga haru otstes osutavad protsentidele kordamiste arvust, mitu korda selline rühm kordamiste jooksul ilmnes. Kõik rühmad, mille ilmnemise protsent on üle 90, on stabiilsed, ja võib väita, et andmestiku põhjal tõepoolest selline rühm selgelt eristub. Kõik rühmad, mille ilmnemise protsent jääb alla 90, on pisut ebaselged, ning nende kohta ei saa tulemuste põhjal lõplikke järeldusi teha. Jooniselt näeb, et kõik moodustatud rühmad on väga stabiilsed. Siin tuleb silmas pidada, et tegemist on kunstliku andmestikuga ning reaalses analüüsis on klassifitseerimistulemused harva nii ilmsed. Dendrogramm näitab, et andmestikust on moodustatud kaks suuremat gruppi, millest esimesse kuuluvad ida-, kirde-, kesk-, lääne-, ranna- ja saarte murre ning teise Tartu, Võru, Seto ja Mulgi murre. Need on kaks suurt rühma, mis omavahel selgelt eristuvad. Kahe suurema rühma sees on omakorda väiksemad, milles omakorda võib veel olla väiksemaid rühmi.

6.5 Geograafilise signaali tugevus

Kvantitatiivse analüüsi andmestiku kvaliteedi mõõtmiseks kasutan kahte mõõdikut. Esimene neist, Cronbachi alfa (Cronbach 1951), mõõdab andmete usaldusväärsust ja näitab, kui suur on vaadeldavate tunnuste omavaheline seos¹². Seda suurust käsitlen töös geograafilise signaali tugevusena, mis näitab, kui usaldusväärseks saab pidada mõõdetava andmestiku põhjal tehtava statistilise analüüsi tulemusi. Cronbachi alfa väärtus on alati nulli ja ühe vahel ning mida kõrgem on Cronbachi alfa väärtus, seda tugevam on uuritavate tunnuste vaheline seos (siinses töös tugevam geograafiline signaal). Sotsiaalteadustes on üldiselt aktsepteeritav piirväärtus 0.7 (Nunnally 1978). Põhjalikuma ülevaate mõõdiku rakendamisest dialektoloogias annab Heeringa (2004: 170–177).

Teise kvaliteedimõõdikuna kasutan kohaliku seosetuse (*local incoherence*) mõõdikut¹³, mis mõõdab, kui sidus andmestik on (Nerbonne & Kleiweg 2007). Mida väiksem on mõõdiku väärtus, seda sidusam ja usaldusväärsem on andmestik. Üldiselt aktsepteeritav piirväärtus on 0,3. Käesoleva töö andmestikus see mõõdik küll väga suure seletusjõuga ei ole, kuna arvutamisel võetakse arvesse ka geograafilisi kaugusi. Kuna selles töös vaadeldakse 10 suhteliselt suurt geograafilist piirkonda, siis geograafiline punkt kaardil, mille põhjal arvutused tehakse on mõnevõrra suvaline. Mõõdik omab suuremat seletusjõudu paljude väiksemate alade vaatlemisel, kus geograafilist kaugust on võimalik täpsemalt määrata.

Sarnaselt klasteranalüüsi rakendamisega kasutan Cronbachi alfa ja kohaliku seosetuse arvutamiseks programmi Gabmap (Nerbonne *et al.* 2011), kuna koefitsiendid arvutatakse klasteranalüüsi andmete sisestamisel automaatselt.

¹² Cronbachi alfa arvutusvalemi ja rakendusviisi kohta Gabmapis vt täpsemalt veebiaadressilt: <<http://www.gabmap.nl/~app/bin/help?s=ca>>

¹³ Kohaliku seosetuse mõõdiku arvutusvalemi ja rakendusviisi kohta Gabmapis vt täpsemalt veebiaadressilt: <<http://www.gabmap.nl/~app/bin/help?s=linc>>

6.6 Verbiühendite produktiivsuse mõõtmisest selles töös

Verbiühendite (ja võimalusel konstruktsioonide) produktiivsuse mõõtmiseks kasutan peatükis 6.2. kirjeldatud kollostruktsiooniliste meetodite võimalusi; varem on meetodit sarnaselt rakendanud Muischnek ja Sahkai (2009; 2010). Meetodi rakendamistehnika on samasugune nagu seda on kirjeldatud peatükis 6.2.

Produktiivsuse kvantitatiivsel mõõtmisel tuli andmestikust lähtudes teha järeleandmisi. Sageli on konkreetseid konstruktsioone kõigi murrete peale ära jagades üsna vähe, seega pole produktiivsuse mõõtmine murde ja konstruktsiooni tasandil mõttekas. Kobaranalüüsile iseloomulikult on siingi andmeid koondatud ning produktiivsust mõõdan kahes murderühmas¹⁴ üle verbiühendite (mitte konstruktsioonide). Konstruktsioonide produktiivsus on võrdne verbiühendi produktiivsusega, juhul kui vastava ühendiga moodustub ainult üks konstruktsioon, s.t et puuduvad homonüümsed konstruktsioonid.

Eri ühendite produktiivsuse mõõtmiseks aga ei mõõda ma nüüd infiniitse verbi morfoloogilise kategooria ja finiiitse verbi lemma vahelist seost, vaid kaasan ka infiniitset vormis esineva verbi enda lemma. Mõõdan statistilist tugevust konstruktsiooni moodustava infiniitse verbi ja finiiitse verbi vahel. Nüüd pole ainult oluline, kui tugevasti need kaks üksust statistiliselt seotud on, vaid oluline on saadud väärtuste jaotumine. Rakendan produktiivsuse mõõtmisel peatükis 6.2 kirjeldatud eristavate kolekseemide analüüsi meetodit, aga erinevusi ei vaadelda siin kui kahe või enama konstruktsiooni semantilisi erinevusi, vaid eristamise aluseks on kaks murderühma. Seega mõõdan infiniitsete verbide seost konkreetse finiiitse verbiga kahes murderühmas. Tulemused näitavad, kui tugevasti vastava finiiitse verbiga on vaadeldavas rühmas seotud konkreetne infiniitsetes positsioonis esinev verb.

Järgin produktiivsuse käsitlust, mille kohaselt näitab verbiühendi (konstruktsiooni) produktiivsust erinevate verbide hulk, mis ühendis (konstruktsioonis) esineda võivad (Croft & Cruse 2004: 309). Siinses töös mõõdan seda siiski kaudselt: mida madalamad on statistilise seose väärtused, seda rohkem on andmestikus erinevaid verbe, mis konkreetse finiiitse verbiga võivad kombineeruda, s.t ühend on vähem kinnistunud (pole püsiühend) ning ühend on eelduste kohaselt produktiivsem. Tabel 8 esitab statistilise seose tugevuse väärtused koos finiiitse verbiga *saama* koosesinevad verbid *tud*-partitsiibis kahes murderühmas. Mida kõrgem väärtus, seda tugevamalt on kaks verbi omavahel seotud, kõrge seose väärtus osutab omakorda sellele, et kaks verbi on üsna püsiühendilaadsed. Mida madalam väärtus (ja mida rohkem madalamaid väärtusi), seda nõrgemalt on kaks verbi omavahel seotud ning seda rohkem verbe võib infiniitsetes positsioonis esineda ehk ühend on produktiivsem. Tabelis 8 on kõrgemad väärtused idarühmas, mis näitab verbide tugevamat seotust. Läänerühmas on väga tugevalt *saama* verbiga seotud vaid *panema* ning seose

¹⁴ Kahe rühma moodustamise aluseid tutvustan peatükkides 7, 8 ja 9.

tugevuse väärtused lähevad üsna kiiresti madalaks ja jaotuvad üsna ühtlaselt, järelikult võib *tud*-partitsiibi vormis olla palju erinevaid verbe ning ükski neist pole finiiitse *saama*-ga väga tugevalt seotud ehk vaadeldav ühend on üsna produktiivne.

Tabel 8. Koos finiiitse *saama*-verbiga esinevad verbid *tud*-partitsiibis ja statistilise seose tugevuse väärtus.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
kolkima	2.44	panema	2.3
pesema	2.24	vedama	0.66
elama	1.98	keetma	0.6
kastma	1.98	andma	0.46
olema	1.73	hakkama	0.46
puistama	1.7	tõmbama	0.46
asutama	1.22	laskma	0.44
harima	1.22	lõikama	0.41
koolitama	1.22	püüdma	0.33
matma	1.22	minema	0.31

Kõik selles ja järgnevates tabelites peatükis 9 esinevad verbid on kirjakeelestatud kujul. Murdekorpuse märgendamisel püütakse anda märgendatavale sõnale n-ö sõnaraamatu vorm, mille tarvis märksõnu püütakse normaliseerida. Analüüsis kasutan neid kirjakeelestatud sõnavorme.

6.7 Kartograafia

Geograafilise keeleuurimise lahutamatu osa on uuritava materjali ja uurimistulemuste visualiseerimine kaartidel. Eesti murdeandmestiku kõige põhjalikumad kaardid on esitanud Saareste (1938a; 1938b). Detailsed levikukaardid on olemas lõuna-eesti murrete erinevate häälikuliste ja vormiliste nähtuste kohta (Toomse 1998).

Eesti murdeandmestiku automaatne visualiseerimine ei ole triviaalne ülesanne. Praeguseks on olemas terve hulk rakendusi, mis võimaldavad automaatselt kaardiandmestikku oma mis tahes uurimistöös hõlpsalt kasutada. Eesti murrete visualiseerimise põhiprobleem on aga, et traditsioonilised murdepiirid ei järgi tänapäeva haldusterritoriaalseid piire ning seetõttu pole võimalik kasutada laialdast kaardiavalikut, mis on koostatud tänapäeva või lähimineviku haldusjaotuste põhjal.

Eesti murrete andmestiku visualiseerimiseks olen välja töötanud automaatse rakenduse, mis võimaldab uurimisandmestiku automaatselt murdealade kaupa esitada. Kõik kaardid on genereeritud automaatselt statistikaprogrammiga R (R Development CoreTeam 2011). Väljatöötatud rakendust on tulevikus võimalik kasutada ükskõik millise dialektoloogilise andmestiku automaatseks esitamiseks. Samuti on raamistiku põhimõtteid võimalik kasutada kõigi vähemuskeelte, k.a eesti lähisugukeelte andmetike visualiseerimiseks ehk selliste keealade visualiseerimiseks, mille rääkijate paiknemine ei järgi tänapäeva haldusterritoriaalseid piire.

Sagedusandmestiku esitamiseks on kasutatud kaarte, kus on esindatud kümme uuritavat murret ja esitatavas konstruktsiooni sagedus on seotud ala tumedusastmega kaardil. Mida tumedam on murdeala kaardil, seda sagedasem on esitatud konstruktsioon sellel alal. Kõige kõrgema konstruktsiooni sagedusega ala on kõige tumedam ning tumedusaste muutub sageduse langusega ja kõige madalama sagedusega ala on kõige heledam.

7 KVANTITATIIVSE KOBARANALÜÜSI TULEMUSED

Järgnevalt esitan kvantitatiivse analüüsi tulemused. Esmalt kirjeldan kvantitatiivse analüüsi läbiviimiseks tehtud eelsamme, mis olid vajalikud heterogeense ja ebastandardse uurimismaterjali võrreldavale kujule viimiseks, järgneb korrespondentsanalüüsi ja hägusa klasteranalüüsi tulemuste kirjeldus. Analüüs on läbi viidud kahe natuke erineva andmestiku peal. Esimene neist on lihtne normaliseeritud sageduste analüüs, kus sisendmaterjaliks on kõik tuvastatud samas osalauses esinenud $V_{fin}+Inf$ ühendid, v.a vähem kui kolm korda esinenud. Teine andmestik kaasab kollostruktsioonilise analüüsi tulemused, s.t et materjal koosneb verbiühenditest, mida esines materjalis vähemalt kolm korda ning mis ületasid FTT-le seatud normaliseeritud kriitilise väärtuse piirväärtuse¹⁵. Kasutatud sagedusandmestikud on esitatud lisades 2 ja 3.

7.1 Sageduste normaliseerimine

Nagu peatükis 4 esitatud murdekorpuse arvandmetest ilmneb, on uuritavad murdealad korpuses esindatud erineva sõnade arvuga. See omakorda teeb keeruliseks sagedusandmete omavahelise võrdlemise: pole võimalik võrrelda keeleüksuse sagedusi, mis on leitud näiteks 1000- ja 100 000-sõnalisest materjalist. On igati loogiline, et 100 000-sõnalisest korpuses esineb tõenäoliselt ka uuritavaid keeleüksusi tunduvalt rohkem kui väiksemas valimis. Et aga sagedusi oleks võimalik omavahel võrrelda, on need normaliseeritud. Normaliseerimine on tehtud keskmise korpuse suuruse põhjal (61 312 sõna). Näiteks supiini ja finitise *minema*-verbi ühendit esines idamurdes 39 korda. Kogu idamurde sõnade arv korpuses oli 43 965. Ühendi normaliseeritud sagedus idamurdes saadakse tehtega $39 \times 61312/43965 = 54$. Ehk siis ühendi lihtsagedus korrutatakse korpuse keskmise sõnade arvuga ning tulemus jagatakse vastava murde sõnade arvuga korpuses. Supiini ja finitise *minema*-verbi normaliseeritud sagedus idamurdes on seega 54. Sel viisil on normaliseeritud kõik vaadeldavate verbiühendite sagedused kõigis murretes eesmärgiga teha need omavahel võrreldavateks. Lisaks on kogu kvantitatiivsest analüüsist eemaldatud $V_{fin}+Inf$ ühendid, mis esinesid korpuses vähem kui kolm korda. Niisiis on edaspidi sagedustest rääkides, mõeldud ainult **normaliseeritud sagedusi**, kui pole mainitud teisiti, ning sagedusi, mille (normaliseerimata) lihtsagedus oli uuritavas materjalis vähemalt kolm.

Korpuslingvistikas on selline sageduste normaliseerimine üldlevinud praktika. Reeglina tehakse seda küll olenevalt materjali suurusest ja töö eesmärkidest kas 100, 1000 jne sõna põhjal, nt mitu korda mingi sõna, vorm, ühend jne 1000 sõna kohta korpuses esines. Siinses töös olen kasutanud

¹⁵ FTT-väärtused on normaliseeritud tehtega $-\log(\text{FTT}, 10)$, selle kohta vaata täpsemalt: (Gries *et al.* 2005: 648).

normaliseerimisbaasina korpuse keskmist suurust ja iga murde sõnade arvu, mis eelduste kohaselt aitab arvestada kõiki neid suurusi ja annab võrreldavamad tulemused. Samasugust normaliseerimispõhimõtet on rakendatud ka konstruktsioonide tähenduste sageduste normaliseerimisel peatükis 9.

7.2 Korrespondentsanalüüsi tulemused

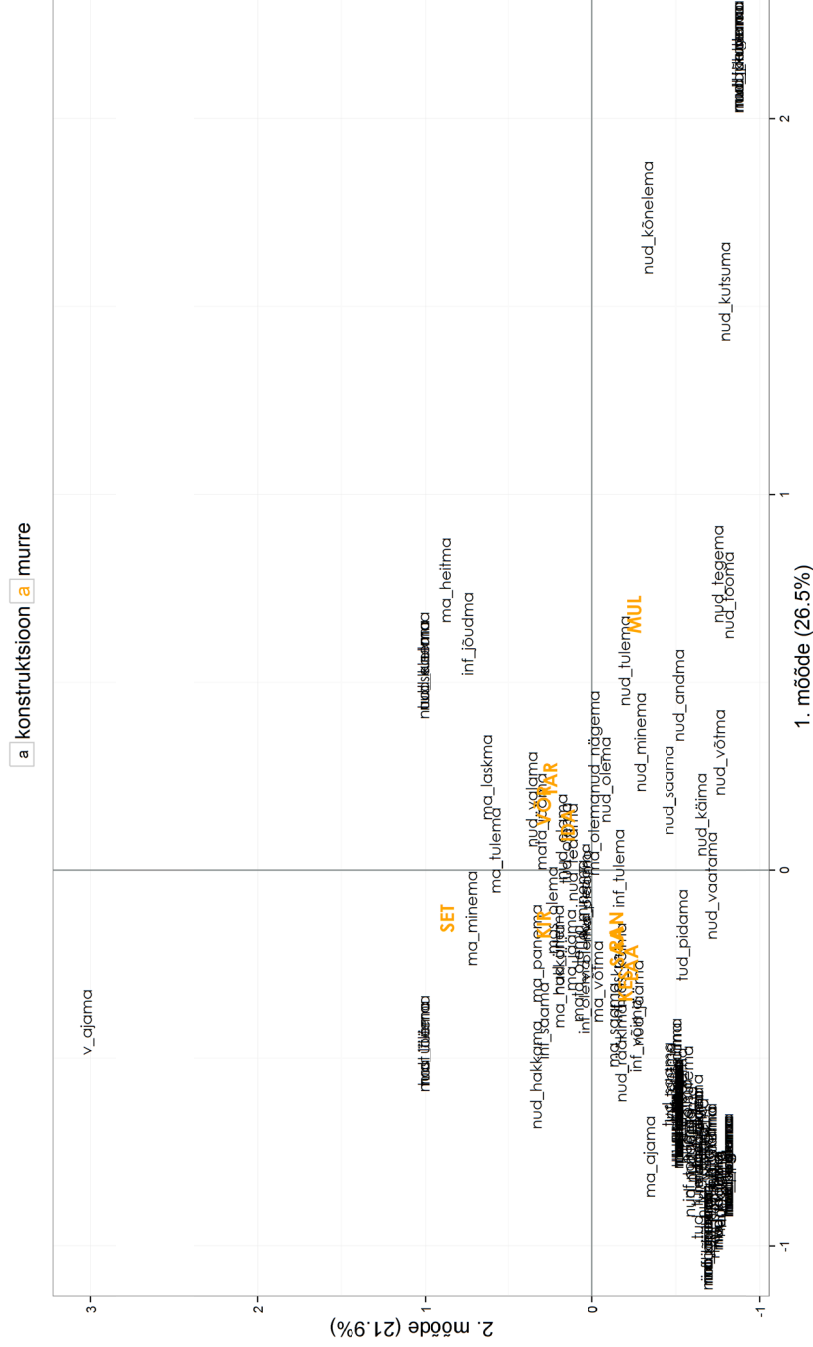
Normaliseeritud sageduste analüüs

Sageduse positsioon keelenähtuste uurimisel ja selgitamisel on mõnevõrra ebaselge (Arppe *et al.* 2010). Siiski on viimasel ajal see hakanud saama järjest rohkem tähelepanu väga erinevate keelenähtuste uurimisel (Bybee 2007; Bybee & Hopper 2001; Divjak & Gries 2012; Gries & Divjak 2012; Szmrecsanyi ilmumas (c) paljude teiste seas). Esimese analüüsi sisendiks on finiiitse verbi ja infiniitse verbi morfoloogilise kategooria ($V_{fin}+Inf$) paaride normaliseeritud sagedused. Esimene analüüs eeldab seega, et kui finiiitne ja infiniitne verb esinevad samas osalauses, moodustavad nad konstruktsiooni ning vastupidi, kaks vaadeldavat vormi moodustavad konstruktsiooni, kui nad esinevad samas osalauses. Selline lähenemine on kindlasti lihtsustatud, kuid analüüsi jätkamist julgustasid andmestiku head kvaliteedimõõdikud (vt täpsemalt ptk 6.5). Cronbachi alfa oli 0,8 ning kohalik seosetus 0,1, mis tähendab, et normaliseeritud sageduste põhjal moodustatud andmestikus oli väga tugev geograafiline signaal (0,8) ning andmestiku hajusus oli madal (0,1).

Sellise lähenemise eelis on kindlasti võimalus kaasata analüüsi võimalikult palju potentsiaalseid konstruktsioone. Samas jääb analüüsi ka palju nn müra, mis on tingitud peamiselt osalausestamise vigadest (vt täpsemalt ptk 5). Kui osalausestaja ei suutnud seada mitme järjestikuse osalause vahele õigeid piire, pidades neid üheks osalauseks, toimub ühendite tuvastamisalgoritmi põhjal verbiühendite ülegenereerimine. Seda probleemi leevendab mõnevõrra madala esinemissagedusega ühendite ($O < 3$) eemaldamine analüüsist. Peale vähem kui kolm korda esinevate ühendite eemaldamist jäi lõplikku andmestikku 120 erinevat ühendit.

Joonis 9 illustreerib korrespondentsanalüüsi (vt meetodi kirjeldust ptk 6.3) tulemusi verbiühendite ja nende normaliseeritud sageduste põhjal. Murded lühenditena on esitatud suurtähelistena tumekollaselt ning ühendid väiketähelistena musta värviga. Mida kaugemal vaadeldavad üksused üksteisest kahemõõtmelisel kaardil paiknevad, seda erinevamad nad on. 1. ja 2. mõõde (x- ja y-telg) tähistavad esimese kahe mõõtme inerts suurust (võrreldav seletatud variatiivsusega), mida kumbki mõõde seletab.

Korrespondentsanalüüsi graafik normaliseeritud sageduste põhjal



Joonis 9. Korrespondentsanalüüsi kaart normaliseeritud sageduste põhjal. Lühendid: IDA – idamurre, KES – keskmurre, KIR – kirdemurre, LÄÄ – läänemurre, MUL – Mulgi murre, RAN – rannamurre, SAA – saarte murre, SET – Seto murre, TAR – Tartu murre, VÕR – Võru murre.

Esimene mõõde eristab selgemalt kahte gruppi. Ühte neist kuuluvad Võru, Tartu ja idamurre ning teise homogeensemata rühma moodustavad kesk-, saarte, ranna- ja läänemurre. Esimese rühmaga on sarnasem Mulgi murre ning teise rühmaga kirde- ja Seto murre. Teine mõõde esitab selgemate gruppidega samad rühmad, kuid nüüd on esimese rühmaga sarnasem kirde ja Seto murre ning Mulgi murre on sarnasem teise rühmaga.

Suur erinevus tundub olevat Mulgi ja Seto murde vahel, mis mõlemad erinevad Tartu ja Võru murdest, kuid väga erineval viisil. Kõige üllatavam on märgatav erinevus Seto ja Võru murde vahel, kuna enamasti loetakse need murded murdeliigendustes samaks (Pajusalu *et al.* 2009).

Korrespondentsanalüüsi kaarti lugemisel tuleb meeles pidada, et kaugused murdealade ja ühendite vahel pole üksüheselt võrreldavad. Korrespondentsanalüüsi graafiku esitamisel moodustatakse esmalt kaks erinevat graafikut, üks ridade jaoks (verbiühendid) ning teine veergude jaoks (murded). Seejärel piltlikult öeldes asetatakse üks graafik teise peale. Ühendite ja murrete vaheline kaugus on üldistus, umbkaudne hinnang, mis üritab siiski uuritavaid üksusi ka ruumiliselt esitada võimalikult võrreldavatena.

Graafikult võib näha ka problemaatilisemaid juhtumeid, mis ilmselgelt ei moodusta konstruktsiooni ja on tingitud osalausestamisvigadest (nud tegema, nud tooma jt), eriti palju on selliseid ühendeid Mulgi murdes (vt ka lisasid 2 ja 3). Selline tulemus pole aga üllatav ja korreleerub hästi varasemate uurimustega. K. Pajusalu (1989) on oma Karksi murraku uurimuses täheldanud, et just selles murrakus on *nud*-partitsiip eriti produktiivne ning võib liituda isegi afiksaaladverbidele. Käesoleva töö uurimismaterjali kogumise ajal oli suurem osa murdekorpuse Mulgi murde tekstide just Karksi murraku alt, seega on *nud*-ühendi rohkus Mulgi murdes seletatav partitsiibi produktiivsusega. Kuna kasutatud algoritm ühendid sel kujul siiski genereeris, pole neid materjalist välja visatud. Nii on võimalik saada terviklik pilt kogu metoodika potentsiaalsetest probleemidest.

Esimese, normaliseeritud sageduste andmestiku põhjal ilmnevad seega selged rühmad, mis osutavad nende murrete suurele sarnasusele verbiühendite kasutuses ja nende kasutussagedustes. Üllataval kombel ei eristu siin selgelt traditsioonilised põhja- ja lõunaeesti murded ning kirde- ja rannamurregi näivad omavahel olevat üsna erinevad. Traditsiooniliselt põhjaeesti murdeks peetud idamurre on siin aga selgelt sarnasem lõunaeesti murretega.

Kollostruktsioonilise analüüsi tulemusi arvestav normaliseeritud sageduste analüüs

Teine korrespondentsanalüüs võtab arvesse ka kollostruktsioonilise analüüsi tulemusi (vt ka ptk 6.2). Sarnaselt eelmise, normaliseeritud sageduste analüüsiga, on selleski analüüsi välja arvatud materjalis vähem kui kolm korda esinenud ühendid. Erinevalt aga eelmisest analüüsist on selles analüüsis välja arvatud veel ühendid, mis ei ületanud FTT-i normaliseeritud kriitilist väärtust 1,3 usaldusnivool 0,05. Niisiis selle analüüsi andmestiku moodustavad verbiühendid, mis said piisavalt kõrge FTT väärtuse, s.t neid võib kollostrukts-

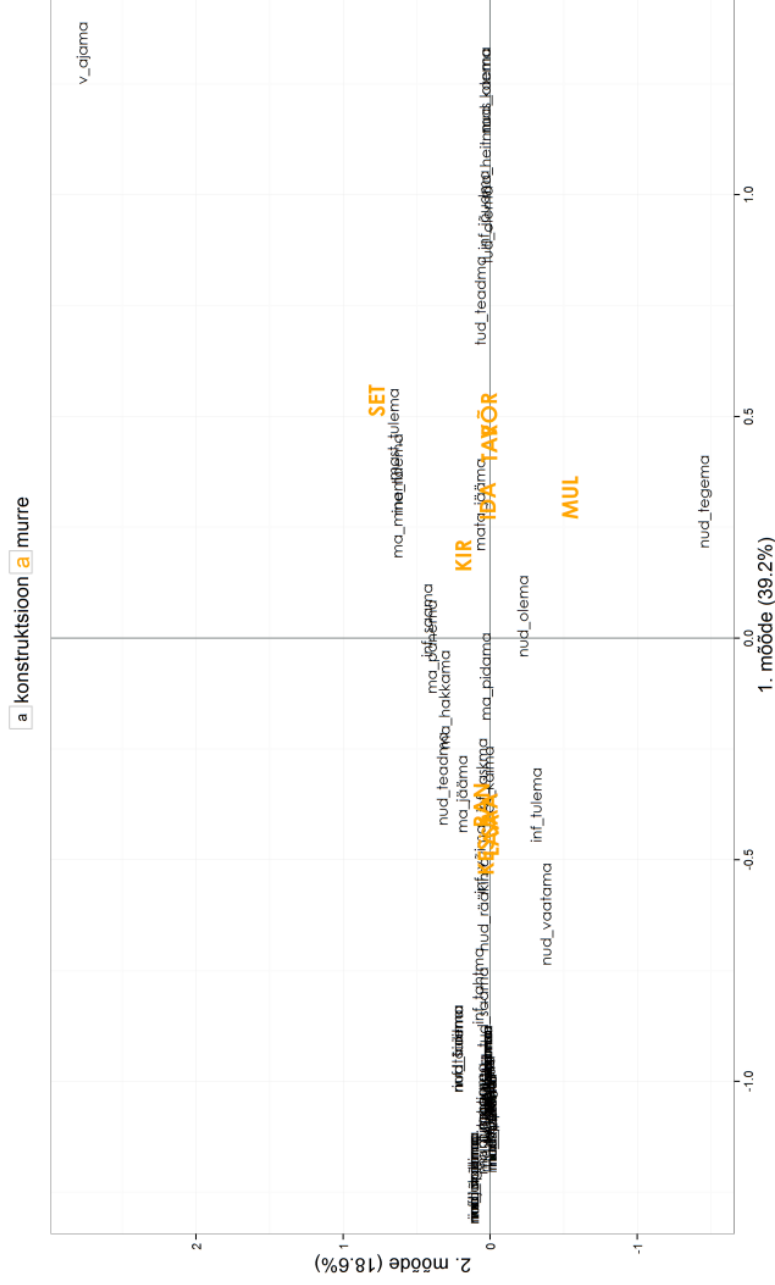
sioonilise analüüsi põhjal pidada statistiliselt konstruktsioonideks. FTT-i kasutatakse siin töös ainult konstruktsioonide tuvastamiseks. Edasi on vaadeldud ainult tuvastatud ühendite (konstruktsioonide) normaliseeritud sagedusi. FTT-i arvutused ei ole tehtud normaliseeritud sageduste põhjal, seega pole FTT-i väärtused murrete lõikes omavahel võrreldavad. Selle analüüsi sammud olid järgmised.

- 1) FTT-i p-väärtused arvutati kõigis murretes eraldi iga $V_{fin}+Inf$ ühendi jaoks. Selle tulemusena valmis kümme eraldi loendit järjestatuna FTT-i normaliseeritud kriitilise väärtuse suuruse järgi. FTT-i väärtused arvutati programmiga Coll.analysis 3.2. (Gries 2007).
- 2) Kriitiline FTT-i väärtus on esitatud 0,05 usaldusnivool. Kui ühendi FTT-i väärtus neile kriteeriumitele ei vastanud, siis ühendit analüüsimaterjali kaasatud ei ole.
- 3) Järgnevalt FTT-i kriitilisi väärtusi enam ei vaadatud. FTT-i on seega kasutatud ainult materjali ekstraheerimiseks.
- 4) Edasi moodustati normaliseeritud sageduste tabelid ühenditele, mida esines materjalis vähemalt kolm korda ning mis ületasid FTT-i kriitilise väärtuse 0,05 usaldusnivool. Niimoodi jäi sellesse analüüsi **57** erinevat ühendit (vt ka lisa 3).

Saadud andmestiku kvaliteedimõõdikud (vt täpsemalt ptk 6.5) olid väga head: Cronbachi alfa 0,8 ning kohalik seosetus 0,1, mis tähendab, et ka selles andmestikus oli väga tugev geograafiline signaal (0,8) ning andmestiku hajusus oli väga madal (0,1). Joonis 10 esitab graafiliselt kollostruktsioonilise analüüsi tulemusi arvestava korrespondentsanalüüsi visuaalsed tulemused. Sellelt jooniselt võib näha, et eelmise analüüsiga võrreldes muutuvad rühmad tunduvalt selgemaks. Kollostruktsioonilised meetodid on aidanud eemaldada suure osa müra (nt paljud *nud*-partitsiibi ühendid).

Kahe mõõtmekoguinerts (võrreldav variatiivsusega) kasvab selles analüüsis samuti (kokku 58%). See osutab tugevamale konstruktsioonide ja murdealade seotusele, s.t et konstruktsioonide kasutus(sagedus) on tingitud murdealast. Joonise vasakul pool moodustavad veelgi selgemalt ühe grupi kesk-, lääne-, ranna- ja saarte murre. Teise grupina eristuvad kõik ülejäänud murded, kuid nende rühmasisene varieeruvus on suurem. Seto ja Mulgi murre on endiselt teineteisest üsna erinevad. Selle analüüsi tulemuste põhjal eristuvad selgelt ida- ja läänepoolsed murded, mis oli aimatav juba ka eelmise analüüsi tulemuste põhjal.

Korrespondentsanalüüsi graafik normaliseeritud sageduste ja FTT väärtuste põhjal



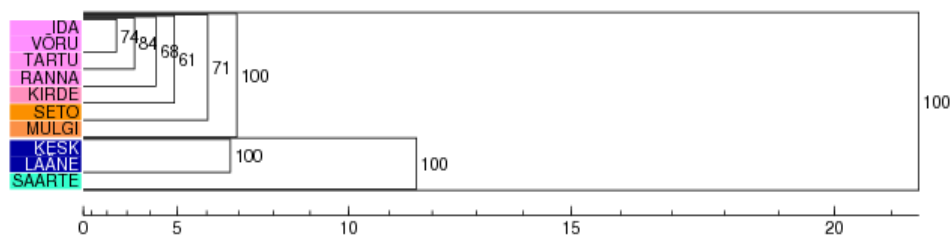
Joonis 10. Korrespondentsanalüüsi kaart kollestruktsioonilise analüüsi tulemuste ja normaliseeritud sageduste põhjal. Lühendid: IDA – idamurre, KES – keskmurre, KIR – kirdemurre, MUL – läänemurre, MUL – Mulgi murre, RAN – rannamurre, SAA – saarte murre, SET – Seto murre, TAR – Tartu murre, VÕR – Võru murre.

On märkimisväärne, et lõuna- ja põhjaeesti murded ei moodusta omavahel selgelt eristuvaid rühmi, nagu seda traditsioonilised murdejaotused esitavad. Ida- ja kirdemurre, mis traditsiooniliselt on liigitatud põhjaeesti murreteks, tunduvad selle analüüsi põhjal olevat vaadeldavate verbiühendite (kasutus)sageduse poolest lähemal lõunaeesti murderühma murretele.

Nii normaliseeritud sageduste kui FTT-analüüsi normaliseeritud sageduste analüüsid osutavad üsna selgelt suuremale ida- ja läänepoolsete murrete vahelisele erinevusele kui põhja- ja lõunaeesti murrete vahelisele erinevusele. Lääne-, kesk-, ranna- ja saarte murre näivad olevat verbiühendite kasutuse ja kasutusageduse põhjal peaaegu identsed. Idamurre on mõlema analüüsi põhjal väga sarnane Tartu ja Võru murdega, FTT-analüüsis on viimastega väga sarnane ka kirdemurre. Mulgi ja Seto murre on nii omavahel kui ka kõigist ülejäänud murretest analüüsides põhjal üsna erinevad. Pean selles osas usaldusväärsemaks teist, FTT-analüüsi, millest nähtub, et mõlemad murded sarnanevad siiski pigem idapoolsete murretega (Tartu, Võru, kirde- ja idamurre) kui läänepoolsetega (kesk-, lääne-, ranna- ja saarte murre).

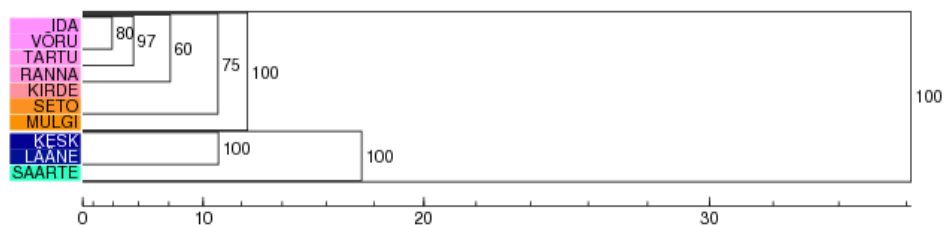
7.3 Hägusa klasteranalüüsi tulemused

Et korrespondentsanalüüsis ilmnenud ida- ja läänepoolsete murrete rühmitumist täpsemalt uurida ja veenduda, et mõne teise meetodiga oleks tulemused sarnased, rakendasin hägusat klasteranalüüsi (meetodi kirjeldust vt ptk 6.4). Hägusa klasteranalüüsi viisin läbi programmiga Gabmap (Nerbonne *et al.* 2011). Analüüsi eesmärk oli saada varieerumisest mitmekülgsemat pilti, kui see vaid üht meetodit rakendades oleks. Andmestikud on täpselt samad, mida kirjeldasin korrespondentsanalüüsi osa juures (vt ka lisa 2 ja 3). Esimene neist arvestab vaid normaliseeritud sagedusi ja ühendeid, mida esines materjalis vähemalt 3 korda ning teine kaasab kollostruktsioonilise analüüsi tulemused ning jätab lisaks madala esinemissagedusega ühenditele välja ka ühendid, mis FTT-testi tulemusena ei osutunud 0,05 usaldusnivool statistiliselt olulisteks.



Joonis 11. Hägusa klasteranalüüsi tulemus normaliseeritud sageduste põhjal.

Joonis 11 esitab hägusa klasteranalüüsi tulemused, kus arvesse võeti ainult verbiühendite normaliseeritud sagedusi. Protsendinumbrid dendrogrammi harudes osutavad sellele, kui suur oli rühma ilmnemise osakaal kogu iteratsiooni-protsessis. Kui väärtus on üle 90 võib rühma pidada väga stabiilseks, madalamad väärtused nii selget signaali ei anna. Sagedusanalüüsis on kaks suuremat rühma väga stabiilsed. Esimese, erakordselt homogeense rühma moodustavad kesk-, lääne- ja saarte murre. Kõik ülejäänud murded moodustavad teise rühma, kuid selle rühma sisene varieerumine on üsna suur ja rühmasised grupeerumised pole eriti suure seletusjõuga (kõiki väiksemaid grupp ilmnes alla 90% protsendi rühmade moodustamise kordadest). Teistest stabiilsema rühma moodustavad ida-, Võru ja Tartu murre, nii nagu see tuli välja ka korrespondentsanalüüsis.



Joonis 12. Hägusa klasteranalüüsi tulemus kollostruktsioonilise analüüsi ja normaliseeritud sageduste põhjal.

Joonis 12 esitab hägusa klasteranalüüsi dendrogrammi, mis arvestab ka kollostruktsioonilise analüüsi tulemusi. Seegi dendrogramm on üsna sarnane eelmisega. Taas eristub väga selgelt grupp, kuhu kuuluvad kesk-, lääne- ja saarte murre. Ülejäänud murded moodustavad teise rühma, kuid rühmasisesed erinevused paistavad olevat siiski suured. Erinevusena võib siin välja tuua eriti tugevalt kokkukuuluvat väiksema rühma, mis moodustub ida-, Võru ja Tartu murdest. Ainult normaliseeritud sageduste klasteranalüüsis see nii selgelt välja ei tulnud.

Mõlemad dendrogrammid esitavad üsna sarnased tulemused, mis suures osas kattuvad korrespondentsanalüüsi tulemustega. Väga sarnased on omavahel kesk-, lääne- ja saarte murre. Ülejäänud murded on omavahel küll sarnasemad, kuid rühmasiselt väiksemaid selgeid grappe ei moodustu. Erandina võib välja tuua ida-, Tartu ja Võru murdest moodustunud üsna homogeense rühma.

7.4 Kvantitatiivse kobaranalüüsi kokkuvõte

Kvantitatiivne kobaranalüüs pakkus huvitavaid tulemusi mitmest aspektist. Esmalt on üsna selged suuremad erinevused idapoolsete ja läänepoolsete murrete vahel kui traditsioonilistes murdjaotustes esitatud põhja- ja lõunaeesti murrete vahel. Selgeid põhja- ja lõunaeesti murrete vahelisi erinevusi statistilises kobaranalüüsis ei ilmnenu. Pigem võib märgata ida- ja kirdemurde suuremat sarnasust lõunaeesti murretega ning nii moodustub selge ida-lääne vaheline erinevus. Peatükk 9 uurib, kas sellised kobaranalüüsi tulemused on ka kuidagi seletatavad erinevate konstruktsioonide tähendusi analüüsides.

Teine olulisem statistilise analüüsi tulemus on lihtsageduse suurepärane toimimine geograafilise varieerumise tuvastamisel. Seega võib väita, et keerulisemaid statistilisi meetodeid rakendamata saab kindla struktuuriga ühendeid kokku lugedes juba üsna hea ruumilise ja lingvistilise pildi. Seda on varem saksa keele prepositsioonist ja verbist moodustatud kollokatsioonide tuvastamisel nentunud ka Krenn ja Evert (2001). Ainult normaliseeritud sagedusandmete analüüs jätab materjali mõistetavalt siiski palju müra ning sagedasemate nähtuste uurimisel aitavad kollostruktsioonilised meetodid selgelt juhuslikku müra vähendada, nii et lõplikku analüüsi jääks võimalikult palju n-ö päriskonstruktsioone.

Korrespondentsanalüüsi tulemused osutasid selgelt, et ühe väga homogeense rühma moodustavad lääne-, kesk-, ranna- ja saarte murre. Hägusa klasteranalüüsi tulemused rannamurde seotust selle rühmaga nii selgelt ei näidanud. Kuna korrespondentsanalüüsi tulemused olid üsna selged ning hägusa klasteranalüüsi puhul polnud rannamurre eriti tugevalt väiksemate rühmadega seotud, arvestan selle järgnevas analüüsis läänepoolsete murrete rühma. Selleks annavad kindlust ka mõned eelnevad uurimused (Uiboed *et al.* 2013). **Idapoolsete murrete rühma** moodustavad siis järgnevas analüüsis **Seto, Tartu, Mulgi, Võru, ida- ja kirdemurre**. **Läänepoolsete murrete rühma** moodustavad **saarte, lääne-, kesk- ja rannamurre**. Mulgi murret võiks taas pidada mõne eelneva uurimuse põhjal (Uiboed 2010a) sarnasemaks läänepoolsete murretega. Siinne kvantitatiivne analüüs osutab küll Mulgi murde erinevusele teistest idapoolsetest murretest, kuid näib nendega siiski sarnasem olevat kui läänepoolsete murretega. Mulgi ja rannamurre on mitmest aspektist üsna huvitavad ja näitavad suhteliselt selgelt üleminekualadele iseloomulikke jooni. See tuleb selgemalt välja järgnevas, konstruktsioonide analüüsis, kus võib märgata nende murrete sarnasust kord ida- kord läänepoolsete murrete rühmaga.

8 KONSTRUKTSIOONIDE ANALÜÜSI PÕHIMÕTTED

8.1 Analüüsitavad verbiühendid

See peatükk kirjeldab konstruktsioonide analüüsi algmaterjali ning põhimõtteid, millest lähtub järgmine, 9. peatükk. Kvalitatiivse iseloomuga konstruktsioonide analüüsi eesmärk on eelkõige kvantitatiivse analüüsi toetamine. Konstruktsioonide analüüsimisega uurin kvantitatiivse analüüsi tulemuste usaldusväärsust ning seda, kas kvantitatiivne analüüs tõepoolest tuvastas verbikonstruktsioone või oli tegemist lihtsalt juhuslikult koosesinevate verbi ja verbivormi paaridega.

Käsitlemist leiavad need verbiühendid (ja nendega moodustunud konstruktsioonid), mis osutusid kvantitatiivses analüüsis statistiliselt oluliseks ja/või olid materjalis piisavalt kõrge esinemissagedusega. Kirjeldamiseks valitud verbiühendid pidid täitma vähemalt kahte kriteeriumi. Esiteks pidid nad esinema automaatselt tuvastatud ühendite loendis ning teiseks pidi nende sagedus kas siis normaliseeritud sageduste tulemuste või FTT-i tulemuste loendis täitma vähemalt ühte järgnevatest sageduskriteeriumitest:

- 1) ühendi absoluutsagedus kõigi murrete koguarvestuses pidi olema vähemalt 100;
- 2) ühendi normaliseeritud sagedus kõigi murrete koguarvestuses pidi olema vähemalt 100;
- 3) ühendi kohta tuvastati kvalitatiivse analüüsi käigus vähemalt 100 lauset.

Viimase kriteeriumi täitmist kontrollisin natuke erineval viisil, kui seda oli ühendite tuvastamise algoritmi põhjal võimalik teha. Kõik nii FTT-analüüsis kui ka sagedusanalüüsis esinenud üendid on ükshaaval läbi vaadatud ning selle kontrolli käigus on kõigi ühendite jaoks otsitud korpusest osalauseid, kus vaadeldav ühend esines. Siin on oluline tähele panna, et näitelausede arv ei pruugi olla vastavuses sagedus- ja FTT-i analüüsi sagedusnumbritega. Seda ebakõla võivad põhjustada erinevad asjaolud. Erinevused sagedusandmestikus võivad tekkida tuvastamisalgoritmi vigadest, s.t et mingil põhjusel (nt osalausepiiride ebakorrektsed tuvastamise tagajärjel) pole ühendeid materjalist leitud. Teiseks genereeris tuvastamisalgoritm ühendeid liialt palju, kui tegemist oli kolme verbi ühenditega. Kvantitatiivses analüüsis on materjalist välja jäetud üendid, mida esines vähem kui kolm korda, konstruktsioonide analüüsi on need üendid kaasatud. Konstruktsioonide analüüsist on välja arvatud üendid, mille tähendust lähikonteksti põhjal analüüsida ei õnnestunud. Konteksti on vaadeldud vajadusel ka laiemalt kui ühe osalause piir. Näitelaused on esitatud ainult osalausekena ja nad on tekstist välja võetud automaatselt osalause põhjal, s.t otsiti vaadeldavat finiiitset verbi ja infiniitset vormi sisaldavaid osalauseid ning leitud osalauseid on kõik käsitsi üle kontrollitud ja analüüsitud. Niisiis peetakse järgnevates peatükkides sagedustest rääkides silmas käsitsi üle kontrollitud konstrukti sisaldavate osalause **normaliseeritud** sagedusi. Normaliseerimine on tehtud sama valemil abil, nagu kvantitatiivses analüüsis (vt ptk 7.1). Analüüsi on kaasatud ka mõned kolme verbi üendid, juhul kui nad

moodustasid koos konstruktsiooni, mille keskmes oli vaadeldav kahe verbi ühend. Ka see on üks lisapõhjus, miks kvantitatiivse analüüsi sagedusnumbrid võivad konstruktsioonianalüüsi omadest erineda.

Sagedastest ühenditest on ainukesena konstruktsioonide analüüsist välja jäetud *nud*-partitsiibi ja finiiitse *olema* verbi ühendid. Selline ühend moodustab peaaegu eranditult mineviku liitaegu ning ta osaleb samuti paljudel juhtudel liitaegades kolme verbi ühendite (*oleksin pidanud nägema*) koosseisus.

8.2 Konstruktsioonide analüüsi alused

Konstruktsioonide analüüsi materjal on kogutud kvantitatiivse analüüsi tulemusi arvesse võttes ning rakendades eelmises peatükis kirjeldatud lisapiiranguid vaadeldava verbiühendi esinemissagedusele. Igat vaadeldavat ühendit sisaldavad osalused on käitsi läbi vaadatud ja kõik verbiühendid on jagatud tähenduste põhjal erinevate konstruktsioonide esindajate vahel. Konstruktsioonid tuvastamisel olen võimaluse korral toetunud eelnevatele eesti kirjakeele ja murrete uurimustele. Ühendid, mille tähendust lähikonteksti põhjal polnud võimalik tuvastada, on materjalist välja arvatud. Seega moodustavad analüüsimaterjali vaid laused, mis vastasid kvantitatiivsetele kriteeriumitele ning mille tähendust oli võimalik lähikonteksti põhjal tuvastada.

Konstruktsioonide analüüs on skemaatiline ning detailsem kirjeldus on esitatud vaid üksiku juhtudel, kui see näis materjali põhjal oluline olevat. Analüüs lähtub konstruktsioonigrammatika eeldusest, et konstruktsioon on vormi ja tähenduse kooslus. Eeldan, et kirjeldatavad ühendid moodustavad koos tähenduse, mis pole ühendi komponentide tähenduste lihtne summa ja kirjeldatava tähenduse kandmiseks peab ühend esinema just sellise struktuurina. Sarnaselt sõnadega võib ka konstruktsiooni puhul sama vorm kanda erinevaid tähendusi ja hoolimata vormiidentsusest võib tegemist olla erinevate konstruktsioonidega, seega võib konstruktsioon olla homonüümne. Samuti võib sama vormiga konstruktsioonil olla mitu suuremal või vähemal määral seotud tähendust ning sellisel juhul on tegemist konstruktsiooni polüseemiaga.

Konstruktsioonide analüüsi peaesmärk oli teha kindlaks, kas kvantitatiivselt tuvastatud üksused on tõepoolest mingit kindlat tähendus kandvad vormikombinatsioonid või on tegemist täiesti juhuslikult koosesinevate vormiüksustega, mille sellisel kujul esiletulek on samuti juhuslik. Konstruktsioonide sisu ja suhteid on analüüsitud skemaatiliselt, seega pole analüüs kindlasti põhjalik ja ammendav. Kuna teemakohased uurimused murretes on siiani väga napid, siis konstruktsioonide analüüsi eesmärk jääb esmaste tulemuste kaardistamise ja üldtendentside kirjeldamise tasemele. Iga konstruktsiooni detailne analüüs jääb edaspidise uurimistöö teemaks.

8.3 Konstruktsioonide analüüsi probleemid ja piirangud

Kvantitatiivselt tuvastatud konstruktsioonikandidaatide grupeerimise tegid eriti keeruliseks väga napid olemasolevad teadmised murrete süntaktilistest ja semantilistest erijoontest; samuti on senistes uurimustes vähe informatsiooni verbiühendite kohta. Järgmises peatükis esitatud konstruktsioonide analüüsile seab piirangud ühelt poolt töö maht, teiselt poolt asjaolu, et töö autoril puudub emakeelne intuitsioon kõigi kirjeldatavate murrete osas, mis teeb keeruliseks kõigi konstruktsioonide, aga ka üksikute verbide tähendusnüansside analüüsimise. Kuna konstruktsioonide kirjeldamise eesmärk on eelkõige kontrollida, kas kollostruktsiooniline meetod toimib ning kas tuvastatud kandidaatpaarid on ka tegelikult konstruktsioonid, ei ole põhjalikum tähendusnüansside ja detailide lisamine konstruktsioonide kirjeldamisel ka tingimata vajalik.

Konstruktsioonide analüüsis ilmneb, et paljude statistiliselt tuvastatud kandidaatpaaride taga peituvad nii homonüümsed kui ka polüseemsed konstruktsioonid, millest mõned on üsna madala esinemissagedusega. Seega toob kvantitatiivne analüüs esile väga madala esinemissagedusega konstruktsioone, millel on kõrgema esinemissagedusega homonüümid. Nii tekib kaks probleemi. Esiteks loob see fakt näilise vastuolu töö peamise eesmärgiga analüüsida sagedasi keelenähtusi. Teiseks võib metoodika analüüsist välja jätta kõrgema esinemissagedusega konstruktsioonid, millel sagedased homonüümid puuduvad. Teine probleem on eesti murrete uurimise hetkeolukorras praegu veel lahendamatu, kuna puudub igasugune info sagedaste verbikonstruktsioonide kohta, mis analüüsist potentsiaalselt välja võivad jääda. Loodetavasti toovad selles osas rohkem selgust põhjalikumad uurimused tulevikus. Esimese probleemi üks võimalikest lahendustest oleks eemaldada analüüsist harvaesinevad konstruktsioonid ning analüüsida vaid iga vormi taga peituvaid sagedasi konstruktsioone. Seda pole järgnevas analüüsis erinevatel põhjustel siiski tehtud. Esiteks pole vastuolu töö eesmärkide ja harvaesinevate konstruktsioonide esitamise vahel nii suur, kui see esmapilgul tunduda võib. Töö eesmärk on küll vaadelda sagedasemaid ühendeid, aga ka neid ühendeid, mille olulisusele osutasid statistilise analüüsi tulemused. Sagedased ja statistiliselt esilduvad ühendid ei pruugi alati samad olla. See asjaolu omakorda toetab ja veel kord põhjendab töö aluseks olevat mitmekülgset metoodikat. Teiseks kui jätta välja verbiühendite taga peituvad harvaesinevad konstruktsioonid, jääks tööst välja uus informatsioon, mis siinse töö tulemusena nüüd olemas on. Kuna analüüsi käigus tuvastati ka harvaesinevaid konstruktsioone, on need järgnevalt ka esitatud, kui sagedasti esinevad verbühendid selliseid konstruktsioone enda taga peidavad. Näiteks võib *da*-infiniitiv ja finiidne *saama*-verb moodustada üsna paljusid erinevaid konstruktsioone, millest sageduse poolest on selgelt tavalisim modaalkonstruktsioon (*poisid said koju minna*). Samas võib selline ühend olla ka passiivi konstruktsiooni vorm (*poisid said õpetaja käest pahan-dada*), mille esindajate sagedus oli materjalis üsna väike. Kuna kõrge sagedusega verbiühend *da*-infiniitiv + *saama* on nii modaalse kui passiivi-

konstruktsiooni vorm, on mõlemad konstruktsioonid järgnevalt ka esitatud. See kaotab näilise vastuolu töö peamise eesmärgiga kirjeldada sagedasi keele-nähtusi, sest kvalitatiivse iseloomuga konstruktsioonianalüüs püüab sellega seletada kvantitatiivseid tendentse.

Kuna töö keskmes on konstruktsioonide tuvastamine kandidaatpaaride hulgas, et kontrollida kollostruktsiooniliste meetodite tulemusi, ei ole märkimisväärset tähelepanu pööratud konstruktsioonide (ja verbide) homonüümia- ja polüseemiasuhetele. Verbide tähenduserinevusi eri murretes on püütud võtta arvesse võimaluste piires. Ei saa eeldada, et käsitletavates konstruktsioonides oleks näiteks finiidne verb, mis sageli talitleb abiverbi-laadselt, kõigis murretes ühtlaselt grammatiseerunud, talitleb funktsionaalselt samamoodi ning kannab kõiki samu tähendusi. Kuigi pean konstruktsioone verbivormide ja tähenduste paarideks, on konstruktsioonide ja verbide suhted ka konstruktsioonigrammatikas laiemalt pisut problemaatilised. Näiteks demonstreerib Goldberg (1995: 31–39), kuidas ühe keske tähendusega ditransitiivne konstruktsioon võib omada erineva tasemega ülekandmistähendusi, nimetades seda siiski üheks ditransitiivseks konstruktsiooniks. Croft (2003) aga demonstreerib, kuidas kõik Goldbergeri esitatud erinevad tähendused on seotud puhtalt verbisemantikaga. Kindlasti on erinevates konstruktsioonides nii finiidse kui ka infiniitse verbi semantikal oluline roll konstruktsiooni tähenduse moodustumisel. Seda pole siinses analüüsis püütud teadlikult ignoreerida. Pigem on selle tagaplaanile jätmine seotud teoreetiliselt lahendamata probleemide ja materjalist tingitud paratamatusega.

Konstruktsioonide analüüsi tulemustes esineb ka juhtumeid, kus finiidne verb ja infiniitne komponent on omavahel süntaktiliselt nõrgalt seotud (näiteks juhtumid kus infiniitne komponent laiendab hoopis mõnd noomenit ning finiidisel verbil on lausungis lihtsalt siduv funktsioon *tal oli soov puhata*). Selliste juhtumite puhul tekib küsimus, kas tegemist on siiski konstruktsiooni mõiste alla kuuluva üksusega ja neile on eraldi tähelepanu juhitud; siinse töö käsitluses nad siiski konstruktsioonideks loetud. Mõõnan, et see on problemaatiline koht analüüsis, mis nõuab põhjalikumat käsitlust tulevikus. Teisalt kui lähtuda konstruktsioonigrammatika põhimõtetest, mille kohaselt keel ehitatakse üles konstruktsioonidest, mis võivad olla erineva üldistusastmega, siis teataval üldistusastmel võivad sellised juhtumid moodustada ühe konstruktsiooni. Sellesse peaksid aga rohkem selgust tooma tulevased uurimused.

8.4 Impersonaalsuse mõistest konstruktsioonide kirjeldamisel

Järgmises, 9. peatükis konstruktsioonide kirjeldamisel kasutan sageli mõistet *impersonaalne*. Impersonaalsus on keeleteaduses üsna laialivalguv ja raskesti defineeritav kategooria. Siewierska (2008: 116) väidab eelnevate tüpoloogiliste ülevaadete põhjal, et impersonaalsust võib käsitleda nii vormilise (kanoonilise subjekti puudumine) kui ka kommunikatiiv-funktsionaalse nähtusena (agendi

defokuseerimine). Impersonaalsust vormiliselt käsitledes võib vastavad konstruktsioonid Siewierska (2008) järgi jagada neljaks: 1) konstruktsioonid, kus subjekt pole täielikult referentsiaalne, 2) mittekanoonilise subjektiga konstruktsioonid, 3) konstruktsioonid subjektiga, mis ei kuulu verbi argumentide hulka ja 4) konstruktsioonid, kus subjekti ei eksplitseerita üldse.

Kommunikatiiv-funktsionaalse lähenemise korral käsitletakse impersonaalsust agendi defokuseerimise kaudu ja agendi mõiste võib sellisel juhul olla väga lai. Agendi defokuseerimise all mõeldakse tavaliselt kas agendi ebaolulisemaks muutmist või selle grammatilise positsiooni madalamale astmele viimist või neid mõlemat (Siewierska 2008: 121). Siinne töö lähtubki pigem kommunikatiiv-funktsionaalsest lähenemisest, kus agenta tõlgendatakse tähenduslikult väga laialt. Käesolevas töös nimetan impersonaalsust ka konstruktsioone, mille põhifunktsioon on mitte pelgalt agendi defokuseerimine, vaid agendi (või laiemalt tegija) mitteväljendamine.

Järgnevalt annan lühikese ülevaate, mida sellega mõtlen, kuna siinses töös erineb selle mõiste kasutus traditsioonilisest. Konstruktsioonide kirjeldamisel ei hõlma impersonaalsuse mõiste vaid tegumoe kategooriat, vaid see on katustermin nähtustele, mida keeleteaduses enamasti erinevatena käsitletakse. Helasvuo ja Vilku (2008: 225) on erinevaid soome impersonaali- ja passiivikonstruktsioone kirjeldades demonstreerinud ilmekalt selle mõiste defineerimise keerukust ning autorid peavad hoopis loomulikumaks impersonaalsust käsitleda konstruktsioonide kontekstis. Sarnasest põhimõttest olen püüdnud lähtuda konstruktsioonide analüüsimisel selleski töös. Impersonaalsed konstruktsioonid käesoleva töö seisukohalt on konstruktsioonid, mille puhul tegevuse sooritaja pole väljendatud (*lapsena sai palju ringi joostud*), pole teada, kas sellega viidatakse geneeriliselt ükskõik kellele (*koolis sai palju õppida*), või isikule viitamine jääb lihtsalt ebaselgeks ning seda pole võimalik tuvastada lähikonteksti põhjal (*vanasti sai kõvasti tantsitud*). Kõiki selliseid juhtumeid hõlmab katuserminina *impersonaalsus*.

Eesti keeles on rohkelt erinevaid isikule viitamise vältimise strateegiaid, millest mõnd on põhjalikumalt vaadelnud Lindström (2010); Jokela (2012) on detailset kirjeldanud erinevaid eesti keele üldisikulisi konstruktsioone kontrastiivselt soome keelega. Eriti sageli ilmneb see nähtus modaalverbiga moodustatavate konstruktsioonide juures. Sarnased juhtumid on üsna tavapärased ka kirjakeeles, kus sellistes konstruktsioonides jääb segaseks viidatav isik, mis küll võib konteksti põhjal tuletatav olla ja vahel viidatakse selgelt ka üldisikule, kuid paljudel juhtudel on ka seda võimatu tuletada või see jääb mitmeti tõlgendatavaks (Penjam 2008: 130–138). Eesti keele akadeemiline grammatika nimetab ilma grammatilise subjektita lauseid vaegisikulisteks lauseteks, mis jagatakse omakorda kolme rühma: üldisikulisteks (nt *Nooruses omandab kõik hõlpsasti.*), umbisikulisteks (nt *Aru ei mõõdeta habeme järgi.*) ja isikuta lauseteks (nt *Õues juba hämardub*) (Erelt et al. 1993: 227–228). Kõik sellised kategooriad on käesoleva töö konstruktsioonide analüüsis koondatud impersonaalsete konstruktsioonide nime alla.

8.5 Ahelverb, perifrastiline verbivorm ja verbiühend

Eesti keele akadeemiline grammatika (Erelt *et al.* 1995: 92) eristab perifrastilisi verbivorme ja perifrastilisi verbe ehk ahelverbe. Perifrastilised verbivormid on EKG käsitluses süntaktiliselt lahutamatud tervikud (*sai kirjutatud, on ilmumas*) ning ahelverbide ehk perifrastiliste verbide komponendid on nõrgemalt seotud (*hakkas kirjutama, läks jooksmas*), kuid samas viidatakse nende kahe eristamise keerukusele. Lisakriteeriumina ahelverbide iseloomustamisel tuuakse välja asjaolu, et selle sisuliseks tuumaks on infiniitne verb ning finiitverb markerib ühendis modaalsust, aspekti, kausatiivsust või viisi (Erelt *et al.* 1993: 19). Siinses töös olen sellisest vahetegemisest loobunud. Verbiühendi mõiste hõlmab siin nii perifrastiliste verbide kui ka verbivormide mõistet. Ahelverbi eristamine teistest verbiühenditest on praktilises analüüsis sageli üsna keeruline kui mitte võimatu. Samuti puudub alus eeldada, et kõigis murretes on verbide ja verbivormide funktsioonid täpselt samad ning käituvad alati sarnaselt. Ahelverbi mõistet kasutan juhul, kui kirjeldan teisi uurimusi, kus sellist vahet on tehtud. Perifrastilise verbi ja verbivormi mõistet kasutan sünonüümselt verbiühendi mõistega, kus ühendi komponendid on omavahel seotud.

8.6 Konstruktsioonide produktiivsus ida- ja läänepoolsetes murretes

Produktiivsuse mõõtmise metoodikat kirjeldasin täpsemalt peatükis 6.6. Enamasti käsitletakse produktiivsust sõne- ja tüübisageduse suhtena (vt ka ptk 3.3 ja 3.4). Keelenähtuste produktiivsus pole kaugeltki üheselt defineeritav mõiste, milles kõik keeleteadlased ühel meelel oleks (erinevate definitsioonide kokkuvõtet vt Barðdal 2008: 9–28). Samuti on esitatud erinevaid viise produktiivsuse mõõtmiseks, millest kõigil on oma puudused (kokkuvõtlikku ülevaadet vt nt Baayen 2009). Samas artiklis on Baayen demonstreerinud, kuidas lihtsa tüübi- ja sõnesageduse suhte mõõtmise tulemused on sõltuvad korpuse suurusest. Kuna erinevate murrete materjali hulk on siinse töö üks suuremaid probleeme, poleks sellise mõõdiku kasutamine mõistlik. Samuti on selle töö aluseks oleva materjali valguses problemaatiline paremaks peetud *hapax legomena* põhiste mõõdikute rakendamine (vt nt Baayen 2009). Olen veendunud, et tüübi- ja sõnesagedust võib pidada üsna adekvaatseks esmase produktiivsuse indikaatoriks, seetõttu olen siinses töös seda sellisena rakendatud (kuid mitte tüübi- ja sõnesageduse omavahelist suhet). Samuti on selline tegutsemisviis kooskõlas töö olulisima eesmärgiga anda uuritavast nähtusest just esmast ülevaatlikku informatsiooni. Järgmises peatükis mõõdan konstruktsioonide produktiivsust kahe murderühma vahel, millest esimesse kuuluvad kvantitatiivse analüüsi põhjal esile tulnud **idapoolsete murrete rühm** (Võru, Seto, Tartu, Mulgi, ida- ja kirdemurre) ning teise rühma moodustavad **läänepoolsed mured** (lääne-, kesk-, ranna- ja saarte murre).

Sellest töös mõõdan siis verbiühendi, mitte konstruktsiooni produktiivsust kahes eraldi murderühmas (ida- ja läänepoolsete murrete rühmas). Verbiühendi produktiivsus on võrdne konstruktsiooni produktiivsusega juhul, kui vastava verbiühendiga moodustub ainult üks konstruktsioon ehk juhtumitel, kui vaadeldaval konstruktsioonil puuduvad homonüümid. Selline valik on tehtud ühelt poolt töö iseloomust lähtuvalt, mille aluseks on kobaranalüütilised põhimõtted, teisalt on teguviis motiveeritud puhtpraktilistest asjaoludest: ühe konstruktsiooni produktiivsuse mõõtmine ühes murdes pole kvantitatiivselt praeguse korpuse suuruse juures mõttekas. Produktiivsuse analüüsitulemused osutavad siiski kaudselt tüübi- ja sõnesageduse suhtele.

9 VERBIÜHENDID EESTI MURRETES

Järgnevad alapeatükid kirjeldavad tuvastatud konstruktsioone, mis osutusid statistiliselt olulisteks ja/või olid uurimismaterjalis piisavalt sagedased (vt ka peatükke 5, 7 ja 8). Analüüsimaterjali kirjeldasin põhjalikumalt eelmises peatükis. Järgmistes alapeatükkides annan esmalt ülevaate kirjeldatavast infiniitse vormi ja finiiitse verbi ühendist ning vastavate konstruktsioonide varasematest käsitlustest põhiliselt eesti kirjakeeles, kuid uurimuste olemasolu korral ka murretes. Seejärel esitan tuvastatud ühendite abil moodustatavate konstruktsioonide detailsema kirjelduse, konstruktsioonide sageduste jmt infoga. Esitan kirjeldatava ühendi abil moodustatavate konstruktsioonide sageduskaardid ning iga ühendi (mitte konstruktsiooni) produktiivsuse statistilise analüüsi tulemused ida- ja läänepoolsete murrete rühma põhjal, seda kvantitatiivse kobaranalüüsi tulemuste põhjalikuma uurimise eesmärgil (vt ptk 7). Produktiivsuse mõõtmise tehnikat kirjeldasin täpsemalt peatükis 6.6. Kogu esitatud ülevaade tuvastatud konstruktsioonidest on väga ülevaatlik. Siinse töö eesmärgiks pole uurida detailselt iga konstruktsiooni semantilisi ja süntaktilisi nüansse, vaid kaardistada skemaatiliselt kahe verbi abil moodustatavate potentsiaalseid konstruktsioone.

Konstruktsioonide sagedustabelite lugemine. Iga kirjeldatava ühendi peatükis on esitatud tuvastatud konstruktsioonid ning nende **normaliseeritud** sagedused. Analüüsivat materjali ja selle kogumise viisi kirjeldasin täpsemalt eelmises peatükis. Konstruktsioonide sagedustabelite lugemisel on oluline meeles pidada, et sagedused on normaliseeritud, mis tähendab, et nende summa ei ole võrdne iga peatüki alguses esitatud leitud näitelausete arvuga. Normaliseerimise eesmärk on teha sagedusandmestikud omavahel võrreldavateks. Tabelites esineb null-lahtreid, kuid samas võib olla näitelausetes konstruktsioon vastavas murdes esitatud. See tuleneb asjaolust, et normaliseerimise tulemusel sai konstruktsiooni esinemisjuhtumite sagedusväärtuseks arv alla 0,5, mis ümardamisel muutub nulliks. Normaliseeritud sagedused võivad tulenevalt arvutusvalemist saada kümnendmurrulise väärtuse, mis on kõik ümardatud täisarvulisteks. Sageduste normaliseerimise ja normaliseerimiseks kasutava valemi kohta vaata täpsemalt peatükki 7.1.

9.1 *da*-infinitiivi konstruktsioonid

da-infinitiiv¹⁶ on infiniitne verbivorm, millel on sarnaselt teiste infiniitsete vormidega nii verbi kui ka substantiivi tunnuseid ja ta suhestab tegevust finiiitse verbivormiga tähistatava väljendatud tegevuse kaudu. Finiiitne verbivorm ja infiniitiiv moodustavad sageli koos ühe süntaktilise terviku. (Erelt 1984: 3–5) Akadeemilise eesti keele grammatika järgi talitleb *da*-infinitiiv lauses seotud laiendina ja täidab subjekti, objekti, adverbiaali, atribuudi ja predikatiivi süntaktilisi funktsioone (Erelt *et al.* 1993: 237).

¹⁶ Edaspidi kasutan sünonüümselt termineid *da*-infinitiiv ja infiniitiiv.

Eesti kirjakeeles on nii *da-* kui ka *ma-*infinitiivi konstruktsioone põhjalikult käsitlenud Penjam (2008). Sarnaselt eesti keele akadeemilise grammatikaga lähtub Penjam konstruktsioonide analüüsimisel ja liigitamisel *da-*infinitiivi süntaktilisest funktsioonist ning konstruktsiooni tähendus on tema käsitluse järgi sellega tugevasti seotud. Rätsep (1978) ja Tauli (1980) on esitanud detailse ülevaate eesti keele lausetüüpidest, kus on esitatud teiste hulgas ka kõikvõimalikke *da-*infinitiiviga esinevaid lausekonstruktsioone. Põhjaliku ülevaate erinevatest infinitiivikonstruktsioonidest kõigis läänemeresoome keeltes on esitanud Saukkonen (1965; 1966). Ühel või teisel viisil on *da-*infinitiivi konstruktsioone palju uuritud eesti kirjakeeles (Kasik 2001; Metslang 2000; Tragel 2003; Tragel & Habicht 2012 jt).

*da-*infinitiiv moodustab sageli modaalkonstruktsioone finiiitsete modaalverbidega (nt *saama*-verb) ja eesti keele alases uurimistöös ongi palju keskendutud just nendele konstruktsioonidele. Põhjalikumalt on modaalsetid *da-*infinitiivi konstruktsioone uuritud kirjakeeles ja läänemeresoome keeltes (Kehayov & Torn-Leesik 2009; Torn-Leesik 2007; Torn-Leesik 2009; Uuspõld 1989). Kehayovi ja Torn-Leesiku (2009: 366) järgi on modaalverbidel läänemeresoome keeltes kaks mustrit: personaalne ja impersonaalne. Käesoleva töö uurimismaterjalist tuleb selgelt välja impersonaalsete mustrite eelistamine.

Murrete kohta on põhjalikumaid ülevaateid siinses töös käsitletava süntaktilise nähtuste kohta tehtud vähe. Üks selleteemalisi töid on Potsepsi (1956) käsikirjaline diplomitöö, mis käsitleb *da-*infinitiivi vormi ja selle ilmnemist liitõeldise koosseisus. See on käesoleva töö autorile teadaolevalt ainuke selleteemaline murdematerjali põhjal tehtud uurimus. Ülejäänud uurimustest võib leida üksikuid tähelepanekuid *da-*infinitiivi või üksikute konstruktsioonide kohta. Näiteks on Univere (1988: 89) täheldanud, et idamurdes kasutatakse *da-*infinitiivi kaudse kõneviisi väljendamiseks. Mõningaid selliseid konstruktsioone on analüüsinud ka Neetar (1988). Kirderannikumurdes on kõnealuseid konstruktsioone käsitlenud Must (1987: 304–306). Harju-Madise murraku *da-*infinitiivi tarinditest on kirjutanud Juhkam (2012: 504; 519–520), läänemurdes on sarnaseid ülestähendusi teinud Juhkam ja Sepp (2000). Tanning (2004) on muu hulgas vaadelnud *da-*infinitiivi ja selle esinemist verbikonstruktsioonides Karksi murrakus. Varem uuritud üksikutest ühenditest ja konstruktsioonidest tuleb täpsemalt juttu vastavate alapeatükkide juures.

9.1.1 *da-*infinitiiv ja finiiitne *saama*

Eesti keeles väljendab verb *saama* paljusid grammatilisi tähendusi ja Trageli (2003) järgi kuulub verb eesti keele tuumverbide hulka. Finiitse verbiga *saama* ja *da-*infinitiivis esineva mentaalse verbi ühendid liigitab omaette rühmaks kogu läänemeresoome keelealal Saukkonen (1966: 8–10), kelle väitel väljendab selline ühend läänemeresoome keeltes tavaliselt modaalust. Saukkoneni (1966: 16) järgi on läänemeresoome keeltes ühendile iseloomulik kahe ühendit moodustava verbi samasubjektilisus ning selles osas on erandina välja toodud

vaid liivi ja eesti keel, kus erandi näidetena *said haugutada, peksta, isa käest* on esitatud juhtumid, mida selles töös käsitlen passiivikonstruktsioonidena (täpsemalt allpool). Põhjaliku diakroonilise ja sünkroonilise ülevaate verbist *saama* ja sellega moodustuvatest konstruktsioonidest eesti kirjakeeles esitavad Tragel ja Habicht (2012), kes omistavad sellisele koos *da*-infinitiiviga esinevale ühendile omamis-kognitiivse, deontilise võimalikkuse, mittedeontilise võimalikkuse, vajalikkuse, tõenäosuse ja passiivi tähenduse. Sellise tähendusega konstruktsioone tuvastasin ka murretes (täpsemalt allpool), kuid erinevaid modaalsuse liike käesolevas töös eristatud ei ole. Kõik mis tahes modaalset tähendust kandvad konstruktsioonid on koondatud ühe, modaalstruktsiooni alla.

Rätsepa lihtlause tüüpe käsitlevas monograafias (1978: 188) esineb *saama* sekundaarsete verbivormidega elementaarlausete ühekohaliste elementaarlausete rühmas $V^1V^2 + da$, kus esimene verb on impersonaalne (*Vaba aega saab kasulikult pruukida*). Sama rühma kahekohaliste elementaarlausete tüüpi kuulub $N+nom V^1V^2+da$ (*Peeter ei saanud tõusta*), mis selles töös moodustavad modaalsuskonstruktsiooni ning $N^1+nom V^1 (N^2+el) V^2+da$ (*Kapten sai mürsukillust haavata*.) moodustavad passiivikonstruktsiooni jne (Rätsep 1978: 189–191). Siinses töös pole vormielemente sellise detailsusega kirjeldatud.

Tauli (1980: 39–40) järgi võivad sellised ühendid esineda modaalühenditena, nagu ka käesoleva töö materjalis kõige sagedamini täheldada võib. Teatud *da*-infinitiivis verbidega võib finitne *saama* Tauli (1980: 138) väitel moodustada püsiühendeid (*lüüa saada, haavata saada, näha saada*). Käesolevas töös jagunevad sellised ühendid enamasti kahe konstruktsiooni vahel: mentaalse verbi ja passiivikonstruktsiooni vahel. Sarnaselt Rätsepaga on ka Tauli (1980: 139) eraldi välja toonud translatiivis noomeniga esineva konstruktsiooni (*Tema eesmärgiks sai raamat kirjutada*.), kuid murdekorpuses selliseid juhtumeid ei leidunud.

Penjam (2008: 130–153) on kirjakeeles liigitanud kirjeldatavad ühendid modaalstruktsiooniks (*Ma ei saa joosta*.), inhoatiivseks *saama*-struktsiooniks (*Sain eksami tulemused teada*.) ja passiivseks *saama*-struktsiooniks (*Veteran sai sõjas haavata*.). Siinse töö materjali põhjal tuvastasin sarnaseid konstruktsioone murretes. Erinev on vaid inhoatiivne konstruktsioon ja sellised esinemisjuhud liigituvad siinses töös mentaalse verbi konstruktsiooni esinemisjuhtudeks. Kinnistunud ühendid (*süüa saada, juua saada*) on samad, mille Penjam (2008: 93) liigitab leksikaliseerunud ühendite rühma. Selles töös moodustavad omaette rühma impersonaalsed kasutusjuhud, mis murretes moodustavad funktsionaalselt ja semantiliselt selgelt eristuva impersonaali-konstruktsiooni.

saama ja *da*-infinitiivi passiivikonstruktsiooni olemasolu murretes toob välja Neetar (1988: 41) ja märgib, et murretes käitub see konstruktsioon sarnaselt kirjakeelega. Samas artiklis esitab ta väite, et lõunaeesti murded eelistavad pigem teonimesid verbi infiniitvormidele, sellega võiks ehk seletada infiniitsete verbiühendite madalamat esinemissagedust just nendes murretes. Lindströmi (2010: 108) järgi on sellekujuline impersonaalikonstruktsioon eriti levinud läänepoolsetes murretes. *da*-infinitiivi ja *saama*-verbi kasutamist eri tähendustes on peetud üldpõhjaeestiliseks ja tavaliseks kirderannikumurdes (Must 1987: 304). Käesoleva töö materjal kinnitab verbiühendi põhjaeestilisust, kuid huvitaval kombel pole see nii sage rannamurdes. *saama* ja *da*-infinitiivi abil väljendatavat passiivi kirderanniku murdes on peetud iseloomulikuks nooremale põlvkonnale ja vanem põlvkond kasutab samas funktsioonis rohkem partitsiibi-konstruktsioone (Must 1987: 306). Varem ei eristatud kirde- ja rannamurret ja need moodustasid ühtse kirderannikumurde, seega pole sagedusandmed eelnevate väidetega vastuolus. Musta monograafia näited on enamasti pärit idapoolsetelt kirderannikumurde aladelt ehk siinse töö käsitluses kirdemurde aladelt¹⁷, millele konstruktsioon näib eriti iseloomulik olevat. Potsepsi (1956) järgi on *saama* koos *da*-infinitiiviga liitõeldise funktsioonis suhteliselt sage kõigis murretes, kuid kõige tavalisemate hulgas ei maini ta seda saarte murdes. Käesoleva töö uurimismaterjal teatud määral kinnitab seda, sest võrreldes teiste põhjaeesti murderühma murretega ei tõuse see saarte murdes tõepoolest sageduse poolest esile.

Kokku tuvastasin murdekorpuses 1349 *da*-infinitiivis verbi ja finiitset *saama*-verbi sisaldavat osalauset. Kollostruktsiooniline analüüs näitas tugevat statistilist seost kahe uuritava üksuse vahel pea kõigis murretes. Üllatuslikult ei saavutanud konstruktsioon 0,05 usaldusnivool piisavalt kõrget statistilise seose tugevuse väärtust, et seda konstruktsiooniks pidada, kesk- ja Mulgi murdes. Analüüs näitas siiski, et konstruktsiooni esineb ka neis murretes. Siit tuleb välja kollostruktsioonilise analüüsi nõrgem pool, kui korpus pole statistilisteks arvutusteks piisavalt suur. Keskmurdes võib madala väärtuse põhjuseks pidada asjaolu, et võrreldes teiste konstruktsioonidega ei olnud see konstruktsioon piisavalt oluline või hajus selle tähtsus teiste, statistiliselt olulisemate konstruktsioonide hulgas. Tabel 9 esitab käesolevas uurimismaterjalis tuvastatud *da*-infinitiivi ja finiitse *saama* abil moodustatavad konstruktsioonid ja nende esinemisjuhtude **normaliseeritud** sagedused.

¹⁷ Aitäh Ellen Niidile sellele tähelepanu juhtimise eest.

Tabel 9. *da*-infinitiivi ja *saama*-verbi konstruktsioonid ning nende esinemisjuhtude normaliseeritud sagedused. Lühendid: imp – impersonaalkonstruktsioon, püsi – püsiühend, ment – mentaalse verbi konstruktsioon, mod – modaalkonstruktsioon, pass – passiivikonstruktsioon.

	imp	püsi	ment	mod	pass
IDA	10	0	4	105	1
KESK	13	5	4	161	2
KIRDE	27	8	5	143	6
LÄÄNE	12	5	3	186	4
MULGI	0	7	9	68	7
RANNA	0	0	3	71	3
SAARTE	3	6	0	73	1
SETO	0	11	6	40	0
TARTU	3	7	0	109	7
VÕRU	2	7	6	76	2

Tabel osutab, et kõige sagedamini väljendavad *da*-infinitiiv ja *saama* mõnd modaaltähendust. Silma torkab erinevus impersonaalse konstruktsiooni kasutusjuhtude sageduses, kus kirdemurdes on see teiste murretega võrreldes vaieldamatult sagedasim. Järgnevalt kirjeldan lühidalt esitatud konstruktsioonide tähendusi.

Modaalkonstruktsioon

saama kuulub koos verbidega *võima* ja *pidama* eesti keele kõige sagedasemate modaalverbide hulka ja on osaliselt kaotanud süntaktilise iseseisvuse. Võrreldes aga modaalverbidega *võima* ja *pidama*, on *saama* kõige vähem grammatiseerunud, s.t selle verbiga pole võimalik väljendada kõiki modaalsuse liike ning verb väljendab tavaliselt dünaamilist võimalikkust. (Erelt 2003: 106) Käesolev töö ei püüa eristada erinevaid modaalsuse liike, mistõttu modaalkonstruktsioonideks siin ja edaspidi on liigitatud kõik konstruktsioonid, mis väljendavad mis tahes modaalsust (vt näiteks Uuspõld 1989). Kehayov ja Torn-Leesik (2009) on käsitlenud modaalverbe läänemeresoome keeltes ja leidnud, et *saama* on üks põhimodaalverbidest kõigis vaatlusalustes keeltes (samas 2009: 368). Torn-Leesiku (2007) järgi esineb modaalverbidel sarnaseid omadusi: nad väljendavad üldiselt võimalikkust või vajalikkust, nad toimivad abiverbidena, kombineerudes mõne põhiverbiga, ning modaalverbidel võib olla erinev, tavaliselt puudulik morfoloogiline paradigma. Kõige tavalisem on murretes *da*-infinitiivi ja *saama*-verbiga moodustatud modaalkonstruktsioon (6).

- (6) nohh naasõl sääanne miiss kiss **iss saa** sukkugi **kävvü** (VÕR)
ma **i saa kraappi** (IDA)
nääd **said** paatti **vädäda** (LÄÄ)

Raskusi tekitas selliste konstruktsioonide eristamine, kus väga selgelt olid olemas nii impersonaalne kui ka modaalne tähendus (7).

- (7) **sai** mittu kord jahh **naerda** (KES)
 sial **sai** neli aastat **tüöttada** (IDA)

Kuna neid kahte tähendust sellistes konstruktsioonides oli praktilises analüüsis enamasti võimatu eristada, olen mõlemat tähendust kandavad *da*-infinitiivi ja *saama* ühendid liigitanud modaalseteks.

Impersonaal ja passiiv

da-infinitiiv ja finitiivne *saama* võivad kirjakeeles väljendada nii impersonaali kui ka passiivi (Erelt *et al.* 1993: 247; Tragel & Habicht 2012) ja sellist vahet olen teinud ka murrete konstruktsioonide analüüsimisel. Sellekujuline impersonaal on Musta väitel (1987: 304) levinud kirderannikumurdes. Impersonaalse konstruktsiooni puhul on oluline märkida, et abiverbina käituv *saama* esineb selles konstruktsioonis ainult imperfekti ainsuse kolmandas pöördes. Muud kohustuslikud elemendid puuduvad (8). Konstruktsioonile võib omane olla siiski ka modaalse tähenduse varjund, kuid reeglina väljendab konstruktsioon mingit korduvat, harjumuspärast tegevust, kus subjekt on geneeriline ja modaalne õnnestumistähendus pigem puudub ja konteksti põhjal on domineeriv impersonaalsus.

- (8) nohh seda poissibõlves **sai** kaa teha ikke t õitsis **käija** (KIR)
 ja kui palju **sae** neid Tartu linna **vedädä** aigemaija sis närvikliinikku
 (IDA)
 siis omigu **sai** kke kella viijest juba niitma **menna** omigune söögivahe
 (LÄÄ)

Lindström (2010: 107–108) on konstruktsiooni analüüsimisel kirjakeeles leidnud, et see on üsna levinud 1. ja 2. isikule viitamise strateegia ning eriti sage selles funktsioonis just läänepoolsetes murretes. Siinse töö materjali põhjal seda nii selgelt väita pole võimalik, kuid konstruktsiooni esineb harvem lõunaeesti murretes.

da-infinitiivi ja *saama* kujulisi passiivikonstruktsiooni esinemisjuhtusid on suhtarvult pea poole vähem kui impersonaalseid konstrukte. Sellekujuline passiiv on murdekorpuses suhteliselt harv nähtus ning uurimismaterjalist tuvastasin ainult üksikud näited. Näiteks Seto murdes ei esinenud mitte ühtegi sellekujulist passiivi. *saama*-passiivi on kirjakeeles peetud dünaamilisemaks kui näiteks verbiga *olema* moodustatavat passiivi (Torn-Leesik 2009: 84).

Impersonaalsest konstruktsioonist eristab passiivi semantilise patsiendina talitleva grammatilise subjekti ühildumine verbiga *saama*. Passiivikonstruktsioon murretes on samakujuline nagu kirjakeeleski (Erelt *et al.* 1993: 247; Penjam 2008: 151–153; Tragel & Habicht 2012). Nende väheste näidete põhjal tundub, et murretes reeglina agenti passiivses konstruktsioonis ei väljendata. Ainult paaris näites on see vormistatud postpositsiooni *käest* abil. *da*-infinitiivis

verb on reeglina transitiivne ja väljendab mõnd negatiivset tegevust (*nuhtlema, pahandama, haavama* jt) (9).

- (9) mina **sai** esä käest **tõrelda** (TAR)
kui ta **ehmadatta saab** (KIR)
aga ussi käest **ammustada** ma küll **eij ole saand** (LÄÄ)

Mentaalse verbi konstruktsioon ja püsiühend

Siinse töö käsitluses mentaalse verbi ja püsiühendi konstruktsioonid on ühtseks püsiühendite rühmaks liigitanud Tauli (1980: 138). Käesolevas töös olen sarnaselt Penjamiga (2008: 93) liigitanud püsiühenditeks vaid verbiga *sööma* ja *jooma* esinevad *da*-infinitiivi ühendid.

Mentaalse verbi (*teadma, nägema, kuulma* jt) ja *da*-infinitiivi ühendid on poolpüsiühendilaadsed üksused, s.t *da*-infinitiivi koosseisus esineb kindlasse tähendusrühma kuuluv verb (10). Selles töös on sellised juhtumid mentaalse verbi konstruktsiooni esindajad.

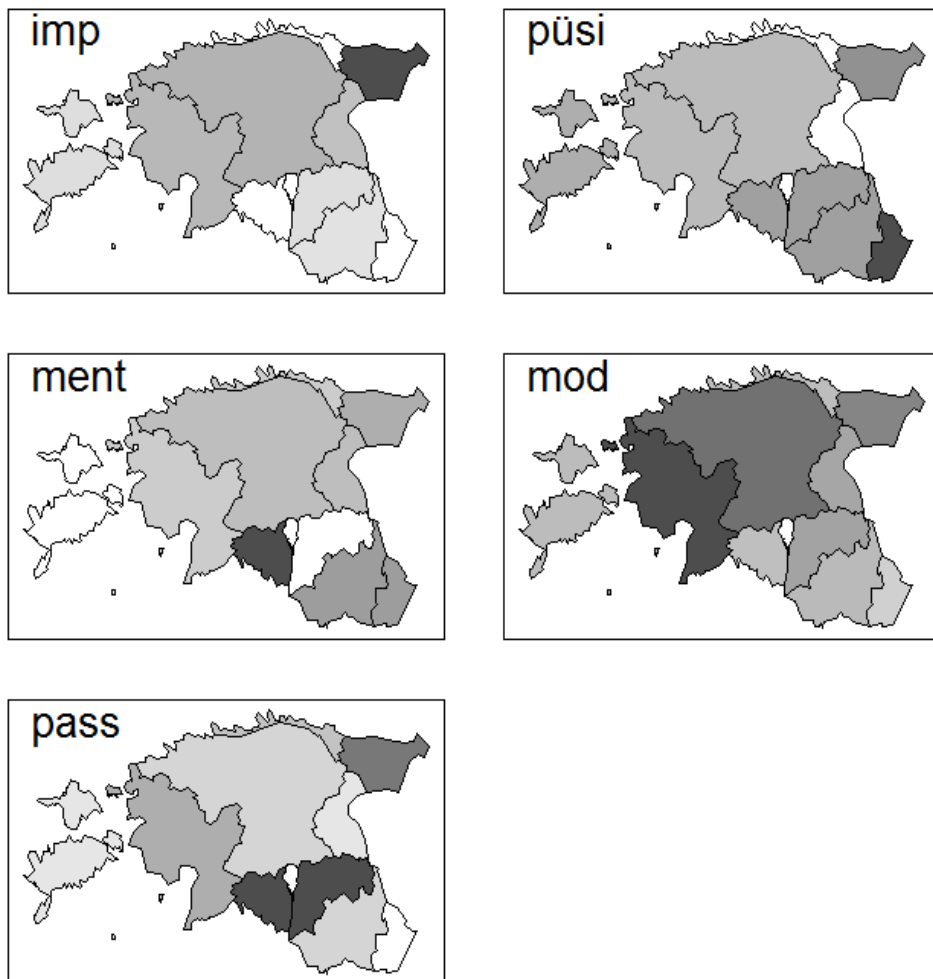
- (10) nohh velekese **saiva tiidäq** (SET)
sial **saeme külmettada** küll (LÄÄ)
üle poole kergemast minu oma nahal **sai tunda** juo (KIR)

Tragel ja Habicht (2012) on sellist konstruktsiooni nimetanud kirjakeelse materjali põhjal omamistähendusega üleminekukonstruktsiooniks omamistähendusega konstruktsiooni ning osalejavälise mittedeontilise modaalkonstruktsiooni vahel.

Eraldi on liigitatud leksikaliseerunud ühendid, kus *da*-infinitivis esinevad verbid *sööma* ja *jooma* (11). Need verbid on koos *saama*-ga esinedes omandanud iseseisva tähenduse, mis vormi poolest on küll *da*-infinitiv, kuid intuitiivselt tajutav pigem nimisõnalise objektina.

- (11) **sai** jõulune aig **juvva** toda (TAR)
süija sai küll seal (SAA)

Nii mentaalse verbi konstruktsiooni esinemisjuhud kui ka püsiühendid on murdekorpuse tekstides suhteliselt harv nähtus ja uurimismaterjalis esinesid vaid üksikud näited. Joonis 13 esitab konstruktsioonide geograafilise sagedusjaotuse. Tumedamad alad tähistavad konstruktsiooni esindajate kõrgemat sagedust ning heledamad alad madalamat sagedust vastavalt sellele nagu need on esitatud tabelis 9 peatüki alguses. Konstruktsioonisageduse põhjalt ei ilmne küll selgeid piirkondi, mis eelistaksid ühe või teise konstruktsiooni kasutamist, kuid iga konstruktsiooni puhul ilmneb murdeid, millele mõni konstruktsioon näib iseloomulikum olevat. Kuna sageduserinevused on siiski suhteliselt väikesed, pole siin väga põhjapanevate järelduste tegemine otstarbekas.



Joonis 13. Finiitse *saama* ja *da*-infinitiivi konstruktsioonide geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Peatükis 6.6 kirjeldatud meetodiga *da*-infinitiivi ja *saama* produktiivsust mõõtes eraldi ida- ja läänepoolsete murrete rühmas (vt ka ptk 8.6) ilmnevad selged erinevused. Idapoolsetes murretes on teatud verbid väga tugevalt *saama* verbiga statistiliselt seotud, s.t mida kõrgem on tabelis esitatud verbi normaliseeritud kriitiline FTT-väärtus, seda püsiühendilaadsema üksuse see finiiitse verbiga *saama* moodustab. Tabel 10 esitab verbiga *saama* kõige tugevamini tõmbuvad *da*-infinitivis esinevad verbid kahes murderühmas¹⁸.

¹⁸ Kõik selles ja järgnevatel tabelites esitatud verbid on kirjakeelestatud kujul ja mitte tekstis esinenud murdelise vormi kujul, selgituseks vt ka ptk 6.6.

Tabel 10. *saama*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
elama	4,08	tõmbama	1,19
kõnelema	3,17	kuivatama	1,08
pagema	2,11	panema	1,08
ütleva	1,81	lõikama	1
kaema	1,58	ajama	0,95
kukkuma	1,58	näitama	0,92
pahandama	1,58	lööma	0,79
töötama	1,58	kuulma	0,77
pidama	1,54	kündma	0,77
teadma	1,34	püüdma	0,77

Analüüsitulemustest ilmneb, et idarühmas on tervelt kümme *da*-infinitiivis verbi, mis on finiiitse *saama*-verbiga väga tugevalt seotud. Läänerühmas selliseid verbe aga pole. See omakorda viitab sellele, et *saama*-verbiga *da*-infinitiivis liituvate verbide hulk on läänerühmas suurem ja seega on verbiühend läänerühmas tunduvalt produktiivsem kui idarühmas. Materjali ebaühtluse tõttu on siin olulisem vaadata normaliseeritud FTT väärtuste jaotumist kui konkreetse verbi kriitilise FTT väärtuse numbrit. Kui väärtused on madalad ja jaotuvad ühtlaselt, siis sellest järeldub, et finiiitse *saama*-verbiga võib kombineeruda rohkem erinevaid verbe. Kui väärtuse tugevus on kõrge ja seostub ainult üksikute verbidega, siis võib sellist kombinatsiooni pidada rohkem püsiühendilaadseks.

9.1.2 *da*-infinitiiv ja finiiitne võima

Modaalverbi *võima* erinevaid tähendusi ja funktsioone on eesti ja teistes läänemeresoome keeltes põhjalikult uurinud Kehayov ja Torn-Leesik (Kehayov & Torn-Leesik 2009; Torn-Leesik 2007). Trageli (2003: 37) järgi kuulub *võima* eesti keele tuumverbide hulka ning on kirjakeeles modaalne abiverb, väljendades dünaamilist, deontilist ja episteemilist modaalust. Taulli (1980: 39–40) esitab vormipaari modaal tähendust kandvana ning toob eraldi välja (1980: 144) tüübi finiiitne *võima* + (*olla* + *nud*-partitsiip), kuid kuna kolme verbi ühendid ei olnud käesoleva töö fookuses, siis selliseid ühendeid eraldi vaadatud pole. Penjami (2008: 124–126) väitel on *võima* kirjakeeles kõige sagedasem *da*-infinitiiviga modaal konstruktsiooni moodustaja.

Potsepsi (1956) järgi esineb *võima* ja *da*-infinitiiv liitöeldisena sagedasti kõigis murretes. Käesoleva töö materjal toetab seda väidet, kuna selline ühend ei tõuse sageduse poolest tugevalt esile üheski murdes, ka ei moodustu selle

põhjal selgelt eristuvaid murderühmi. Mõnevõrra esineb ühendit rohkem keskmurdes ning kõige vähem Seto ja kirdemurdes.

Analüüsitava materjalis oli 461 finiiitse verbi *võima* ja *da*-infinitiivi sisaldavat osalauset. **Normaliseeritud** konstruktsionid murrete kaupa esitab tabel 11.

Tabel 11. *da*-infinitiivi ja *võima*-verbi modaalkonstruktsiooni (mod) esinemisjuhtude normaliseeritud sagedused.

	mod
IDA	45
KESK	70
KIRDE	29
LÄÄNE	50
MULGI	50
RANNA	41
SAARTE	47
SETO	14
TARTU	37
VÕRU	28

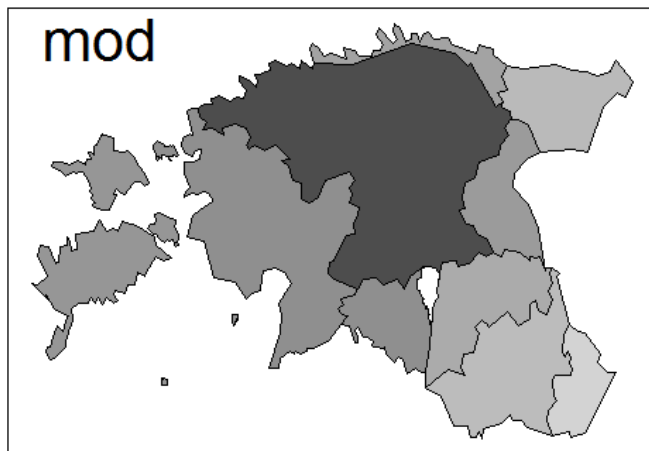
Kollostruktsioonilises analüüsis tõusis ühend statistiliselt olulisena esile vaid ida-, saarte, kesk- ja läänemurdes. Teistes murretes see 0,05 olulisuse nivool kriitilist FTT väärtust ei ületanud. Seega poleks ainult kollostruktsioonilise analüüsi tulemusi arvesse võttes ühendit murdekorpuse materjali peal tuvastatud.

Modaalkonstruktsioon

Materjali analüüs näitas, et selline vormipaar kannab kõigis murretes ainult modaalselt tähendust (12). Erinevaid modaalsuse liike eristatud pole. Samaselt kõigi teiste modaalkonstruktsioonidega murretes on sellelgi konstruktsioonil tugevalt impersonaalne iseloom (vt ka ptk 8.4). Konstruktsioon esineb sageli üldisikulisena ja finiiitse verbi ja infiniitse vormiga tähistatud tegevuse sooritajat püütakse tõrjuda tagaplaanile.

- (12) palju **võib** inimine veel **elada** (KIR)
 hmm neid **võib** viis tükki ikka ära **juuwa** (SAA)
 et mina **võin** sulle **juhata** (IDA)

Joonis 14 esitab eesti murrete finiiitse *võima* ja *da*-infinitiivi modaalkonstruktsiooni geograafilise sagedusjaotuse, kus tumedamad alad tähistavad konstruktsiooni kõrgemad esinemissagedused ning heledamad alad madalamat esinemissagedust. Konstruktsioon on sagedasem keskmurdes, teistes murretes on selle sagedus üsna ühtlane. Taas ei eristu geograafiliselt selgelt suuremaid piirkondi.



Joonis 14. Finiitse *võima* ja *da*-infinitiivi modaalkonstruksiooni geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 12 esitab finiiitse *võima*-verbiga statistiliselt kõige tugevamini tõmbuvad *da*-infinitiivis verbid ida- ja läänepoolsetes murretes. Kahe rühma lõikes võib näha, et verbiga *võima* tugevamini tõmbuvaid verbe on rohkem idapoolsetes murretes, kuid nähtavasti on modaalverb *võima* üsna ühtlaselt produktiivne mõlemas rühmas.

Tabel 12. *võima*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
laitma	1.61	käima	1.52
elama	1.11	rääkima	1.3
juhatama	1.07	olema	0.79
kaema	1.07	laskma	0.6
kuulma	1.07	võtma	0.58
nurisema	1.07	panema	0.56
ütleva	0.75	ajama	0.45
tõstma	0.69	näitama	0.45
sööma	0.62	sõitma	0.45
hõikama	0.53	saama	0.4

Kuna *võima*-verb ja *da*-infinitiiv moodustavad murretes vaid modaalkonstruktsioone, saab tabelis esitatut käsitleda kui konstruktsiooni produktiivsust. Seega võib väita, et *võima*-modaalkonstruktsioon on üsna produktiivne nii idakui ka läänepoolsete murrete rühmas.

9.1.3 *da*-infinitiiv ja finiidne *tohtima*

Tauli (1980: 40) esitab *tohtima* modaallühendite rühmas, mille vorm on sama nagu murreteski: finiidne *tohtima* + *da*-infinitiiv. Samas rühmas esitab sellise vormipaari ka Rätsep (1978: 188). Eraldi rühmana toob Rätsep (1978: 189) välja veel N+nom + V₁ V₂+da (*Saal ei tohi olla suur.*). Sellisel kujul võib konstruktsioon esineda ka murretes, kuid käesolevas töös pole Rätsepa esitatud kahte rühma eristatud, sest ühendi tähendus on murretes sama ja nominatiivis subjekt pole murretes selle konstruktsiooni kohustuslik komponent. Sarnaselt kirjakeelega (Penjam 2008: 124) on ka murretes *tohtima* modaalkonstruktsioonid, võrreldes modaalverbidega *saama* ja *võima* konstruktsioonidega, tunduvalt harvem esinevad.

Tuvastasin murdekorpusest 206 *da*-infinitiivi ja finiidse *tohtima* ühendit sisaldavat osalause. Sarnaselt modaalkonstruktsiooniga, mis moodustub verbiga *võima*, kannavad ka *da*-infinitiiv ja finiidne *tohtima* murretes ainult modaalse tähendust. Tabel 13 esitab konstruktsiooni esinemisjuhtude **normaliseeritud** sagedused murretes.

Tabel 13. *da*-infinitiivi ja *tohtima*-verbi modaalkonstruktsiooni (mod) esinemisjuhtude normaliseeritud sagedused.

	mod
IDA	20
KESK	25
KIRDE	24
LÄÄNE	21
MULGI	29
RANNA	6
SAARTE	22
SETO	3
TARTU	20
VÕRU	21

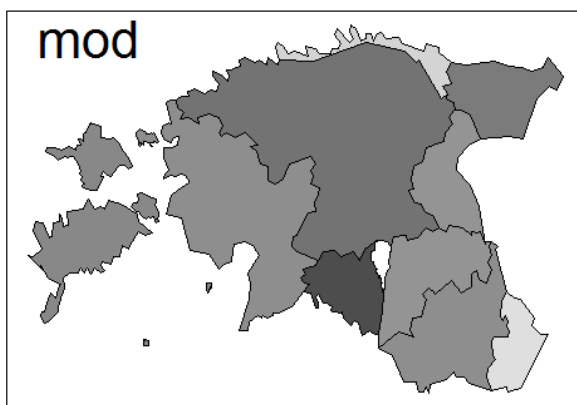
Kollostruktsioonilise analüüsi tulemusena osutus ühend statistiliselt oluliseks vaid saarte ja keskmurdes.

Modaalkonstruksioon

Sarnaselt kirjakeelega (Penjam 2008: 135) on ka murretes konstruksioon tugevalt eituslembene. Materjalist ei tuvastanud ma ühtegi ühemõtteliselt jaatavas keskkonnas esinevat *tohtima* konstrukti. Konstruksioon esines pea alati eitavana või kui verbivorm oli jaatav, siis siis küsimuse vormis, mille vastus on eeldatavalt eitav (13), nagu näite esimeses lauses.

- (13) kuradi muudu ma haigõss **tohtsõ jäädä** tütrik (VÕR)
nii jälle et **ess tohi** mette palav **olla** (SAA)
ei tõhi pasunat **aada** (IDA)

Joonis 15 esitab finiiitse *tohtima* ja *da*-infinitiivi konstruksiooni geograafilise sagedusjaotuse. Sagedusjaotus on selle konstruksiooni puhul teistega võrreldes märgatavalt erinev. Silma torkab konstruksiooni esindajate väga madal esinemissagedus Seto ja rannamurdes, samas kui teistes murretes on see suhteliselt ühtlane. Üllatuslikult näitab kõrgeimat normaliseeritud sagedusnumbrit siin Mulgi murre. Potseps (1956) mainib *tohtima* verbi *da*-infinitiiviga liitõeldise koosseisus vaid läänemurdes, kuid käesoleva töö uurimismaterjal teiste murretega võrreldes sellist arvulist ülekaalu ei kinnita. Ka ei ilmne siin tugevaid rühmaerinevusi.



Joonis 15. Finiiitse *tohtima* ja *da*-infinitiivi modaalkonstruksiooni geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 14 esitab ida- ja läänepoolsete murrete verbid, mis *da*-infinitiivis tõmbuvad kõige tugevamini finiiitse *tohtima*-verbiga. Siingi on ühendi produktiivsus võrdne konstruksiooni produktiivsusega, kuna *tohtima*-verb moodustab murretes *da*-infinitiiviga vaid modaalkonstruksioone. Selline konstruksioon on üsna produktiivne kõigis murretes ning kahe murderühma vahel märkimisväärsed erinevusi ei ilmne.

Tabel 14. *tohtima*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
viima	1.26	panema	1.06
võtma	1.11	olema	0.96
kõnelema	0.84	andma	0.85
kündma	0.84	tooma	0.85
tulema	0.81	näitama	0.63
jätma	0.8	käima	0.44
tegema	0.78	keelama	0.42
jääma	0.59	lõikama	0.42
minema	0.55	püüdma	0.42
hoidma	0.42	seisma	0.42

9.1.4 *da*-infinitiiv ja finiiitne *tulema*

Kirjakeeles samuti eesti keele tuumverbide hulka (Tragel 2003: 31–32) kuuluva *tulema*-verbi grammatilisi funktsioone on eesti kirjakeeles uuritud üsna põhjalikult (Penjam 2008; 2005; 2006a; 2006b). Finiitsena koos *da*-infinitiiviga esineb *tulema*-verb enamjaolt modaalkonstruksioonides (Penjam 2006a; Uuspõld 1989). Tauli (1980: 137) esitab sellise ühendi vormirühmas $V_{3sg}+da$ ja väidab, et ühendi võib asendada *pidama* ja *ma*-infinitiivi vormiga. Rätsep (1978: 188) esitab sekundaarsete verbivormidega ühekohaliste elementaarlausete rühmas ($N+ad$) V^1V^2+da tüübi (*Taavetil tuleb kiri ära viia.*), kuhu kuulub ka verb *tulema*. Sellisel kujul väljendab modaalkonstruksioon kohustust ja modaalsus vormistub siin kirjakeeles konstruksioonitähendusena (Pajusalu *et al.* 2004: 30–31). EKG järgi (1993: 258) on *tulema* verbil perifrastilise verbi osana algust väljendava faasi-verbi funktsioon, kuid ainukese näitena on esitatud lause *Lehm tuli liipsma*. Murdematerjalist *tulema* faasiverbi funktsioon ei ilmnenud.

Murretes pole selle verbiühendi kohta eriti palju tähelepanekuid tehtud. Potseps (1956) mainib *tulema* verbi *da*-infinitiiviga liitõeldise sagedase komponendina kõigis murretes, mõnevõrra vähem läänemurdes. Käesoleva töö uurimismaterjali põhjal võiks kaudselt selliseid järeldusi teha, kuna keskmurdes on ühendi esinemissagedus märgatavalt kõrgem kõigist teistest murretest, samas pole see läänemurdes teiste murretega võrreldes märkimisväärselt madalam.

Tabel 15 esitab *da*-infinitiivi ja finiiitse *tulema* verbiga moodustunud konstruksioonid ja nende konstruktide **normaliseeritud** sagedused. Kokku tuvastasin materjalist 245 sellist verbiühendit sisaldavat osalause ja enamjaolt väljendatakse sellise vormiga modaalsust. Kollostruktsioonilise analüüsi tulemusena ei osutunud ühend statistiliselt oluliseks ida-, kirde-, Seto ja Tartu murdes.

Tabel 15. *da*-infinitiivi ja *tulema*-verbi konstruktsioonid ning nende esinemisjuhtude normaliseeritud sagedused. Lühendid: atrib – atributiivkonstruktsioon, finaal – finaalkonstruktsioon, mod – modaalkonstruktsioon.

	atrib	finaal	mod
IDA	0	0	18
KESK	0	1	53
KIRDE	0	0	17
LÄÄNE	0	0	30
MULGI	0	0	27
RANNA	0	0	13
SAARTE	0	0	14
SETO	0	0	23
TARTU	0	1	11
VÕRU	0	0	13

Saukkoneni (1965: 152) järgi ei pruugi selline ühend olla läänemeresoomelist päritolu ja selle puudumine soome murretes viitab tema arvates ühendi võõrapärasusele. Eesti murretes on ühendi kasutus selgesti olemas, kuid mitte eriti sage.

Modaalkonstruktsioon

Sarnaselt teiste modaalverbikonstruktsioonidega on ka *tulema* modaalkonstruktsioon sagedasti murretes impersonaalne¹⁹ (14), s.t isikut sageli ei markeerita. Verb on konstruktsioonis alati ainsuse 3. isiku vormis.

- (14) **tuli** pidusid **teha** (LÄÄ)
 saega ikke **tuleb leigatta** (KES)
 aa nüüt **tulõp õppi** (TAR)

Nagu kirjakeeles, nii ka murretes võidakse selles konstruktsioonis modaalset agenti väljendada adessiiviga (Penjam 2006b), ent selle näiteid esines murde- materjalis üksikuid (15).

- (15) aga *sull* **tuleb lahkuda** meije maeast (LÄÄ)
 nii et *mehel* **tuli** kohe külast abi **kutsuda** (LÄÄ)
 ku pailu *mull* **tuli** ka sinna **vedäda** (MUL)

Finaal- ja atributiivkonstruktsioon

Materjalis leidus kolm ühendit, mis ei kandnud modaalsustähendust. Sellised ühendid esindasid finaal- ja atributiivkonstruktsiooni. Finaalkonstruktsioonis (16) pole finiitne *tulema* ja *da*-infinitiiv omavahel süntaktiliselt seotud ning *da*-infinitiivis verb väljendab siin otstarvet, eesmärki.

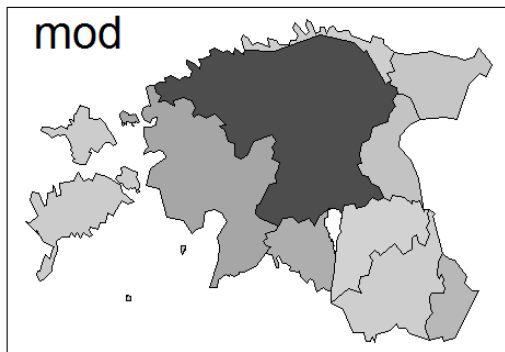
¹⁹ Vt ka ptk 8.4.

(16) noo juba viimasel aeal **tullid** poodi ka lambid **müija** (KES)

Finaalkonstruksiooni näites on *tulema* tõlgendatav ka algust väljendava faasi-verbina ning kogu ühendi tähendus oleks ümbersõnastatav kujule *hakati müüma*.

Atributiivses konstruksioonis esineb *da*-infinitiiv grammatilise subjekti täiendina. Nii finaali- kui ka atributiivkonstruksiooni puhul tekib küsimus, kas neid ikkagi saab nimetada verbikonstruksioonideks, sest finiiitne ja infiniitne verb ei moodusta siin koos ühte tähendust ning *da*-infinitiivis verb on pigem seotud nimisõnaga. Kuna tehniliselt sellised ühendid tuvastati, siis on nad siin ka välja toodud, kuigi verbikonstruksioonideks saab neid tõepoolest pidada mõõndustega.

Joonis 16 esitab *da*-infinitiivi ja finiiitse *tulema* abil moodustatava konstruksiooni geograafilise sagedusjaotuse. Tumedamad alad tähistavad kõrgemat kasutussagedust ning heledamad alad madalamat kasutussagedust. Finaali- ja atributiivkonstruksiooni esines materjalis vaid üksikjuhtumitel ja nende kohta pole sagedusjaotust esitatud (vt ka tabelit 15 peatüki alguses). *tulema*-modaalkonstruksioon näib olevat iseloomulikum keskmurdele ning muudes murretes on selle kasutussagedus üsna ühtlane. Pisut kõrgemad on sagedused veel lääne-, Mulgi ja Seto murdes.



Joonis 16. Finiitse *tulema* ja *da*-infinitiivi modaalkonstruksiooni geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 16 esitab *da*-infinitiivis esinenud verbid, mis tõmbuvad statistiliselt kõige tugevamini finiiitse *tulema* verbiga samas osalauses esinedes. Läänerühmas pole ühtegi verbi, mille statistiline väärtus lubaks oletada ühe kindla verbi eelistamist, seega tundub nende väärtuste põhjal, et läänepoolsetes murretes on ühend üsna produktiivne. Idarühmas tõmbuvad tugevamini verbid *vedama* ja *määrima*, mis lubaks eeldada nende verbide ja *da*-infinitiivi püsiühendilaadsust, konstruksioonide analüüsi põhjal seda siiski väita ei saa ning kõrged väärtused on juhuslikud. Mõlemas rühmas on *tulema* sellises kontekstis esinedes üsna produktiivne, kombineerudes paljude erinevate verbidega.

Tabel 16. *tulema*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
vedama	2.28	võtma	0.62
määrima	1.7	hoidma	0.56
maksma	1.18	viima	0.51
taguma	1.13	kuivatama	0.42
käima	0.89	laskma	0.42
minema	0.6	lõikama	0.42
elama	0.56	rabama	0.42
jahutama	0.56	tooma	0.42
jääma	0.56	panema	0.36
kloppima	0.56	kakkuma	0.28

9.1.5 *da*-infitiiv ja finiidne *laskma*

Verb *laskma* kuulub Trageli (2003: 51) järgi eesti keele tuumverbide hulka. Verbi kausatiivseid tähendusi on põhjalikult kirjeldanud Kasik (2001), kuid modaalse ja kausatiivse tähenduse eristamise raskusele, mis oli ilmne ka selles töös, on viidanud Tragel (2003: 51–52). Kasik (2001) jagab *laskma* verbiga moodustatud kausatiivkonstruktsioonid nelja rühma: võimaldav (*Peremees laseb majal laguneda.*), kuratiivne (*Mees laskis fotograafil endast pilti teha.*), direktiivne (*Tädi laseb teid kõiki tervitada.*) ja instrumentaalne (*Tüdruk laskis käitel käia.*). Penjami (2008:85–86) liigituses moodustavad *da*-infitiiv ja finiidne *laskma* võimaldamis-kohustamiskonstruktsiooni, Rätsepa (1978: 189–190) järgi kuulub *laskma* sekundaarsete verbivormidega lausetüüpides paljudesse eri rühmadesse, mis erinevad suuresti vaid mõnes lausetüübis esineva noomeni käände või fraasi poolest. Käesolevas töös pole konstruktsioone nii detailset vaadeldud, kuigi Rätsepa kirjeldatud tüüpidele vastavaid näiteid võib leida murretestki.

Käesoleva töö uurimismaterjali põhjal olen eristanud kaht erinevat konstruktsiooni: modaalset ja instrumentaalset. *da*-infinitiivi ja finiidset *laskma* ühendit sisaldavaid osalauseid oli materjalis kokku 238. Tabel 17 esitab konstruktide **normaliseeritud** sagedused murrete kaupa. Enamjaolt moodustab kirjeldatav ühend modaalse tähendusega (võimaldamine ja kohustus) konstruktsioone.

Tabel 17. *da*-infinitiivi ja *laskma*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: instrument – instrumendikonstruktsioon, mod – modaalkonstruktsioon.

	instrument	mod
IDA	0	35
KESK	0	20
KIRDE	0	11
LÄÄNE	1	23
MULGI	0	34
RANNA	0	7
SAARTE	0	24
SETO	0	34
TARTU	0	30
VÕRU	0	28

Kollostruktsioonilise analüüsi tulemusena osutus ühend statistiliselt oluliseks vaid saarte, kesk-, ida- ja läänemurdes.

Modaalkonstruktsioon

Modaalkonstruktsioon võib väljendada kellegi kohustamist millegi tegemiseks, s.t agent annab kellelegi selge käsu või kohustuse *da*-infinitiiviga väljendatava tegevuse sooritamiseks (17), kuid kohustatav ei pruugi olla eksplitsiitselt väljendatud.

- (17) küll me **laseme** sulle peksta **anda** (SAA)
laskis sedavisi settul **laska** oma obese maha (IDA)
maa **lasin** oma mälettuse järel **teha** rätsep leikkab sedasi viistakki (RAN)

Modaalkonstruktsioon võib kanda ka võimaldamistähendust, kus kohustatuse aste on nõrgem. Väljendatakse situatsiooni, kus keegi, miski või mingi olukord on loonud võimalused või eeldused *da*-infinitiiviga väljendatava tegevuse toimumiseks (18). „Võimaldaja” võib olla ka mingi olukord, nähtus vmt. Seda konstruktsiooni käsitleb Kasik (2001: 106–107) lubava kausatiivse konstruktsioonina.

- (18) ta **lask** iks üise **magada** (TAR)
kui see ristija koeo **lasti tulla** (LÄÄ)
lask raadiul **ütteldä** (VÖR)

Nii nagu kirjakeele kuratiivseks kausatiivikonstruktsioonis (Kasik 2001:107–108) on ka murretes vastava modaalkonstruktsiooni *laskma da*-infinitiiviga väljendatava tegevuse subjektid erinevad. Niisiis võib neid kahe erineva – kohustava ja võimaldava – tähendusega konstruktsiooni variante pidada ka

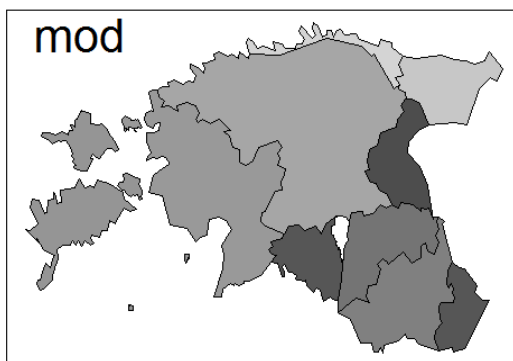
homonüümseteks konstruktsioonideks. Siinse käsitluse kohaselt on nad modaalse konstruktsiooni variandid.

Instrumentikonstruktsioon

Kuigi instrumentaalse tähendusega konstruktsiooni näiteid oli vaid üks läänemurde alalt, olen selle liigitanud eraldi konstruktsiooniks põhjusel, et sellise tähenduse väljendamisel on selgelt kohustuslik mingi instrumendi või instrumentaalsena kujuteldava eseme väljendamine (19). Selline konstruktsioon on Kasiku (2001: 112) järgi selgelt kausatiivne.

- (19) kui **lase käppud käijä** siis seal oli suur pappi aav (LÄÄ)

Kuigi selleski konstruktsioonis on võimalik modaalne tõlgendus, on siin kausatiivsus selgemini esil kui modaalkonstruktsioonis. Joonis 17 esitab finiiitse *laskma* ja *da*-infinitiivi modaalkonstruktsiooni geograafilise sagedusjaotuse. Instrumentaalkonstruktsiooni kohta kaarti esitatud pole, kuna selle esinemine materjalis oli marginaalne.



Joonis 17. Finiitse *laskma* ja *da*-infinitiivi modaalkonstruktsiooni geograafiline sagedus-esitus.

laskma-modaalkonstruktsiooni põhjal joonistub murreteüleselt võrreldes teiste modaalkonstruktsioonidega (*võima*, *tohtima*) erinev pilt. See modaalkonstruktsioon näib olevat iseloomulikum ida- ja lõuna- ja idalõuna murretele, samas kui kirde- ja rannamurdes kasutatakse konstruktsiooni teistega võrreldes vähem. Eelpool kirjeldatud modaalkonstruktsioonid paistsid sageduse poolest silma eelkõige põhja- ja läänepoolsetes murretes.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 18 esitab finiiitse *laskma* verbiga statistiliselt tugevamini tõmbuvad verbid *da*-infinitiivi vormis. Erinevalt paljudest teistest produktiivsustabelitest, on siin kõrgeima väärtusega verb (*tulema*) hoopis läänerühmas. Siiski on rohkem tugevamini tõmbuvaid verbe idarühmas, mis lubab jällegi oletada ühendi

suuremat produktiivsust läänepoolsetes murretes. Verbisemantilisi erinevusi kahe rühma vahel selgelt välja tuua pole võimalik.

Tabel 18. *laskma*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
viima	1.72	tulema	2.13
minema	1.47	tooma	1.04
hapnema	1.37	võtma	1.04
panema	1.14	jahtuma	0.82
olema	0.93	jooksma	0.8
magama	0.68	kasvama	0.8
petma	0.68	lugema	0.8
tõstma	0.68	laskma	0.62
kuivama	0.6	ligunema	0.62
käima	0.56	peksma	0.53

9.1.6 *da*-infinitiiv ja finitiivne *olema*

olema ja *da*-infinitiivi konstruktsioonide liigitamise põhimõtete väljatöötamine oli kahtlemata selle töö keerukaim. Nende konstruktsioonide tõlgendamise keerukusele kirjakeeles viitab ka Penjam (2008: 50–51). *olema* kuulub varasemate käsitluste järgi eesti keele tuumverbide hulka ning koos *da*-infinitiiviga võib see väljendada kohustust (Pajusalu *et al.* 2004: 34–35; Tragel 2003: 27–28; Uuspõld 1989). Nii Tauli (1980) kui ka Rätsepa (1978) järgi jagunevad sellised ühendid mitme erineva verbiühendite rühma vahel ning mõlemad autorid võtavad oma liigitustes arvesse ka muid potentsiaalse konstruktsiooni elemente peale finitiitse ja infiniitse verbi. Käesolevas töös pole konstruktsioonide vormi poolt sellise detailsusega analüüsitud, kuid mainitud autorite liigitused on paljuski kattuvad allpool esitatuga. Konstruktsioone, mis moodustuvad *olema* ja *da*-infinitiiviga, on eesti ja soome keeles nimetatud koopulaga kogejakonstruktsioonideks (Jaakola 2003). Verb *olema* võib kõigis läänemeresoome keeltes infiniitiviga väljendada erinevaid tähendusi: tegevuse eesmärki, kohta, aega. Konstruktsioonid ei ole alati selged ja üheselt mõistetavad. (Saukkonen 1965: 159)

Kuigi kollostruktsioonilise analüüsi käigus osutus *da*-infinitiivi ja *olema* ühend statistiliselt kõigis murretes ebaoluliseks, pani ühendi kõrge esinemissagedus vaatama, kas ühendiga väljendatakse regulaarselt kindlaid tähendusi. Kokku tuvastasin uurimismaterjalist 805 *da*-infinitiivi ja *olema* verbi ühendit sisaldavat osaluset ning ühend moodustab arvukalt erinevaid konstruktsioone. Tabel 19 esitab *da*-infinitiivi ja finitiitse *olema*-verbiga moodustunud konstruktsioonide tähendused ja nende esinemisjuhtude **normaliseeritud** sagedused murretes.

Tabel 19. *da*-infinitiivi ja *olema*-verbi konstruktsioonid ning nende esinemisjuhtude normaliseeritud sagedused. Lühendid: lahk – lahkstruktsioon, prag – pragmaatilise funktsiooniga konstruktsioon, ment – mentaalse verbi konstruktsioon, maa – vahemaad väljendav konstruktsioon, mod – modaalkonstruktsioon, om – omamistähendusega konstruktsioon, omadus – omadust väljendav konstruktsioon, tarvis_vaja – modaalad-verbidega *tarvis* või *vaja* modaalkonstruktsioon, öt – öeldistäite-konstruktsioon.

	tarvis/ vaja								
	prag	maa	mod	om	öt	lahk	ment	omadus	
IDA	4	8	15	14	8	22	0	13	0
KESK	2	1	24	17	2	40	3	45	3
KIRDE	0	2	8	3	5	12	0	6	0
LÄÄNE	2	0	23	15	6	22	0	43	0
MULGI	2	2	18	2	7	23	0	15	0
RANNA	3	3	16	6	6	20	0	6	0
SAARTE	2	1	29	14	4	11	0	54	1
SETO	3	0	31	6	3	28	0	0	0
TARTU	1	0	28	7	3	18	0	4	0
VÖRU	2	2	25	18	2	24	0	17	1

Modaalkonstruktsioon ja *tarvis/vaja* konstruktsioon

Modaalse tähendusvarjundi võib suuremal või vähemal määral omistada kõigile *da*-infinitiivi ja finiidse verbi *olema* abil moodustatud konstruktsioonidele. Modaalseks konstruktsiooniks olen siinses töös aga nimetanud ühemõtteliselt kohustust (20) või võimalikkust (21) väljendavat konstruktsiooni, milles modaalse agenta väljendatakse (kui see lauses esineb) alalütlevas käändes.

- (20) noorik **olli ehitada** (KES)
neil **oli** oma rent **maksa** põllu iist ja vakkamaade pealt (MUL)
aga teise oomikku **oli** jälle pool kolm **minna** (SAA)
- (21) siis **es ole** änam poes küündlad **saaca** (SAA)
naase **eij ol** latsi kohege **jättäq** (VÖR)
ja selle kohta **põle** änam midagi **ütelda** (LÄÄ)

Näites (20) on lauseid võimalik ümber sõnastada teiste modaalverbide abil (*tulema*, *pidama*) ning näites (21) on võimalik lisada lausesse omadussõna *võimalik*. Ühendi *saada olema* (21) võiks liigitada ka täiesti omaette püsiühendiks, kuid kuna sellisel püsiühendil on kindlasti ka modaalne tähendus, siis olen sellised juhtumid arvanud modaalkonstruktsiooni esindajateks.

Saukkoneni (1965: 163–165) järgi on ülal esitatud näidete kohustuse tähendus *olla* ja *da*-infinitiivil välja arenenud vaid karjala aunuse, lüüdi, vepsa, saami ja ka eesti keeles ning ta peab seda saksa, vene ja mõnedes murretes ka läti keele mõjuliseks. Veel toob ta välja eraldi rühmana ühendid, kus *olema* verbil

võib olla täiesti verbi *pidama* tähendus (s.t väljendab kohustuslikkust) (22), mis on tema hinnangul samuti vene keele mõjuline ja üldiselt iseloomulikum idapoolsetele läänemeresoome keeltele.

- (22) kuiss hundõlõ **oll üldäq** (SET)
aga teise oomikku **oli** jälle pool kolm **minna** (SAA)

Eraldi konstruktsioonina võib käsitleda modaalkonstruktsiooni, kus kohustuslike elementidena lisanduvad modaaladverbid *tarvis* ja *vaja* (23). Sellised juhtumid moodustavad omaette rühma ka EKG (Erelt *et al.* 1993: 238) järgi ning Penjami (2008: 51) järgi kuuluvad need kirjakeeles netsessiivkonstruktsiooni esindajate hulka. Huvitav on, et kirjakeeles väga tavaline adessiivne või allatiivne argument, mida on kirjeldanud Jaakola (2003), oli käesoleva töö uurimismaterjalis pigem harv nähtus ning modaalset agenti murretes selles konstruktsioonis enamasti ei väljendata.

- (23) ega siis **ei ole** korraga *tarvis* nii **ahmida** mitte (KES)
siss **eij olõ vaja võtta** masinat endä ütten (VÕR)
kass meil *vaija* siin nõnda kavva **olla omm** (MUL)

Kui agenti üldse väljendati, oli see adessiivis; allatiivis agendiga juhtumeid uurimismaterjalis ei esinenud. Selle konstruktsiooni puhul on huvitav ka modaaladverbide *tarvis* ja *vaja* geograafiline distributsioon. See on üks nähtus, mis väga selgelt tõmbab piiri taas ida- ja läänepoolsete murrete vahele ning traditsiooniline põhja- ja lõunaeeesti murrete vaheline piir on tunduvalt hägusam: *tarvis* esineb sagedamini lääne- ja põhjapoolsetes murretes, *vaja* aga domineerib Lõuna-Eestis ja idamurdes (Lindström & Uiboed 2013; Lindström & Uiboed valmimas).

Omamistähendusega konstruktsioon

Omamistähendusega konstruktsiooni kasutatakse, et osutada mingile esemele, omadusele või võimele, mis on vajalik *da*-infinitiiviga väljendatava tegevuse sooritamiseks (24).

- (24) **ei ole** omal änäm julgust üläl **olla** (LÄÄ)
sie **oli** mõisa asi **teha** (KES)
seppa pojäl **õli** siis ikke mõtte seda möldri tüttard ära **võtta** (IDA)

Konstruktid on ümbersõnastatavad translatiivis deverbaali abil (*kellelgi või millelgi on mingi asi/omadus/võime vmt millegi tegemiseks*). *da*-infinitiiv on siin atributiivses rollis ja mõnevõrra problemaatiline on selle konstruktsiooni lugemine finiiitse ja infiniitse verbi konstruktsioonide hulka. Pöördeline verb *olema* esineb siin koopulana ning *da*-infinitiivis verb laiendab hoopis noomenit või noomenifraasi (*julgus üleval olla, mõisa asi teha, mõte ära võtta*), seega on ühend sarnane omajakonstruktsiooniga. Kuna vaadeldav ühend esineb suhteliselt sageli, võib oletada, et see on siiski teatud määral kinnistunud.

Mentaalse verbi konstruktsioon

Mentaalse verbi konstruktsioon koosneb finiiitsest *olema* verbist ja mentaalse tähendusega *da*-infinitiivis verbist (25). Kuigi tähenduse poolest ei erine see konstruktsioon palju näiteks omamis- ja omaduskonstruktsioonist, siis kirjelduse ühtlustamise eesmärgil (vt sünonüümset *saama* ja *da*-infinitiivi konstruktsiooni eespool) on seda käsitletud neist konstruktsioonidest eraldi.

- (25) **põle** änam seekst asja **kuulda oln** (LÄÄ)
 töösehh külähh no siist **omm nättäq** treppi pealt (VÕR)
 kõik **oli** siis mõisal **tiada** (KIR)

Kuigi konstruktsioonide tähendused on nii mentaalsete *olema*- kui ka *saama*-ühendite puhul sarnased, tundub uurimismaterjali põhjal, et *olema*-verbiga mentaalse verbi konstruktsioonid on vähem kinnistunud kui sama konstruktsioon verbiga *saama* moodustunult, s.t *olema* ja *da*-infinitiiv mentaalse konstruktsiooni puhul ei ole tajutavad nii kinnistunud ühendina kui finiiitse *saama* verbi kujuline mentaalne konstruktsioon. Selliseid mentaalse verbi *olema* konstruktsioone on omaette rühmaks pidanud ka Saukkonen (1965: 167–168), kes peab sellist nähtust läänemeresoome keeltest iseloomulikuks just eesti keelele. Eesti keele akadeemiline grammatika nimetab selliseid ühendeid ahelverbideks (Erelt *et al.* 1993: 246). Penjami (2008: 139) järgi kuuluvad ühendid enamjaolt passiivse ja impersonaalse *olema*-konstruktsiooni esindajate hulka.

Vahemaad väljendav konstruktsioon

Eraldi konstruktsiooni moodustab vahemaa väljendamine finiiitse *olema* ja *da*-infinitiivis verbiga. Sellisel konstruktsioonil on fikseeritud vorm ja tähendus (26). Konstruktsioon väljendab liikumisverbide *minema* ja *tulema* abil vahemaa, kus või kui kaugel miski paikneb.

- (26) Kolju ots jahh sie **oli** siit vähe länne puole siit **minna** Ülgasi rand
 (RAN)
 ilusamb maja **omm** kura siist sinnäq jaama poolõ **minnäq** (VÕR)
 omm e sält Tindi puult **tulla** või sält Ainest müüdä (MUL)

Kirjakeelse materjali põhjal nimetab Penjam seda konstruktsiooni kohamäärus-konstruktsiooniks ja kirjeldab seda järgnevalt:

„Tegemist on konstruktsiooniga, kus alus märgib kirjeldusobjekti, mille asukohta selgitatakse, predikaadiks on seisundiverb (*olema, jääma*) ning kohamääruseks *da*-infinitiivitarind, mille abil väljendatakse objekti kaugust vaatluspunktist mõõdetuna sinna jõudmiseks vajaliku tegevuse ulatuse abil.” (Penjam 2008: 100).

Täpselt samuti võib kirjeldada konstruktsiooni ka murretes. Tähelepanuväärne on, et kui Penjam (2008: 100) on kirjakeelses uurimismaterjalis tuvastanud ainult ühe sellise konstrukti, siis suhteliselt väiksest murdekorpuse materjalist

tuvastasin konstruktsiooni esindajaid seitseteist. See lubab oletada, et konstruktsioon on eriti iseloomulik kas kõnekeelele või murretele. Geograafilise distributsiooni põhjal võib konstruktsiooni pidada iseloomulikumaks idapoolsetele murretele.

Öeldistäite-konstruktsioon

Öeldistäite-konstruktsiooni iseloomustab mõne adjektiivi, adverbi või noomeni abil *da*-infinitiiviga sooritatavat tegevust (27). Konstruktsioon on tähenduse poolest osaliselt kattuv kirjakeelse hinnangukonstruktsiooniga (Penjam 2008: 57). Murdematerjali põhjal selles konstruktsioonis isikut ei väljendatud.

- (27) eentte ses **oli** kogoni tore **magada** (LÄÄ)
siin **on** hää **magada nugastada** (RAN)
selle nu **omm** rassõ pidosit **pittäkki** nu väega (SET)

Lahkkonstruktsioon, pragmaatilise funktsiooniga konstruktsioon ja omadust väljendav konstruktsioon

Lahkkonstruktsiooni (28) esinemisjuhtusid tuvastasin materjalist kaks, kuid kuna tegemist on piisavalt erilise vormilise juhtumiga, ei olnud autori arvates õigustatud selle liitmine olemasolevate konstruktsioonidega.

- (28) sie **oli** vanasti mõisnikku tarkkus sis sedasi **ehittada** ja **olla** (KES)

Konstruktsioon moodustub justkui kahest eraldi seisvast osalausest, kuid finiidne *olema* on siin taas pigem koopula rollis. Siingi on *da*-infinitiiv noomenit laiendav ning finiidse verbiga *olema* pole infiniitne komponent eriti tugevalt seotud.

Pragmaatilise funktsiooniga konstruktsiooni (29) kohustuslik lisaelement on küsisõna *mis*. Kõnekeeles on sellised hüüatused või retoorilised küsimused üsna levinud. Konstruktsioonil on selgelt eriline suhtlusfunktsioon, mis eristab seda teistest *olema*-verbi *da*-infinitiivi konstruktsioonidest. Konstruktsioonis on *da*-infinitiivi positsioon vabalt täidetav ja põhitähendus tekib vormilise, tähendusliku ja pragmaatilise terviku kaudu.

- (29) ei *miss* sialt ikka **varastada oli** (KES)
miss säl **juttustatta ongi** (RAN)
noohh *mis* tuu siss **oll sõitta** tuu jao (VÕR)

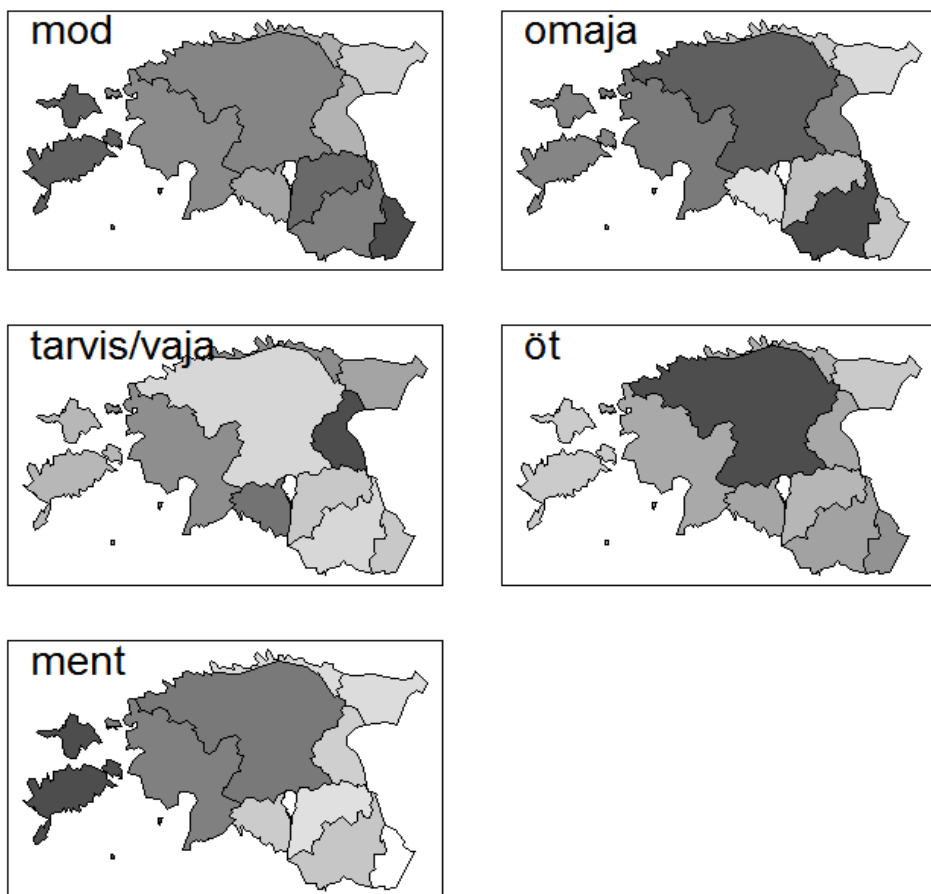
Omadust väljendava konstruktsiooni esinemisjuhtusid on materjalis vaid mõned üksikud (30). Siingi talitleb finiidne *olema* koopulana ning tähelepanu keskpunktis olevad kaks verbi on omavahel süntaktiliselt nõrgalt seotud.

- (30) tuu **ol** siss sitta **heittäq** vigõl (VÕR)
ja **kraasida** kraasid **olid** jo (KES)

da-infinitiivis verbivorm laiendab noomenit ja konstruktsiooni võib siin ümber sõnastada translatiivis noomeni abil (*kraasimiseks mõeldud kraasid*). Seega on

da-infinitiiv siin otsatarvet väljendava funktsiooniga. Saukkonen (1966:56) on sellise konstruktsiooni kujunemist pidanud läänemeresoome keeltest omaseks just eesti keelele ja ta väidab, et kaugematest sugulaskeeltest on *da*-infinitiivil selline kasutuskontekst välja kujunenud veel mordva keeltes. Kindlasti väärivad esitatud juhtumid edasiuurimist rohkem *da*-infinitiivi funktsioonide kontekstis.

Joonis 18 esitab finiiitse *olema* ja *da*-infinitiivi konstruktsiooniesindajate geograafilise sagedusjaotuse. Kaardid on esitatud vaid sagedasemate konstruktsioonide kohta. Selle rühma modaalkonstruktsioon näib olevad eriti iseloomulik Seto murdele. Harvem tuleb seda esile ida- ja kirderanniku murretes. Omamistähendusega konstruktsioon toob esile Võru ja keskmurde. Modaalised konstruktsioonid adverbidega *tarvis* ja *vaja* esinevad sagedamini idamurdes. Õeldistäite-konstruktsioon on kõige iseloomulikum keskmurdele. Mentaalsed konstruktsiooni esindajad tulevad sagedamini esile saarte, kesk- ja läänemurdes ning muudel murdealadel esineb seda väga harva või üldse mitte.



Joonis 18. Finiitse *olema* ja *da*-infinitiivi konstruktsioonide geograafiline sagedusesitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 20 esitab ida- ja läänepoolsetes murretes verbiga *olema* tugevamini tõmbuvad *da*-infinitiivis verbid. Pilt on suhteliselt sarnane finiiitse verbi *saama* ja *da*-infinitiivi ühendite omaga. Siiski tundub, et verbi *olema* puhul pole ka idarühmas niivõrd tugevaid verbielistusi. Mõlemas rühmas on tugevamini tõmbuvate verbide seas mentaalset võimet või tegevust väljendavaid verbe. Idarühmas on esindatud paljud verbid, mis on lõunaeestilised, kuna kõik lõunaeesti murded kuuluvad just sellesse rühma.

Tabel 20. *olema*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
tundma	2.06	teadma	1.48
kaema	1.42	viima	1.46
kõnelema	1.42	müüma	1.15
elama	1.1	tooma	1.08
jooma	0.96	tarvitama	0.9
jätma	0.94	vaatama	0.89
ketrama	0.94	võtma	0.85
koolema	0.94	nägema	0.82
lööma	0.94	kuulma	0.77
mõskma	0.94	jahvatama	0.72

Idarühma kõrgeima FTT-väärtusega verbid *tundma* ja *kaema* ning läänerühma *teadma* viitavad mentaalse verbi konstruktsioonide püsiühendilaadsusele, seega võib ka statistilise analüüsi põhjal pidada põhjendatuks selle konstruktsiooni eristamist muudest.

9.1.7 *da*-infitiiv ja finiiitne *tahtma*

Kõigis läänemeresoome keeltes väljendab *tahtma* tüüpiliselt soovi, kavatsust ja tahtmist (Saukkonen 1966: 18). Trageli (2003: 45) järgi kuulub *tahtma* eesti keele tuumverbide hulka ja ta toob selleks ühe põhjendusena välja verbi modaaltähenduse leviku ka elututele agentidele. Rätsepa (1978: 189) järgi kuulub *tahtma* sekundaarsete verbivormidega elementaarlausete ühekohaliste elementaarlausete rühma tüüpi N+nom. V¹V²+da, mis on ka murretes selle ühendi puhul levinuim tüüp. Verbi *tahtma* tähenduse muutumist ja abiverbi funktsioonide omandamist on kirjakeeles põhjalikult kirjeldanud Habicht jt (2010) ning *tahtma* abiverbifunktsioonide omandamist on vaadelnud Erelt ja Metslang (Erelt 2009; Erelt & Metslang 2009).

Murretes on verbiühendi kohta tehtud mõned üksikud tähelepanekud. Potsepsi (1956) järgi esineb *tahtma* *da*-infinitiiviga liitõeldise koosseisus sagedasti saarte, Võru ja Seto murdes, vähem keskmurdes ning teiste murrete juures pole ta seda eraldi välja toonud. Neetari (1988) järgi on väljendatakse sel viisil kaht tähendust: ühelt poolt elusolendi soovi või tahet midagi teha ning teisalt 3. pöördes *tahtma* tähistab *da*-infinitiivis väljendatud tegevuse sooritamise vajalikkust, kusjuures tema järgi võib *da*-infinitiivi positsioonis esineda ka mõni teonimi, mis ei muuda konstruktsiooni tähendusi. Esimest konstruktsiooni (soovi ja tahet väljendav) peab ta omaseks kogu eesti keele alale. Teist lausetüüpi (tegevuse sooritamise vajalikkus) *da*-infinitiiviga on Neetari järgi vähem kirja pandud saarte murdes. Eraldi toob Neetar välja vormi *tahtma* + *da*-infinitiiv + *saada*, mida ta peab omasemaks läänepoolsetele murretele. Konstruktsiooni läänepärasust kinnitavad ka Juhkam ja Sepp (2000: 29).

Kokku tuvastasin uuritavas materjalis 558 *da*-infinitiivi ja finiidset verbi *tahtma* sisaldavat osalause. Konstruktsioone analüüsides jagasin need nelja erinevat tähendust kandvaks. Tabel 21 esitab konstruktsioonid ja nende esinemisjuhtude **normaliseeritud** sagedused murrete kaupa.

Tabel 21. *da*-infinitiivi ja *tahtma*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: mod – modaalkonstruktsioon, proks – proksimatiiv-konstruktsioon, soov – soovi väljendav konstruktsioon, vastupanu – vastupanu väljendav konstruktsioon.

	mod	proks	soov	vastupanu
IDA	0	1	81	3
KESK	2	0	66	1
KIRDE	1	2	27	2
LÄÄNE	5	0	49	1
MULGI	0	0	82	0
RANNA	7	0	44	0
SAARTE	2	0	50	0
SETO	0	0	31	0
TARTU	1	0	47	0
VÕRU	0	0	39	0

Kollostruktsioonilises analüüsis ei osutunud ühend statistiliselt oluliseks Võru, Tartu, Seto, Mulgi ja kirdemurdes.

Soovi väljendav konstruktsioon

Kõige sagedamini moodustavad *da*-infinitiiv ja finiidverb *tahtma* soovimistähendusega konstruktsiooni, milles soovitavat tegevust tähistab *da*-infinitiivis verb (31).

- (31) no mes sa ulluga **taht tehä** (IDA)
 a perähh **taht** yks naist **võtta** (VÖR)
 krahv **ei taha andada** enamb tagasi (KIR)

Soovija on väljendatud nominatiiviga ja nõuab enamasti elus osalist, mõnikord võib see olla siiski ka elutu, ent ka siis väljendab osa-terviku suhete kaudu sisuliselt elus osalist (32).

- (32) a mõis **tahi** oma tiobäiväd **saada** (KIR)
 vot südä **tahi süijä** (KIR)

Niisiis tuleneb murretes soovi väljendava konstruktsiooni tähendus suuresti tahtma verbi semantikast ning *da*-infinitiivis verb talitleb objektina, millele finiitverbiga tähistatav „tahtmine” on suunatud.

Penjami (2008: 75) liigituse põhjal esindavad sellised juhtumid kirjakeeles soovimis-kavatsemiskonstruktsiooni. See on murretes tüüpilisim *da*-infinitiivi ja finiitse *tahtma*-verbi kasutusviis, muid konstruktsioone võis tuvastada vaid üksikutes lausetes. Käesolevas töös pole vahet tehtud kavatsus- ja soovimis-tähendusel, nagu kirjakeele puhul on teinud Habicht jt (2010), samad autorid viitavad artiklis ka ise nende kahe tähenduse eristamise raskustele.

Vastupanu väljendav konstruktsioon

Vastupanu väljendav konstruktsioon on vormi poolest alati eitav. Selles puudub soovimis- või kavatsemistähendus ning grammatiline subjekt viitab elutule osalisele (33).

- (33) ega ta siin iast **kasva ei taha** (IDA)
 aga sis paber **ei taha** ka ästi **põleda** (KES)
 ikka kaa ikka kaa ega **äi tahagid unuda** (SAA)

Habicht jt (2010: 135) on sama konstruktsiooni nimetanud õnnestumismo-daalseks konstruktsiooniks, Saukkoneni (1966: 22) järgi väljendatakse sel viisil tegevuse toimumise mittesovimist.

Konstruktsioon väljendab murdematerjali põhjal mingite asjaolude ebasoodsust *da*-infinitiivis väljendatava tegevuse toimumiseks.

Modaal- ja proksimatiivkonstruktsioon

Modaalkonstruktsioonis väljendab *tahtma* verb *da*-infinitiivis verbiga tähistatava tegevuse sooritamise vajadust või kohustust (34). Grammatiline nominaatiivis subjekt on alati elutu. Habicht jt (2010: 131) järgi väljendab selline konstruktsioon tänapäeva eesti keeles modaalset osalejavolist vajalikkust.

- (34) et sie asi **tahtub** ige **oijendada** (RAN)
 e nattukke rohkem kül ta²⁰ **aududa tahtis** kui kana oma muna (LÄÄ)
 ja tüü **tahtsõ tettä** (TAR)

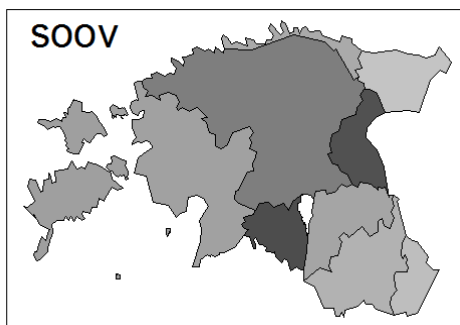
²⁰ ta = hanemuna

Tauli (1980: 138) järgi moodustab selline lausetüüp rühma verb aktiivis + *da*-infinitivis transitiivne verb + nominatiivis noomen. Penjam (2008: 92) nimetab sarnast konstruktsiooni sobilikkuskonstruktsiooniks.

Proksimatiivi kujunemist kirjakeeles on vaadelnud Erelt ja Metslang (Erelt 2009; Erelt & Metslang 2009), kelle järgi väljendab proksimatiiv mingi oleviku- või minevikusündmuse peatset algamist, kuid lahtiseks jääb, kas sündmus realiseeru(s)b või mitte. Proksimatiivkonstruktsiooni juhtumeid *da*-infinitivi ja *tahtma*-verbiga tuvastasin uurimismaterjalist kolm (35). Näite (35) kaks esimest lauset tõenäoliselt avertiivsed, s.t et *da*-infinitiviga väljendatav tegevus tõenäoliselt aset ei leia.

- (35) ja jällä ta **tahab** paergu piaaegu et kaela **kukkuda** (IDA)
ja jälle uuwesta ühegõrra **tahatsime** lapsed **upputada** (KIR)
ja **tahtsid** siis tugevasti tuld **võttada** otsa (SAA)

Joonis 19 esitab finiiitse *tahtma* ja *da*-infinitivi soovi väljendava konstruktsiooni geograafilise sagedusjaotuse. Kuna teisi konstruktsioone esines selles rühmas väga vähe, siis nende jaotust esitatud pole.



Joonis 19. Finiitse *tahtma* ja *da*-infinitivi soovi väljendava konstruktsiooni geograafiline sagedusesitus.

Tumedamad alad tähistavad taas kõrgemat sagedust ning heledamat madalat. Eriti sage on konstruktsioon Mulgi ja idamurdes, teistes murretes on sagedus pisut madalam. Selgelt eristuvaid suuremaid piirkondi soovi väljendava konstruktsiooni puhul ei näi olevat.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 22 esitab kahes murderühmas finiiitse verbiga *tahtma* tugevamini tõmbuvad verbid *da*-infinitivi vormis. Pilt on mõlemas rühmas suhteliselt sarnane, tugevamini tõmbuvaid verbe on rohkem idapoolsetes murretes, mille põhjal võiks eeldada idamurretes madalamat produktiivsust. Samas on selles rühmas rohkem verbe, mis ei allu agendi tahtele (*surema*, *olema*), mille põhjal võiks eeldada, et siin on vähem levinud muidu sage soovi väljendav konstruktsioon.

Tabel 22. *tahtma*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
kõnelema	2.62	jooma	1.82
sööma	2.39	saama	1.82
surema	1.74	käima	0.92
olema	1.48	hakkama	0.85
kuulutama	1.3	abielluma	0.8
andma	1.25	panema	0.72
kuduma	0.87	hoidma	0.6
kuulama	0.87	keema	0.6
saatma	0.87	kinkima	0.6
teenima	0.87	kuulma	0.6

9.1.8 *da*-infitiiv ja finiiitne jõudma

Nii Rätsepa (1978: 188) kui Tauli (1980: 136) järgi kuulub finiiitne *jõudma* ja *da*-infitiiv lausetüüpi V^1V^2+da . Penjam (208: 77) on sellise vormiga väljendatavat nimetanud soovimis-kavatsemiskonstruktsiooniks. Murretes pole sellist ühendit autorile teadaolevalt käsitletud. Tabel 23 esitab finiiitse verbiga *jõudma* ja *da*-infitiivi suutlikkuskonstruktsiooni ja seda sisaldavate osalausset **normaliseeritud** sagedused. Kollostruktsioonilises analüüsis osutus ühend konstruktsioonina statistiliselt oluliseks vaid ida- ja Tartu murdes. Kokku oli konstrukte korpuses 132.

Tabel 23. *da*-infitiivi ja *jõudma*-verbi suutlikkust väljendava konstruktsiooni (suut) esinemisjuhtude normaliseeritud sagedused.

	suut
IDA	21
KESK	9
KIRDE	12
LÄÄNE	12
MULGI	16
RANNA	7
SAARTE	4
SETO	31
TARTU	21
VÕRU	27

Suutlikkuskonstruktsioon

jõudma ja *da*-infinitiivi suutlikkuskonstruktsiooni semantika on murretes tingitud täielikult *jõudma*-verbi semantikast ja konstruktsioon võib väljendada füüsilist suutlikkust *da*-infinitiivis väljendatava tegevuse sooritamiseks (36).

- (36) aga vat **ess jõvva** enämp **kõndi** (TAR)
aga kus ta **jõus** neid ärä **vedädä** nii (MUL)
kiss pagan seda aeda **jõudis** siis nii paelu nii tihedaks **teha** (LÄÄ)

Samuti võib konstruktsiooni jõudmistähendus olla mentaalset laadi, väljendades mõne võime või oskuse vajalikkust, olemasolu või puudumist *da*-infinitiivis verbiga tähistatud tegevuse sooritamiseks (37).

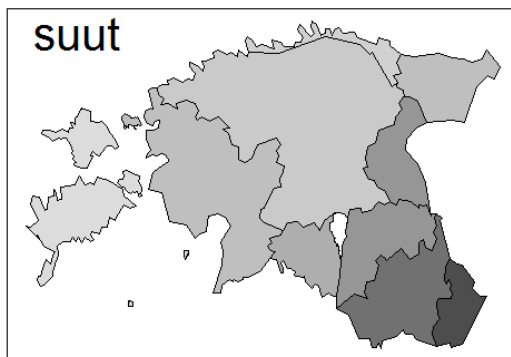
- (37) kess soo unõnäo mullõ ärr **jõvvass juudustada** (SET)
mina **ei jõuwa** enamb **lugeda** (KIR)
ei jõudnd ää **oodatta** (IDA)

Veel võib jõudmine olla tingitud mingi ressursi olemasolust või puudumisest (38).

- (38) et siss **jõvvavad** meele aridust **anda** (IDA)
vallakool **ei ei jõun õppettada** (KES)
a kess tuul **aiga jõudsõ** siss sai ja **osta** (SET)

da-infinitiivis väljendatud tegevuse sooritamiseks vajalik ressurs ei pea olema konstruktis eksplitsiitselt väljendatud, vaid see võib olla implitsiitselt tuletatav. Eri tähendusvarjundite põhjal võib oletada küll *jõudma*-verbi mõningast tähenduse laienemist, kuid tähenduserinevused on nähtavasti siiski veel liiga väikesed, et nende põhjal võiks rohkem konstruktsioone eristada.

Joonis 20 esitab finiiitse *jõudma* ja *da*-infinitiivi suutlikkuskonstruktsiooni geograafilise sagedusjaotuse. Konstruktsioon näib olevat iseloomulikum lõuna-eesti murderühma murretele ja idamurdele. Analüüs osutas küll väikestele, aga siiski märgatavatele erinevustele lõunaeesti murretes, kus *jõudma*-konstruktsioon väljendas sagedamini mentaalset jõudmist kui füüsilisest ressursist tulenevat jõudmist.



Joonis 20. Finiitse *jõudma* ja *da*-infinitiivi suutlikkuskonstruktsiooni geograafiline sagedus-
esisitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 24 esitab finiiitse verbiga *jõudma* tugevamini tõmbuvad verbid *da*-infinitiivi vormis. Erinevalt mõnest muust ühendist pole siin nii selget FTT-väärtuste erinevust kahe rühma vahel. Pigem võiks nende väärtuste põhjal väita, et verb on enam-vähem võrdse produktiivsusestmega mõlemas rühmas. Ida-rühmas on tugevamini tõmbuv verb *olema*, mis kinnitab selle verbi ja *da*-infinitiivi püsiühendilaadsust, nagu seda eelnevalt kirjeldasin. Läänerühma kõige tugevamini tõmbuv *lõikama* pole lingvistiliselt seletatav, ilmselt on tegemist materjali eripäraga, mis ilmneb statistilises analüüsis, kuna verb esineb ainult ja mõned korda läänerühma murretes. Verbisemantiliselts kaks rühma ei erine, kuid näiteks idarühmas on esindatud verbid (*maksma* ja *ostma*), mis viitavad ressursist tingitud suutlikkusele.

Tabel 24. *jõudma*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
olema	1.49	lõikama	1.28
võtma	1.05	lööma	0.85
maksma	0.84	seletama	0.85
viima	0.84	sööma	0.85
andma	0.63	vedama	0.82
harutama	0.42	tegema	0.59
lugema	0.42	ajama	0.42
ostma	0.23	hangeldama	0.42
jutustama	0.21	jooksma	0.42
kitkuma	0.21	juhtima	0.42

9.1.9 *da*-infinitiiv ja finiiitne *oskama*

oskama verb on tuntud kõigis läänemeresoome keeltes ning väljendab tüüpiliselt oskamis- ja suutmistähendust (Saukkonen 1966: 46–53). Rätsep (1978: 189) on *oskama* verbi ja *da*-infinitiivi liigitanud sekundaarsete verbivormidega kahekohaliste elementaarlausete rühma kujul N+nom. V¹ V²+da (*Peeter ei oska ujuda*). Tauli (1980: 136–137) järgi talitleb *da*-infinitiiv koos finiiitse *oskama*-verbiga objektitaoliselt ning tegevus on suunatud just *da*-infinitiivis verbile. Eesti kirjakeeles võib finiiitne *oskama* väljendada võimalikkusmodaalsust ja modaliseeritavat situatsiooni tähistab sihitise funktsioonis *da*-infinitiiv (Erelt et al. 1993: 185). Penjam (2008: 77) on sellised ühendid koondanud soovimiskavatsemiskonstruktsiooni mõiste alla.

Murdekorpuse materjalist tuvastasin 137 *da*-infinitiivi ja finiitset *oskama*-verbi sisaldavat osalauset ning sellise ühendiga väljendatakse *da*-infinitiiviga tähistatava tegevuse oskamist või mitteoskamist, niisiis tingib selle konstruktsiooni semantika suuresti finiitses positsioonis esinev *oskama*-verb. Tabel 25 esitab oskamiskonstruktide **normaliseeritud** sagedused kõigis murretes.

Tabel 25. *da*-infinitiivi ja *oskama*-verbi oskamist väljendava konstruktsiooni (oskam) esinemisjuhtude normaliseeritud sagedused.

	oskam
IDA	13
KESK	13
KIRDE	20
LÄÄNE	22
MULGI	0
RANNA	25
SAARTE	16
SETO	0
TARTU	6
VÕRU	1

Kollostruktsioonilise analüüsi käigus osutus ühend konstruktsioonina statistiliselt oluliseks vaid saarte ja läänemurdes.

Oskamiskonstruktsioon

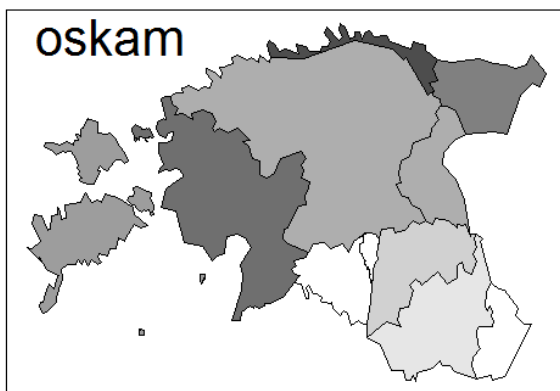
Oskamiskonstruktsioon väljendab mingit võimet, mentaalset või füüsilist ressursi *da*-infinitiivis väljendatud tegevuse sooritamiseks (39). Agenti väljendatakse nominatiiviga.

- (39) **oska** juu rääkkimesega kõikki välla **kueottada** (LÄÄ)
 kui nad kuradid **ep oska** see tanu Muhu tanu pähe **panna** (SAA)
 nuu üts kiä külä päl kess siss **osass mängi** (VÕR)

Esines üksikuid näiteid, kus konkreetne oskamistähendus on natuke laienenud. Selliseid juhtumeid võiks ümber sõnastada fraasiga *millegi peale tulema* (40).

- (40) ju nii vana et **oskas** mehele **minna** (KES)
mõned teesed nooremad lollimad **ei oskagi** nii **kipputa** ühti kui see
vana tark (LÄÄ)
ja isa tuld siis **eij oskand vahtidagi** (RAN)

Joonis 21 esitab finiiitse *oskama* ja *da*-infinitiivis verbiga väljendatavate oskamiskonstruktsiooni geograafilise sagedusjaotuse. Üldse mitte või väga vähe esineb konstruktsioon lõunaeesti murderühma murretes. Osaliselt on see seletatav kindlasti *oskama*-verbi põhjaeestilisusega, lõunaeesti murretes kasutatakse tõenäoliselt samas kontekstis sageli verbi *mõistma*. Huvitav on ka konstruktsiooni kõrgem esinemissagedus ranna-, kirde- ja läänemurdes, samas kui kesk- ja saarte murdes selle sagedus sama kõrgele ei tõuse.



Joonis 21. Finiitse *oskama* ja *da*-infinitiivi konstruktsiooni geograafiline sagedusesitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 26 esitab finiiitse *oskama* verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid. Selle ühendi puhul on kahe murderühma vaheline pilt sarnasem kui teiste ühendite puhul. Kummaski rühmas ei ilmne ühtegi väga tugevasti tõmbuvat verbi ning näib, et selles osas on *oskama* üsna ühtlaselt produktiivne mõlemas rühmas. Kummaski rühmas esineb kümne tugevamini tõmbuva verbi hulgas rohkem mentaalset suutlikkust väljendavaid verbe.

Tabel 26. *oskama*-verbiga tugevamini tõmbuvad *da*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
seletama	1.04	ütleva	1.09
arvestama	0.71	kuduma	0.48
kõnelema	0.71	laskma	0.39
kündma	0.71	rääkima	0.34
niitma	0.71	küsima	0.19
rehkendama	0.71	luiskama	0.19
taguma	0.71	minema	0.19
lugema	0.59	panema	0.19
hakkama	0.45	sööma	0.19
keetma	0.45	tantsima	0.19

9.1.10 Kokkuvõte

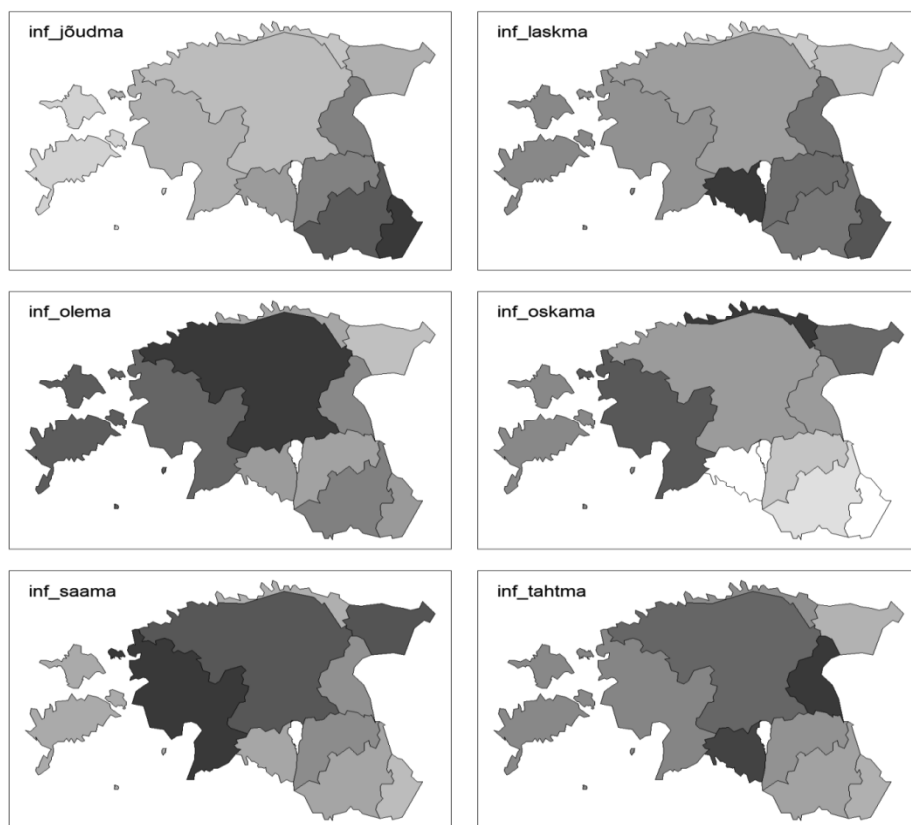
Joonised 22 ja 23 esitavad kvantitatiivse analüüsi põhjal verbiühendite sagedused murdealadel. Siin on sagedused esindatud kobaranalüüsile iseloomulikult koondatuna ja konstruktsioonide tähendusi eristatud pole. Seega osutavad joonised verbiühendite, mitte konstruktsioonide sagedusele. Tumedamad alad tähistavad kõrgemat sagedust, heledamad alad madalamat sagedust.

jõudma-ühendid näivad olevat sagedasemad lõunaeesti murretes ja idamurdes, kõrgeim on sagedus Seto murdes. Sarnane on muster ka *laskma*-ühendite puhul, kuid siin on sagedus kõrgeim Mulgi murdes. *olema*-ühendid näivad olevat eriti iseloomulikud keskmurdele ning ülejäänud aladel on sagedusjaotus üsna sarnane. *oskama*-ühendeid eelistavad kirde-, ranna- ja läänemurre ning lõunaeesti murderühma murretes on sellised ühendid harvad. *saama*-verbi ja *da*-infinitiivi ühendid on iseloomulikumad lääne-, kesk- ja kirdemurdele ning *tahtma*-ühendid Mulgi ja idamurdele. *tohtima*-ühendid on sagedasemad Mulgi murdes ning muudest murretest harvem esineb neid Seto ja rannamurdes. *tulema*- ja *võima*-verbi ning *da*-infinitiivi ühendid on sagedased keskmurdes ja teistes murretes leidub neid vähem.

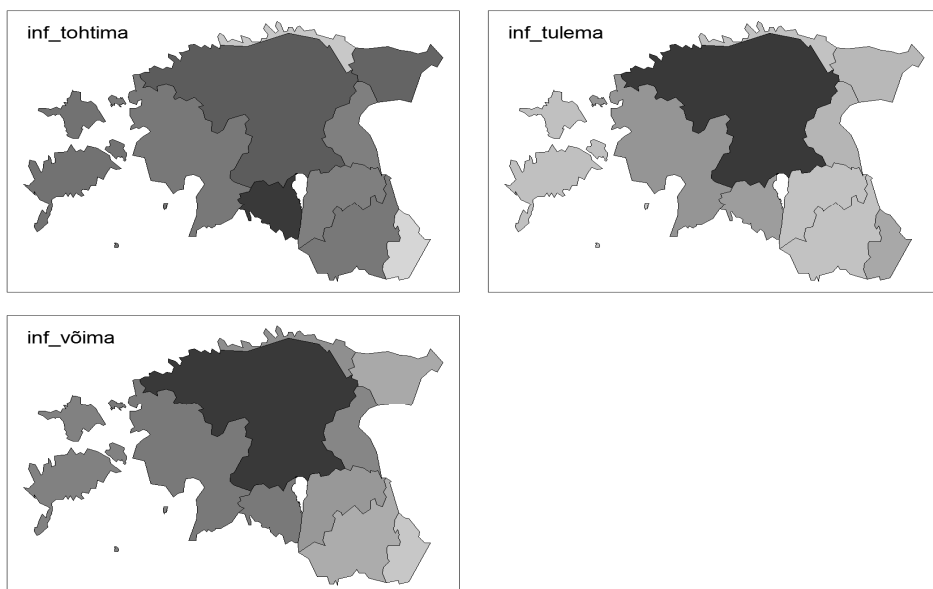
Sageduskaartide põhjal ei tundu murded ühe konkreetse ühendi ja selle sageduste poolest mingeid selgeid rühmi moodustavat. Siin ilmneb kobardialektoloogia suur eelis. Murded moodustavad harva üksikute keelenähtuste põhjal selgelt eristuvaid grupe, aga kui kokku panna rohkem ja erinevamaid nähtusi, saab murderühmade kohta üldisemalt tunduvalt rohkem informatsiooni ning rühmaerinevused tulevad selgemini esile nii, nagu see ilmnes kvantitatiivses kobaranalüüsis peatükis 7, mis kaasas analüüsi kõik teatud kriteeriumite alusel tuvastatud verbiühendid ning analüüsis neid koos. Tulemused esitati kvantitatiivses kobaranalüüsis koondandmete põhjal, mitte üksikute ühendite põhjal, ja tulemuste põhjal moodustusi selged murdegrupid.

Huvitav on ka võrrelda eespool esitatud konstruktsioonide sageduskaarte verbiühendite sageduskaartidega. Ilmneb, et mitte alati pole need sarnased. Seega võib oletada, et paljudes murretes esinevad küll samad verbiühendid, kuid nende abil väljendatavad tähendused on mõnevõrra varieeruvad. Konstruktsioonide analüüs tõi esile veel huvitavaid tendentse. Kui tüüpilised modaalkonstruktsioonid modaolverbidega *saama*, *võima*, *tohtima* on üsna tavalised põhjaeesti murretes, eriti keskmurdes, siis näiteks *tahtma*, *olema* (*tarvis*, *vaja*), *laskma* modaalsete konstruktsioonide puhul selline jagunemine nii selge ei ole. See lubab oletada, et lõunaeesti ja idapoolsetes murretes on tüüpilisemad teistsugused modaalsuse väljendamisega viisid kui tüüpiline modaolverb + *da*-infinitiiv. See tõik väärrib kindlasti põhjalikumat uurimist tulevikus.

Kuigi harvem tuleb ette, et mingi tähendusega konstruktsioon on iseloomulik üksiku(te)le murretele, siis kindlamalt võib väita, et erinevused ilmnevad selgelt konstruktsiooni esindajate kasutussageduses. Seega on uurimistulemused kooskõlas näiteks Kortmanni (2010: 842) väitega, mille kohaselt süntaktiliste nähtuste varieerumine on pigem kasutussageduse küsimus kui mingi kindla keelenähtuse puudumine või olemasolu kindlas murdes.



Joonis 22. *da*-infinitiivi ja finitiitse verbi ühendite sagedused (1).



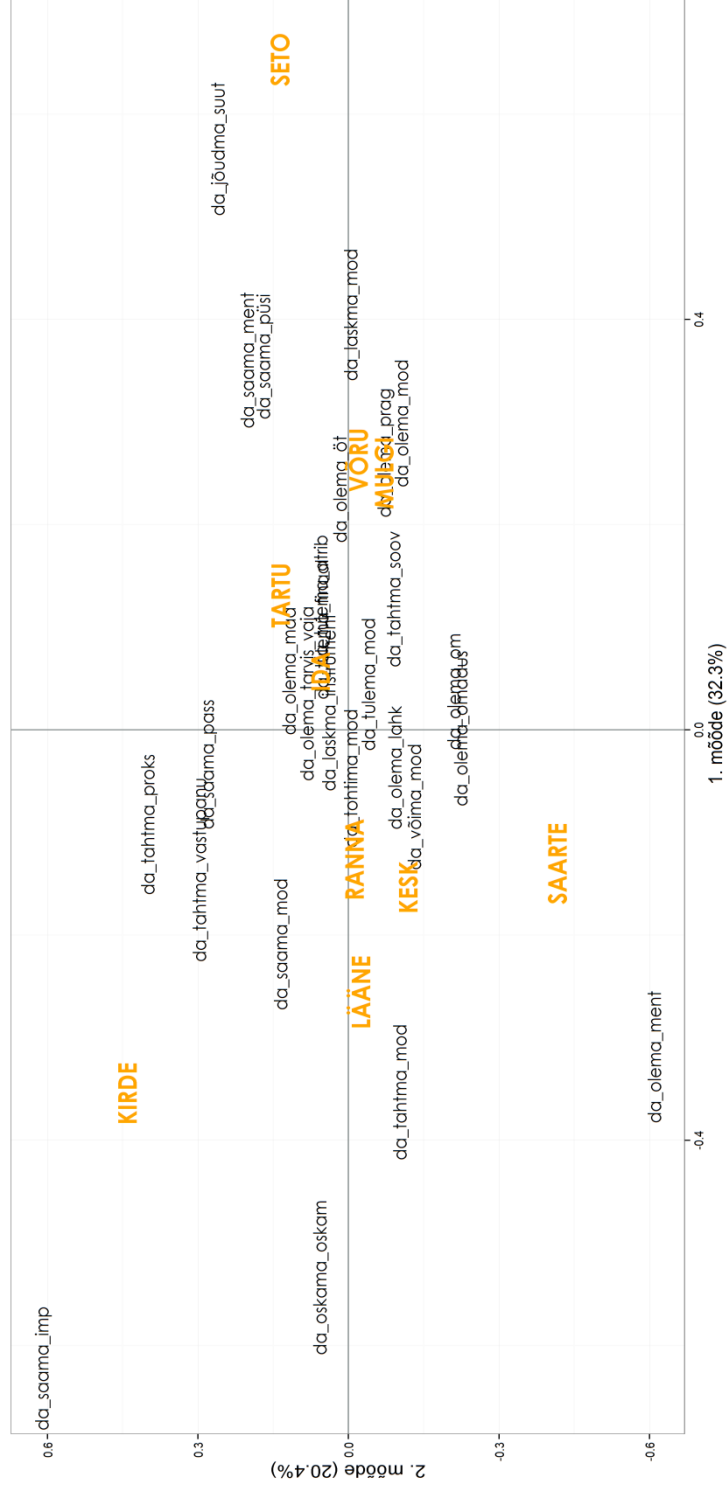
Joonis 23. *da*-infinitiivi ja finiitse verbi ühendite sagedused (2).

Joonis 24 esitab *da*-infinitiivi ja finiitse verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafiku konstruktsioonianalüüsi tulemuste põhjal. Graafikul esitatud kaks kõige seletavat mõõdet seletavad üle 50% andmestiku variatiivsusest. Esimese mõõtme põhjal eristuvad traditsiooniline põhja- ja lõunaeesti murderühm, erinevuseks on vaid idamurre, mis näib olevat sarnasem lõunaeesti murretega. Teine mõõde eristab muudest murretest selgelt kirde-, saarte ja Seto murde. Kogu joonisel tekib neist kolmest murdest huvitav kolmnurk: kolm murret näivad olevat erinevad nii omavahel kui ka kõigist teistest murretest.

jõudma-, *laskma*- ja erinevad püsiühendikonstruktsioonid näivad olevat iseloomulikumad lõunaeesti murretele. *saama*-impersonaalikonstruktsioon on iseloomulik kirdemurdele. Modalkonstruktsioonidest on lõunaeesti murretele (ja idamurdele) tüüpilisemad *olema*- ja *laskma*-verbiga moodustatavad variandid, samas kui sünonüümne *saama*-konstruktsioon näib olevat põhjaeestilisema levikuga. Huvitav on aga, et mentaalse verbi konstruktsioonide puhul on *olema* ja *saama* oma põhja- ja lõunaeestilised positsioonid vahetanud. Suurem osa konstruktsioonidest on koondunud graafiku keskele, mis tähendab, et need ei näita tugevamaid geograafilisi eelistusi.

Korrespondentsanalüüs: da-infinitiivi konstruktsioonid

a konstruksioon a murre



Joonis 24. *da*-infinitiivi ja finiitse verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafik.

da-infinitiivi konstruktsioonide analüüs osutab selgelt idamurde sarnasusele lõunaeesti murretega ja see erineb traditsioonilistest murdejaotustest, mille järgi idamurre kuulub põhjaeesti murrete hulka. Muus osas jagunevad mured traditsiooniliselt põhja- ja lõunaeesti murreteks. Esimeses rühmas erineb teistest suurel määral kirdemurre ning lõunaeesti rühmas Seto murre.

Tabel 27 esitab veel kord ülevaatlikult materjalist tuvastatud *da*-infinitiivi ja finitiise verbi konstruktsioonid koos näidetega.

Tabel 27. Finiitse verbi ja *da*-infinitiiviga moodustatavad konstruktsioonid eesti murretes.

VORM	TÄHENDUS	NÄIDE
<i>da</i> -infinitiiv		
jõudma	suutlikkus	<i>ei jõudnd ää oodatta</i> (IDA)
laskma	instrumendikonstruktsioon	<i>kui lase käppud käija</i> siis seal oli suur pappi aav (LÄÄ)
	modaalne	<i>küll me laseme sulle peksa anda</i> (SAA)
olema	lahkkonstruktsioon	<i>sie oli</i> vanasti mõisnikku tarkkus sis sedasi <i>ehitada ja olla</i> (KES)
	vahemaa	<i>ilusamb maja omm</i> kura siist sinnäq jaama poolõ <i>minnäq</i> (VÕR)
	mentaalne	<i>tõõsehh külähh no siist omm nättäq</i> treppi pealt (VÕR)
	modaalne	<i>noorik olli ehitada</i> (KES)
	omamistähendus	<i>jahh oll</i> aigu latsi ka <i>hoolidaq</i> jahh (VÕR)
	omadus	<i>ja kraasida</i> kraasid <i>olid</i> jo (KES)
	pragmaatiline	<i>miss sialt ikka varastada oli</i> (KES)
	tarvis/vaja modaalsus	<i>siss eij olõ</i> vaja <i>võtta</i> masinat endä ütten (VÕR)
olema	öeldistäite-konstruktsioon	<i>eentte ses oli</i> kogoni tore <i>magada</i> (LÄÄ)
oskama	oskamine	<i>nuu üts kiä külä päl kess siss osass mängi</i> (VÕR)
saama	impersonaal	<i>nohh seda poissibõlves sai</i> kaa <i>teha</i> ikke t õitsis <i>käija</i> (KIR)
	mentaalne	<i>sial saeme külmettada</i> küll (LÄÄ)
	modaalne	<i>ma i saa kraappi</i> (IDA)
	passiiv	<i>mina sai</i> esä käest <i>tõrelda</i> (TAR)
	püsiühend	<i>süija sai</i> küll seal (SAA)
tahtma	modaalne	<i>ja tüü tahtsõ tettü</i> (TAR)
	proksimatiiv	<i>ja jällä ta tahab</i> paergu piaaegu et kaela <i>kukkuda</i> (IDA)
	soovimistähendus	<i>krahv ei taha andada</i> enamb tagasi (KIR)
	vastupanu tähendus	<i>ega ta siin iast kasva ei taha</i> (IDA)

VORM	TÄHENDUS	NÄIDE
tohtima	modaalne	<i>kuradi muudu ma haigõss tohtsõ jäädä tütrik (VÖR)</i>
tulema	atributiivne konstruktsioon finaaltähendus modaalne	<i>siis tulnd möisast käsk teule minna (SAA)</i> <i>noo juba viimasel aeal tullid poodi ka lambid müüja (KES)</i> <i>tuli pidusid teha (LÄÄ)</i>
võima	modaalne	<i>et mina võin sulle juhattatta (IDA)</i>

Konstruktsioonide analüüs osutas, et kvantitatiivse analüüsi käigus tuvastatud konstruktsioonikandidaadid tõepoolest moodustavad keelekonstruktsioone ja kaks vaadeldavat vormi ei esine koos juhuslikult. Tulemused näitasid, et enamasti peidavad sellised ühendid enda taga nii homonüümseid kui polüseemseid konstruktsioone. Homonüümsete konstruktsioonide puhul on tavaline, et üks neist on märkimisväärselt kõrgema esinemissagedusega. Sellest tulenevalt võimaldab selline analüüs tuvastada ka harva esinevaid konstruktsioone, mis tulevad esile oma kõrge esinemissagedusega homonüümi tõttu. Kuigi eesti kirja-keeles tehakse vahet perifrastilistel, ahelverbidel ja verb + komplement ühenditel (Erelt *et al.* 1993: 18–20; Erelt 2013: 49–55), siis käesolevas töös seda tehtud ei ole, sest nende kõigi eristamise kriteeriumid pole kirjakeeleski kuigi selged.

Paljud kirjeldatud finiiitsetest verbidest on kindlasti omandamas või juba omandanud selgeid abiverbide funktsioone, kuid seda erinevates murretes kindlasti erineval tasemel. Siinse töö lähtekoht on Bolingeri (1980) väide, mille kohaselt kõigil verbidel, mis võtavad infiniitse komplemendi, on potentsiaali areneda abiverbiks. Paljud sellistest verbidest jäävad oma arengus abiverbistumise protsessi algusesse, teised võivad muutuda täiesti grammatilisteks üksusteks. Murretes käituvad selgelt (modaalsete) abiverbidena, nii nagu kirja-keeleski, *olema*, *saama*, *tohtima*, *võima*, samas kui näiteks *jõudma* ja *oskama* on nähtavasti sellise potentsiaalse arengu algusjärgus. Seega võiks vaadeldud verbid paigutada kontiinumile, mille ühe ääre moodustavad selgelt grammatilisi funktsioone täitvad verbid ning kontiinumi teise äärde jääksid just sellised verbid nagu *oskama* ja *jõudma*. Viimaste puhul on siiski vähemalt teatud murretes märgata mõningat tähenduse laienemist abstraktsema suunas. Ülejäänud verbid võiks paigutada kusagile nende kahe rühma vahele. Kindlasti ei ole pilt samasugune kõigis murretes. Näiteks väga erinevaid konstruktsioone moodustavad *olema*- ja *saama*-verb osutavad selgelt, kuidas homonüümsete konstruktsioonide geograafiline distributsioon võib olla väga erinev. Kui koondada ühte kõik *saama*-verbi ja *da*-infinitiivi ühendid, siis näivad need olevat sageduse poolest eriti iseloomulikud lääne-, kesk- ja kirdemurdele. Sama vormiga moodustatavate eri konstruktsioonide puhul on pilt aga erinev. Näiteks on *saama*-verbi impersonaalkonstruktsioon selgelt väga tavaline kirdemurdes, samas kui modaalstruktsiooni puhul on murreteülene pilt üsna homogeenne.

olema-verbi ja *da*-infinitiivi konstruktsioonide puhul on pilt veelgi mitmekülgsem. Ühendi sagedus on kõrgeim keskmurdes, samas kui ülejäänud murre-

tes jaotuvad sagedused ühtlaselt. Sellekujuline öeldistäite-konstruktatsioon on iseloomulikum keskmurdele, *tarvis/vaja*-modaalkonstruktatsioon aga idamurdele. Mentaalse verbi konstruktatsioonid näivad olevad selgelt lääne-eestilised.

Need asjaolud osutavad vaadeldavate finiiitsete verbide erinevatele funktsioonidele eri murretes, aga ka erinevate konstruktatsioonide erinevale geograafilisele distributsioonile. See kinnitab, et võimatu on praeguste teadmiste juures eristada ahel- ja perifrastilisi verbe, kuid kindel on see, et selline jaotus pole samasugune kõigis murretes.

9.2 *ma*-infinitiivi konstruktatsioonid

ma-tegevusnimi ehk supiin on üks kahest eesti keele infinitiivist. *ma*-tegevusnimele võivad liituda mõned käändelõpud: inessiiv (-*mas*), elatiiv (-*mast*), abessiiv (-*mata*) ning translatiiv (-*maks*) ja kõik need vormid väljendavad erinevaid grammatilisi tähendusi. Supiini illatiivil ehk *ma*-tegevusnimel on finaalne tähendusvarjund ning see võib esineda sihtkoha (*Ta läks sööma.*), rektsooniadverbiaali (*Ta sunnib mind sööma.*), adverbiaalatribuudi (*sööma minek*), liitpredikaadi komponendi (*Ta hakkas minema.*) ja iseseisva öeldisena (*Ruttu hambaid pesema!*). Finaalse varjundiga on ka *mas*-vorm ning see võib esineda lauses asukohaadverbiaalina (*Kõik olid parajasti marju korjamas.*), rektsooniadverbiaalina (*Nägime neid uttu kadumas.*), adverbiaalatribuudina (*söömas käimine*) ning võib koos finiiitse *olema*-verbiga perifrastilise verbivormi koosseisus väljendada aspektitähendust (*Ma olin just väljumas.*). (Erelt *et al.* 1995: 65) Kõik supiinivormid võivad koos teatud finiiitsete verbidega esinedes moodustada perifrastilisi verbe ja/või verbivorme, siinse töö kontekstis erinevaid konstruktasioone. Käesoleva töö uurimismaterjalis on vaatluse all supiini illatiivi vorm ehk *ma*-tegevusnimi ning selle inessiivivorm ehk *mas*-vorm. Teiste supiini käändevormidega esinenud konstruktasioone oli murdematerjalis vähe ja kvantitatiivne analüüs nendega moodustatud konstruktasioone esile ei toonud, seega pole neid materjali analüüsimisel eraldi vaadeldud. Kirjakeeles on erinevaid supiinikonstruktasioone vaadeldud üsna põhjalikult. Penjam (2008) esitab oma konstruktasioonigrammatilises monograafias *ma*-infinitiividega moodustatavate konstruktatsioonide kirjeldused nii tänapäeva kui ka vanemas kirjakeeles. Erinevaid verbikonstruktasioone, milles *ma*-infinitiiv on üks moodustajatest, on vaadelnud erineva põhjalikkuse astmega paljud autorid (Erelt 2009; Erelt & Metslang 2009; Kasik 2001; Kehayov & Torn-Leesik 2009; Metslang 1993a; Mägiste 1936; Pajusalu *et al.* 2004; Rätsep 1978; Tauli 1980; Tragel 2003; Tragel & Habicht 2012 jne). Kõigist varasematest käsitlustest tuleb täpsemalt juttu vastavate konstruktatsioonide kirjeldamisel järgnevates alapeatükkides.

ma-infinitiivi ja finiiitse verbiga moodustunud konstruktatsioonide kohta on murretes tähelepanekuid siiani tehtud vähem. Neetar (1988: 42) on vaadelnud teonimetarindeid murretes ja leidnud, et finiiitne *olema* koos *ma*-infinitiivi inessiiviga väljendab kirderannikumurdes ja põhjaeesti murretes nii kestusega tegevusi (*õli väljal künmān*) kui ka üleminekut ühest olukorrast teise (*vana*

maja on längakile ja viltu – ümmer kukkumas). Huvitaval kombel toob Neetar samas välja, et ka Mulgi murre on põhjakeesti murderühmaga selles osas sarnane. Sarnase tendentsi on välja toonud ka murrete ühendverbide uurimus (Uihoaed 2010a). Veel märgib Neetar, et Tartu ja Võru murdes võib selline ühend väljendada vaid kestusega tegevust või protsessi (*naistõrahvass oll enne tanh kapstõid kitskman*). Tanning (2004: 123–124) on oma põhjalikus Karksi murraku ülevaates välja toonud mõned *ma*-infinitiivi funktsioonid. Erinevad murdeülevaated esitavad küll põhjalikke kirjeldusi nii infinitiivide kui ka partitsiipide vormivarieeruvusest, kuid napilt on infot infinitsete vormide funktsioonide ja veelgi vähem nende abil moodustatud konstruktsioonide kohta.

9.2.1 *ma*-infiniitiv ja finiitne *hakkama*

Eesti keele akadeemilise grammatika järgi kuulub *hakkama*-verb algust väljendavate faasiverbide hulka, mis moodustab ahelverbe koos supiinivormiga (Erelt *et al.* 1993: 20). Trageli (2003: 42) järgi kuulub *hakkama* ka eesti kirjakeele tuumverbide hulka. Rätsep (1978: 188, 191) liigitab supiini ja *hakkama* ühendid sekundaarsete verbivormidega ühe- ja kahekohaliste elementaarlausete hulka kujul V^1V^2+ma (*hakkab tibutama*) ja $N+nom. V^1V^2+ma$ (*Buss hakkas liikuma.*). Tauli (1980: 39–40; 133) liigitab finiitset *hakkama*-verbist ja infinitsetest verbist koosnevad ühendid modaalühendite gruppi (*Lõpuks hakati laulma.*). Finiitse *hakkama* ja supiinivormi modaaltähendusele osutab ka EKG (*Laps hakkas kukkuma.*) (Erelt *et al.* 1993: 184).

hakkama konstruktsioone on eesti kirjakeeles uuritud üsna põhjalikult. *hakkama*-verbist ahelverbide koosseisus on põhjaliku diakroonilise ülevaate andnud Piiraja (2010). Penjam (2008) on sellised verbiühendid enamjaolt liigitanud alguskonstruktsioonide esindajate hulka ning kirjeldanud konstruktsiooni nii tänapäeva kui ka vanemas kirjakeeles. Tuleviku väljendamisega võimalusi verbiga *hakkama* on vaadelnud Metslang (näiteks 1994a), kelle järgi väljendab *hakkama*-tulevik tulevikus algavaid, jätkuvaid või korduvaid aktiivseid sündmusi (Metslang 2006). R. Pajusalu jt (2004: 37–38) on eraldi rõhutanud veel selle konstruktsiooni neutraalse alguse aspekti.

Murretes pole finiitse *hakkama* ja supiini funktsioone eriti kirjeldatud. Univere (1966) on põhjalikult vaadelnud *hakkama* ja *nakkama* tähendusi, keskendumata oluliselt konstruktsioonidele, mida nende verbidega moodustatakse.

Käesoleva töö kvantitatiivses analüüsis olen ühtlustanud lõunaeestilise *nakkama* põhjakeestilisele kujule *hakkama* ja selliseid ühendeid on vaadeldud koos, kuigi nende kahe verbi etümoloogiaid on erinevateks peetud (Uiho 2010; Univere 1966). Kuna käesolevas töös vaadeldavate konstruktsioonide seisukohalt etümoloogilised erinevused olulist tähendust ei oma ning ka konstruktsioonide koosseisus käituvad verbid suhteliselt sarnaselt, siis võib verbide ühtlustatud kujule viimist põhjendatuks pidada²¹. Kvalitatiivse analüüsi tule-

²¹ Käesolevas töös nimetatakse nii *nakkama* kui ka *hakkama* konstruktsioone *hakkama*-konstruktsioonideks.

muste kirjeldamisel on näitelausetes esitatud verbi tegelik esinemiskuju tekstis, vastavalt siis kas *nakkama* või *hakkama*.

Tuvastasin oma uurimismaterjalis 1421 osalauset, kus esinesid nii *h(n)ak-kama* kui ka *ma*-infinitiiv. Kollostruktsioonilise analüüsi tulemusena osutus ühend statistiliselt oluliseks kõigis murretes. Tabel 28 esitab murretes tuvas-tatud *hakkama*-konstruktsioonid ning nende esinemisjuhtude **normaliseeritud** sagedused.

Tabel 28. *ma*-infinitiivi ja *hakkama*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: algus – algust väljendav konstruktsioon, püsi – püsiühend.

	algus	püsi
IDA	120	1
KESK	160	0
KIRDE	131	0
LÄÄNE	136	0
MULGI	34	0
RANNA	168	0
SAARTE	209	1
SETO	108	0
TARTU	81	0
VÕRU	103	0

Finiitse *hakkama* ja supiini abil väljendatakse murretes enamasti tegevuse algusfaasi ja verbiühend võib moodustada ka püsiühendeid. Järgnevalt annan neist lühikese ülevaate.

Algust väljendav konstruktsioon

Finiitse *hakkama* ja *ma*-infinitiivi abil väljendatakse murretes sarnaselt kirja-keelega enamjaolt mingi sündmuse algust (41).

- (41) kess koloosis **akkab** piäd **kandma** (IDA)
 akkavad mulle viha **kandama** (KIR)
 siss **nakka** mina nüt vanna inemist säält **koristama** viil (TAR)

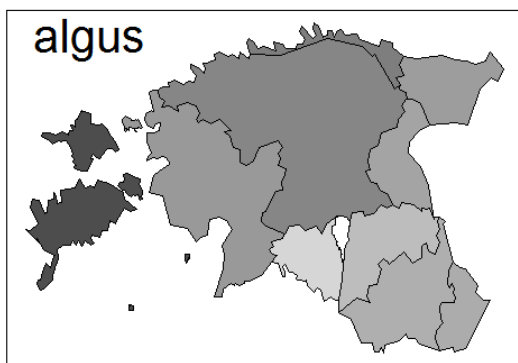
Selles töös pole *ma* eristanud tuleviku- ja algustähendust, mille eristamise keerukusele on viidanud teisedki sama konstruktsiooni käsitlenud autorid (Penjam 2008: 107; Piiraja 2010: 54), kes on konstruktsiooni analüüsinud tunduvalt põhjalikumalt, kui käesoleva töö eesmärgiks on seatud. Sageli on mõle-mad tähendused konstruktsioonis teineteisega tihedalt põimunud.

Püsiühendid

Püsiühendites ei puudu algustähendus, kuid töö autori arvates on sellised esinemisjuhud piisavalt erinevad kinnistunud ühte tähendusse. Püsiühendeid tuvastasin materjalist kolm (42).

- (42) **akka** vai **nutma** (IDA)
siis lehmad **akkavad** piima **lüksma** (SAA)

Kuigi sellesse rühma kuuluvaid ühendeid on vähe, on nad erineva tähendusega. Esimene, idamurde osalause näites (42) on kindla suhtlusfunktsiooniga lausung, millel on antud juhul ka kohustuslik element sõna *või*. Kuigi konstruktsioonis on olemas ka algustähendus nagu teistel vaadeldava ühendiga moodustunud konstruktsioonidel, on sellel juhtumil siiski selgelt eriline pragmaatiline funktsioon, mis eristab seda neutraalsest algust väljendavast konstruktsioonist. Teises, saarte murde konstruktis *lüksma hakkama* on selgelt olemas algus- ja tulevikutähendus, kuid autori arvates on ühend piisavalt kinnistunud tähendusega, et seda teistest eristada. Joonis 25 esitab finiiitse *hakkama* ja *ma*-infinitiivi alguskonstruktsiooni geograafilise sagedusjaotuse. Konstruktsioon on sagedasem saarte murdes, kuid märkimisväärseid sageduserinevusi ei näi murretes selle konstruktsiooni kasutuses olevat. Mõnevõrra harvem kasutatakse seda Mulgi murdes. Seega selgelt eristuvaid suuremaid rühmi selle konstruktsiooni puhul ei eristu. Kuna püsiühendeid oli materjalis vaid mõni üksik, siis nende kohta visuaalseid andmeid esitatud ei ole.



Joonis 25. Finiiitse *hakkama* ja *ma*-infinitiivi alguskonstruktsiooni geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 29 esitab finiiitse verbiga *hakkama* tugevamini tõmbuvad verbid *ma*-infinitiivis ida- ja läänepoolsete murrete rühmas. Pilt näib suhteliselt sarnane mõlemas rühmas ning statistiliselt tugevamini seotud verbe on rohkem ida-

rühmas. Statistiliste seoste tugevuste põhjal võiks oletada vastavate paaride püsiühendilaadsust, kuid konstruktsioonide analüüs seda siiski ei kinnita. Taas on rohkem tugevamini tõmbuvaid verbe idarühmas, kuid selgeid semantilisi erinevusi rühmade vahel ei ilmne.

Tabel 29. *hakkama*-verbiga tugevamini tõmbuvad *ma*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
pesema	3.7	vedama	2.44
kandma	3.25	püüdma	1.32
peksma	2.39	andma	1.08
ehitama	1.87	panema	1
kaevama	1.73	laskma	0.99
kartma	1.73	minema	0.97
kõnelema	1.73	tegema	0.95
röökima	1.73	käärima	0.94
tahtma	1.71	tantsima	0.94
nutma	1.58	käima	0.92

9.2.2 *ma*-infinitiiv ja finiiitne *minema*

Rätsepa (1978: 191–192) järgi kuulub supiini ja *minema* ühend sekundaarsete verbivormidega kahekohaliste elementaarlausete hulka kujul N+nom. V¹V²+ma (*Aeg läks uitamisega kaduma.*). Rätsep esitab oma liigituse suurema detailsuse astmega, kui seda käesolevas töös teen, ning tema järgi jaotub ühend olenevalt muudest elementidest eri rühmade vahel. *minema* kuulub eelnevate käsitluste kohaselt samuti eesti keele tuumverbide hulka ja Tragel (2003: 34–36) peab verbi grammatiseerumise potentsiaali suureks. Eesti keele akadeemilise grammatika järgi võivad sellised ühendid kuuluda seotud supiinitarindite hulka, mis väljendavad millegi või kellegi ümberpaigutumist ruumis ning lisaks kohatähendusele on sellisel tarindil ka finaalne tähendus (Erelt *et al.* 1993: 252–253). Samuti võib *minema* moodustada supiiniga perifrastilise verbi(vormi), kus finiiitne *minema* talitleb algust väljendava faasiverbina (Erelt *et al.* 1993: 258–259).

Urimismaterjalist tuvastasin 450 finiiitse *minema* ja *ma*-infinitiivi sisaldavat osalauset. Kollostruktsioonilises analüüsis osutus ühend statistiliselt oluliseks kõigis murretes. Tabel 30 esitab finiiitse verbi *minema* ja *ma*-infinitiivi abil moodustatavad konstruktsioonid ja nende esinemisjuhtude **normaliseeritud** sagedused kõigis murretes.

Tabel 30. *ma*-infinitiivi ja *minema*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: afadv – afiksaaladverbikonstruktsioon, algus – algust väljendav konstruktsioon, lok – lokatiivse tähendusega konstruktsioon, püsi – püsiühend.

	afadv	algus	lok	püsi
IDA	7	10	35	0
KESK	9	6	33	2
KIRDE	3	3	27	0
LÄÄNE	3	4	33	0
MULGI	0	2	20	0
RANNA	7	3	38	0
SAARTE	2	4	32	0
SETO	14	9	85	0
TARTU	1	6	41	0
VÕRU	9	6	34	0

Eesti keele akadeemiline grammatika eristab kohta ja otstarvet väljendavaid määruslikke *ma*-infinitiivi tarindeid (Erelt *et al.* 1993: 252–254), kuid käesolevas töös pole neil vahet tehtud samadel põhjustel, millele viitab ka Penjam (2008: 158): praktilises analüüsis osutub selliste määruste vaheline piir ähmaseks.

Järgnevad alajaotised kirjeldavad uurimismaterjalist tuvastatud *ma*-infinitiivi ja finiiitse *minema* verbi abil moodustatavaid konstruktsioone.

Lokatiivne konstruktsioon

Lokatiivse konstruktsioon hõlmab kohatähenduse selle kõige laiemas mõttes, kus finiiitse verbiga väljendatakse supiiniga tähistatavas tegevusse suundumist (43). Lokatiivne konstruktsioon on murretes selgelt kõige tavalisem, mida *ma*-infinitiivi ja finiiitse *minema*-verbiga moodustatakse, nii nagu see on kirja-keeleski (Pajusalu & Orav 2008: 110). Murretes on selle konstruktsiooni grammatiline subjekt alati agentiivne.

- (43) aga õde **läks** kua mehele üheksadeistuaastaselt seda ilma iad elu
suama (KES)
 pannkuokkidega **läks** seda tite **vaatama** (RAN)
 ega neläbäiv ödagodi **lätsivä** üttekokko **heegeldämmä** (VÕR)

Eriti torkab silma lokatiivsete konstruktide kõrge esinemissagedus Seto murdes, samas kui kõigis teistes murretes on sagedused ühtlaselt madalamad. See lubab oletada, et Seto murdes on finiiitse *minema* ja supiinivormi ühend suhteliselt vähe grammatiseerunud, väljendades ülekaalukalt vaid lokatiivset-finaalset tähendust ning ühend pole veel laienenud teisi tähendusi kandma.

Sellist lokatiivset konstruktsiooni kirjakeeles on nimetatud liikumis-sobivus-konstruktsiooniks (Penjam 2008: 159–165). Liikumistähendus on sellistel konstruktsioonidel murreteski, mille tingib juba *minema*-verbi semantika.

Algust väljendav konstruktsioon

Algust väljendav supiini ja finiiitse *minema* konstruktsioon on sünonüümne *hakkama* verbiga moodustatud algust väljendava konstruktsiooniga, s.t et vastavates konstruktides on võimalik alati finiiitne *minema* asendada finiiitse *hakkama* verbiga (44).

- (44) et söss **läävät** jo ilmaq ilosale **minemä** ku (VÕR)
kui **läks** sie ülemine kord **põlema** (IDA)
nii ett **keema** juo **lähäp** ja (RAN)

Konstruktsioonile iseloomulik joon on elutu subjekti ilmnemine. EKG järgi (Erelt *et al.* 1993: 258) on sellised ühendid perifrastilised verbid, kus *minema* on algust väljendav faasiverb. Lokatiivsest konstruktsioonist eristabki algus-konstruktsiooni elutu subjekti ilmnemine, samuti on siin liikumistähendus tajutav abstraktsemal tasandil kui seda on verbi *minema* tüüpiline liikumistähendus.

Afiksaaladverbikonstruktsioon ja püsiühendid

Perifrastilistes verbivormides võivad liikumisverbid *minema* ja *tulema* käituda sarnaselt afiksaaladverbiga *ära* (Erelt *et al.* 1993: 259; Penjam 2008: 195). Penjam (2008: 195) on oma doktoritöös sellised ühendeid nimetanud leksikaliseerunud ühenditeks. Murretes võib verbiühend esineda sama funktsiooniga ja nii moodustub afiksaaladverbikonstruktsioon. Supiinivormis *minema* verb on antud juhul selgelt sünonüümne afiksaaladverbiga *ära* (45).

- (45) mees **läks** nõnna tänuga **minemä** (KES)
sis pühabä õhta **läksime minema** (SAA)
lätsi Mintkahe **minemä** (VÕR)

Konstruktsiooni iseloomustab veel tegevuse sooritaja elusus, vaid paaril korral oli tegevuse sooritaja esitatud kollektiivsubstantiivi fraasiga (46).

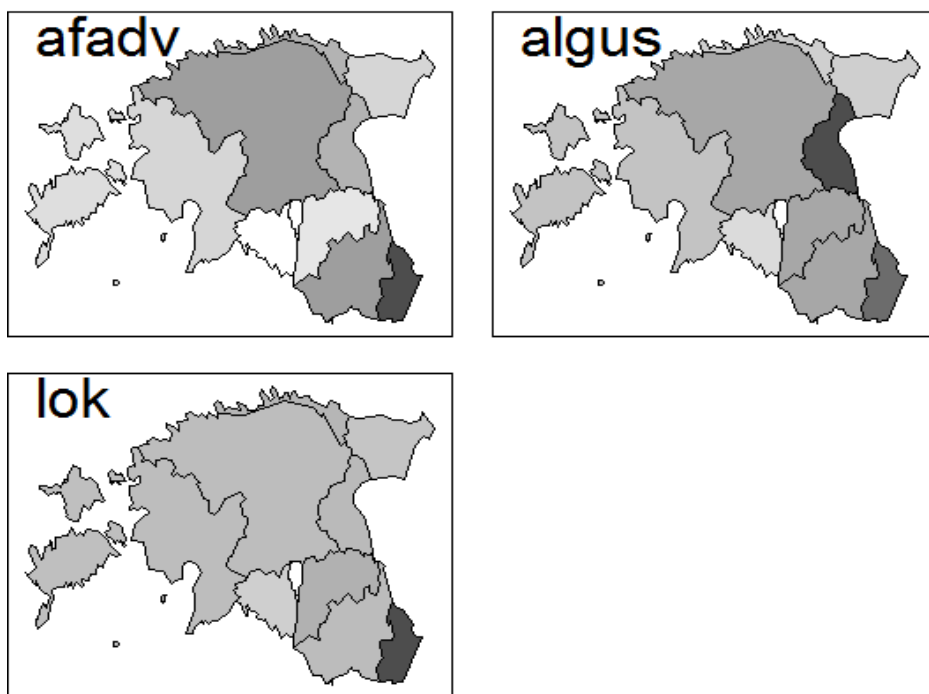
- (46) suo nüid kogu maa **lähäb minema** (RAN)
juba pulm **lähäb** Kodavere juba **minema** (IDA)

Paari näite puhul (**lätsi** Mintkahe **minemä** (VÕR), juba pulm **lähäb** Kodavere juba **minema** (IDA)) on üsna tugevalt olemas siiski veel algustähendus ning supiinivormis *minema* pole siin veel nii tugevalt leksikaliseerunud. See lubab oletada, et liikumisverbi afiksaaladverbistumine on tihedalt seotud lokatiivse tähendusega ning alguse väljendamine on grammatiseerumise järgmine aste.

Püsiühendite rühma moodustavad neli ühendit saarte ja keskmurdest. Siia kuuluvad finiiitse verbi *minema* ja supiinivormis *kaduma* ühendid (47). Penjam (2008: 196) on sellise ühendi liigitanud leksikaliseerunud ühendite rühma, Tauli (1980: 133) järgi kuulub selline ühend V+ma vormiga püsiühendite rühma.

- (47) mies **läks kaduma** nagu sõea aeal (KES)
mool **läks** see **kaduma** siis (SAA)

Joonis 26 esitab *ma*-infinitiivi ja finiiitse *minema*-verbi konstruktsioonide geograafilise sagedusjaotuse. Kaarti pole esitatud püsiühendite kohta, mida esines materjalis vaid paaril korral. Esitatud konstruktsioonide puhul torkab silma mõnevõrra erinev geograafiline jagunemine, kui seda eelnevate konstruktsioonide põhjal märgata võis. Supiini ja *minema*-verbi erinevaid konstruktsioone näivad märkimisväärselt rohkem kasutavad Seto ja idamurre. Afiksaaladverbikonstruktsioon on eriti iseloomulik Seto murdele. Sama näib olevat ka lokatiivse konstruktsiooniga, kus ükski teine ala ei näi sageduse poolest eristuvat. Algust väljendav konstruktsioon kõige tüüpilisem ida- ja Seto murdele.



Joonis 26. Finiitse *minema* ja *ma*-infinitiivi konstruktsioonide geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 31 esitab finiiitse verbiga *minema* tugevamini tõmbuvad *ma*-infinitiivis verbid ida- ja läänepoolsete murrete rühmas. Eriti torkab silma verbi *kaema* tugev statistiline seos idarühmas. See on kindlasti suurel määral seletatav lõuna-eesti murrete kuulumisega idarühma, läänerrühmas on lõunaeeetilise verbi esine-mine teoreetiliselt välistatud. Sellest tulenevalt saab vormipaar kõrge statistilise

seose tugevuse väärtuse. Ülejäänud statistikute väärtused jagunevad mõlemas rühmas suhteliselt ühtlaselt, mis lubab eeldada verbi *minema* ühtlast produktiivsust erinevate konstruktsioonide moodustamisel. Selgeid verbisemantilisi erinevusi kahe rühma vahel ei ilmne.

Tabel 31. *minema*-verbiga tugevamini tõmbuvad *ma*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
kaema	7.54	tooma	1.28
otsima	1.85	laskma	1.16
põlema	1.54	laulatama	1.16
kutsuma	1.33	hallitama	0.93
kuulama	1.16	hauduma	0.93
jooksma	0.77	kaduma	0.93
proovima	0.77	vedama	0.93
rabama	0.77	viima	0.84
rääkima	0.77	magama	0.79
võtma	0.75	vaatama	0.75

9.2.3 *ma*-infinitiiv ja finiidne *tulema*

Sarnaselt teiste analoogsete ühenditega kuulub ka *ma*-infinitiivi ja finiidse *tulema* ühend Rätsepa (1978: 191–193) liigituse järgi sekundaarsete verbi-vormidega kahekohaliste elementaarlausete rühma (nt *Tudengid tulid klubisse kohtuma ministriga.*), mille ta liigitab omakorda kitsamateks erinevate kohustuslike elementidega alamrühmadeks. Siinses töös lähtun vaid kahe verbi ühendist ning muid komponente nii detailset ei analüüsi. *tulema* kuulub eelnevate käsitluste kohaselt eesti keele tuumverbide hulka (Tragel 2003: 31–32). Sarnaselt *minema*-konstruktsioonidega moodustab ka *tulema*-verb supiiniga Penjami (2008: 159–165) järgi liikumis-sobivuskonstruktsiooni. EKG järgi (1993: 252; 258) võib supiini vormis verb finiidse *tulema* verbiga koos esinedes olla kohamääruse funktsioonis ja ühend võib esineda perifrastilise verbina, kus finiidne *tulema* väljendab faasiverbina mingi protsessi või tegevuse algust. Sarnaselt verbiga *minema* võib ka supiinivormis *tulema* finiidse *tulema*-verbiga koos esinedes saada afiksaaladverbi *ära* tähenduse.

Kokku tuvastasin uurimismaterjalist 157 finiidse *tulema* ja supiini ühendit sisaldavat osalauset. Kollostruktsioonilises analüüsis ei osutunud ühend konstruktsioonina statistiliselt oluliseks Mulgi, kirde- ja läänemurdes. Tabel 32 esitab finiidse *tulema* ja supiinivormi konstruktsioonid nende esinemisjuhtude **normaliseeritud** sagedused kõigis murretes.

Tabel 32. *ma*-infinitiivi ja *tulema*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: afadv – afiksaaladverbikonstruktsioon, püsi – püsiühend, lok – lokatiivse tähendusega konstruktsioon.

	afadv	püsi	lok
IDA	1	0	29
KESK	1	1	14
KIRDE	0	0	12
LÄÄNE	0	0	10
MULGI	0	0	11
RANNA	0	0	25
SAARTE	0	0	10
SETO	0	0	31
TARTU	4	0	17
VÕRU	1	0	17

Ühendi sagedus on kõrgem Seto, ranna- ja idamurdes ning ülejäänud murretes jaguneb sagedus suhteliselt ühtlaselt. Järgnevalt kirjeldan konstruktsioone detailsemalt.

Lokatiivse tähendusega konstruktsioon

Sarnaselt eelmises peatükis kirjeldatud *minema*-konstruktsioonidega on ka *tulema*-konstruktsioonide tavalisim tähendus lokatiivne, s.t et *ma*-infinitiivis verbil on lokatiivne ja/või finaalne tähendus ning finiidne *tulema* osutab liikumisele selle tegevuse või eesmärgi suunas (48). Tegevuse sooritaja esineb grammatilise subjektina ning on alati elus.

- (48) keda te **tulette** siin **sõimama** (IDA)
ja tää **tuli** siis oma meeleärma muule **rääkmä** (LÄÄ)
ja tämä **tuleb** oma lollakkalle viel **õtsimaje** (KIR)

Penjami kirjakeele käsitlese kohaselt (2008: 159–165) moodustab ühend liikumis-sobivuskonstruktsiooni. EKG (1993: 252) järgi on supiinivorm sellises kontekstis kohamäärus. Konstruktsiooni semantika tingib siis nii kirjakeeles kui ka murretes suuresti liikumistähendusega finiidne verb.

Püsiühend ja afiksaaladverbikonstruktsioon

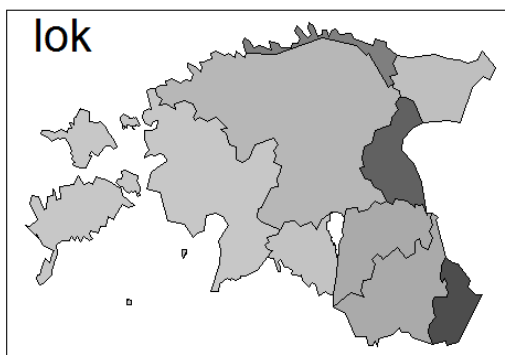
Sarnaselt *minema*-verbiga võib ka *tulema* väljendada koos supiiniga mingi protsessi või tegevuse algust. Sellise tähendusvarjundiga on siin püsiühendiks liigitatud '*lõpsma tulema*'. Ainukese sellise juhtumi tuvastasin keskmurdest (49). Püsiühendiks sellist paari nimetanud ka Tauli (1980: 133), samas kui EKG-s (1993: 258) on taoline ühend ainukese näitena algust väljendavate ühendite rühmas.

(49) ka nüüd teist kord **tuli lüpsma** (KES)

Nii nagu supiinivormis *minema* koos finiiitse *minema*-verbiga võib ka supiini-vormis *tulema* esineda finiiitse *tulema*-verbiga afiksaaladverbiühendina, kus supiinivormis *tulema*-verb on sünonüümne adverbi *ära* tähendusega (50). Sellesse rühma kuulub küll tunduvalt vähem ühendeid kui verbi *minema* puhul.

(50) sial **tulime tulema** (IDA)
ss **tullime** kellä ühe aigu kodu **tulõma** (TAR)

Kirjakeeles on selliseid ühendeid peetud samuti leksikaliseerunuteks (Erelt *et al.* 1993: 259; Penjam 2008: 195–196). Joonis 27 esitab finiiitse *tulema*-verbi ja supiiniga moodustatud lokatiivse konstruktsiooni geograafilise sagedusjaotuse. Kui üldiselt on sagedused kõigis murretes üsna sarnased, siis Seto, ida- ja rannamurdes esineb selle konstruktsiooni esindajaid mõnevõrra rohkem. Mingil põhjusel on konstruktsioon eriti iseloomulik just neile murretele.



Joonis 27. Finiiitse *tulema* ja *ma*-infinitiivi lokatiivse konstruktsiooni geograafiline sagedusesitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 33 esitab finiiitse *tulema*-verbiga tugevamini tõmbuvad *ma*-infinitiivis verbid ida- ja läänepoolsete murrete rühmas. Nagu *minema*-konstruktsioonide puhul on ka siin tugevamini tõmbuv supiinivormis verb idarühmas *kaema*, mis on jällegi seletatav lõunaeeesti murrete kuulumisega idarühma. Kuid huvitav on, et läänerrühmas on sama verb põhjaeestilisel kujul statistilise tugevuse poolest esikohal. Muude verbide puhul on pilt suhteliselt ühtlane, mis viitab taas verbi *tulema* suurele produktiivsuseastmele mõlemas rühmas. Selgelt ei joonistu välja rühmadevahelisi erinevusi supiinivormis verbi semantika põhjal.

Tabel 33. *tulema*-verbiga tugevamini tõmbuvad *ma*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
kaema	2.54	vaatama	1.58
ajama	1.26	küsima	1.2
tulema	1	võtma	1.05
mängima	0.93	katsuma	0.89
käima	0.62	rabama	0.59
talitama	0.62	teenima	0.59
ostma	0.52	vahtima	0.59
hirmutama	0.31	elama	0.55
kaaluma	0.31	otsima	0.45
kaevama	0.31	kauplema	0.3

9.2.4 *ma*-infinitiiv ja finiidne *panema*

Finiitse *panema* ja *ma*-infinitiivi ühend kuulub Rätsepa (1978: 197–199) järgi sekundaarsete verbivormidega kolmekohaliste elementaarlausete rühma (*Niisugused andmed panevad lugeja pead vangutama.*). Tauli (1980: 133) järgi kuuluvad sellised ühendid vormirühma V+*ma* ning *panema* koos *ma*-infinitiiviga võib esineda tähenduses „hakkama”, lisaks on Tauli mõned *panema*-ühendid eraldi välja toonud püsiühendite rühmas. Verb *panema* kuulub eesti keele tuumverbide hulka (Tragel 2003: 46–48) ning väljendab kirjakeeles koos supiiniga esinedes sageli kausatiivset algust (Pajusalu *et al.* 2004: 38). Finiitse *panema* ja *ma*-infinitiiviga moodustatud analüütilisi kausatiivseid konstruktsioone on põhjalikumalt analüüsinud Kasik (2001). Eesti keele akadeemilise grammatika (Erelt *et al.* 1993: 252;258–259) järgi võib *ma*-infinitiiv finiidse *panema*-verbiga kohamäärusena ja faasiverbina väljendada ahelverbi koosseisus algust.

Uurimismaterjalis tuvastasin 350 *ma*-infinitiivi ja finiidse *panema* ühendit sisaldavat osalauset. Tuvastatud konstruktsioonid ja nende esinemisjuhtude **normaliseeritud** sagedused on murrete kaupa esitatud tabelis 34 Kollostruktsioonilises analüüsis osutus ühend statistiliselt oluliseks kõigis murretes.

Tabel 34. *ma*-infinitiivi ja *panema*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: algus – algust väljendav konstruktsioon, kaus – kausatiivkonstruktsioon, püsi – püsiühend.

	algus	kaus	püsi
IDA	3	22	0
KESK	2	44	0
KIRDE	3	9	0
LÄÄNE	3	51	0
MULGI	2	14	0
RANNA	1	22	0
SAARTE	1	32	1
SETO	0	48	0
TARTU	3	21	0
VÕRU	1	27	0

Selliste ühendite abil võib murretes moodustada kolm erinevat konstruktsiooni: algus- ja kausatiivkonstruktsiooni ning püsiühendeid. Sageduselt suurima rühma moodustavad kausatiivkonstruktsioonide esindajad ja ülejäänud kahe rühma kvantitatiivne osakaal on analüüsis marginaalne. Järgnevalt kirjeldan konstruktsioone detailsemalt.

Kausatiivkonstruktsioon

Kasiku (2001: 87) järgi väljendavad analüütilised kausatiivid, mille hulka kuulub ka *panema* + *ma*-infinitiiv, kaudset ja mittekontaktset põhjustamist, samas kui leksikaalsed kausatiivid väljendavad kontaktset kauseerimist. Sarnaselt käituvad analüütilised kausatiivid ka murdematerjali põhjal. Selles konstruktsioonis on *ma*-infinitiiviga tähistatava tegevuse põhjustaja eksplitsiitselt väljendatud või kontekstist ühemõtteliselt tuletatav (51).

- (51) kadakkad kadakkadoppi **panen** patta **kiema** (KES)
jahh ega mei vägisi **tienima ei pane** (RAN)
nuid **pantti mineme** kyik Riija poolõ (VÕR)

Penjam (2008: 166–173) on sellised juhtumid jaotanud kahe kausatiivkonstruktsiooni vahel: esimesel juhul on *ma*-infinitiivis verb lauses adverbiaal ning teisel juhul moodustavad *panema* ja supiinivorm ahelverbi (2008: 190–194). Käesolevas töös sellist vahet pole tehtud, kuna nende kahe funktsiooni eristamine vaadeldavates konstruktsioonides on suhteliselt keeruline, samuti ei mõjuta need kaks asjaolu siinse töö autori arvates konstruktsiooni üldisemat tähendust. Pajusalu jt (2004: 38) järgi väljendavad sellised ühendid kausatiivset algust ning algus ja põhjustamine on siin tihedalt seotud. Seda väidet toetab ka murdematerjal, kuigi neid tähendusi pole siin püütud eristada.

Kirjakeeles on sarnase konstruktsiooni puhul täheldatud, et infiniitvorm väljendab enamasti inimese emotsionaalseid, füsioloogilisi ja mentaalseid tegevusi (Penjam 2008: 192–193), kuid murdekorpuse materjalist selliseid tendentse ei ilmnenu. Selliseid näiteid esines üksikjuhtumitena (52).

(52) et e nii külm saand et **pane** keha **värisema** (LÄÄ)

Tauli (1980: 133) on küll ühendi *põlema panema* liigitanud püsiühendiks, kuid selles töös käsitletlen neid juhtumeid kausatiivkonstruktsiooni esindajatena.

Juhkam (2009) on kirjeldanud eestirootsi refleksiivverbi mõju eesti murrete verbikasutusele ja toonud välja ühe huvitava konstruktsiooni, mis koosneb finiidest verbist, refleksiivpronoomenist ja supiivvormis verbist. Konstruktsiooni leviala piirab ta saarte ja Lääne-Eestiga ning mainib, et selle levik on ebahütlane (2009: 66). Siinse töö materjalist tuvastasin verbiühendite koosseisus vaid ühe sellise konstrukti Ridala kihelkonnas (53).

(53) ja **panid** jälle *ended* merele **minema** (LÄÄ)

Juhkam selle konstruktsiooni leviku kirjeldamisel Ridalat eraldi kihelkonnana välja ei too ja selle poolest on huvitav, kuidas murdekorpuse suhteliselt väike materjal võimaldab tuvastada väga eripäraste ja harva esinevate nähtuste seni täheldamata levikualasid.

Algust väljendav konstruktsioon ja püsiühendid

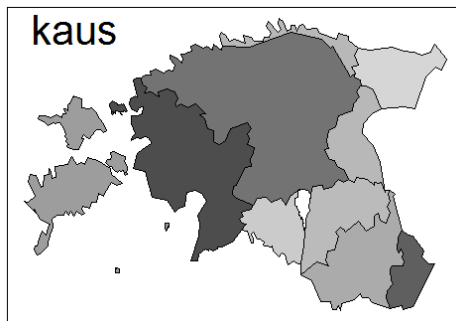
Algust väljendaval konstruktsioonil puudub selge kausatiivne tähendus ning supiivvormis verb väljendab liikumist (54). Verb *panema* on konstruktiiv asendatav verbiga *hakkama*, ilma et põhitähendus oluliselt muutuks. Konstruktsiooni moodustavate verbide subjekt on alati sama.

(54) jaahh targad sõnad **pane**vad **menema** (LÄÄ)
panin silades **tulema** (KES)
siss **pa**ind kõik looma üitstõsile otsa **minema** (TAR)

Püsiühendi esindajaid tuvastasin vaid kaks kujul „hakkama panema” saarte murdes, mis väljendavad tähendust „valmis seadma” (55). Sama ühend kuulub püsiühendite rühma ka Tauli analüüsis (1980: 133).

(55) no ja kui ta **pane**b nee asjad sedasi **akma** (SAA)
jaa ruut **pandi** seal **akkama** siis omingust (SAA)

Joonis 28 esitab finiidse *panema*-verbi ja supiivi kausatiivkonstruktsiooni geograafilise sagedusjaotuse. Huvitaval kombel tõusevad siin esile lääne-, kesk- ja Seto murre. Ülejäänud murretes on konstruktsiooni kasutussagedus üsna hütlane.



Joonis 28. Finiitse *panema* ja *ma*-infinitiivi kausatiivkonstruktsiooni geograafiline sagedus-esisitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 35 esitab finiiitse *panema*-verbiga tugevamini tõmbuvad supiiniivormis verbid ida- ja läänepoolsetes murretes. Mõlemas rühmas on pilt sarnane, mis viitab verbi *panema* suurele produktiivsusele kummaski rühmas. Idarühmas on küll verbil *põlema* kõrge statistilise seose tugevuse väärtus finiiitse *panema*-verbiga, mille põhjal võib eeldada nende kahe elemendi püsiühendilaadsust, nagu on seda märkinud Tauli (1980: 133). Selgeid semantilisi erinevusi supiiniivormis verbide põhjal kahe murderühma vahel ei näi olevat.

Tabel 35. *panema*-verbiga tugevamini tõmbuvad *ma*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
põlema	2.45	ligunema	1.94
istutama	1.15	keema	1.29
kastma	1.15	käima	0.8
niitma	1.15	käärima	0.68
magama	1.11	leekima	0.68
minema	0.71	kuivama	0.64
andma	0.57	kasvama	0.59
hammustama	0.57	hauduma	0.41
kirjutama	0.57	lõikama	0.41
naerma	0.57	tooma	0.41

Uurimismaterjali põhjal võib öelda, et *panema*-verb on omandanud selged abiverbi funktsioonid kõigis murretes ja osaleb koos supiinivormiga mitmete grammatiliste tähenduste, eelkõige kausatiivsuse väljendamisel. Ülejäänud konstruktsioonide esiletulek murdematerjalis on küll üsna marginaalne.

9.2.5 *ma*-infinitiiv ja finiidne *pidama*

Eesti keele akadeemilise grammatika (Erelt *et al.* 1993: 259) järgi võib finiidne *pidama* koos supiinivormiga moodustada ahelverbe, mis väljendavad modaal-sust ja evidentsiaalsust. *pidama* on levinud modaalverb kõigis läänemeresoome keeltes (Kehayov & Torn-Leesik 2009), verbi modaaltähendusi eesti keeles on kirjeldanud Uuspõld (1989). Erelt (2001) on põhjalikumalt vaadelnud nii modaalseid kui ka postmodaalseid *pidama* tähendusi ja funktsioone. Saukkonen (1965: 113–134) on oma monograafias põhjalikult kirjeldanud *pidama*-konstruktsioonide levikut ja struktuuri läänemeresoome keeltes, sh ka eesti keeles. Eesti keeles kuulub verb eesti keele tuumverbide hulka (Tragel 2003: 32–34). Tauli jagab sellised ühendid erinevate rühmade vahel. Harva esinevate liitvormide rühma kuulub finiidse *pidama* imperfekti vorm ja supiin (*Sa pidid virk olema.*) või impersonaalse supiini ühend (*Seal pidi iga päev lauldama.*), mis asendab kvotatiivi preesensit, kus *pidama*-verbi on abiverbi funktsioonis (Tauli 1980: 39). Samas kuulub finiidne *pidama* ja supiini ühend ka modaal-ühendite rühma (*Tal peaks raha olema.*) (Tauli 1980: 39–40). Impersonaalses tähenduses esinevad *pidama* ainsuse kolmanda isiku vormi ja impersonaalse supiini vormi ühendid on Tauli (1980: 59) esitanud eraldi rühmana ja toonud välja konstruktsiooni erinevad sõnajärgemallid. Konstruktsioon võib väljendada koos *pidama* konditsionaalivormiga *ma*-infinitiiviga väljendatava tegevuse vähetõenäosust (Tauli 1980: 233). Lisaks võib sama *pidama* vorm moodustada küsilauseste erinähtusi (*Mis ma peaksin peale hakkama?*) nii supiini kui ka supiini impersonaalse vormiga ning sellisel juhul on ühendil tundevarjund (Tauli 1980: 302). Penjami (2008: 174–183) järgi jagunevad sellised ühendid modaal- ja kvotatiivikonstruktsioonide esindajateks.

Eesti murretes on põhjalikult kirjeldanud erinevaid kvotatiivikonstruktsioone Kask (1984). Kvotatiivikonstruktsioon tõuseb esile ka siinse töö uurimis-materjalis. Näiteks toob Kask (1984: 262–264) välja, et Järvamaal, Harjumaal ja Virumaal on selliste analüütiliste kvotatiivikonstruktsioonide (*ta peab olema maal*) kasutus üsna levinud. *ma*-infinitiiv koos *pidama* erinevate vormidega võib moodustada kvotatiivikonstruktsioone (*Tõnu piab aige olema*) põhiliselt põhjaeesti murderühmas ja on eriti levinud näiteks saartel (Kask 1984: 268–269). Kirderanniku murderühmas on kaudse kõneviisi levik üldiselt ebaühtlane, kuid teiste kvotatiivi väljendusvõimaluste kõrval on ka seal erinevaid *peab*-konstruktsioone (*nämäd pidid üästi laulama*) (Kask 1984: 269–270). Lõunaeesti murretes on sellised konstruktsioonid üsna harv nähtus (1984: 270), seda kinnitab ka käesoleva töö materjal (vt põhjalikumalt allpool). Konstruktsiooni esinemist idamurdes on täheldanud Univere (1988: 89).

Kokku esines uurimismaterjalis 1408 finiiitse *pidama* ja supiinivormiga osalauset. Tabel 36 esitab kõik finiiitse *pidama* ja supiiniga moodustuvad konstruktsioonid ja nende esinemisjuhtude **normaliseeritud** sagedused. Kollostruktsioonilise analüüsi tulemusena osutus ühend statistiliselt oluliseks kõigis murretes.

Tabel 36. *ma*-infinitiivi ja *pidama*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: avert – avertiivkonstruktsioon, kvot – kvotatiivikonstruktsioon, kvot_ips – impersonaalne kvotatiivikonstruktsioon, mod – modaalkonstruktsioon, proks – proksimatiivkonstruktsioon.

	avert	kvot	kvot_ips	mod	proks
IDA	1	11	0	73	1
KESK	1	16	2	149	0
KIRDE	0	6	2	181	0
LÄÄNE	0	13	3	144	0
MULGI	0	16	0	167	0
RANNA	3	4	1	90	0
SAARTE	2	10	1	166	0
SETO	0	0	0	62	0
TARTU	0	1	0	57	0
VÕRU	0	1	0	74	0

Sellise ühendiga moodustavad konstruktsioonid on murdematerjali põhjal jagatud kuueks. Kõige tavalisem kõigis murretes on modaalkonstruktsioon. Kvotatiivi ja impersonaalse kvotatiivi näol on tegemist küll üsna sarnase tähendusega konstruktsioonidega, kuid supiini eri vormi tõttu on impersonaalsed juhtumid tavalise kvotatiivikonstruktsiooni esindajatest siiski eraldatud.

Modaalkonstruktsioon

Finiitne *pidama* koos supiiniga väljendab murretes enamasti modaalsust (56), nii nagu see on levinud kirjakeeleski. Konstruktsioon on üsna tavaline kõigis murretes ja sagedamini väljendatakse modaalsust sel viisil kesk-, kirde-, lääne-, saarte ja Mulgi murdes. Selle konstruktsiooni põhjal ei jookse levikupiir traditsioonilise põhja- ja lõunaeesti murrete vahel, nagu võib täpsemalt näha allpool esitatud geograafilise leviku kaartil.

- (56) si poeg õõ ega sii **pidi** kaa kaunis noorest peast **abielluma**
(MUL)
siis pruut **piti** neile ikä ühele siis niiskese tasku rätte jälle siis
kinkkima vai **andamaje** (KIR)
aga noo juu sedä **pidi** koa **kartma** (LÄÄ)

Sarnaselt paljude teiste modaalkonstruksioonidega välditakse ka *pidama*-konstruksioonis isikule viitamist (vt ka ptk 8.4) ja isik on eksplitsiitselt väljendatud harva.

Kvotatiivi- ja impersonaalse kvotatiivikonstruksioon

Kvotatiivikonstruksioonid jagunevad kaheks. Tavaline kvotatiivikonstruksioon väljendab lihtsalt evidentsiaalsust (57). Praktilises analüüsis oli enamasti keeruline eristada modaalselt tähendust kvotatiivsest ja konstruksioonis modaalne ja evidentsiaalne tähendus on murretes selle konstruksiooni puhul tugevalt seotud.

- (57) seal **pidi** vesilillisi **olema** (LÄÄ)
jaa ikki ta ikki ta **piass** võibolla kiili ka **mõistma** küllalt (MUL)
siis **pidi** kehva vilja aasta **tulema** (SAA)

Impersonaalses kvotatiivikonstruksioonis on *ma*-infinitiivis verb impersonaali tunnusega (58). Konstrukte esines materjalis suhteliselt vähe, peamiselt põhjaeesti murretes.

- (58) mõttes nehh **pidi nõijuttama** et kala (RAN)
siis midagit **pidi** ikka sönnä pääle **pandama** (SAA)
seal **peab** püssiga maha **lastama** (LÄÄ)

Avertiiv- ja proksimatiivkonstruksioonid

Postmodaalse tähendusega avertiiv- ja proksimatiivkonstruksioonide esindajaid oli materjalis üksikuid, kuid siiski selgelt olemas ka nii väikse materjali hulgas, kui seda on murdekorpuse oma. Avertiivkonstruksiooni esindajaid leidis põhiliselt põhjaeesti murderühmas (59). Huvitav on seejuures, et kirjakeele materjalist sarnases uurimistöös ei tuvastatud ühtegi avertiivkonstruksiooni (Penjam 2008: 178). *pidama*-verbi minevikuvormi ja *ma*-infinitiivi on peetud ainukeseks avertiivi väljendusvõimaluseks eesti kirjakeeles (Erelt 2001; Erelt & Metslang 2009). Murretes seda kategooriat uuritud ei ole.

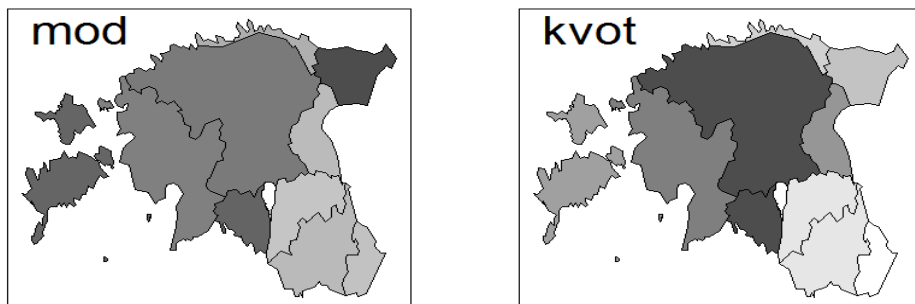
- (59) ku mina ema **pidi minestama** ära (IDA)
õhk **pidi** suust sisse **minema** (KES)
ett e et **pidi** inge kinni **panema** (RAN)

Proksimatiivi väljendamisvõimalusi erinevalt avertiivist on eesti keeles tunduvalt rohkem (Erelt & Metslang 2009). Proksimatiivkonstrukte *pidama* verbiga esines murdekorpuse materjalis ainult üks ja seda idamurdes (60).

- (60) ja mina **pidin kukkuma** (ja kukkusingi) (IDA)

Joonis 29 esitab finiiitse *pidama*-verbi ja *ma*-infinitiivi sagedasemate, modaalse ja kvotatiivikonstruksiooni geograafilise sagedusjaotuse. Mõlema konstruksiooni geograafiline jagunemine on huvitav. Modaalse konstruksiooni sagedus on kõrgeim kirdemurdes ning üsna sageli ilmneb seda ka kesk-, lääne-, saarte ja Mulgi murdes, samas kui ülejäänud murretes on see ühtlaselt madalama

sagedusega. Kvotatiivikonstruktsioon on pea olematu Seto, Võru ja Tartu murdes, suhteliselt harva esineb seda kirderanniku murretes, samas on see väga sage Mulgi ja keskmurdes. Mõlema konstruktsiooni põhjal rühmituvad murded traditsioonilisest murdejaotusest erinevalt.



Joonis 29. Finiitse *pidama* ja *ma*-infinitiivi konstruktsioonide geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 37 esitab finiiitse *pidama* verbiga statistiliselt tugevamini tõmbuvad verbid *ma*-infinitiivis ida- ja läänepoolsete murrete rühmades. Taas on rohkem tugevamini tõmbuvaid verbe idarühmas, mis osutab verbi *pidama* väiksemale produktiivsusele selles rühmas. Selgelt ei joonistu välja rühmadevahelisi verbi-semantilisi erinevusi.

Tabel 37. *pidama*-verbiga tugevamini tõmbuvad *ma*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
kandma	3.17	laskma	2.31
viima	2.61	kasvama	1.08
magama	1.53	kuduma	0.77
minema	1.39	näitama	0.77
austama	1.09	rääkima	0.68
alluma	1.05	sööma	0.68
eitama	1.05	ajama	0.61
hüppama	1.05	passima	0.61
kõnelema	1.05	viskama	0.61
murdma	1.05	tulema	0.6

9.2.6 ma-infinitiiv ja finiidne saama

Finiitset verbist *saama* koos *da*-infinitiiviga esinedes on juba põhjalikumalt juttu olnud peatükis 9.1.1 ja siin keskendun finiidse *saama* esinemisele koos supiinivormiga. Sellise ühendiga väljendatavat tulevikutähendust on kirjakeeles eelnevalt käsitletud suhteliselt põhjalikult (vt näiteks Metslang 1994a; Metslang 1994b; Mägiste 1936) ja konstruktsiooni teisi tähendusi mõnevõrra vähem (näiteks Tragel & Habicht 2012). Rätsepa (1978: 153) lihtlausetüüpide liigituses esinevad sellised ühendid kahekohaliste elementaarlausete rühmas (*Mehed said ühte majja elama.*) ja sekundaarsete verbivormidega kahekohaliste elementaarlausete rühmas (*Mehed said sakslaste käest minema.*) (1978: 192). Tauli (1980: 38–39) on ühendi välja toonud harvaesinevate verbivormide rühmas ühendina, mis väljendab tulevikku (*See ei saa kunagi nii olema.*) ja siirdumistähendusega konstruktsioonide ja püsiühendite rühmas (*Laupäeva viimase tunni saavad nad hakkama temata.*) (1980: 133).

Finiitse *saama* ja *ma*-infinitiivi ühendeid esines uurimismaterjalis suhteliselt vähe – kõigis murretes kokku vaid 39 ühendit sisaldavat osalauset. Kollostruktsioonilise analüüsi käigus ei osutunud ühend konstruktsioonina statistiliselt oluliseks üheski murdes. Olgugi et selliseid esinemisjuhtusid leidis materjalis vähe, moodustab ühend üsna erinevate tähendustega konstruktsioone. Konstruktsioonide esinemisjuhtude **normaliseeritud** sagedused kõigis murretes esitab tabel 38. Ühtegi konstruktsiooni esinemisjuhtu ei tuvastatud Mulgi murdes.

Tabel 38. *ma*-infinitiivi ja *saama*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: fut – tulevikukonstruktsioon, kaus – kausatiivikonstruktsioon, püsi – püsiühend, fut_ipf – imperfekti tuleviku konstruktsioon, õnnest – õnnestumistähendusega konstruktsioon.

	fut	kaus	püsi	fut_ipf	õnnest
IDA	3	0	0	0	0
KESK	1	1	1	1	1
KIRDE	2	0	3	2	2
LÄÄNE	2	0	2	1	1
MULGI	0	0	0	0	0
RANNA	1	1	3	1	0
SAARTE	0	0	0	0	1
SETO	0	0	0	0	3
TARTU	0	0	0	3	3
VÕRU	0	0	0	0	2

Tulevikku väljendavad konstruktsioonid

Finiitse *saama* ja *ma*-infinitiivi abil väljendatavat tulevikukonstruktsiooni on kirjakeeles kirjeldatud kõige põhjalikumalt. Mägiste (1936) on vaadelnud *saama*-futuuri kasutust ja levikut nii eesti kui ka sugulaskeeltes ja märkinud, et konstruktsioon võib olla eesti keele oma algupära ning saksa keele mõju on selle kasutusala vaid laiendanud. Lisaks osutab ta (1936: 24) *saama*-verbi tulemist väljendavale tähendusele, mis on levinud teisteski läänemeresoome ja kaugemates sugulaskeeltes. Eesti keele akadeemilise grammatika (Erelt *et al.* 1993: 32–33) järgi on see vormipaar samuti üks tuleviku väljendamise võimalusi, ent põhjalikumalt konstruktsiooni kasutust ja esinemist EKG ei kirjelda. Metslang (1994b) on analüüsinud tuleviku väljendamise vahendeid eesti kirjakeeles ja leiab, et konstruktsioon on siiski saksapärase.

Käesolevas töös on eristatud kahte liiki tulevikukonstruktsioone. Esimene on tüüpiline *ma*-infinitiivi ja *saama* abil moodustuv konstruktsioon, mida kõik eelnevad allikad on tulevikukonstruktsioonina kirjeldanud (61). Konstruktsioon viitab tulevikus asetleidvale sündmusele ning finitiivne *saama* on olevikuvormis. Grammatiline subjekt võib olla nii elus kui elutu.

- (61) kui tema kuulsad nime saa **saad** iest ja tagant **üidma** (KES)
tulevankred **saavad sõitma** (LÄÄ)
et vihtlemine **saab akkama** ja (IDA)

Teine tulevikukonstruktsioon moodustub imperfektivormis *saama*-verbiga ning viitab selgelt mingile sündmuse tulevikulisusele mineviku sündmuse kontekstis (62). Ka selle konstruktsiooni grammatiline subjekt võib olla nii elus kui ka elutu.

- (62) ka siss vist **sai** kerikku koht **olema** nii säälbbool külles sääl (KES)
se **sai** üks lollukkane **olema** (RAN)
ja keskpääväss inemse jälle põllutühü **sai** muut **tegemä** jälle (TAR)

Selle konstruktsiooni käsitlemine on mõnevõrra problemaatiline. Keem (1973: 609) on selliseid *saama* + *olema* verbi ühendeid pidanud uuemaks potentsiaali väljendamise võimaluseks, nende näidete puhul on Keema väiteid raske kinnitada, kuigi esimeses (62), keskmurde näites on selline potentsiaali tõlgendus võimalik.

Õnnestumiskonstruktsioon

Õnnestumiskonstruktsioon väljendab *ma*-infinitiiviga tähistatava tegevuse õnnestumist mingi takistuse ületamisega (63). Konstruktsiooni õnnestumistähendust kirjakeeles on detailsemalt analüüsinud Tragel ja Habicht (2012). See konstruktsioon on murdematerjalis tugevalt lõimunud tulevikutähendusega ning tähenduste eristamine üksikutele juhtumitele võib olla keeruline ning pole üheselt mõistetav.

- (63) et **saa** ma ruttõmbade **minemä** (VÕR)
ja sedasi nad **said** siis sõnna **elama** (SAA)
sai teid **kaema** ja (SET)

Erinevalt tulevikukonstruksioonidest saab õnnestumiskonstruksiooni grammatailine subjekt olla vaid elus.

Kausatiivkonstruksioon ja püsiühendid

Kausatiivkonstruksioon oli materjalis esindatud vaid kahe näitelausega (64). Konstruksiooni iseloomustab finiiitse ja infiniitse verbi erisubjektilisus, s.t. verbiga *saama* ja *ma*-infinitiivis verbiga tähistatud tegevusel on erinevad sooritajad. Ka selle konstruksiooni puhul on tugevalt olemas õnnestumistähendus nagu eelnevalt kirjeldatud õnnestumiskonstruksioonil. Metslang (1994b) on viidanud konstruksiooni kausatiivsele tähendusele kirjakeeles.

- (64) sa **i saa** teda enam **käima** (RAN)
ei **saand** enam seda **pidama** (KES)

Püsiühendite rühma moodustavad konstruksioonid, kus supiinivormis esineb verb *hakkama* (65).

- (65) rautseppa tööga **saan** **akkama** (SAA)
küll ma neega **saan** isegi **akkama** (KIR)
aga nüüd üks mees **sai** sellega **akkama** (LÄÄ)

Penjam (2008: 195) järgi kuulub selline ühend kirjakeeles leksikaliseerunud ühendite rühma ning Tauli (1980: 133) on esitanud verbipaari püsiühendina. Sarnaselt kirjakeelega väljendab selline püsiühend murretes toimetulemist, millegagi hakkama saamist.

Kirjeldatud finiiitse *saama* ja supiini ühendite jaoks pole tehtud eraldi ei geograafilise sagedusjaotuse kaarte ega ka murderühmade statistilise tugevuse ja produktiivsuse analüüsi, kuna ühendeid esines materjalis selleks liiga vähe.

9.2.7 *ma*-infinitiiv ja finiiitne *olema*

Eesti keele akadeemilise grammatika järgi kuuluvad *olema* ja *ma*-infinitiivi verbiühendid supiinitarindite hulka, kus supiinivorm esineb koha- või sõltuvusmäärusena. Supiinivorm kohamäärusena võib finiiitse *olema* verbiga väljendada sobivust koos mõne sellist tähendust väljendava adjektiivi või adverbiga (*on sobiv/kohane/paslik*). (Erelt *et al.* 1993: 254) Sõltuvusmäärusena, lisaks mainituile, võivad supiinivormis verbi laiendada võimelisust, valmistumist ja valmisolekut, nõusolekut, kohustust jmt modaalpredikaate (Erelt *et al.* 1993: 254). Tauli (1980: 211) on eraldi välja toonud vormirühma *Volema* + (AdjN/Adv+MA), kus supiinivorm väljendab tegevust, milles adjektiivnoomeni või adverbiga väljendatu ilmneb (*Ta on esimene seda tegema.*). Penjami (2008: 160–161) liigituses moodustab selline vorm liikumis-sobivuskonstruksioone, kus semantiliselt predikaadiks võib olla ka *olema*-verbi ja adjektiivi või adverbü ühend ning kirjakeele materjalis olid sellisel juhul sagedasemaks konstruksioonis esinevateks *olema*-ühenditeks *valmis olema* ja *nõus olema*.

Tabel 39 esitab finiiitse *olema* ja supiinivormiga moodustatavate konstruktsioonide esinemisjuhtude **normaliseeritud** sagedused. Selliseid ühendeid esines murdekorpuses suhteliselt vähe, kõigis murretes kokku vaid 33 osalauset. Ühtegi sellist ühendit sisaldavat osalauset ei tuvastatud Seto murdest. Kollostruktsiooniline analüüs ühtegi sellist ühendit statistiliselt olulise konstruktsioonina üheski murdes ei tuvastanud, kuid konstruktsioon oli piisava sagedusega normaliseeritud sageduste analüüsis.

Tabel 39. *ma*-infinitiivi ja *olema*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: mod – modaalkonstruktsioon, omadus – omadust väljendav konstruktsioon, proks – proksimatiivkonstruktsioon.

	mod	omadus	proks
IDA	0	0	1
KESK	1	2	1
KIRDE	0	8	0
LÄÄNE	0	4	0
MULGI	0	7	0
RANNA	0	3	0
SAARTE	0	2	0
SETO	0	0	0
TARTU	0	1	0
VÕRU	0	1	0

Omadust väljendav konstruktsioon

Suurem osa supiini ja *olema*-verbi ühendeid moodustab omadust väljendavaid konstruktsioone (66). Sarnaselt on selliseid tähendusi analüüsitud ka kirjakeeles (Pajusalu & Orav 2008: 110). Omadust väljendavale konstruktsioonile murde-materjalis on iseloomulik kolmanda kohustusliku elemendi olemasolu, milleks võib olla mõni adverb, adjektiiv või substantiiv, mis iseloomustab *ma*-infinitiiviga väljendatavat tegevust.

- (66) aga sie **õli kange magama** (KIR)
 see **oli kange noomima** oma rahvast (LÄÄ)
 ja mia kohe **ole loll minema** kaa ja (MUL)

Huvitav on selle konstruktsiooni juures see, et enamasti on *ma*-infinitiivi laiendavaks sõnaks kas siis noomeni või adjektiivina sõna *kange*. Selle põhjal võiks eeldada kogu konstruktsiooni suhteliselt kindlat kasutuskonteksti ja püsiühendilaadust. Teistest murretest sagedamini oli konstruktsiooni esinemisjuhtusid kirde-, lääne- ja Mulgi murdes.

Penjami (2008: 165) järgi on selles konstruktsioonis subjektina vormistatud tegija enamjaolt inimene. Kuigi käesolevas materjalis on näiteid vähe, ei saa murrete põhjal sama kinnitada. Juba nii väikse materjali peal tuleb välja piisavalt näiteid, kus subjekt võib olla elutu ja mitteagentiivne (67), seega murretes selline piirang ei kehti.

- (67) see teine sort **oli** täis e ea piha **akma** (SAA)
miss **ol** nii kõvemb vasta **pidämä** (VÕR)
eks se esimine küinar **old** ikke visa **menema** juo (KES)

Modaal- ja proksimatiivkonstruktsioonid

Modaal- ja proksimatiivkonstruktsiooni esinemisjuhtusid oli materjalis vähe. Modaalstruktsiooni esinemisjuhte tuvastasin vaid kolm saarte ja keskmurdes (68). Üksikud näited on siiski huvitavad, sest selline modaalsuse väljendusviis tundub kirjakeeles olevat võimatu. Kuna näitelauseid on vähe, siis on raske konstruktsiooni põhjalikumalt kirjeldada, kuid esitatud näidetes esineb finitiivne *olema* selgelt modaalses tähenduses, mis on sünonüümne *pidama*-verbiga.

- (68) endal enne **oli** kodu ära **kaaluma** puntti **panema** (KES)

Eraldi olen analüüsinud proksimatiivkonstruktsiooni, mida tuli materjalis ette vaid kahel korral, ida- ja keskmurdes (69). Selles konstruktsioonis esineb *ma*-infinitiiv funktsioonis, kus tavaliselt kirjakeeles ja sageli ka murretes esineb *ma*-infinitiivi inessiivivorm.

- (69) siis **olivad vaidlema** nagu (IDA)
juba enne ku se puhastus **on tulema** (KES)

Kuna *olema* ja *supiini* konstruktsioonide esinemisjuhtusid oli materjalis vähe pole nende kohta tehtud geograafilist sageduskaarti ega ka produktiivsuse analüüsi.

9.2.8 *ma*-infinitiiv ja finitiivne *jääma*

Verb *jääma* on tuntud kõigis läänemeresoome keeltes ja verbil on esialgselt olnud liikumist väljendav tähendus (Saukkonen 1965: 65), kuid eesti ja liivi keele põhjal on esitatud väiteid, et verbi esialgsed tähendused võivad olla *siirduma*, *muutuma* ja *tulema* (Genetz 1900). Eesti keele akadeemilise grammatika (Erelt *et al.* 1993: 258) järgi kuulub *jääma* koos *supiiniga* esinedes tegevuse algusfaasi väljendavate verbide hulka (*Laps jäi magama.*), kuid ühend võib väljendada ka mingi tegevuse või protsessi jätkuvust. Viimast tähendust peab EKG siiski alguse erijuhuks.

Eesti kirjakeeles on põhjalikult kirjeldatud verbi *jääma* muutumistähendust (Pajusalu 1994; Pajusalu & Orav 2008). *jääma*-verb kuulub Trageli (2003: 39–40) järgi eesti keele tuumverbide hulka ning ta on oletanud, et just 3. isiku lihtminevikus esinedes väljendab ühend duratiivset ja inhoatiivset aspekti.

Supiiniga koos esinedes väljendab *jääma* kirjakeeles passiivset algust, s.t agendil puudub kontroll tegevuse üle, mille algamist supiiniga väljendatakse (Pajusalu *et al.* 2004: 39). Rätsep (1978: 153) on finiiitse *jääma* ja supiini ühendid liigitanud kahekohaliste elementaarlausete rühma (*Mehed jäid ühte tuppä elama.*) ja sekundaarsete verbivormidega kahekohaliste elementaarlausete hulka kujul N+nom. V¹ V²+ma (*Ants jäi mõtlema.*) (1978: 191–193).

Käesoleva töö uurimismaterjalist tuvastasin 106 osalauset, kus esines finiiitne *jääma*-verb ja supiinivorm. Analüüsis ei ole ma püüdnud eristada duratiivset ja inhoatiivset tähendust ning kõik tuvastatud näited esindavad algust väljendavat konstruktsiooni. Tabel 40 esitab konstruktide **normaliseeritud** esinemissagedused murretes.

Tabel 40. *ma*-infinitiivi ja *jääma*-verbi algust väljendava konstruktsiooni (algus) esinemisjuhtude normaliseeritud sagedused.

	algus
IDA	11
KESK	12
KIRDE	14
LÄÄNE	9
MULGI	7
RANNA	16
SAARTE	9
SETO	6
TARTU	16
VÕRU	8

Kollostruktsioonilise analüüsi käigus ei osutunud ühend statistiliselt oluliseks ida-, Mulgi, Seto ja Võru murredes.

Algust väljendav konstruktsioon

EKG järgi kuulubki finiiitne *jääma* koos supiiniga esinedes algust väljendavate faasiverbide hulka (Erelt *et al.* 1993: 258). Veel on *jääma* verbi käsitletud kui passiivset algust väljendavat verbi, mis tähendab, et agendil puudub supiiniga väljendatava tegevuse üle kontroll (Pajusalu *et al.* 2004: 39). Algust väljendavate konstruktide juhtumeid murretes on esitatud näites (70).

- (70) siss **jäid** kohta **pidama** (SAA)
 mõni **jäeb kuulama** (KES)
 veski **jäi seisma** (IDA)

Tauli (1980: 133) on nimetanud ühendeid *seisma jääma* ja *magama jääma* püsiühenditeks. Siinses töös sellist liigitust järgitud pole, kuid tuleb tõdeda, et algust väljendava konstruktsiooni esindajate hulgas on *seisma* ja *magama* ning viimasega sünonüümne *tukkuma* sagedasti esinevad verbid. See võiks viidata vaadeldava ühendi püsiühendilaadsusele. Siinses töös olen need siiski liigitanud algust väljendavate konstruktsioonide esinemisjuhtudeks, kuna ühendil on selgelt olemas ka inhoatiivne tähendus. Teema väärib kindlasti põhjalikumat uurimist tulevikus.

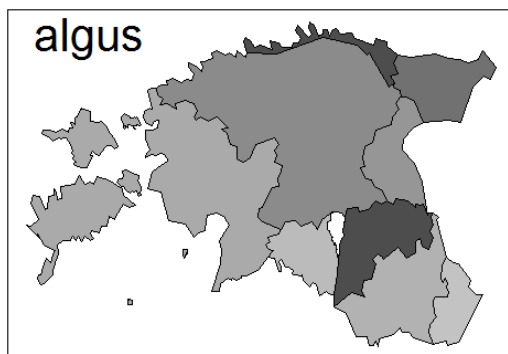
Kirjakeeles on *jääma*-verbi duratiivset aspekti vaadelnud Tragel (2003: 40). Murreteski esineb juhtumeid, kus konstruktil on olemas seisundi või protsessi jätkumise tähendus (71). Alguse ja kestvuse tähenduse eristamine on neil juhtumitel siiski üsna keeruline, seega on vahetegemisest siin loobutud.

- (71) ja nuq ku **jääss** meil tynõ **kaoma** (SET)
 iks nii et vihma kätte **ess jääq mädänemä** (VÕR)
 tä nii kui **jäi põlgama** seda oma leivavanemaid (LÄÄ)

Püsiühendilaadseks võiks selles rühmas pidada paar korda esinenud ühendit *kaduma jääma* (72), kuna ühendil tundub olevat kinnistunum tähendus.

- (72) ja kõikkõga **jäi kaduma** (SAA)
 ja nuq ku **jääss** meil tynõ **kaoma** (SET)

Joonis 30 supiini ja *jääma*-verbi algust väljendava konstruktsiooni geograafilise sagedusjaotuse. Selgelt eristuvaid suuremaid rühmi siin ei ilmne, küll aga näib konstruktsiooni kasutus olevat sagedasem Tartu ja rannamurdes. Ülejäänud murretes sageduserinevused märkimisväärselt üksteisest ei erine.



Joonis 30. Finiitse *jääma* ja *ma*-infinitiivi alguskonstruktsiooni geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 41 esitab finiiitse *jääma* verbiga tugevamini tõmbuvad supiivormis verbid ida- ja läänepoolsete murrete rühmas. Idarühmas on selgelt tugevamini tõmbuv *magama*, mis läänerrühmas isegi ei esine kümne kõige tugevamini tõmbuva verbi hulgas. See lubab eeldada, et *magama* ja *jääma* ühend on tõepoolest püsiühendilaadsem ning seda just idapoolsete murrete rühmas. Läänepoolsete murrete rühmas on seevastu tugevamini tõmbuv *seisma*-verb, mis Tauli järgi (1980: 133) moodustab koos finiiitse *jääma*-verbiga samuti püsiühendi. Muus osas on verbid mõlemas rühmas üsna nõrgalt statistiliselt finiiitse *jääma*-verbiga seotud ning selle põhjal võib oletada, et *jääma* on ühtlaselt produktiivne mõlemas rühmas.

Tabel 41. *jääma*-verbiga tugevamini tõmbuvad *ma*-infinitiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
magama	2.78	seisma	2.03
põdema	0.83	elama	1.14
tolknema	0.83	pidama	0.61
jooma	0.41	õitsema	0.21
kannatama	0.41	hapnema	0.21
kasvama	0.41	kandma	0.21
kukkuma	0.41	keetma	0.21
lugema	0.41	korraldama	0.21
mādanema	0.41	kuduma	0.21
pūsimā	0.41	kuulama	0.21

9.2.9 *ma*-infinitiivi inessiiv ja finiiitne *kāima*

Eesti keele akadeemilise grammatika järgi võib *mas*-tarind esineda kohamāārusena, laiendades teiste hulgas ka intransitiivverbi *kāima* ning tarindil on lisaks lokatiivsele ka finaalne tähendus (*Ta kāis metsas marju korjamas.*). Samuti väljendab konstruktsioon siirdumist, mis *kāima* verbiga koos esinedes on eksplitsiitselt väljendatud. (Erelt *et al.* 1993: 255) Tauli (1980: 134) esitab vaid vormipaari V+*mas*, kus finiiitse verbina esinevad *istuma*, *kāima*, *olema*, *seisma*, analüüsimate ühendite tähendusi. Rātsepa (1978: 193) eesti keele lihtlausetüüpide liigituses kuulub ühend -*mas+kāima* sekundaarsete verbi-vormidega kahekohaliste elementaarlausete rühma (*Peeter kāis järve āāres kala pūūdmas.*).

Finiitse *käima* ja *mas*-vormi konstruktsiooni tähendusi pole ka eesti kirjakeeles kuigi palju kirjeldatud, on tehtud vaid üksikuid tähelepanekuid. Tragel (2003: 48) toob eesti keele tuumverbide käsitluses välja *ma*-infinitiivi inessiivi ja finiiitse *käima* ühendi lokaliseerituma ja staatilisema tähenduse, võrreldes supiini illatiiviga, samas osutab ta ka ühendi habituaalse aspekti tähenduse ilmnemisele. Niisiis on kirjakeeles ühendiga väljendatav tegevus korduv ja tavaline, mille tingib suuresti verbi *käima* semantika.

Siinse töö uurimismaterjalist tuvastasin 298 *mas*-vormi ja finiiitset *käima*-verbi sisaldavat osalauset. Kollostruktsioonilise analüüsi tulemusena ei osutunud ühend statistiliselt oluliseks vaid Seto murdes. Tabel 42 esitab finiiitse *käima* ja supiinivormi inessiiviga moodustatava lokatiivse konstruktsiooni esinemisjuhtude **normaliseeritud** sagedused kõigis murretes.

Tabel 42. *ma*-infinitiivi inessiivi ja *käima*-verbi lokatiivse konstruktsiooni (lok) esinemisjuhtude normaliseeritud sagedused.

	lok
IDA	36
KESK	25
KIRDE	29
LÄÄNE	40
MULGI	18
RANNA	32
SAARTE	48
SETO	6
TARTU	4
VÕRU	8

Kuna praktilises analüüsis on sageli keeruline eristada konstruktsiooni lokatiivset ja habituaalset tähendust, pole seda selles töös tehtud ning kõik vaadeldava verbiühendi esinemisjuhtumid on liigitatud lokatiivse tähendusega konstruktsiooni esindajateks. Konstruktsiooni habituaalne tähendusvarjund on siiski sageli tajutav.

Lokatiivne konstruktsioon

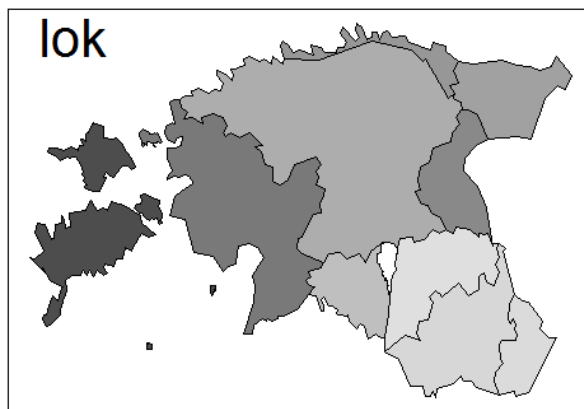
Lokatiivse konstruktsiooni puhul on supiinivormis inessiiv tõlgendatav kohamäärusena (73).

- (73) poig **käve** herneht **vahtmahh** (SET)
käisimi Pikkäsillal ja Liival ja **jooman** (TAR)
ja **käisitte** e seda e mõisa kätte **saamas** (RAN)

Sellest hoolimata võib lokatiivse konstruktsiooni esindaja edasi anda habituaalset ja korduvat tegevust ning lokatiivsus ja habituaalsus pole teineteist välistavad, sellise tähendusvarjundi tingib juba verbi *käima* semantika ise. Habituaalne tähendus ilmneb konstruktsioonis selgemini, kui tegevuse korduvust mingil muul moel eksplitseeritakse (74), nagu esitatud näidetes *iga päev*, *kaks päeva nädalas*, *mõnikord mitmes kohas*.

- (74) ikke iga kevade **käis** koolis **katsumas** (IDA)
 ja isa **käis** siis õõ õõn õõ tiubäivi **tegemas** kaks päeva nädalas ja
 (KES)
 aga vanad eided ikke **käisiväd** mõnigõrd mitmes kõhas puteliga
 õtsimas pojale naist (KIR)

Joonis 31 esitab finiidse *käima* ja *ma*-infinitiivi inessiivi lokatiivse konstruktsiooni geograafilise sagedusjaotuse. Sagedus on kõrgeim saarte murdes, järgnevad lääne- ja idamurre. Sageduse poolest ebatüüpilisim näib olevat konstruktsioon enamikele lõunaesti murretele.



Joonis 31. Finiitse *käima* ja *ma*-infinitiivi inessiivi konstruktsiooni geograafiline sagedusesitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 43 esitab finiidse verbiga *käima* tugevamini tõmbuvad *ma*-infinitiivi inessiivis verbid ida- ja läänepoolsete murrete rühmas. Statistiliselt tugevamini tõmbuvaid verbe on tunduvalt rohkem idapoolsete murrete rühmas ning läänepoolsete murrete rühmas on väga tugevalt tõmbunud vaid verb *püüdma*. See võib olla seletatav murdealade geograafilise paiknemisega, kus merendusteemaliste tekstide osakaal on suurem. Läänepoolsetes murretes, kus verbiühend on selgelt produktiivsem, on infiniitses positsioonis esinevate verbide hulk suurem ja tugevamaid eelistusi ühe või teise verbi suhtes (v.a *püüdma*) ei esine. Selgeid verbisemantiliselt erinevusi kahe rühma vahel ei näi olevat.

Tabel 43. *käima*-verbiga tugevamini tõmbuvad *ma*-infinitiivi inessiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
otsima	2.15	püüdma	2.51
jooma	1.97	ostma	0.78
kaema	1.97	tantsima	0.78
mõskma	1.97	käima	0.66
hoidma	1.31	lõugutama	0.66
lugema	1.31	lastima	0.55
vaatama	0.95	tegema	0.52
jahvatama	0.9	korjama	0.44
laulatama	0.66	tooma	0.38
imema	0.65	niitma	0.36

9.2.10 *ma*-infinitiivi inessiiv ja finiidne *olema*

Rätsepa (1978: 193) järgi jagunevad finiidse *olema* ning *ma*-infinitiivi inessiivi ühendid sekundaarsete verbivormidega kahekohaliste elementaarlausete hulgas kolme eri tüübi vahel, millele kõigile on iseloomulik subjekti olemasolu (*Kontsert oli lõppemas.*). Tauli (1980: 134) järgi kuulub see ühend (*Ta oli jalutamamas.*) samasse rühma koos *mas*-vormi ja finiidse *käima* ühenditega. Sarnaselt eelmises peatükis kirjeldatud *käima*-ühenditega on *mas*-vormil EKG järgi ka koos finiidse *olema*-verbiga koos esinedes finaali-lokatiivne tähendus (*Ta on söömas.*), mis täidab sellisel juhul lauses kohamääruse funktsiooni (Erelt *et al.* 1993: 255). Erinevalt *käima*-ühenditest võib aga *ma*-infinitiivi inessiivi ja finiidse *olema* ühend EKG järgi moodustada perifrastilise verbivormi, mis võib väljendada kestva protsessi (progressiivne aspektitähendus: *Kellad on helisemas.*) või sündmuse eelfaasi (proksimatiivitähendus: *Kontsert on algamas.*) (Erelt *et al.* 1993: 259; Erelt 1987: 45–46). Selle ühendi progressiivset tähendust on põhjalikult analüüsinud Metslang (1993a; 1993b; 1993c) ja leidnud, et tarindi grammatiseerumine eesti kirjakeeles on veel pooleli. Proksimatiivi ehk oleviku või mineviku sündmuse eelfaasi kategooriat on vaadelnud Erelt ja Metslang (2009) ning toonud teiste hulgas välja *olema* + *mas*-vormi kui ühe proksimatiivi (*Vene väed on Gorist lahkumas.*) väljendamise võimaluse eesti keeles. Proksimatiivi kategooria jätab markeerimata, kas mingi oleviku või mineviku sündmus toimub või mitte ning finiidne *ma*-infinitiivi inessiivivormis verb on punktuaalse tähendusega (Erelt 2009: 46), kuigi järjest rohkem võivad selles positsioonis esineda ka tegevusverbid (Metslang 1993a; 1993b; 1993c).

Selle konstruktsiooni varieerumist murretes on põgusalt analüüsinud Neetar (1988). Tema tähelepanekute kohaselt (1988: 42–44) väljendab kirjeldatav ühend Mulgi, kirderanniku ja põhjaeesti murretes nii kestust kui ka üleminekut

ühast olukorrast teise, samas kui Tartu ja Võru murdes väljendatakse sel viisil vaid kestusega tegevust või protsessi. Peatse alguse väljendusvahendina domineerib nendes murretes Neetari järgi adessiivne teonimi *mas*-konstruktsiooni üle.

Uurimismaterjalist tuvastasin 169 supiinivormi inessiivi ja finiidset verbi *käima* sisaldavat osalauseid, mille jagasin viie erineva konstruktsiooni esindajate vahel. Tabel 44 esitab tuvastatud konstruktsioonid ja nende esinemisjuhtude **normaliseeritud** sagedused. Kollostruktsioonilise analüüsi käigus osutus ühend konstruktsioonina statistiliselt oluliseks vaid Võru murdes.

Tabel 44. *ma*-infinitiivi inessiivi ja *olema*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: lok – lokatiivse tähendusega konstruktsioon, omadus – omadust väljendav konstruktsioon, progr – progressiivkonstruktsioon, proks – proksimatiivkonstruktsioon, püsi – püsiühend.

	lok	omadus	progr	proks	püsi
IDA	11	0	1	0	1
KESK	6	1	4	0	7
KIRDE	11	0	0	2	8
LÄÄNE	5	0	2	2	5
MULGI	2	0	0	2	5
RANNA	15	0	4	1	3
SAARTE	7	0	3	1	3
SETO	9	0	0	0	6
TARTU	3	0	9	0	0
VÕRU	13	0	1	0	17

Lokatiivse tähendusega konstruktsioon

Enamasti esindab finiidne *olema*-verb koos supiini inessiiviga lokatiivset konstruktsiooni (75).

- (75) toist kümme jälle siis **olid** e **puhkamass** (RAN)
ma **kullõmahh olõ** (VÕR)
nigu nõu **küsümäs** siis kusagil **oli** (SAA)

Supiini inessiivis verb on siin tõlgendatav kohamäärusena. Konstruktsioon on sarnane eelmises alajaotuses kirjeldatud lokatiivse supiini ja *käima*-verbi konstruktsiooniga, kuid tulenevalt finiidse verbi semantikast puudub siin dünaamiline tähendus.

Progressiivi- ja proksimatiivkonstruktsioon

Progressiivse ehk kestva aspekti väljendamine *mas*-vormi ja *olema* abil on murretes suhteliselt harv nähtus, kuid siiski selgelt olemas (76).

- (76) ja teene **olli** koa **tahtmas** (SAA)
no sõda **olli** **käeman** (TAR)
siis **oli** siin ligidal **elamas** inimene (KES)

Progressiivsete konstruktsioonide sageduse varieerumine murdeti on märkimisväärne. Ühtegi esinemisjuhtu ei tuvastanud ma Mulgi, Seto ja kirdemurdest, samas kui Tartu murdes domineerib just progressiivse aspekti väljendamine. Progressiivse aspekti tähenduse tõlgendus ilmneb küll infiniitses positsioonis esinevast verbist, millest tingitult on konstruktsioonile keerulisem omistada lokatiivset tähendust. Murretes näib selle tähenduse laienemine olevat küll üsna algusjärgus ning tuleb harva esile.

Proksimatiivkonstruktsiooni (77) eristab mingi oleviku- või mineviku-sündmuse peatse algamise tähenduse ilmumine, kusjuures konstruktsioon ei markeeri seda, kas sündmus leiab lõpuks aset või mitte (erinevalt avertiivist, mis markeerib sündmuse mittetoimumist).

- (77) ja siis ku vihm **oli tulemas** no (KIR)
ja edasi ikka **oli vaomas vaomas** ja maha **kukkumas** (LÄÄ)
et nüid **omm** üits igävene nuhtluss tall ikki **tulemen** (MUL)

See konstruktsioon ilmnes ainult kirde-, lääne-, ranna-, saarte ja Mulgi murretes ning esinemisjuhtusid oli vähe.

Omadust väljendav konstruktsioon ja püsiühend

Omadust väljendava konstruktsiooni esinemiskordi (78) oli materjalis kolm, saarte ja keskmurdes. Nii analüüsi ühtlustamise eesmärgil (vt ka finiiitse *olema*-verbi ja supiini konstruktsioone) kui ka piisavalt erineva funktsiooni ilmlemise tõttu käsitlen seda siin siiski eraldi konstruktsioonina.

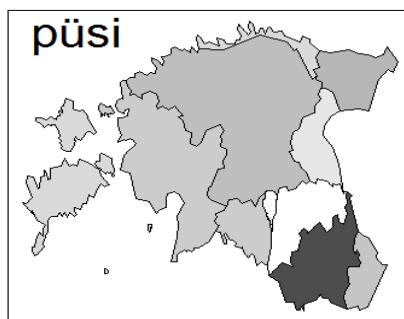
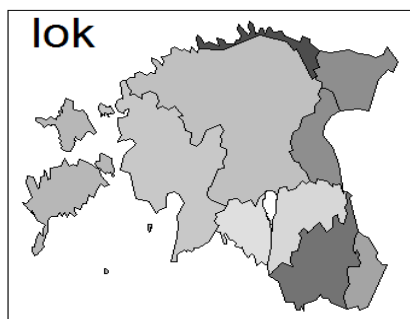
- (78) no mina **olin** alles ikka vist *akkaja* **käimas** (KES)
no **oless** *nõus* **minemas** (KES)
ma **ss ole** *isus* **tulemas** ühtid (SAA)

Konstruktsioon on nii vormilt kui tähenduselt sarnane *ma*-infinitiivi ja finiiitse *olema*-verbiga moodustatud omadust väljendava konstruktsiooniga, mida kirjeldasin peatükis 9.2.7. Ainuke erinevus näibki väheste näidete põhjal otsustades olevat infiniitses vormis. Põhitähendust ei kanna siin mitte infiniitne vorm, vaid finiiitne *olema* koos adverbi või substantiiviga. Konstruktsiooni kolmas element iseloomustab *mas*-vormiga väljendatavat tegevust või selleks vajalikku omadust.

Püsiühendi moodustavad finiiitne *olema* ja *ma*-infinitiivi inessiivi vormis *olema*-verb (79). Eksistentsiaalse tähendusega ühend rõhutab subjekti referendi olemasolu. Subjekt võib olla nii nominatiivis kui ka partitiivis.

- (79) kõikki kauppassi ültse mida **olemas onn** (KIR)
üks tanu **onn olemas** (SAA)
ja puuluidsaq pidosih kahh puuluidsaq kahvlitt **ess oleq olemahkiq**
(VÕR)

Kuigi sellist püsiühendit esineb pea kõigis murretes, näib see eriti iseloomulik olevat Võru murdele. Joonis 32 esitab finiiitse *olema*-verbi ja supiini inessiiviga moodustuvate püsiühendite ja lokatiivse konstruktsiooni geograafilise sagedusjaotuse. Kuna teisi konstruktsioone tuli esile väga harva, siis nende kohta eraldi kaarte esitatud pole. Lokatiivse konstruktsiooni puhul torkab silma, et see on pisut iseloomulikum idapoolsetele murdealadele. Kõrgeim on sagedus küll rannamurdes; ülejäänud aladel on selle sagedus ühtlaselt madalam. Püsiühendi esinemissagedus on kõrgeim Võru murdes ja muudel aladel seda märkimisväärselt ei esinenud.



Joonis 32. Finiitse *olema* ja *ma*-infinitiivi inessiivi konstruktsioonide geograafiline sagedusesitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 45 esitab finiiitse *olema* verbiga tugevamini tõmbuvad supiini inessiivis esinevad verbid ida- ja läänepoolsete murrete rühmas. Idarühma kuuluvad traditsiooniliselt lõunaesti murderühma liigitatud murded, millele on iseloomulik püsiühendi sage kasutus, see selgitab *olema*-verbi esinemist selles rühmas. Läänerühmas pole ühtegi verbi, mille statistilise seose tugevuse mõõdik võiks viidata mõne ühendi püsiühendilaadsusele, seega näib, et ühend on selles rühmas produktiivsem. Idarühmaski on ühendi produktiivsus üsna ühtlane ja kõrge. Rühmas on vaid kaks verbi, mille statistilise seose tugevuse väärtus on kõrgem: *istuma* ning *teenima* on siin finiiitse *olema* verbiga tugevamalt seotud.

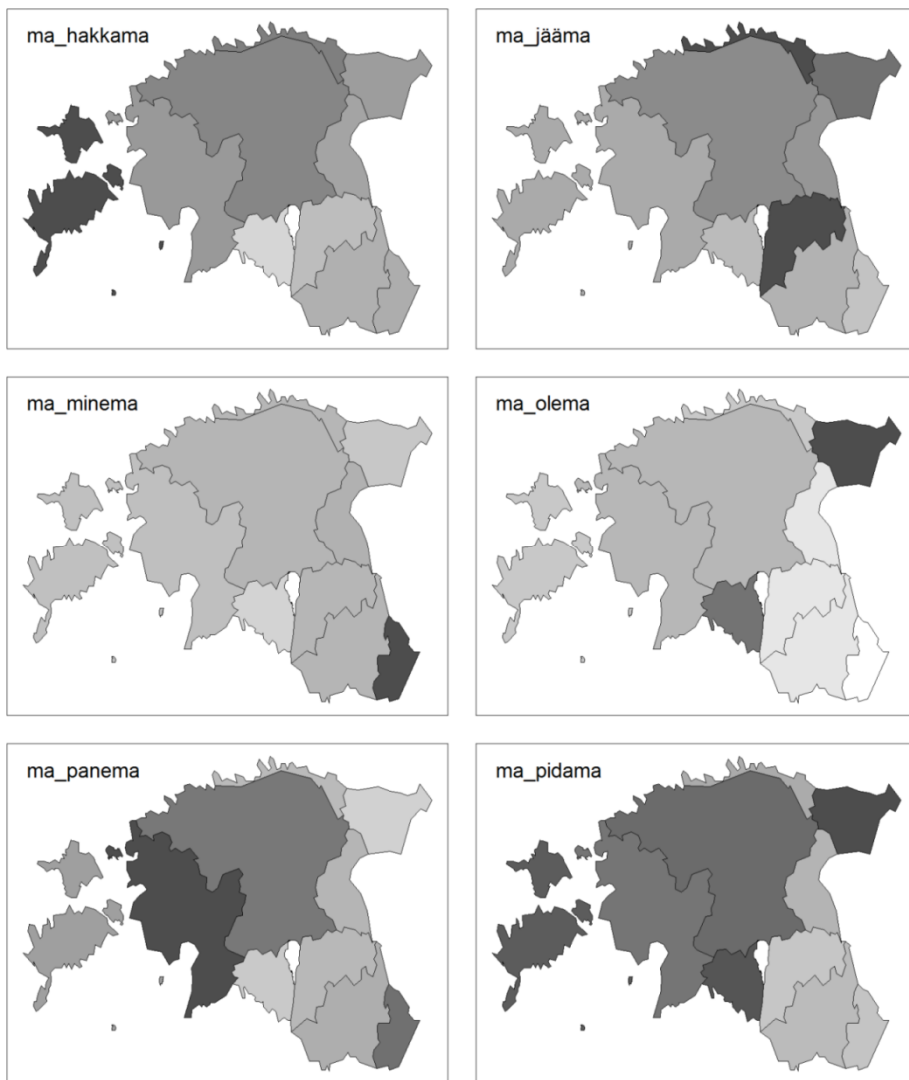
Tabel 45. *olema*-verbiga tugevamini tõmbuvad *ma*-infinitiivi inessiivi vormis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
istuma	2.45	käima	1
teenima	1.34	elama	0.67
olema	0.92	minema	0.67
kakkuma	0.8	püüdma	0.67
mõskma	0.8	tegema	0.67
õmblema	0.8	laadima	0.45
hulkuma	0.4	laulatama	0.45
jalutama	0.4	liikuma	0.45
kuivama	0.4	niitma	0.45
kuulama	0.4	tahtma	0.45

9.2.11 Kokkuvõte

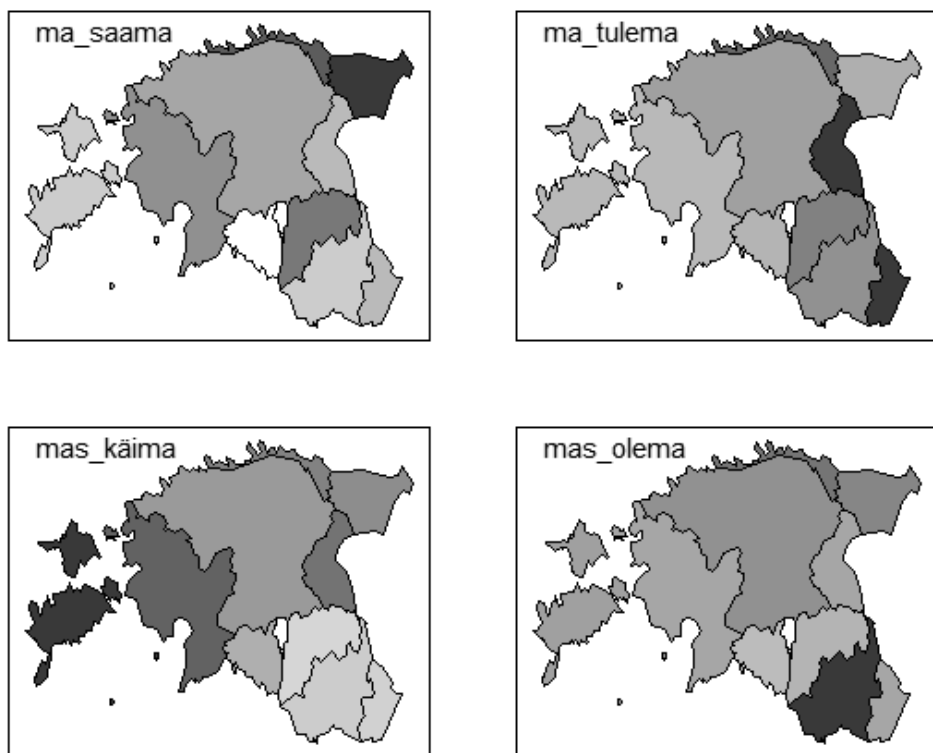
Joonised 33 ja 34 esitavad visuaalselt kvantitatiivse analüüsi põhjal tuvastatud finiiitse verbi ja *ma*-infinitiivi ühendite geograafilise sagedusjaotuse. Supiini ja *hakkama*-verbi ühendi sagedus on eriti kõrge saarte murdes ja harvem esineb seda Mulgi murdes. Ülejäänud murded näivad sagedusmuustrite poolest olevat üsna sarnased. *jääma*-verbi ja supiini ühendeid esineb eriti palju Tartu ja rannamurdes. *minema*-ühendi suurem sagedus torkab silma Seto murdes, samas kui ülejäänud murdealadel esineb seda suhteliselt vähe. *olema*-verbi ja supiini ühendid on iseloomulikumat kirde- ja Mulgi murdele, ent näiteks Seto murdes ei esine seda üldse. Supiin koos finiiitse *panema*-verbiga on eriti sage lääne-murdes ning sagedasem ka kesk- ja Seto murdes ning ülejäänud murdealadel on sagedused üsna ühtlased. *pidama*-ühendite põhjal eristub selgemalt kaks rühma, millest kõrgema sagedusega rühma kuuluvad Mulgi, kirde-, kesk-, lääne- ja saarte murre ning madalama sagedusega rühma moodustavad Seto, Võru, Tartu, ida- ja rannamurre.

Sarnaselt *da*-infinitiivi ühenditega ei moodustu siingi üksikute verbiühendite lõikes homogeensemaid rühmi. Erandiks võib pidada *pidama*-ühendeid, kus selgemalt eristusid kõrgema ja madalama esinemissagedusega rühm.



Joonis 33. *ma*-infinitiivi ja finiidse verbi ühendite sagedused (1).

Supiinivorm koos *saama*-verbiga on sage kirde- ja rannamurdes ning üldse ei esinenud seda Mulgi murdes. *tulema*-verb koos supiiniga on eriti sage Seto ja idamurdes ning ülejäänud murretes jaotuvad sagedused ühtlaselt. *käima*-verbi ja supiini inessiivi ühendid on teistest sagedasemad saarte murdes ning *mas*-vormi ja *olema*-verbi ühendeid esineb rohkem Võru murdes.



Joonis 34. *ma*-infinitiivi ja finiidse verbi ühendite sagedused (2).

Joonis 35 esitab supiini ja finiidse verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafiku konstruktsioonide analüüsi tulemuste põhjal. Joonis esitab kaks esimest, seletavat mõõdet, mis kokku seletavad üle 60% andmestiku variatiivsusest.

Esimese mõõtmepõhjal eristuvad kaks gruppi, millest esimesse kuuluvad Mulgi, kirde-, kesk-, lääne- ja saarte murre ning teise ida-, ranna-, Võru, Tartu ja Seto murre. Selgelt eristuvaid homogeenseid gruppe siiski ei teki ning kahe nimetatud grupi sisene varieerumine on üsna suur. Huvitav on siin see, et traditsiooniliselt lõunaeesti murderühma kuuluvad Seto ja Mulgi murre on esimese mõõtmepõhjal üksteisest kõige erinevamad. Võru ja Tartu murre näivad olevat väga sarnased, aga eristuvad samuti olulisel määral Mulgi murdest. Esimese mõõtmepõhjal tekib visuaalselt teatav kontiinum, mille üheks servaks on Mulgi murre, järgnevad üleminekud vasakult paremale kirdemurdelt lääne-, saarte ja keskmurdele, millele järgnevad ranna- ja idamurre, Võru ja Tartu murre ning kontiinumide teises ääres on Seto murre.

Teise mõõtmepõhjal on Mulgi ja Seto murre sarnasemad, moodustades ühtsema grupi koos Võru, Tartu ja kirdemurdega. Saarte, ranna- ja idamurre koos suhteliselt keskel paiknevate lääne- ja keskmurdega moodustavad teise grupi.

Korrespondentsanalüüs: ma-infinitiivi konstruktsioonid

a konstruksioon a murre



Joonis 35. *ma*-infinitiivi ja finitiise verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafik.

Seto, Võru ja Tartu murdele näivad olevat iseloomulikumad afiksaaladverbi- ja lokatiivsed konstruktsioonid. Lokatiivne supiini inessiiv näib olevat siiski rohkem lääne-eestilise levikuga. Algust väljendav *hakkama*-konstruktsioon tundub olevat tugevalt seotud rannamurdega. Mulgi murret eristab teistest *olema*-verbiga moodustatav omadust väljendav konstruktsioon. Erinevad *pidama*-konstruktsioonid näivad olevat iseloomulikumad läänepoolsetele murretele ja kirdemurdele.

Erinevalt *da*-infinitiivi konstruktsioonidest, kus selgemini eristusid traditsioonilised põhja- ja lõunaeesti murderühmade vahelised erinevused, on *ma*-infinitiivi konstruktsioonide levik sarnasem kvantitatiivse kobaranalüüsi tulemustega ning pigem erinevad omavahel ida- ja läänepoolsed mured. Erandlikum näib siin olevat Mulgi murre, mis erineb suurel määral teistest lõunaeesti murderühma murretest, aga ka kõigist ülejäänud murretest. Eriti sarnased supiini konstruktsioonide kasutuse poolest näivad olevat ranna- ja idamurre ning Tartu ja Võru murre. Kvantitatiivse analüüsi tulemuste võrdluses on oma kohad vahetanud ranna- ja kirdemurre, millest esimene näib siinses analüüsis sarnasem olevat idapoolsete murretega. Supiinkonstruktsioonide kasutusmuster on heterogeensem kui *da*-infinitiivi konstruktsioonide puhul. Vähem on konstruktsioone, mis on kuhjunud graafiku keskele ja konkreetsete konstruktsioonide distributsioon on tugevamalt seotud geograafiaga.

Tabel 46 esitab veel kord kokkuvõtlikult *ma*-infinitiivi ja finitiitse verbi konstruktsioonid eesti murretes.

Tabel 46. Finitiitse verbi ja *ma*-infinitiiviga moodustatavad konstruktsioonid eesti murretes.

VORM	TÄHENDUS	NÄIDE
<i>finiitverb</i>	<i>ma</i> -infinitiiv	
hakkama	algus püsiühend	<i>kess koloosis akkab piäd kandma</i> (IDA) <i>siis lehmad akkavad piima lüpsma</i> (SAA)
jääma	algus	<i>veski jäi seisma</i> (IDA)
minema	afiksaaladverbikonstruktsioon algus lokatiivsus püsiühend	<i>sis pühabä õhta läksime minema</i> (SAA) <i>et sõss läävüt jo ilmaq ilosale minemä ku</i> (VÖR) <i>pannkuokkidega läks seda titte vaatama</i> (RAN) <i>mies läks kaduma nagu sõea aeal</i> (KES)
olema	modaalsus omadust väljendav proksimatiiv	<i>endal enne oli kodu ära kaaluma</i> (KES) <i>aga sie õli kange magama</i> (KIR) <i>siis olivad vaidlema nagu</i> (IDA)

VORM	TÄHENDUS	NÄIDE
panema	algus kausatiivsus püsiühend	<i>panin silades tulema</i> (KES) <i>jahh ega mei vägisi tienima ei pane</i> (RAN) <i>no ja kui ta paneb nee asjad sedasi akma</i> (SAA)
pidama	avertiiv kvotatiiv impersonaalne kvotatiiv modaalsus proksimatiiv	<i>ku mina ema pidi minestama ära</i> (IDA) <i>seal pidi vesilillisi olema</i> (LÄÄ) <i>siis midagit pidi ikka sönnna pääle pandama</i> (SAA) <i>sii pidi kaa kaunis noorest peast abielluma</i> (MUL) <i>ja mina pidin kukkuma ja kukkusingi</i> (IDA)
saama	tulevik imperfektiivne tulevik kausatiivsus püsiühend õnnestumine	<i>kui tema kuulsad nime saa saad iest ja tagant üidma</i> (KES) <i>se sai üks lollukkane olema</i> (RAN) <i>ei saand enam seda pidama</i> (KES) <i>küll ma neega saan isegi akkama</i> (KIR) <i>et saa ma ruttõmbade minemä</i> (VÕR)
tulema	afiksaaladverbikonstruktsioon püsiühend lokatiivsus	<i>sial tulime tulema</i> (IDA) <i>ka nüid teist kord tuli lüpsma</i> (KES) <i>keda te tulette siin sõimama</i> (IDA)
ma-infinitiivi inessiiv		
käima	lokatiivsus	<i>käisimi Pikkäsillal ja Liival ja jooman</i> (TAR)
olema	lokatiivsus omadus progressiiv proksimatiiv püsiühend	<i>ma kullõmahh olõ</i> (VÕR) <i>no mina olin alles ikka vist akkaja käimas</i> (KES) <i>no sõda olli käeman</i> (TAR) <i>ja edasi ikka oli vaomas vaomas ja maha kukkumas</i> (LÄÄ) <i>neid tarkkasi onn olemas</i> (LÄÄ)

Sarnaselt *da*-infinitiivi konstruktsioonidega on ka supiinikonstruktsioonidele iseloomulik madala esinemissagedusega homonüümsete konstruktsioonide esiletulek ja *ma*-infinitiivi konstruktsioone esineb ka üldiselt vähem kui erinevaid *da*-infinitiivi konstruktsioone. Teistest selgelt grammatilisem on *pidama*-verb, mis enamasti osaleb küll modaalkonstruktsiooni moodustamisel, kuid üsna levinud on ka sellekujuline analüütiline kvotatiiv. Supiinikonstruktsioo-

nide juures on veel huvitavad verbid *saama* ja *olema*, millega moodustunud konstruktsioonide sagedus on materjalis küll madal, kuid erinevate konstruktsioonide arv jällegi üsna kõrge. Supiiniga koos esinevad vaadeldud finiitsed verbid on küll märgatavalt vähem grammatiseerunud ja moodustavad pigem verbi ja infiniitse laiendi kombinatsioone, kuid siingi on see eri murrete lõikes erinev. Selgelt grammatilisemad näivad olevat *saama*, *hakkama*, *olema* ja *pidama*. Finiitsed *jääma* ja *käima* on piiratuma grammatilise kasutusega. Supiinikonstruktsioonide puhul on kõrgem erinevate püsiühendite arv ning need püsiühendid näivad olevat eriti iseloomulikud lõunaeesti murretele.

9.3 *tud*-partitsiibi konstruktsioonid

Siinses töös käsitletavate *tud*-partitsiibi ja finiitsete *olema* ning *saama* perifrastiliste vormidega väljendatakse eesti kirjakeeles valdavalt impersonaali ja passiivi. Traditsiooniliselt eristatakse pea kõigis läänemeresoome keeltes personaali ja impersonaali kategooriat, passiivi eristamine on lisandunud hiljem (Erelt 2003: 216–219). EKG (1993: 30–31) järgi on üks nii impersonaali kui passiivi väljendamise viise verbide *saama* ja *olema* finiitsed vormid ja *tud*-paritsiip, selliste konstruktsioonide ja nende muude tähendustega tegelevadki kaks järgmist alapeatükki.

Impersonaali ja passiivi eristamine ei ole alati kerge. Siinses töös olen püüdnud järgida Torn-Leesiku (2009) poolt esitatud kriteeriume. Impersonaalikonstruktsioone on võimalik moodustada nii transitiivsete kui intransitiivsete verbidega, kuid verb eeldab valentsist tulenevalt elusat tegijat subjekti rollis, mis impersonaalis on välja jäetud. Kui verb on transitiivne, on semantiline patsient impersonaalis grammatilise objekti rollis. Passiivikonstruktsioone saab moodustada ainult transitiivsetest verbidest ning semantiline patsient on tõstetud grammatilise subjekti rolli, mis on nimetavas käändes ning ühildub abi-verbina talitleva *saama*- või *olema*-verbiga. Selline lähenemine on loomulikult äärmiselt lihtsustatud ning praktiline analüüs näitas, et seda pole võimalik alati järgida. Eriti tekitavad probleeme impersonaali liitajavormid, mis on identsed passiivi oleviku ja mineviku ajavormidega (vt nt Torn-Leesik 2009: 85). Lisaks on *saama*-verb koos *tud*-partitsiibiga esinedes üsna levinud isikule viitamise vältimise strateegia (Lindström 2010: 108–110), sellised juhtumeid on siinses töös käsitletud samuti impersonaalikonstruktsioonina.

Helasvuo ja Vilku demonstreerivad soome keele näitel ilmekalt, kui varieeruvad ja mitmekülgsed võivad erinevad passiivi- ja impersonaalikonstruktsioonid olla ning tõdevad, et ühes keeles võib esineda erinevaid konstruktsioone, mida kõiki võib nimetada impersonaalisteks (2008: 235). Kuna sinne töö tegeleb kümne murdega, võib eeldada, et olukord erinevates murretes on erinev, just sellel põhjusel olen püüdnud lähtuda võimalikult palju vormikriteeriumist ja analüüs on üsna skemaatiline. Olen analüüsist välja jätnud laused, mille puhul oli konstruktsiooni määramine lähikonteksti põhjal võimatu. Kindlasti on *tud*-partitsiibi konstruktsioonide analüüs selle töö kõige subjektiivsem.

R. Pajusalu jt (2004: 42) järgi väljendavad nii *olema* kui *saama* koos *tud-*partitsiibiga terminatiivset aspekti ehk sündmuse lõppemist või lõpetatust, selliseid lisatähendusi võib omistada ka perifrastilistele passiivi- ja impersonaali konstruktsioonidele. Eesti keeles on passiivi ja impersonaali kategooriad ning nende vahetõrge viimastel aastatel leidnud laialdast käsitlemist (Lindström & Tragel 2007; Rajandi 1999; Torn-Leesik 2007; Torn-Leesik & Vihman 2010), kuid sellest hoolimata pole kategooriate eristamise kriteeriumid selged.

Erinevate passiivikonstruktsioonide kohta on murretes tehtud tunduvalt rohkem tähelepanekuid kui teiste selles töös käsitletavate konstruktsioonide kohta. Passiivi erijoonte kohta nii murretes kui ka lähisugulaskeeltes on hea ülevaate andnud Alvre (1993). Must ja Univere on märkinud *tud-*partitsiibi ja *saama* konstruktsiooni laialdast levikut keskmurdes, kus sellisel viisil väljendatakse impersonaalsust, tegevuse lõpetatust ja imperfekti. Sama konstruktsiooni eitav vorm väljendab keskmurde põhjamurrakutes mingi tegevuse sooritamiseks võimaluse puudumist. (Must & Univere 2002: 349–350) Wiedemann ([1875]; 2011: 527) aga väidab, et *saama* ja *tud-*partitsiip moodustavad isikulise passiivi ja nähtus on levinum lõunapoolsetel aladel, aga Kesk-Eestis väljendatakse ühendi abil vaid neid tähendusi, mida muul viisil pole võimalik edasi anda. Eraldi toob Wiedemann ([1875]; 2011: 528) välja sama vormipaari esinemise umbisikulisel tähenduses koos adessiiviga nn Tartueesti murdes. Juhkam (2012: 519–520) nimetab samuti *saama* + *tud-*ühendit põhjaeestiliseks ja tema järgi väljendab ühend isikulise ja umbisikulise tegumoes tegevuse lõpetatust. Kirderannikumurdes väljendab ühend samuti impersonaalsust ja tegevuse lõpetatust (Must 1987: 301–304), viimast funktsiooni on peetud iseloomulikuks ka läänemurdele (Juhkam & Sepp 2000).

9.3.1 *tud-*partitsiip ja finiidne *olema*

Tauli (1980: 65–66) järgi kuulub *olema* 3sg+*tud*+objekt nominatiivis omaette infiniidiühendite rühma ning vormiga väljendatakse passiivi ja impersonaali mineviku liitaegu. Eraldi on ta (1980: 91–92) esitanud juhtumi, kus ühendis lisandub adessiivis noomen, mis väljendab tegijat (*mul on töö tehtud*) või kellel midagi on keelatud või lubatud (*meil ei ole lubatud arvet maksta*). Suurema rühma moodustavad Tauli (1980: 144–145) järgi lihtsalt infiniidiühendid, mis koosnevad verbist ja *tud-*partitsiibist, kus viimane väljendab mingit seisundit, kuhu keegi on sattunud või asetatud ning samuti võib partitsiip täita adjektiivivi või adverbi funktsiooni (*juuksed olid kammitud*). Viimane konstruktsioon on sama, mida EKG-s (Erelt *et al.* 1993: 30–31) nimetatakse seisundipassiiviks.

Kokku analüüsisin 1370 *tud-*partitsiibi ja finiidset *olema* ühendit sisaldavat osalauseid ning tuvastasin kolm erinevat konstruktsiooni: impersonaali-, passiivi- ja possessiivse perfekt konstruktsioonid. Tuvastatud konstruktsioonide esinemisjuhtude **normaliseeritud** sagedused murrete kaupa esitab tabel 47.

Tabel 47. *tud*-partitsiibi ja finiiitse *olema*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: imp – impersonaalkonstruktsioon, pass – passiivikonstruktsioon, possperf – possessiivne perfekt.

	imp	pass	possperf
IDA	38	71	20
KESK	46	85	6
KIRDE	32	63	11
LÄÄNE	36	80	9
MULGI	80	84	12
RANNA	41	52	4
SAARTE	68	91	12
SETO	26	45	20
TARTU	33	82	14
VÕRU	51	68	23

Kollostruktsioonilise analüüsi käigus ei osutunud finiiitse *olema* ja *tud*-partitsiibi ühend statistiliselt oluliseks ranna-, saarte, kesk- ja läänemurdes. Järgnev osa kirjeldab iga konstruktsiooni detailsemalt.

Passiivikonstruktsioonid

Passiivikategooria ja -konstruktsioonid on viimastel aastatel eesti kirjakeeles leidnud laialdaselt käsitlemist (Torn-Leesik 2009; Lindström & Tragel 2007; Torn-Leesik & Vihman 2010; Alvre 1993 jt). Konstruktsiooni on analüüsitud paljudest eri aspektidest, kuid kategooria ise on eesti keeles siiski veel suhteliselt ebaselge. Käesolevas töös ei ürita ma esitada omapoolset passiivikategooria kirjeldust, vaid püüan, toetudes teiste analüüsidele, seda konstruktsiooni eristada muudest, autori arvates erinevat tähendust kandvatest konstruktsioonidest murretes.

Passiivikonstruktsioonis on semantiline patsient grammatilise subjekti positsioonis ja ühildub finiiitse *olema*-verbiga (80). Agent kas on konstruktsioonist välja tõrjutud või seda väljendatakse agentadverbiaali abil (näiteks (80) esimene näitelause rannamurdest).

- (80) et vat sie **on** Lienini **ehittattud** (RAN)
 siis **olid** pised sitsipluused **tehtud** (SAA)
 niid vanad ajad **on** kõik muudetud ja eedesipitte **lökkättud** (IDA)

Sellisel moodustatav passiivikonstruktsioon näib olevat üsna tavaline kõigis murretes, mõnevõrra vähem on konstruktsiooni esinemisjuhtusid ranna- ja Seto murdes. Konstruktsioonil on kõigis murretes tugev resultatiivsuse tähendus ning vahel rõhutatakse resultatiivsust eraldi afiksaaladverbiga (81).

- (81) **onn jägattud ära** taludelle se Nurga kuusik (IDA)
 noo nüid ta **onn** küll juba **ostettud ära** (KES)
 syss **umm** hobõsõq kabi **ärä** ras ee **ravvuttõt** (VÕR)

Impersonaalikonstruksioon

Impersonaalikonstruksioonis (82) võib semantiline patsient olla vormistatud grammatilise objektina. Lisaks saab impersonaalikonstruksiooni moodustada ka intransitiivsetest verbidest.

- (82) ja siis **on imestattud** seda saege (IDA)
 näid **oll** jo **otsit** (VÕR)
 mind **pole** mette **küsittud** muud (SAA)

Konstruksioon on kohati tõlgendatav nii üld- kui umbisikulisena. Semantiline agent jäetakse eksplitseerimata ning näitelausestes ei ilmne see ka lähikontekstist. Eriti iseloomulik on konstruksioon Mulgi ja saarte murdele.

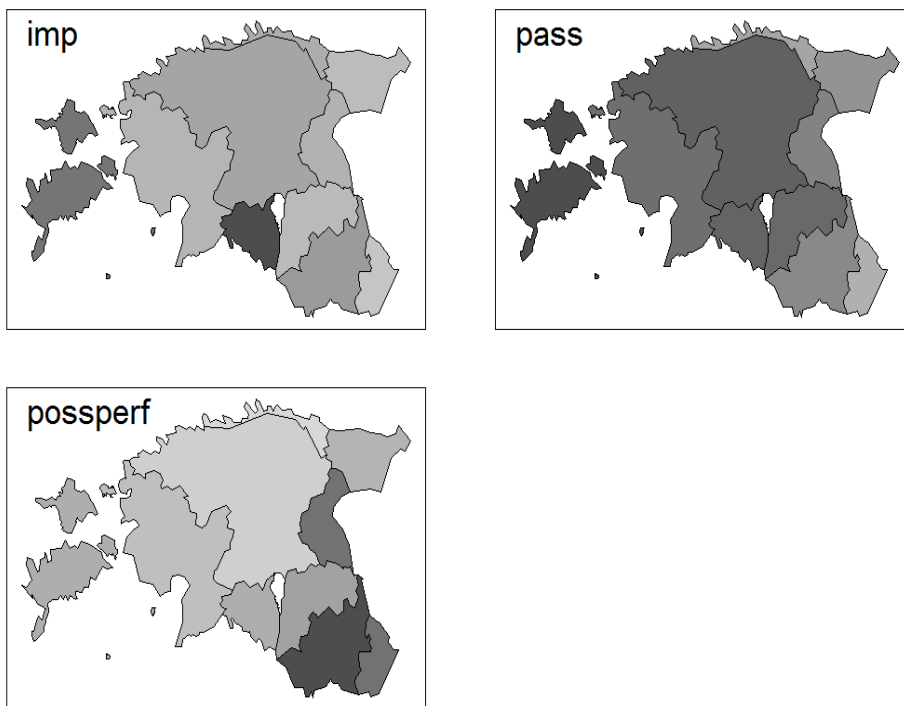
Possessiivse perfekti konstruksioon

Possessiivse perfekti konstruksiooni kirjakeeles on põhjalikult käsitlenud Lindström ja Tragel (2010; 2007). Samakujulisest impersonaalist ja passiivist eristab seda konstruksiooni agendi väljendamine adessiivivormiga ja siingi on tugev resultatiivsustähendus (83).

- (83) ja niibalju **õli** juba *mull* käsidööd ära **tehtud** (IDA)
 noo prostoi rõiva üks **olli** *umal* **koedu** (VÕR)
Salmõl **oll** ka **ostõt** purgitäis mett (SET)

Adessiivis agent paikneb kõige tüüpilisemalt lause algul, aga mitte alati. Semantiline patsient võib olla sarnaselt passiivile vormistatud grammatilise subjektina, aga võib ka puududa (sarnaselt impersonaalile).

Joonis 36 esitab finiiitse *olema*-verbi ja *tud*-partitsiibi konstruksioonide geograafilise sagedusjaotuse. Impersonaalikonstruksioon näib olevat eriti iseloomulik Mulgi murdele ja teistes murretes on selle sagedused mõnevõrra madalamad. Passiivikonstruksioon on suhteliselt ühtlase levikuga, mõnevõrra vähem esineb seda lõuna- ja idapoolsetel aladel. Possessiivne perfekt on see-eest tugevalt lõuna- ja idaeestiline ja läänepoolsetele aladele on konstruksioon vähem iseloomulik.



Joonis 36. Finiitse *olema* ja *tud*-partitsiibi konstruktsioonide geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

Tabel 48 esitab finiiitse *olema*-verbiga tugevamini tõmbuvad *tud*-partitsiibis verbid. Pilt on mõlemas rühmas suhteliselt sarnane, kuid tugevamini tõmbuvaid verbe on rohkem idarühmas, mis lubab oletada *olema*-verbi suuremat produktiivsust *tud*-partitsiibiga kombineerudes läänerrühmas. Kuigi kollostruktsioonilises analüüsis läänepoolsetes murretes ei osutunud ühend isegi statistiliselt oluliseks, et seda konstruktsiooniks pidada, ilmnes kvalitatiivsest analüüsist konstruktsioonide laialdane kasutus läänepoolsetes murretes. Siiski tundub, et *olema*-verb vaadeldavas ümbruses on suhteliselt ühtlase produktiivsusestmega mõlemas rühmas. Suuri verbisemantilisi erinevusi kahe rühma vahel ei ilmne.

Tabel 48. *olema*-verbiga tugevamini tõmbuvad *tud*-partitsiibis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
tapma	4.11	siduma	1.83
tikkima	3.02	kuduma	1.6
istutama	2.16	nägema	1.27
käänama	1.72	keetma	1.22
mõõtna	1.29	laulatama	1.01
otsima	1.29	lõikama	0.97
sugema	1.29	puurima	0.81
vahtima	1.29	tappima	0.81
keelama	1.18	käima	0.79
matma	1.15	ehtima	0.7

Sarnaselt kirjakeelega on *olema* koos *tud*-partitsiibiga koos esinedes murretes üsna produktiivne ühend, osaledes kirjakeelega sarnaste tähenduste väljendamisel (impersonaal, passiiv, possessiivne perfekt).

9.3.2 *tud*-partitsiip ja finiidne *saama*

tud-partitsiibi ja finiidse *saama* ühend ja konstruktsioonid on murrete lõikes selgelt kõige varieeruvamad, ka on just selle konstruktsiooni kohta tehtud murretes kõige rohkem tähelepanekuid (Alvre 1993; Must 1987; Must & Univere 2002; Juhkam & Sepp 2000; Juhkam 2012; Wiedemann 2011; [1875]). Ühendi abil väljendatavaid erinevaid tähendusi on üsna põhjalikult käsitletud kirjakeeles (nt Lindström & Tragel 2010; Lindström & Tragel 2007; Tragel & Habicht 2012). Murrete käsitlustes on ühendit peetud eriti tüüpiliseks kesk- ja läänemurdele (Alvre 1993; Juhkam & Sepp 2000: 28–29; Juhkam 2012: 519–520; Must & Univere 2002: 349). Siinse töö analüüsitulemused kinnitavad ühendi põhjaestilist levikut.

Kokku analüüsisin 1190 *tud*-partitsiibi ja *saama* ühendit sisaldavat osalauset. Kollostruktsiooniline analüüs paljastas vormielementide tugeva statistilise seose kõigis murretes, v.a Seto murdes. Seto murdes oli ühendi esinemissagedus üleüldiselt madal. Analüüsi käigus tuvastasin neli konstruktsiooni, mida moodustatakse *tud*-partitsiibi ja finiidse verbi *saama* abil: passiivi-, impersonaali-, resultatiivsus- ning possessiivse perfekti konstruktsioonid. Konstruktsioonide analüüs näitas selget ida- ja läänepoolsete murrete vahelist erinevust. Tabel 49 esitab tuvastatud konstruktsioonid ja nende esinemisjuhtude **normaliseeritud** sagedused kõigis murretes eraldi. Järgnev osa kirjeldab iga konstruktsiooni detailsemalt.

Tabel 49. *tud*-partitsiibi ja finiiitse *saama*-verbi konstruktsioonid ja nende esinemisjuhtude normaliseeritud sagedused. Lühendid: imp – impersonaalikonstruktsioon, pass – passiivikonstruktsioon, possperf – possessiivne perfekt, result – resultatiivsus-konstruktsioon.

	imp	pass	possperf	result
IDA	15	14	0	0
KESK	139	55	0	5
KIRDE	11	8	0	0
LÄÄNE	102	69	1	1
MULGI	14	5	0	2
RANNA	135	36	0	7
SAARTE	121	52	1	6
SETO	6	0	0	3
TARTU	3	9	1	0
VÕRU	15	1	0	1

Impersonaalikonstruktsioon

Kõige sagedamini väljendatakse finiiitse *saama* verbi ja *tud*-partitsiibiga murretes impersonaali (84). Tähendust on analüüsitud sarnaselt *olema*-verbi ühenditega, mida kirjeldasin eelmises alajaotuses. Impersonaali moodustavad konstruktsioonid, kus verbiühendiga sooritatava tegevuse agent on eksplitseerimata. Finiitne *saama* on alati ainsuse kolmandas pöördes.

- (84) siis **sai** maha **tõstetud** (LÄÄ)
 kui magama **sai** **eidettut** (RAN)
 aga kaloosi töö! no **sai** ju ikka **käidud** (KES)

Konstruktsioon on selgelt põhjaeestiline, sagedused on kõige suuremad saarte, kesk-, lääne- ja rannamurdes. Lindström (2010) on käsitlenud seda konstruktsiooni kirjakeeles ja leidnud et see on väga levinud 1. ja 2. isikule viitamise vältimise strateegia.

Passiivikonstruktsioon

Passiivi väljendamine finiiitse *saama* ja *tud*-partitsiibi abil (85) on vormi poolest sarnane vastava *olema*-konstruktsiooniga. Siingi on oluline abiverbina talitleva finiiitse *saama* ühildumine grammatilise subjekti funktsioonis oleva semantilise patsiendiga.

- (85) leppad **said** kõik ära juurteg tükki juured ära **võetud** (KES)
 ja kivid **said** sinna peale **pandud** (LÄÄ)
 et suu **saab** **õppettatud** (IDA)

Konstruksioon on selgelt põhjateaduslik levikuga. Võrreldes impersonaaliga on see sagedasem idamurdes, kuid näiteks Seto murdest ei tuvastatud mitte ühtegi sellekujulist passiivikonstruktsiooni esinemisjuhtu.

Alvre (1993: 408) väitab, et Kagu-Eestis *saama*-konstruktsioonides *tud*-partitsiip esineda ka translatiivis. Seda võib siinse töö materjali põhjal kinnitada. Nii Seto kui Võru murdes olid kõik *tud*-partitsiibid *saama*-verbiga koos esinedes translatiivis ning ka Tartu murdes oli vaid üks lause, kus translatiivivorm puudus. Erandina võib siin välja tuua ühe Seto murde näite, kus translatiivis *tud*-partitsiip puudus ning sel juhul oli tegemist resultatiivsuskonstruktsiooniga (vt allpool). Seega võiks ettevaatlikult oletada, et translatiivse partitsiibi kasutuse tingib konstruktsioon.

Possessiivne perfekt ja resultatiivsuskonstruktsioon

saama ja *tud*-partitsiibi kujuline possessiivne perfekt (86) on sarnane sama tähendust kandva *olema*-verbi abil moodustatud konstruktsiooniga, ent abiverbi rollis talitleb siin *saama*. Konstruktsiooni eristavaks elemendiks on adessiivis esinev semantiline agent.

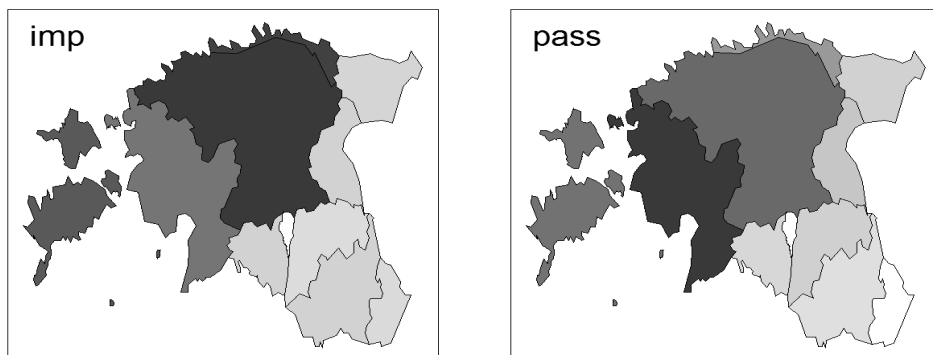
- (86) see **sai** jälle **kaaritud mool** paergus (SAA)
aga änam **mul tehtud ei saa** (LÄÄ)
sää **sai meil tüütledüss** (TAR)

saama-verbi abil moodustatud possessiivperfekti konstruktsioon on murretes tunduvat vähem levinud kui sama tähendusega *olema*-verbi konstruktsioon, selle töö materjalist tuvastasin vaid üksikud esinemisjuhud. Konstruktsiooni tuvastasin vaid saarte, lääne-, ja Tartu murdes, teistes murretes ei esinenud ühtegi *saama* ja *tud*-partitsiibiga moodustavat possessiivse perfekt konstrukti.

Resultatiivsuskonstruktsiooni moodustavad finiiitse *saama* ja *tud*-partitsiibi ühendid, kus grammatiline subjekt ja semantiline agent langevad kokku. Konstruktsioon väljendab mingit saavutatust või tulemust (87).

- (87) seal et *sa* rendi **maksetud suad** (KES)
sis **sai me küll esi küdset** leivä (SET)
ja et *temä velleboig* **saass** kangruss **õppit** (MUL)

Joonis 37 esitab finiiitse *saama*-verbi ja *tud*-partitsiibi impersonaali- ja passiivikonstruktsioonide esinemisjuhtude geograafilise sagedusjaotuse. Possessiivse perfekt ja resultatiivsuskonstruktsioonide madala esinemissageduse pärast pole neile eraldi kaarte tehtud. Impersonaalikonstruktsioon on selgelt iseloomulik kesk- ja rannamurdele. Idapoolsetes ja lõunaeeesti murretes tuvastati konstruktsiooni esinemisjuhtusid harva. Passiivikonstruktsioon on kõige kõrgema esinemissagedusega läänemurdes. Sarnaselt impersonaalikonstruktsiooniga esineb sedagi konstruktsiooni väga harva või üldse mitte lõuna- ja idapoolsetes murretes.



Joonis 37. Finiitse *saama* ja *tud*-partitsiibi konstruktsioonide geograafiline sagedus-esitus.

Ida- ja läänepoolsete murrete vahelised erinevused

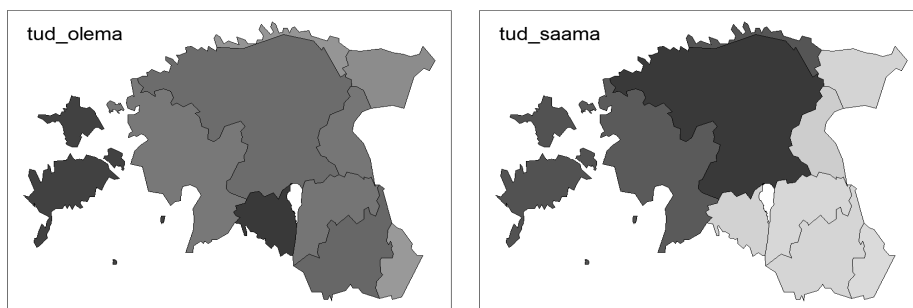
Tabel 50 esitab finiiitse *saama*-verbiga statistiliselt tugevamini tõmbuvad verbid *tud*-partitsiibis. Ootuspäraselt on ühend produktiivsem läänepoolsetes murretes, mida lubavad oletada madalad statistilise seose tugevuse väärtused selles rühmas. Konstruktsioon näib olevat suhteliselt produktiivne ka idarühmas, kus tugevamalt seotud verbe on siiski rohkem. Suuri verbisemantilisi erinevusi selle ühendi puhul kahe rühma vahel ei ilmne.

Tabel 50. *saama*-verbiga tugevamini tõmbuvad *tud*-partitsiibis verbid ida- ja läänepoolsete murrete rühmas.

IDARÜHM		LÄÄNERÜHM	
VERB	FTT	VERB	FTT
kolkima	2.44	panema	2.3
pesema	2.24	vedama	0.66
elama	1.98	keetma	0.6
kastma	1.98	andma	0.46
olema	1.73	hakkama	0.46
puistama	1.7	tõmbama	0.46
asutama	1.22	laskma	0.44
harima	1.22	lõikama	0.41
koolitama	1.22	püüdma	0.33
matma	1.22	minema	0.31

9.3.3 Kokkuvõte

Joonis 38 esitab *tud*-partitsiibi ja finiiitsete verbide *olema* ja *saama* ühendite sageduskaardid konstruktsioone eristamata. Tumedamad alad tähistavad kõrgemaid sagedusi ning heledamad alad madalamaid sagedusi. Vasakpoolisel kaardil on *olema*-ühendid, mille puhul ei tundu murretevahelised erinevused olevat märkimisväärsed. Kõrgeimad on nende ühendite sagedused Mulgi ja saarte murdes, teistest aladest on natuke madalamad sagedused Seto ja rannamurdes. Parempoolne kaart esitab *saama*-ühendid, kus ida- ja läänepoolsete murrete vahelised erinevused on ilmsed. Kõrgeim on ühendi kasutussagedus keskmurdes, järgnevad lääne-, saarte ja rannamurre. Kirde-, ida-, Mulgi, Tartu, Võru ja Seto murdes näib selle ühendi sagedus olevat ühtlaselt madalam.



Joonis 38. *tud*-partitsiibi ja finiiitse verbi ühendite sagedused.

Joonis 39 esitab korrespondentsanalüüsi graafiku *tud*-partitsiibi ja finiiitsete *saama* ja *olema* verbidega moodustatavate konstruktsioonide analüüsi tulemuste põhjal. Kaks esimest mõõdet seletavad kokku üle 90% analüüsitava andmestiku variatiivsusest, seega on nende seletusjõud märkimisväärne. Esimese mõõtmepõhjal eristuvad selgelt kaks gruppi. Vasakule jääb rühm, mis koosneb lääne-, kesk-, saarte ja rannamurdest ning teise, parempoolse rühma moodustavad ida-, kirde-, Tartu, Seto, Võru ja Mulgi mured. Seega on ida- ja läänepoolsete murrete eristumine nende konstruktsioonide puhul kõige selgejoonelisem ja korreleerub tugevalt kvantitatiivse kobaranalüüsi tulemustega. Teise, märkimisväärselt vähemolulise mõõtmepõhjal eristuvad lääne-, ida-, kirde- ja Tartu murre ning teise rühmana ranna-, saarte, Mulgi, Võru ja Seto murre. Finiiitse *saama*-verbidega moodustatavad konstruktsioonid näivad selgelt iseloomulikumat olevat läänepoolsetele murretele. Erandina võib välja tuua possessiivse perfektikonstruktsioonid, mis geograafiliselt eelistavad idapoolseid mureid. Kuigi kahe rühma vahelised erinevused on ilmsed, tuleb konstruktsioonide analüüsi kvantitatiivsel tõlgendamisel olla ettevaatlik. Nagu verbiühendite sagedusi esitavatel kaartidel näha võis, eristasid *saama*-ühendid selgelt ida- ja läänerrühma, samas kui *olema*-ühendite puhul suuri murretevahelisi sageduserinevusi ei ilmnunud. Seega mõjutab *saama*-konstruktsioonide suurem eristumine ka kõigi *tud*-partitsiibi konstruktsioonide tähenduste korrespondentsanalüüsi tulemusi.

Korrespondentsanalüüs: tud-partitsiibi konstruktsioonid



Joonis 39, tud-partitsiibi ja finitse verbiga moodustatavate konstruktsioonide korrespondentsanalüüsi graafik.

Tabel 51 esitab veel kord kokkuvõtlikult kõik finiiitsete verbidega *olema* ja *saama* ning *tud*-partitsiibiga moodustatavad konstruktsioonid eesti murretes.

Tabel 51. Finiitse verbi ja *tud*-partitsiibiga moodustatavad konstruktsioonid eesti murretes.

VORM	TÄHENDUS	NÄIDE
finiitverb	<i>tud</i> -partitsiip	
olema	impersonaal passiiv possessiivne perfekt	<i>ja siis on imestatud seda saege</i> (IDA) <i>siis olid pisised sitsipluused tehtud</i> (SAA) <i>Salmõl oll ka ostõt purgitäis mett</i> (SET)
saama	impersonaal passiiv possessiivne perfekt resultatiivsus	<i>kui magama sai eiddettut</i> (RAN) <i>et suu saab õppettettut</i> (IDA) <i>sääl sai meil tüüttedüss</i> (TAR) <i>et sa rendi maksettud suad</i> (KES)

Finiitne *olema*- ja *saama*-verb on *tud*-partitsiibiga koos esinedes omandanud selgelt abiverbifunktsioonid, s.t et mõlemad verbid osalevad produktiivselt erinevate grammatiliste tähenduste (impersonaal, passiiv) väljendamisel. Siin ilmnevad ka mõningased huvitavad tendentsid ida- ja läänepoolsete murrete rühmas. Konstruktsioonid, mis moodustuvad *saama*-verbiga, on selgelt lääne-eestilise iseloomuga, erinevad possessiivse perfekti konstruktsioonid on iseloomulikud aga idapoolsetele murretele. Possessiivse perfekti puhul ilmneb kõige selgemini ka konstruktsioonide sünonüümia. Materjalis esines näiteks kaks sama tähendusega possessiivse perfekti konstruktsiooni, millest ühel juhul esineb finiiitses positsioonis *saama*-verb (*meil sai kaks oinast ära tapetud*) ja teisel juhul verb *olema* (*Salmel oli mett ostetud*). Possessiivse perfekti konstruktsioon on selgelt idaeestiline; erinevad *saama*-konstruktsioonid on väga sagedased läänepoolsetes murretes. Sellest tulenevalt on *saama*-kujulist possessiivset perfekti materjalis väga vähe, ent *olema*-kujuliste possessiivse perfekti konstruktsioonide esinemissagedus on üsna kõrge ja seda just eriti ida-Võru ja Seto murdes.

9.4 Konstruktsioonide analüüsi kokkuvõte

9.4.1 Erinevused varasematest käsitlustest

Järgnev osa esitab veel kord kokkuvõtvalt kogu konstruktsioonide analüüsi tulemused. Konstruktsioonide tuvastamisel toetusin võimaluse korral varasematele murde- ja kirjakeele uurimustele, kuid analüüsi eesmärk ei olnud materjali sobitamine etteantud raamidesse. Konstruktsioonide lõplik liigitus erineb kohati suuremal, kohati väiksemal määral siiani tehtud kirjakeele konstruktsioonide analüüsides esitatust. Järgnev osa toob välja olulisemad erinevused, millele varasemates uurimustes pole tähelepanu juhitud või mida on käsitletud teistsugusel viisil.

da-infinitiivi ja finiiitse *tulema* konstruktsioonide juures olen eristanud atributiivset ja finaalsset konstruktsiooni sama vormiga moodustuvatest modaal-konstruktsioonidest. Finaalkonstruktsiooni puhul polnud verbid omavahel süntaktiliselt seotud, kuid moodustasid siiski tähendusliku ja vormilise terviku (*noo juba viimasel aeal tullid poodi ka lambid müüja* (KES)). Konstruktsioonil on ka inhoatiivne tähendus ning *tulema* käitub siin tegevuse algusfaasi tähistava faasiverbina. Atributiivses konstruktsioonis esines *da*-infinitiiv grammatilise subjekti täiendina (*siis tulnd mõisast käsk teule minna* (SAA)). Mainitud konstruktsioonid on siiski väga madala esinemissagedusega homonüümid modaalse konstruktsiooniga, mis murretes on *da*-infinitiivi ja finiiitse *tulema*-verbi tavalisim tähendus.

Kõige rohkem senistest käsitlustest erineb ehk finiiitse *olema*-verbi ja *da*-infinitiiviga moodustuvate konstruktsioonide liigitus. Ühendi modaal-tähendus on ilmne nii murretes kui kirjakeeles. Erilisemad konstruktsiooni esinemisjuhud on sellised, kus *olema* esines sünonüümselt *pidama*-verbiga (*kuiss hundõlõ oll üldäq* (SET)). Omamistähendusega konstruktsioonile on iseloomulik mingi omaduse, ressursi või võime väljendamine *da*-infinitiiviga tähistatud tegevuse sooritamiseks (*ei ole omal änam julgust üläl olla* (LÄÄ)). Õeldistäite-konstruktsiooni eristab omadust väljendavast konstruktsioonist *da*-infinitiiviga tegevuse iseloomustamine (*siin on hüä magada nugastada* (RAN)). Erilise konstruktsiooni moodustavad püsiühendid, mis on nii vormilised, tähenduslikud kui pragmaatilised tervikud (*ei miss sialt ikka varastada oli* (KES)). Kohustusliku elemendina lisandub siin küsisõna *mis*, ning konstruktsioonil on selge suhtlusfunktsioon, s.t muudab konstruktsiooni tähenduse eitavaks ehk 'sealt ei olnud midagi varastada'. Omadust väljendava konstruktsiooni moodustavad *da*-infinitiivi ja finiiitse *olema* ühendid, kus verbid pole omavahel eriti seotud. Finiiitse *olema* on sellisel juhul täielikult koopula funktsioonis ning konstruktsiooni keskme moodustab noomen, mida laiendab *da*-infinitiivi vormis verb (*tuu ol siss sitta heittäq vigõl* (VÕR)).

ma-infinitiivi ja finiiitse *saama* konstruktsioonide esinemisjuhtusid oli materjalis vähe, kuid sellest hoolimata olid need uurimismaterjali ühed huvitavamad. Erilisemad konstruktsioonid selles rühmas on mineviku sündmuse tuleviku väljendamisega konstruktsioon (*se sai üks lollukkane olema*

(RAN)) ja püsiühendiks liigitatud vormiline, tähenduslik ja pragmaatiline tervik (nt *kas sa saat minema* (TAR)).

Supiini inessiiv käitub koos finiiitverbiga *olema* sarnaselt kirjakeelega. Huvitavamad sellekujulised konstruktsioonid murretes on omadust väljendav konstruktsioon (*no oless nõus minemas* (KES)), mis tähenduselt on sarnane samanimelise supiini illatiivi ja finiiitse *olema* vormilise konstruktsiooniga. Püsiühendi moodustavad finiiitse *olema* ja supiini inessiivi vormis *olema* ühendid, konstruktsioonil on eksistentsiaallause funktsioon (*neid tarkkasi onn olemas* (LÄÄ)).

tud-partitsiibi ja finiiitsete *olema* ja *saama* verbidega moodustunud konstruktsioonid murretes olid sarnased eesti kirjakeele uurimustes varem esitatutega. Huvitavam oli nende konstruktsioonide puhul geograafilise distributsiooni muster.

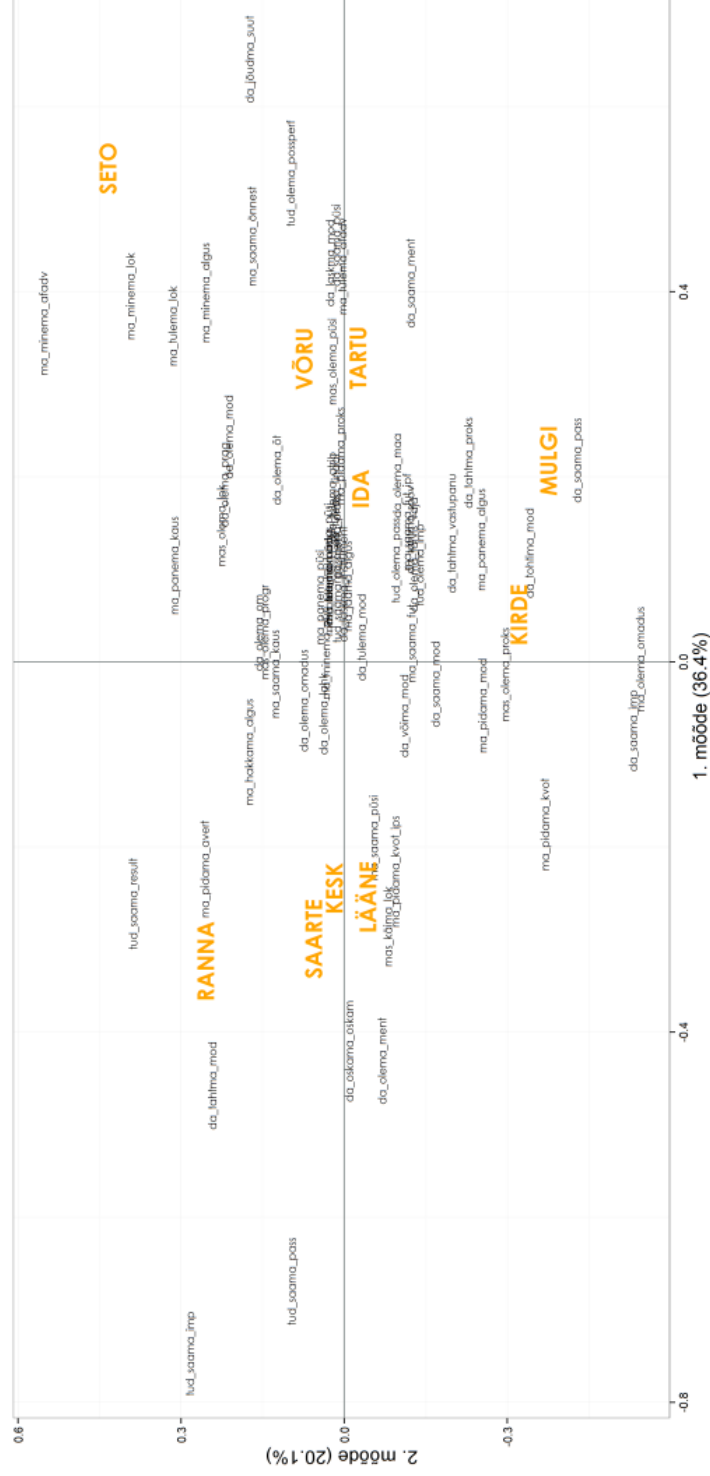
9.4.2 Konstruktsioonide kvantitatiivne kobaranalüüs

Järgnevalt esitan erinevate konstruktsioonide kvantitatiivse kobaranalüüsi tulemused. Siinne analüüs erineb peatükis 7 esitatust, kus kvantitatiivne kobaranalüüs kirjeldas ainult erinevate finiiitse ja infiniitse verbi ühendite varieerumist, pööramata tähelepanu nendega väljendatavatele tähendustele. Käesoleva peatüki analüüs esitab tulemused tuvastatud konstruktsioonide põhjal ning arvestab erinevate konstruktsioonide esinemisjuhtude sagedusi. Eelnevalt analüüsisin erinevate infiniitsete vormidega moodustatud konstruktsioone eraldi. Siinse analüüsi eesmärk on koondada kõigi infiniitse vormiga moodustatud konstruktsioonide ning nende esinemisjuhtude sagedused samasse kvantitatiivsesse kobaranalüüsi.

Joonis 40 esitab kõigi tuvastatud konstruktsioonide ja nende esinemisjuhtude sageduste põhjal korrespondentsanalüüsi visuaalsed tulemused. Joonisel esitatud kaks esimest, kõige olulisemat mõõdet seletavad üle 56% kogu konstruktsiooniandmestiku variatiivsusest. Tulemused on üsna sarnased 7. peatükis esitatud kvantitatiivse kobaranalüüsi tulemustega. Esimese mõõtmepõhjal moodustuvad kaks selgemat gruppi. Esimesse kuuluvad ranna-, saarte, kesk- ja läänemurre, teise grupi moodustavad kirde-, ida-, Mulgi, Tartu, Võru ja Seto murre. Esimene grupp on selgelt homogeensem, teise grupi sisesed murretevahelised erinevused näivad olevat suuremad. Teise mõõtmepõhjal eristuvad selgemalt omavahel sarnasemad ranna- ja Seto murre ning kirde- ja Mulgi murre. Ülejäänud murded selle mõõtmepõhjal ei erine oluliselt n-ö keskmisest murdest ning on ka omavahel üsna sarnased.

Korrespondentsanalys: $V_{\text{fin}} + \text{fin konstruksioonid}$

a konstruksioon a murre



Joonis 40. Finiitse ja infiniitse verbi konstruktsioonide korrespondentsanalüüsi graafik.

Konstruksioonide analüüsi põhjal tekib kontiinum, mille ühes ääres on ranna-, saarte, kesk- ja läänemurre, üleminekuala moodustavad kirde-, ida- ja Mulgi murre ning kontiinumi teise äärde jäävad ülejäänud lõunaeesti murded, kontiinumi teises ääres on Seto murre. Suurem osa konstruksioone on graafikul üsna keskpunkti lähedal, mis tähendab, et nende konstruksioonide kasutuse poolest murrete vahel märkimisväärsed erinevusi pole. Tulemust mõjutavad rohkem *tud*-partitsiibi passiivi- ja impersonaalkonstruksioonid, mis on selgelt iseloomulikud läänepoolsetele murretele, samas kui possessiivse perfekt ja liikumisverbidega moodustuvad konstruksioonid näivad olevat iseloomulikud idapoolsetele murretele, eriti aga Seto murdele. Huvitaval kombel erineb murretes mentaalse verbi konstruksioonide levik. Kui *saama*-verbi ja *tud*-partitsiibi konstruksioonid on eriti iseloomulikud läänepoolsetele murretele, siis sama finitiivse verbi ja *da*-infinitiivi mentaalne konstruksioon näib olevat selgelt iseloomulik idapoolsetele murretele. Mentaalne *da*-infinitiiv ja *olema*-verbi konstruksioon aga on selgelt tavalisem Lääne-Eestis. Erinevad püsiühendid (näiteks *da*-infinitiiv ja finitiivsete verbide *olema*, *saama*, *jõudma* püsiühendid) näivad olevat iseloomulikud idapoolsetele, eriti aga lõunaeesti murderühma murretele.

Konstruksioonikasutuse kobaranalüüsi põhjal ilmnevad üsna selged ida- ja läänepoolsete murrete vahelised erinevused, mis on üsna sarnased kvantitatiivsete kobaranalüüsi tulemustega, mida kirjeldasin peatükis 7. Üksikute infiniitsete vormide konstruksioonide analüüsi põhjal sellised erinevused nii selged ei olnud (vt peatükke 9.1.10, 9.2.11 ja 9.3.3). *da*-infinitiivi konstruksioonide analüüsis olid erinevused suuremad traditsiooniliste lõuna- ja põhjaeesti murrete vahel, kus idamurre oli siiski sarnasem lõunaeesti murretega. Kirdemurre aga erines olulisel määral teistest põhjaeesti murretest. *ma*-infinitiivi konstruksioonide puhul ilmnes pigem ida- ja läänepoolsete murrete erinevus, kus Mulgi ja kirdemurre sarnanesid rohkem lääne-, kesk- ja saarte murdega. Ida- ja rannamurre moodustasid üleminekuala ülejäänud lõunaeesti murderühma murrete ja esimese grupi (Mulgi, kirde-, lääne-, kesk- ja saarte murre) vahel. Sellised eri lingvistiliste tasandite pisut erinevad tulemused on kobaranalüüsi põhimõtete taustal ootuspärased ja demonstreerivad hästi kobaranalüüsi eeliseid. Üksikute nähtuste vaatlemisega on harva võimalik tuvastada murdegrupe. Samuti võivad murded erinevate nähtuste põhjal rühmituda täiesti erineval viisil. Kobaranalüüs aga võimaldab saada terviklikuma pildi üldisemal tasandil ning arvesse on võimalik võtta nii kõigi üksikute keelenähtuste levikut kui ka nende koosmõju.

10 KOKKUVÕTE

Käesolev väitekirj käsitles eesti murrete verbikonstruktsioonide poolautomaatse tuvastamise võimalusi ja väljatöötatud meetodi abil tuvastatud verbikonstruktsioone. Töö põhieesmärgid jagunesid kaheks. Esimene, olulisem eesmärk oli murdekorpusega töötamise metoodika ja verbikonstruktsioonide poolautomaatse tuvastamise võimaluse loomine. Töö püüdis leida parima viisi verbi-konstruktsioonide tuvastamiseks eesti murrete korpusest. Rakendasin tavalist sagedusel põhinevat metoodikat ning võrdlusena kasutasin keerulisema, konstruktsioonigrammatikale tugineva statistilise analüüsi meetodit, kollostruktsioonilist analüüsi. Töö teine olulisem eesmärk oli saada ülevaatlik ja süstemaatiline pilt eesti murretest ühe süntaktilise nähtuse põhjal. Töö teises, kvalitatiivsema iseloomuga konstruktsioonide analüüsi osas uurisingi, milliseid konstruktsioone statistiliste meetoditega tuvastatud verbiühendid eesti murretes moodustavad. Nii töö kvantitatiivsed kui ka kvalitatiivsed tulemused kinnitasid, et uuritud verbiühendid ja -konstruktsioonid varieeruvad murrete lõikes olulisel määral ning selgemalt eristuvad ida- ja läänepoolsete murrete rühm. Esimesse neist kuulusid Seto, Võru, Mulgi, Tartu, ida- ja kirdemurre ning läänepoolsesse rühma saarte, lääne-, kesk- ja rannamurre.

10.1 Töö tulemused varasemate uurimuste kontekstis

Murdesüntaks on siiani veel suhteliselt väheuuritud valdkond ning seda mitte ainult eesti keeleteaduses, vaid ka rahvusvaheliselt. Veelgi vähem on tehtud katseid ebestandardse keele, nagu seda on ka murded, automaattöötuseks. Seega annab käesolev töö märkimisväärse panuse nii murdesüntaksi kui ka ebestandardse keele automaatse töötlemise võimaluste arendamisse.

Süntaktiline varieerumine on raskesti tajutav ja selle uurimiseks vajaliku andmestiku mahu suurus on siin olulisem, kui näiteks häälikulise või leksikaalse varieerumise uurimiseks (Kortmann 2010; Labov 1991: 277). Siinse uurimuse tulemused kinnitavad seda väidet. Praegust murdekorpuse materjali võib pidada küll juba üsna piisavaks kvantitatiivsete süntaktiliste uurimuste jaoks, kuid eripärasemate nähtuste detailsemaks kirjeldamiseks jäi materjali kohati väheks. Uurimuse tulemused on kooskõlas Kortmanni (2010: 842) väitega, et süntaktilistel nähtustel on laiem piirkondlik levik ning harva on need piiratud mingi väiksema geograafilise alaga. See tuli selgelt välja töö 9. peatükis, mis käsitles erinevaid verbikonstruktsioone. Mingi konstruktsiooni piirang ühe murdealaga oli pigem erand kui reegel ning selliste erandite ilmnemist võib pidada tõenäolisemalt materjali puudujäägiks kui geograafiliseks eripäraseks. Murretevahelised erinevused ilmnesisid sagedusandmestikes, mis paljastasid erinevate sagedusastmetega murdealade eelistusmustreid ühe või teise konstruktsiooni suunas.

Kuigi sageduse mõju keelenähtuste selgitamisel on tekitanud palju vaidlusi, võib siinse töö tulemuste põhjal väita, et teatud vormiüksuste koosinemise lihtsageduse põhjal on varieerumismustrite tuvastamine võimalik. Seda illustreerivad põhjalikumad kvantitatiivse analüüsi tulemuste selgitused peatükis 7. Szmrecsanyi (ilmumas (c)) on väitnud, et sagedusandmestik peegeldab selgemini lingvistilist reaalsust kui dialektoloogias traditsiooniliselt rakendatav atlasandmestik ning sagedusmüra andmestikus on lingvistiline reaalsus, millega tuleb tegeleda ja mida peab arvesse võtma. Sellest põhimõttest lähtus ka siinne uurimus. Esiteks ei seadnud ma korpusandmestikule piiranguid, vaid kaasasin analüüsi kõik uuritavad vormid ja vormipaarid, isegi kui tegemist polnud ilmselgelt potentsiaalse konstruktsioonikandidaadiga. Kui murdeatlased esitavad sageli tulemused selle põhjal, kas mingi vorm, sõna vmt on mingil alal olemas, s.t pole vahet, kas see on tuvastatud üks või tuhat korda, siis sageduste kaasamine esitab infot oluliselt detailsemal tasandil. Seda demonstreeris ka käesolev töö. Murretevahelised erinevused süntaktilisel tasandil ilmnevad pigem erinevas sagedusastmes kui mingi vormi esinemise või mitte-esinemise kaudu. Ka eesti murrete netsessiivkonstruktsioonide pilootuuringu tulemused (Lindström & Uibo 2013) illustreerivad kõnekalt, kuidas sagedusandmestik võimaldab saada varieerumisest tunduvalt põhjalikumalt informatsiooni kui traditsiooniline atlasandmestik.

Kvantitatiivses analüüsis rakendasin kollostruktsioonilisi meetodeid (Stefanowitsch & Gries 2003), mis on statistilistes konstruktsioonilistes uurimustes laialdast kasutamist leidnud. Küll on meetodit märkimisväärselt vähem rakendatud varieerumise uurimiseks, kuid üksikud selleteemalised tööd kinnitavad, et meetod täidab ka seda eesmärki üsna edukalt (nt Gries & Mukherjee 2010). Meetodit on kirjeldatud põhjalikumalt peatükis 6.2 ja tulemused on esitatud peatükis 7. Meetod osutus üsna edukaks sagedusmüra eemaldamisel, see tähendab, et meetodi rakendamise tulemusena oli võimalik edasisest analüüsist eemaldada palju mittekonstruktsioone. Teisalt jäid selle tulemusena välja ka paljud n-ö päriskonstruktsioonid, mille tuvastas töö kvalitatiivse iseloomuga konstruktsioonide analüüs peatükis 9. Kollostruktsiooniliste meetodite rakendamiseks tundus aga siiski uurimismaterjal jäävat pisut väikseks, sest paljud konstruktsioonid ei tulnud kvantitatiivses analüüsis esile. Kollostruktsioonilised meetodid sobivad suurepäraselt eeldefineeritud konstruktsioonierinevuste tuvastamiseks (Gries & Stefanowitsch 2004; Stefanowitsch & Gries 2003; Stefanowitsch & Gries 2005 jpt). Kui aga pole täpselt selge, millised konstruktsioonid andmestikus esineda võivad, nii nagu selles töös, jääb ainult kollostruktsiooniliste meetodite rakendamisest väheks. Seda puudujääki korvas töö kvalitatiivsem analüüsiosa, kus uurisin kõiki kvantitatiivse analüüsi tulemusena tuvastatud konstruktsioone ning vaatasin, milline on nende grammatiline ja semantiline funktsioon. Sellisel viisil oli võimalik kvantitatiivsele, morfoloogilisel vormil põhinevale analüüsile (vt ptk 7) lisada kvalitatiivselt tuvastatud konstruktsioonide liigitus ning omakorda nende konstruktsioonide esinemisjuhtude sageduste analüüs. Korrespondentsanalüüs osutus edukaks murrete-

vaheliste erinevuste tuvastamisel, nagu seda on kinnitanud paljud eelnevad uurimused (Cichocki 2006; Keune *et al.* 2005; Wilson 2005).

Kvantitatiivset kobaranalüüsi sellisel kujul nagu selles väitekirjas pole autorile teadaolevalt varem rakendatud. Tööd võib pidada konstruktsioonigrammatikal põhinevaks lähenemiseks kobardialektoloogiale, kus samas kvantitatiivses raamistikus analüüsitakse nii vormi kui ka tähendust. Siinne töö erineb mõneti traditsioonilisest kobaranalüüsist (Nerbonne & Kleiweg 2007; Szmrecsanyi 2013), mille eesmärk on kaasata analüüsi võimalikult palju eri tasandi keelenähtusi. Selle töö analüüs vaatab laialt võttes küll ühte süntaktilist nähtust – kahe seotud verbi paare, kuid samas on töösse kaasatud kõik korpusest automaatselt tuvastatud ühendid ja analüüsimaterjalile pole uurija lisapiiranguid seadnud. Nagu töötulemused kinnitavad, on vaadeldavate ühendite näol tegemist piisavalt heterogeense nähtusega, et sellel kobaranalüüsi rakendada. Niisiis on töö ikkagi kobardialektoloogilise iseloomuga, kus analüüsi kaasatakse kogu korpusest tuvastatud materjal. Lisaks sellele on kvantitatiivselt analüüsitud ka kvalitatiivse analüüsi tulemusena tuvastatud konstruktsioonide esinemisjuhtude sagedusi (vt peatükke 9.1.10, 9.2.11, 9.3.3, 9.4). Sellist analüüsi võiks nimetada kihiliseks kobaranalüüsiks, kus esmalt analüüsitakse ühendite vormi, seejärel nendega tähistatavaid funktsioone, misjärel uuritakse kvantitatiivselt nende funktsioonide esinemiskordi. Seejärel vaadeldakse konstruktsioonide esinemisjuhtusid kõigepealt konstruktsioonide infiniitsete osiste kaupa (peatükid 9.1.10, 9.2.11, 9.3.3) ning lõpuks koondatakse ühte analüüsi erinevate ühendite ja nendega moodustunud konstruktsioonide esinemisjuhtude sagedused (peatükk 9.4.2). Selline nn kihiline analüüs on geograafilise varieerumise uurimisel uudne ja vääriks kindlasti arendamist tulevikus.

Töö kvalitatiivsema suunitlusega osa annab suure panuse seni vähe uuritud eesti murdesüntaksi ning kitsamalt eri verbikonstruktsioonide varieerumise uurimisse. Varem on murretes esinevate verbikonstruktsioonide kohta tehtud üksikuid tähelepanekuid (põhjalikumad neist nt Juhkam 2009; Kask 1984; Keem 1961; Neetar 1988; Potseps 1956; Uibo *et al.* 2013). Siinne töö lisab uut teadmist juba kirjeldatud ja seni kirjeldamata verbikonstruktsioonide kohta kõigis eesti murretes. Töö loob ühe süntaktilise varieerumise süstemaatilise ja kõiki murdeid hõlmava pildi.

Kollostruktsioonilisi meetodeid rakendasin 9. peatükis Muischneki ja Sakhai (2009; 2010) eeskujul ka konstruktsioone moodustavate verbiühendite produktiivsuse mõõtmiseks. Siin ei olnud oluline statistilise seose tugevuse väärtus, vaid väärtuste jaotumine ühendite lõikes. Mida tugevamini finiidne ja infiniitne verb olid statistiliselt seotud, seda püsiühendilaadsemaks neid võis pidada. Kui väärtused polnud eriti kõrged ja jagunesid verbide lõikes ühtlasemalt, võiks eeldada, et tegemist on üsna produktiivsete üksustega. Tulemused näitasid, et produktiivsuse mõõtmiseks tulevikus peaks kindlasti rakendama nii väikse materjali peal pisut keerulisemaid ja kvalitatiivsema suunitlusega meetodeid. Produktiivsuse erinevused eriti selgelt ei ilmnunud, küll aga tulid erinevates analüüsides hästi esile ühendid, mida eelnevas kirjanduses ja ka siinses analüüsis püsiühenditena on esitatud. See julgustab tulevikus meetodit kindlasti

selleks otstarbeks rohkem rakendama ja saadud tulemusi võib pidada usaldusväärseteks.

Töö kõige huvitavam tulemus eesti dialektoloogia jaoks on kindlasti suure erinevuse ilmnenemine ida- ja läänepoolsete murrete vahel. Kui siiani on kõige olulisemaks peetud põhja- ja lõunaeeesti murrete vahelisi häälikulisi ja morfoloogilisi erinevusi (Pajusalu *et al.* 2009), siis selle töö tulemuste põhjal on süntaktiliselt tunduvalt olulisem ida- ja läänepoolsete murrete vaheline erinevus. Tulemus on huvitav sellegi poolest, et ka varem on erinevate keelenähtuste uurimisel jõutud ida- ja läänepoolsete murrete erinevust kinnitavate järeldusteni. Näiteks Lippus ja Pajusalu on (2008) uurinud murdetausta mõju eesti keele väljete tajumisele ning leidnud, et piirkondlikud erinevused on kõige suuremad just ida- ja läänepoolsete murrete vahel. Läänepoolsetelt murdealadelt pärit olevad inimesed on märkimisväärselt tundlikumad põhitooni suhtes välteerinevuste tajumisel, samas kui idapoolsete murrete taustaga informantide puhul polnud autorite väitel põhitoonil vältetajumisele eriti märgatavat mõju. Selliste erinevuste ilmnenemist põhjendavad Lippus ja Pajusalu keelekontaktidega. Nende esitatud murretevaheline piir on üsna sarnane siinses töös tuvastatud ida- ja läänepoolsete murrete vahelise piiriga. Varasem kõiki murdeid hõlmav ühendverbide uurimus (Uiboaed 2010a) jõuab samuti sarnaste tulemusteni, kus idapoolsed murded kasutavad vähem ühendverbe kui läänepoolsed murded. Pall (1982: 247) on viidanud idamurde häälikulisele ja morfoloogilisele sarnasusele lõunaeeesti murrete ja kirderanniku murdega. Lindström jt (2009) on uurinud 1. isikule viitamise varieerumist eesti murretes ning autorid osutavad samuti ida- ja läänepoolsete murrete vahelisele erinevusele. Niisiis ei saa siinse töö tulemusi kindlasti pidada juhuslikeks või uuritava nähtuse eripärast tulenevateks.

10.2 Töö tulemuste olulisus

Väitekirj annab suure panuse murdesüntaksi uurimisvaldkonda laiemalt, kuna tööga on lisandunud andmeid nii eesti murretest kui ka teadmisi süntaktilisest varieerumisest ja selle uurimise problemaatikast. Lisaks panustab töö suurel määral metodoloogia arendamisse ning mittestandardse keele automaattöötluse võimaluste avastamisse.

Uurimuse üheks eesmärgiks oli püüd ühendada tänapäevased keeleteaduslikud kirjeldusvahendid ja dialektoloogia. Konstruktioonigrammatika pole siiani olnud dialektoloogias üldlevinud lähtekohaks, kuid viimastel aastatel on üksikud tööd selles raamistikus juba ilmunud (nt Hollmann & Siewierska 2007; Siewierska & Hollmann 2007). Lähenemisviis on suhteliselt paindlik, võimaldades kirjeldusest välja arvata erinevaid elemente, kui nende kohta puudub info (Fried 2008), lähenemisviis võimaldab luua uusi, seni kirjeldamata kategooriad, kui need tulevad analüüsis esile (Sahkai 2011), ning võimaldab käsitleda kõiki nähtusi nii keelevariandi- kui ka konstruktioonikeskselt (Croft 2001). Varieerumise uurimisel on kõigi nende aspektide rakendamisvõimalus hädavajalik, seega sobib konstruktioonigrammatika suurepäraselt just heterogeense

varieeruvusmaterjaliga töötamiseks. Siinse töö seisukohalt oli vajalik käsitleda konstruktsioone väga skemaatilisel tasandil ja seda eesmärki täitis konstruktsioonigrammatika hästi. Konstruktsioonigrammatika sobivust murrete uurimiseks on analüüsinud Östman ja Trousdale (2013), kes möönavad, et kuigi lähenemisviisil on küll veel palju lahendamata probleeme, sobib raamistik varieerumise uurimiseks siiski üsna hästi.

Töö annab omapoolse panuse seni keeleteaduses palju vaidlusi tekitanud sageduse mõju seletamisse. Tulemused kinnitavad, et keelenähtuse koosseisemissagedus on tähenduslik ja võib uuritava nähtuse kohta anda üsna palju olulist lingvistilist ja metalingvistilist informatsiooni. Sageduse põhjal on võimalik saada üsna hea ülevaade analüütilise keelenähtuse varieerumisest. Huvitava kõrvalsaadusena tõid töö tulemused selles osas esile homonüümsete konstruktsioonide tuvastamisprobleematika. Paljud töös käsitletud konstruktsioonidest tulid esile vaid seetõttu, et neil on olemas sagedasti esinevad homonüümid. Ühelt poolt oli sel viisil võimalik tuvastada madala esinemissagedusega konstruktsioone, teisalt jäid tõenäoliselt välja mõned konstruktsioonid, millel sagedased homonüümid puudusid. Kindlasti on see üks koht, kus metoodikat konstruktsioonide tuvastamisel tulevikus edasi arendama peaks, samas aga nõuab see rohkem eelteadmisi selle kohta, millised potentsiaalsed konstruktsioonid murretes veel olemas võivad olla.

Uurimuse peaeesmärk oli kvantitatiivsete meetodite arendamine ja testimine eesti murrete korpuse materjaliga töötamiseks. Kvalitatiivse suunitlusega konstruktsioonide analüüsi osa oli kvantitatiivseid tulemusi täiendava iseloomuga. Kui kvantitatiivsed tulemused tõid esile üsna selgelt murretevahelistele potentsiaalsetele konstruktsioonide kasutuserinevustele, siis kvalitatiivsema iseloomuga analüüs võimaldas neid erinevusi vaadelda detailsemalt. Kobaranalüüsile iseloomulikult ei ole erinevate konstruktsioonide varieerumismustrid alati samasugused, kuid tulemuste liitmine annab varieerumisest laiemat ja üldistavat pildi. Nii kvantitatiivse kui kvalitatiivse kobaranalüüsi tulemused olid sarnased, mis tõestab, et kvantitatiivsete meetoditega on võimalik tuvastada murretevahelisi erinevusi üsna usaldusväärselt isegi juhul, kui kvalitatiivselt pole selge, millised on täpselt erinevused konstruktsioonide tähendustes.

Verbiühendite ja nendega moodustatud konstruktsioonide kasutus varieerub murdeti suurel määral. Läänepoolsed murded näivad kasutavat erinevaid verbiühendeid rohkem. Lisaks näitab töö eesti murrete materjali põhjal, kuidas uue lingvistilise tasandi – süntaksi – kaasamine võib esitada märkimisväärselt erineva murdejaotuse. See on oluline tulemus, mida peaks tulevikus erinevate süntaktiliste nähtuste põhjal edasi uurima.

Siinse töö tulemused ja väljatöötatud metoodika on aluseks 2013. aastal Tartu ülikoolis alanud murdesüntaksi projektile EstDiaSyn. Seega on töö tulemused kohe rakendatavad pikaajalise ja põhjaliku projekti teenistusse. Uurimuse käigus väljatöötatud andmestiku kogumise ja korrastamise skriptid ning visualiseerimisvõimalused on edaspidi kasutatavad projekti töös. Siinne uurimus on andnud suure panuse projekti eeltöösse ja -teadmistesse.

10.3 Töö kitsaskohad ja edasised uurimissuunad

Töötulemused pakuvad huvi ja on uudsed, kuid sellest hoolimata on töös probleeme, mis jäid (osaliselt) lahendamata ja mis mõjutavad töö tulemusi. Kuigi peatükis 5 kirjeldatud verbiühendite tuvastamise meetod oli piisav, täitmaks selle töö eesmärke, on seda võimalik olulisel määral parandada. Probleemaatiline on süntaksianalüsaatori toimetulek vigadega ja tulevikus võiks üritada arendada puhtalt morfoloogial põhinevat tuvastamisalgoritmi. Kuna murdekorpuse on morfoloogiliselt käsitsi märgendatud, on selles esinevate süstemaatiliste vigade osakaal marginaalne; ning morfoloogilise märgenduse kvaliteet on oluliselt parem sellest, mida süntaksianalüsaator murdetekstide analüüsimisel võimaldab esitada.

Keeleuurimisel, eriti just murrete ja vana kirjakeele uurimisel, tuleb sageli ette, et keelematerjali hulk on piiratud ning et seda pole praktilistel põhjustel võimalik juurde koguda. Analüüsida tuleb materjali, mis on olemas. Sellised piirangud kehtivad ka eesti murrete uurimisel. Sagedusandmete normaliseerimine oli vajalik, et võrrelda erineva sõnade arvuga murrete materjali, ja selline toimimisviis on keeleteaduslike andmestikega töötamisel levinud praktika, kuid andmete normaliseerimise probleem on siiski suur. Seda enam on üllatav, et sellega pole keeleteaduses eriti laialdaselt tegeletud. Sageli lahendatakse olukord suhteliste sageduste arvutamisega või mingi arvulise baasi peal normaliseerimisega (nähtus 1000, 10 000 jne sõna kohta). Käesolev uurimus normaliseeris sagedusi keskmise korpuse suuruse põhjal (täpsemalt peatükis 7.1), mille eesmärgiks oli võtta arvesse nii kogu korpuse kui ka erinevate alamkorpuste suurust. Töö autor ei väida, et üks või teine normaliseerimisviis on oluliselt parem, kuna seda pole siiani eriti põhjalikult uuritud. Normaliseerimine seab andmestikule teatavad eeldused, mis ei pruugi olla lingvistiliselt motiveeritud. Kas on olemas mingi lingvistiline alus, et keelenähtused esinevad erinevatel aladel võrdselt ja suhtarvude leidmine on põhjendatud? Kas peaksime üldse sagedusandmeid normaliseerima või võime hoopis eeldada, et korpuse erinevast suurusest hoolimata tulevad sagedasemad lingvistilised tendentsid ka erineva suurusega materjali peal reaalsust peegeldavates proportsioonides esile? Siinses töös rakendatud korpuse keskmise arvestamine on eelduste kohaselt pisut parem lahendus kui väga abstraktne (nt 1000 sõna kohta) normaliseerimine, kuid kindlasti ei pruugi see olla parim viis. Sagedusandmete normaliseerimine on kindlasti probleem, millega tuleb tulevikus põhjalikumalt tegeleda ning erinevate normaliseerimisviiside mõjusid andmestikule ja kvantitatiivse analüüsi tulemustele peaks detailsemalt uurima.

Siinse töö konstruktsioonide analüüs oli väga skemaatiline, kuid osutas sellest hoolimata huvitavatele tendentsidele. Kindlasti peaks tulevikus erinevaid konstruktsioone ja nende kasutuspiiranguid põhjalikumalt analüüsima, et saada erinevate murrete kohta oluliselt detailsemat infot. Sellise varieerumisandmestiku ühtsete ja põhjalikumate analüüsipõhimõtete väljatöötamine jääb kindlasti suuremaks eesmärgiks tulevikus.

10.4 Lõpetuseks

Käesolev doktoritöö uuris verbikonstruktsioonide varieerumist eesti murretes ning laiendas meie teadmisi seni vähe uuritud murdesüntaksist ja eesti murrete süntaktilisest varieerumisest. Töö olulisemateks eesmärkideks olid erinevate statistiliste meetodite abil süntaktilise varieeruvuse tuvastamine ning sellise varieerumise iseloomu analüüs. Töö tulemused osutasid märkimisväärsele murretevahelisele erinevusele. Saadud tulemustel põhinev murdeliigendus ei sarnanenud siiani häälikulistel, morfoloogilistel ja leksikaalsetel andmetel tugineva murdejaotustega. Suuremad olid süntaktilised erinevused ida- ja läänepoolsete murrete vahel ning traditsioonilised põhja- ja lõunaeesti murrete rühmad erinevates analüüsides nii selgelt välja ei tulnud. Töö tulemused kinnitavad süntaktilise mõõtmise olulisust murdejaotuste esitamisel, aga ka seda, et nii uued statistilised meetodid kui ka uue lingvistilise tasandi kaasamine analüüsi pakub uudset ja huvitavat lingvistilist informatsiooni. Töö tulemused laiendasid teadmisi nii murretevaheliste erinevuste kui ka erinevate konstruktsioonide kohta, mida seni pole ka kirjakeeles eriti põhjalikult uuritud. Töö osutas kvalitatiivsete ja kvantitatiivsete meetodite kombineerimise olulisusele ning näitas, kuidas mõlemad meetodid võivad teineteist täiendades esile tuua seni tuvastamata varieerumismustreid ning keelelisi nüansse.

KIRJANDUS

- Abdi, Hervé & Lynne J. Williams. 2010. Correspondence Analysis. Neil Salkind (ed.), *Encyclopedia of Research Design*, 267–278. Thousand Oaks, CA: Sage.
- Alvre, Paul. 1993. Mõningaist passiivi erijoontest. *Keel ja Kirjandus* 7. 404–409.
- Anderson, Gregory D. S. 2006. *Auxiliary verb constructions*. New York: Oxford University Press.
- Anderwald, Liselotte & Benedikt Szmrecsanyi. 2008. Corpus linguistics and dialectology. Merja Kytö & Anke Lüdeling (eds.), *Corpus Linguistics. An international handbook*, 1126–1139. Berlin/New York: Walter de Gruyter.
- Arppe, Antti, Gaëtanelle Gilquin, Dylan Glynn, Martin Hilpert & Arne Zeschel. 2010. Cognitive Corpus Linguistics: five points of debate on current theory and methodology. *Corpora* 5(1). 1–27.
- Auer, Peter. 2004. Non-standard evidence in syntactic typology – Methodological remarks on the use of dialect data vs spoken language data. Bernd Kortmann (ed.), *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, vol. 153, 69–92. Berlin: Mouton de Gruyter.
- Baayen, R. Harald. 1994. Derivational productivity and text typology. *Journal of Quantitative Linguistics* 1(1). 16–34.
- Baayen, R. Harald. 2008. *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. Harald. 2009. Corpus linguistics in morphology: morphological productivity. Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics. An international handbook*, 900–919. Berlin: Walter de Gruyter.
- Baayen, R. Harald. 2011. Corpus linguistics and naive discriminative learning. *Revista Brasileira de Linguística Aplicada* 11(2). 295–328.
- Barbiers, Sjef & Hans Bennis. 2007. The Syntactic Atlas of the Dutch Dialects. *Nordlyd* 34(1). 53–72.
- Barbiers, Sjef, Leonie Cornips & Jan Pieter Kunst. 2007. The Syntactic Atlas of the Dutch Dialects: A corpus of elicited speech and text as an on-line dynamic atlas. *Creating and digitizing language corpora* 1. 54–90.
- Barðdal, Jóhanna. 2008. *Productivity: evidence from Case and Argument Structure in Icelandic*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Barlow, Michael & Suzanne Kemmer (eds.). 2000. *Usage Based Models of Language*. Stanford: CSLI Publications.
- Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.
- Benzécri, Jean-Paul. 1973. *L'analyse des données*. Paris: Dunod.
- Berthele, Raphael. 2004. The typology of motion and posture verbs: A variationist account. Bernd Kortmann (ed.), *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, vol. 153, 93–126. Berlin: Mouton de Gruyter.
- Biber, Douglas. 1995. *Dimensions of register variation: a cross-linguistic comparison*. Cambridge: Cambridge University Press.
- Biber, Douglas & Susan Conrad. 2009. *Register, genre, and style*. Cambridge: Cambridge University Press.
- Bisang, Walter. 2004. Dialectology and typology – An integrative perspective. Bernd Kortmann (ed.), *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, vol. 153, 1–10. Berlin: Mouton de Gruyter.

- Bisang, Walter. 2007. Categories that make finiteness: discreteness from a functional perspective and some of its repercussions. Irina Nikolaeva (ed.), *Finiteness: theoretical and empirical foundations*, 115–137. Oxford: Oxford University Press.
- Boas, Hans C. 2013. Cognitive Construction Grammar. Graham Trousdale & Thomas Hoffmann (eds.), *The Oxford Handbook of Construction Grammar*, 233–252. Oxford: Oxford University Press.
- Bod, Rens, Jennifer Hay & Stefanie Jannedy. 2003. Introduction. Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic Linguistics*, 1–10. Cambridge/Massachusetts/London: The MIT Press.
- Bolinger, Dwight. 1980. Wanna and the Gradience of Auxiliaries. Gunter Brettschneider & Christian Lehmann (eds.), *Wege zur Universalien Forschung: Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*, 292–299. (Tübinger Beiträge Zur Linguistik 145). Tübingen: Narr.
- Bybee, Joan. 2007. *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- Bybee, Joan. 2010. *Language, usage, and cognition*. Cambridge: Cambridge University Press.
- Bybee, Joan & David Eddington. 2006. A usage-based approach to Spanish verbs of “becoming.” *Language* 82(2). 323–355.
- Bybee, Joan & Paul J. Hopper. 2001. *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins Publishing.
- Chambers, John Kenneth & Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: The MIT press.
- Cichocki, Wladyslaw. 2006. Geographic Variation in Acadian French /r/: What Can Correspondence Analysis Contribute toward Explanation? *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing* 21(4). 529–541.
- Cristofaro, Sonia. 2007. Deconstructing categories: finiteness in a functional-typological perspective. Irina Nikolaeva (ed.), *Finiteness: theoretical and empirical foundations*, 91–114. Oxford: Oxford University Press.
- Croft, William. 2001. *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William. 2003. Lexical rules vs. Constructions: A False Dichotomy. Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in Language*, vol. 243, 49–68. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Croft, William & David Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Cronbach, Lee J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 3(16). 297–334.
- Divjak, Dagmar & Stefan Th. Gries. 2012. *Frequency effects in language representation*. Berlin: De Gruyter Mouton.
- EMK. 2011. *Eesti murrete korpus*. <http://www.murre.ut.ee/murdekorpus/>.
- Erelt, Mati. 1984. *da-infinitiivi süntaktilised funktsioonid tänapäeva eesti keeles*. Soome-eeesti kontrastiivseminar (Helsingi, 17.–20.10.84). Preprint KKI-24. Tallinn: Eesti NSV TA.
- Erelt, Mati. 1987. *Sekundaartarindid eesti keeles*. Preprint KKI-50. Tallinn: Eesti NSV TA.

- Erelt, Mati. 2001. Some notes on the grammaticalization of the verb *pidama* in Estonian. Mati Erelt (ed.), 7–25. *Estonian: Typological Studies V*. Publications of the Department of Estonian of the University of Tartu 18. Tartu: University of Tartu.
- Erelt, Mati (ed.). 2003. *Estonian language*. Tallinn: Estonian Academy Publishers.
- Erelt, Mati. 2009. Prospektiiv ja retrospektiiv eesti keeles. *Emakeele Seltsi aastaraamat 2008* (54). 44–53.
- Erelt, Mati. 2013. *Eesti keele lauseõpetus: Sissejuhatus. Õeldis*. Helle Metslang (ed.), Tartu Ülikooli eesti keele osakonna preprintid 4. Tartu: Bookmill.
- Erelt, Mati, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael & Silvi Vare. 1993. *Eesti keele grammatika II. Süntaks. Lisa: kiri*. Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut.
- Erelt, Mati, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael & Silvi Vare. 1995. *Eesti keele grammatika I. Morfoloogia. Sõnamoodustus*. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut.
- Erelt, Mati & Helle Metslang. 2009. Some Notes on proximative and avertive in Estonian. *Linguistica Uralica* 45(3). 178–191.
- Everitt, Brian & Torsten Hothorn. 2011. *An Introduction to Applied Multivariate Analysis with R*. New York: Springer.
- Everitt, Brian S., Sabine Landau, Morven Leese & Daniel Stahl. 2011. *Cluster Analysis*. 5th revised edition. Chichester: Wiley-Blackwell.
- Evert, Stefan. 2004. The Statistics of Word Cooccurrences: Word Pairs and Collocations. PhD Dissertation, Institut für maschinelle Sprachverarbeitung, University of Stuttgart. Published in 2005, URN urn:nbn:de:bsz:93-opus-23714. Available from <http://www.collocations.de/phd.html>.
- Evert, Stefan. 2008. Corpora and collocations. Merja Kytö & Anke Lüdeling (eds.). *Corpus Linguistics. An international handbook*, 1212–1248 Berlin/New York: Walter de Gruyter.
- Evert, Stefan & Brigitte Krenn. 2001. Can we do better than frequency? A case study on extracting PP-verb collocations. *Proceedings of the ACL Workshop on Collocations*, 39–46. Toulouse, France.
- Fillmore, Charles J. 1968. The case for case. Emmon Bach & Robert T. Harms (eds.), *Universals in linguistic theory*, 1–88. New York: Holt, Rinehart & Winston.
- Fillmore, Charles J. 1982. Frame Semantics. *Linguistics in the Morning Calm*, 111–137. (The Linguistic Society of Korea). Seoul: Hanshin.
- Fillmore, Charles J. 1988. The mechanisms of “Construction Grammar.” *Proceedings of the annual meeting of the Berkeley Linguistics Society* 14. 35–55.
- Fillmore, Charles J. 1989. Grammatical Construction Theory and the familiar dichotomies. Rainer Dietrich & Carl Friedrich Graumann (eds.), *Language processing in social context*, 17–38. Amsterdam: Elsevier Science Publisher.
- Fillmore, Charles J., Paul Kay & Catherine O’Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of Let alone. *Language* 64(3). 501–538.
- Filppula, Markku, Juhani Klemola, Marjatta Palander & Esa Penttilä. 2005. Introduction. Markku Filppula, Juhani Klemola, Marjatta Palander & Esa Penttilä (eds.), *Dialects across Borders*, vii–xii. Amsterdam, Netherlands: John Benjamins Publishing Company.
- Fried, Mirjam. 2008. Constructions and Constructs: Mapping a Shift between Predication and Attribution. Alexander Bergs & Gabriele Diewald (eds.), *Constructions and Language Change*, 47–79. Berlin, Germany: Mouton de Gruyter.

- Fried, Mirjam. 2010. Introduction: from instances of change to explanations of change. Mirjam Fried, Jan-Ola Östman & Verschuieren (eds.), *Variation and change: pragmatic perspective*, 1–16. Amsterdam: John Benjamins Publishing Company.
- Fried, Mirjam. 2010. Construction Grammar. *Handbook of syntax*. Berlin: Walter de Gruyter.
- Fried, Mirjam & Jan-Ola Östman. 2004. Construction Grammar: A Thumbnail Sketch. Mirjam Fried & Jan-Ola Östman (eds.), *Construction Grammar in a Cross-Language Perspective*, 11–86. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Genetz, Arvid. 1900. Minkätähden jäädä sana vaatii tulosijoja? *Virittäjä: Journal of the Kotikielen Seura (Society for the Study of Finnish)* 1. 27–28.
- Goba, Karlis. 2006. Mark. Murdekorpuse morfoloogilise märgendamise abiprogramm. Tartu Ülikool.
- Goebel, Hans. 1982. *Dialektometrie: Prinzipien und Methoden des Einsatzes der Numerischen Taxonomie im Bereich der Dialektgeographie*. Wien: Österreichische Akademie der Wissenschaften.
- Goebel, Hans. 2006. Recent Advances in Salzburg Dialectometry. *Literary and Linguistic Computing: Journal of the Association for Literary and Linguistic Computing* 21(4). 411–435.
- Goebel, Hans, Siegfried Selberherr & Wolf-Dieter Rase. 1984. *Dialektometrische Studien: anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*. Tübingen: Niemeyer.
- Goldberg, Adele E. 1995. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago press.
- Goldberg, Adele E. 2006. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
- Greenacre, Michael. 2007. *Correspondence analysis in practice*. 2nd ed. Boca Raton Fla.: CRC Press.
- Gries, Stefan Th. 2006. Exploring Variability within and between Corpora: Some Methodological Considerations. *Corpora: Corpus-Based Language Learning, Language Processing and Linguistics* 1(2). 109–151.
- Gries, Stefan Th. 2007. *Coll.analysis 3.2. A program for R for Windows 2.x*.
- Gries, Stefan Th. & Dagmar Divjak. 2012. *Frequency effects in language learning and processing*. Berlin: De Gruyter Mouton.
- Gries, Stefan Th., Beate Hampe & Doris Schönefeld. 2005. Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16(4). 635–676.
- Gries, Stefan Th. & Joybrato Mukherjee. 2010. Lexical gravity across varieties of English: An ICE-based study of n-grams in Asian Englishes. *International Journal of Corpus Linguistics* 15(4). 520–548.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis A corpus-based perspective on “alternations.” *International Journal of Corpus Linguistics* 9(1). 97–129.
- Grieve, Jack. 2009. *A corpus-based regional dialect survey of grammatical variation in written Standard American English*. San Francisco AZ: Northern Arizona University PhD Dissertation. Manuscript.
- Habicht, Külli, Pille Penjam & Ilona Tragel. 2010. Kas tahtma tahab abiverbiks? *Eesti ja soome-ugri keeleteaduse ajakiri ESUKA / Journal of Estonian and Finno-Ugric Linguistics JEF* 2. 115–146.

- Haspelmath, Martin. 1989. From Purposive to Infinitive – a Universal Path of Grammaticization. *Folia Linguistica Historica* 10(1–2). 287–310.
- Haspelmath, Martin. 1994. Passive Participles across Languages. Barbara Fox & Paul J. Hopper (eds.), *Voice: Form and Function*, 151–177. Amsterdam: John Benjamins Publishing Company.
- Heeringa, Wilbert Jan. 2004. *Measuring dialect pronunciation differences using Levenshtein distance*. Groningen: Rijksuniversiteit Groningen. PhD dissertation.
- Heine, Bernd. 1993. *Auxiliaries: cognitive forces and grammaticalization*. New York/Oxford: Oxford University Press.
- Helasvuo, Marja-Liisa & Maria Vilkuna. 2008. Impersonal is personal: Finnish perspectives. *Transactions of the Philological Society* 106(2). 216–245.
- Hernández, Nuria. 2010. *Personal pronouns in the dialects of England. A corpus study of grammatical variation in spontaneous speech*. Freiburg: Albert-Ludwigs-Universität Freiburg. PhD dissertation.
- Hilpert, Martin. 2006. Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2(2). 243–256.
- Hollmann, Willem B. & Anna Siewierska. 2007. A construction grammar account of possessive constructions in Lancashire dialect: Some advantages and challenges. *English Language and Linguistics* 11(2). 407–424.
- Hopper, Paul J. & Elizabeth Closs Traugott. 2003. *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press.
- Jaakola, Minna. 2003. Kokijarakenteen kontrastointia. *VIRSU. Viro ja suomi – kohdekielet kontrastissa* (13). Lähivertailuja. 167–177.
- Jokela, Hanna. 2012. *Nollapersoonalause suomessa ja virossa: tutkimus kirjoitetun kielen aineistosta*. Turku: Turun yliopiston julkaisuja.
- Juhkam, Evi. 1983. Rootsipärane objekt eesti murretes. *Keel ja Kirjandus* 3. 122–125.
- Juhkam, Evi. 1998. *Eesti-rootsi murdekontaktid*. Tallinn: Tallinna Pedagoogikaülikooli filoloogiateaduskond. Käsikirjaline magistritöö.
- Juhkam, Evi. 2009. Eestirootsi refleksiivverbi mõjust eesti murrete verbikasutusele. Marja Kallasmaa & Vilja Oja (toim.), *Kodu keel ja keele kodu. Pühendusteos Helmi Neetarile 75. sünnipäevaks 29. jaanuaril 2009*, 57–69. Tallinn: Eesti Keele Sihtasutus.
- Juhkam, Evi. 2012. *Harju-Madise murrak*. Mari-Liis Kalvik & Helmi Neetar (toim.). Tallinn: Eesti Keele Sihtasutus.
- Juhkam, Evi & Aldi Sepp. 2000. *Eesti mured VIII. Läänemurde tekstid*. Anu Haak (toim.). Tallinn: Eesti Keele Instituut.
- Kalmus, Mervi. 2009. *Isiku ühildumine eesti murretes*. Tartu Ülikooli eesti keele õppe-ool. Käsikirjaline magistritöö.
- Kasik, Reet. 2001. Analytic causatives in Estonian. Mati Erelt (ed.), 77–122. *Estonian: Typological Studies V*. Publications of the Department of Estonian of the University of Tartu 18. Tartu: University of Tartu.
- Kask, Arnold. 1984. *Eesti mured ja kirjakeel*. Tallinn: Valgus.
- Kay, Paul & Charles J. Fillmore. 1999. Grammatical Constructions and linguistic generalizations: The What's X Doing Y? construction. *Language* 75(1). 1–33.
- Keem, Hella. 1961. Vanapäraseid objektireksioone Nõo murrakus. *Keel ja Kirjandus* 9. 544–545.
- Keem, Hella. 1970. *Eesti mured III. Tartu murde tekstid*. Mari Must, Salme Nigol & Mari Must (toim.). Tallinn: Valgus.
- Keem, Hella. 1973. Võru murde vahekorra teiste lõunaestli murretega. *Keel ja Kirjandus* 10. 605–611.

- Keem, Hella & Inge Käsi. 2002. *Eesti murded VI. Võru murde tekstid*. Aldi Sepp (toim.) Tallinn: Eesti Keele instituut.
- Kehayov, Petar & Reeli Torn-Leesik. 2009. Modal verbs in Balto-Finnic. Björn Hansen & Ferdinand Haan de (eds.), *Modals in the Languages of Europe*, vol. 44, 363–401. Berlin/New York: Mouton de Gruyter.
- Kemmer, Suzanne & Michael Barlow. 2000. Introduction: a usage-based conception of language. Michael Barlow & Suzanne Kemmer (eds.), *Usage Based Models of Language*, vii–xxviii. Stanford: CSLI Publications.
- Keune, Karen, Mirjam Ernestus, Roeland van Hout & R. Harald Baayen. 2005. Variation in Dutch: From Written MOGELIJK to Spoken MOK. *Corpus Linguistics and Linguistic Theory* 1(2). 183–223.
- Klaus, Anneliis. 2009. *Eitus eesti murretes*. Tartu: Tartu Ülikooli eesti keele õppetool. Käsikirjaline magistritöö.
- Kleiweg, Peter, Therese Leinonen, Charlotte Gooskens, Rinke Colen & John Nerbonne. 2011. *Gabmap – web application that visualizes dialect variations*.
- Koit, Enn. 1963. Eitus saarte murdes. *Emakeele Seltsi toimetised. Nonaginta: Johannes Voldemar Veski 90. sünnipäevaks 27. juunil 1963* (6). 139–147.
- Koptjevskaja-Tamm, Maria. 1999. Finiteness. Keith Brown & Jim Miller (eds.). *Concise Encyclopedia of Grammatical Categories*. 146–149. Oxford: Elsevier.
- Kortmann, Bernd. 2004. Introduction. Bernd Kortmann (ed.), *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, vol. 153, 1–10. Berlin: Mouton de Gruyter.
- Kortmann, Bernd. 2010. Areal variation in syntax. Peter Auer & Jürgen E. Schmidt (eds.), 837–864. *Language and Space: Theories and Methods: An international handbook of linguistic variation*. Berlin/New York: De Gruyter Mouton.
- Krikmann, Arvo & Karl Pajusalu. 2000. Kus on keskmurde keskpunkt? Jüri Viikberg (ed.), *Inter dialectos nominaque. Pühendusteos Mari Mustale 11. novembril 2000*, 131–172. Tallinn: Eesti Keele Sihtasutus.
- Kunst, Jan Pieter & Franca Wesseling. 2011. The Edisyn search engine. *Oslo Studies in Language* 3(2). 63–74.
- Kuteva, Tania. 2001. *Auxiliation: an enquiry into the nature of grammaticalization*. Oxford: Oxford University Press.
- Labov, William. 1991. The boundaries of a grammar: inter-dialectal reactions to positive any more. Peter Trudgill & John Kenneth Chambers (eds.). *Dialects of English: studies in grammatical variation*. 273–288.
- Lakoff, George. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago (IL): University of Chicago Press.
- Langacker, Ronald Wayne. 1987. *Foundations of cognitive grammar*. Stanford, California: Stanford University Press.
- Lebart, Ludovic, André Salem & Lisette Berry. 1998. *Exploring textual data*. Dordrecht: Kluwer Academic Publishers.
- Lindström, Liina. 1997. *Võru murde suulise kõne erijooni*. Tartu: Tartu Ülikooli eesti keele õppetool. Käsikirjaline bakalaureusetöö.
- Lindström, Liina. 2001. Eesti murrete korpuse iseloomustus argivestlusega võrrelduna. Reet Kasik (ed.), *Keele kannul. Pühendusteos Mati Ereli 60. sünnipäevaks 12. märtsil 2001*, 212–221. Tartu Ülikooli eesti keele õppetooli toimetised 17. Tartu: Tartu Ülikooli Kirjastus.
- Lindström, Liina. 2005. *Finiitverbi asend lauses. Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles*. Tartu: Tartu Ülikool.

- Lindström, Liina. 2009. Ülevaade eesti murrete korpusest. <http://www.murre.ut.ee/assets/EMK.PDF> (7. jaanuar 2012).
- Lindström, Liina. 2010. Kõnelejale ja kuulajale viitamise vältimise strateegiaid eesti keeles. *Emakeele Seltsi aastaraamat 2009* (55). 88–118.
- Lindström, Liina, Mervi Kalmus, Anneliis Klaus, Liisi Bakhoff & Karl Pajusalu. 2009. Ainsuse 1. isikule viitamine eesti murretes. *Emakeele Seltsi aastaraamat 2008*(54). 159–185.
- Lindström, Liina & Kaili Müürisep. 2009. Parsing Corpus of Estonian Dialects. *Proceedings of the NODALIDA 2009 workshop Constraint Grammar and robust parsing*, vol. 8, 22–29. Northern European Association for Language Technology.
- Lindström, Liina & Karl Pajusalu. 2003. Corpus of Estonian Dialects and the Estonian Vowel System. *Linguistica Uralica* 4. 241–257.
- Lindström, Liina & Ilona Tragel. 2007. Eesti keele impersonaali ja seisundipassiivi vahekorrast adesiivargumendi kasutamise põhjal. *Keel ja Kirjandus* 3. 352–553.
- Lindström, Liina & Ilona Tragel. 2010. The possessive perfect construction in Estonian. *Folia Linguistica* 44(2). 371–399.
- Lindström, Liina & Kristel Uihoaed. valmimas. Syntactic variation in *need*-constructions in Estonian dialects.
- Lindström, Liina & Kristel Uihoaed. 2013. Variation in neccessive constructions in Estonian dialects. Paper presented at the Grammar and Context IV: New Approaches to the Uralic Languages. 6th–8th of June 2013. Tartu.
- Lindström, Liina, Eva Velsker, Ellen Niit & Karl Pajusalu. 2009. Mees “man”, aeg “time” and other frequent words in the Corpus of Estonian Dialects. Marja Kallasmaa & Vilja Oja (eds.), *Kodukeel ja keele kodu. Home Language and the home of a language. Pühendusteos Helmi Neetarile 75. sünnipäevaks 29. jaanuaril 2009*. Eesti Keele Instituudi toimetised 13. Tallinn: Eesti Keele Sihtasutus.
- Lippus, Pärtel & Karl Pajusalu. 2008. Regional variation in the perception of Estonian quantity. *Nordic prosody. Proceedings of the Xth conference, Helsinki*, 151–157.
- Lonn, Varje & Ellen Niit. 2002. *Eesti murded VII. Saarte murde tekstid*. Evi Juhkam (toim.). Tallinn: Eesti Keele Instituut.
- Maas, Utz. 2004. “Finite” and “nonfinite” from a typological perspective. *Linguistics* 42(2). 359–385.
- Macaulay, Ronald. 2010. Dialect. Mirjam Fried, Jan-Ola Östman & Jef Verschueren (eds.), *Variation and Change: Pragmatic Perspectives*, 61–72. Amsterdam: John Benjamins Publishing Company.
- Manning, Cristopher D. 2003. Probabilistic Syntax. Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic Linguistics*, 289–341. Cambridge/Massachusetts/London: MIT Press.
- Matras, Yaron. 2004. Typology, dialectology and the structure of complementation in Romani. Bernd Kortmann (ed.), *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, vol. 153, 277–304. Berlin: Mouton de Gruyter.
- Mendoza-Denton, Norma, Jennifer Hay & Stefanie Jannedy. 2003. Probabilistic Sociolinguistics. Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic Linguistics*, 97–138. Cambridge/Massachusetts/London: MIT Press.
- Mets, Mari. 2010. *Suhtlusvõrgustikud reaajas: võru kõnekeele varieerumine kahes Võrumaa külas*. Tartu: Tartu Ülikooli Kirjastus.
- Metslang, Helle. 1993a. Kas eesti keeles on olemas progressiiv? *Keel ja Kirjandus* 7. 326–334.
- Metslang, Helle. 1993b. Kas eesti keeles on olemas progressiiv? *Keel ja Kirjandus* 8. 410–417.

- Metslang, Helle. 1993c. Kas eesti keeles on olemas progressiiv? *Keel ja Kirjandus* 9. 470–476.
- Metslang, Helle. 1994a. Eesti ja soome – futuurumita keeled? *Keel ja Kirjandus* 9. 534–547.
- Metslang, Helle. 1994b. Eesti ja soome – futuurumita keeled? *Keel ja Kirjandus* 10. 603–616.
- Metslang, Helle. 2000. Reflections on the development of a Particle in Estonian. 59–86. *Estonian: Typological Studies IV*. Publications of the Department of Estonian of the University of Tartu 14. Tartu: University of Tartu
- Metslang, Helle. 2006. Predikaat ajastut kogemas. *Keel ja Kirjandus* 9. 714–727.
- Moisl, Hermann. 2009. Using Electronic Corpora in Historical Dialectology Research: The Problem of Document Length Variation. Marina Dossena & Roger Lass (eds.), *Studies in English and European Historical Dialectology*, 67–90. Bern, Switzerland: Peter Lang.
- Muischnek, Kadri & Heete Sahkai. 2009. Using collocation-finding methods to extract constructions and to estimate their productivity. Magnus Sahlgren & Ola Knutsson (eds.), *Workshop on extracting and using constructions in NLP. SICS Technical Report T2009:10*, 22–27.
- Muischnek, Kadri & Heete Sahkai. 2010. Liitpredikaadid leksikoni-grammatika kontiinumil. *Eesti ja soome-ugri keeleteaduse ajakiri. Keele rajad. Pühendusteos professor Helle Metslangi* 60. sünnipäevaks 1(2). 295–316.
- Muižniece, Liena, Helle Metslang & Karl Pajusalu. 1999. Past participle finitization in Estonian and Latvian. Mati Ereht (ed.), 128–157. *Estonian: Typological Studies III*. Publications of the Department of Estonian of the University of Tartu 11. Tartu: University of Tartu
- Mukherjee, Joybrato. 2007. Mahlberg, Michaela (2005). English General Nouns. A Corpus Theoretical Approach. *International Journal of Corpus Linguistics* 12(1). 131–147.
- Murumets, Sirje. 1982. Eesti keeleala murdelisest liigendusest “Väikese murdesõnastiku” põhjal. *Keel ja Kirjandus* 1. 11–17.
- Murumets, Sirje. 1983. Eesti keeleala murdelisest liigendusest “Väikese murdesõnastiku” põhjal, II. *Keel ja Kirjandus* 11. 615–623.
- Murumets, Sirje. 1984. Regional Vocabulary and Lexical Regions in Estonia. Hans Goebel (ed.), *Dialectology*, 225–253. (Quantitative Linguistics (QLing): 21). Bochum: Brockmeyer.
- Must, Mari. 1987. *Kirderannikumurre. Häälikuline ja grammatiline ülevaade*. Tallinn: Valgus: Eesti NSV Teaduste Akadeemia, Eesti Keele ja Kirjanduse Instituut.
- Must, Mari. 1995. *Eesti murded V. Kirderannikumurde tekstid*. Helmi Viires & Eevi Ross (toim.). Tallinn: Eesti Teaduste Akadeemia. Eesti Keele Instituut.
- Must, Mari & Aili Univere. 2002. *Põhjaeesti keskmurre: häälikulisi ja morfoloogilisi peajooni*. Eesti Keele Instituudi toimetised 10. Tallinn: Eesti Keele Instituut.
- Must, Mari. 1965. *Eesti murded II. Keskmurde tekstid*. (Ed.) Aili Univere & Helmi Viires. Tallinn: Eesti NSV TA Keel ja Kirjanduse Instituut.
- Mägiste, Julius. 1936. Eesti *saama*-futuuri algupärast ja tarvitamiskõlblikkusest. *Eesti Keel* 3. 65–92.
- Müürisep, Kaili. 2000. *Eesti keele arvutigrammatika: süntaks*. Tartu: Tartu Ülikooli Kirjastus.
- Neetar, Helmi. 1963. Ühest aluse ja öeldise ühildumise juhust eesti murretes. *Emakeele Seltsi toimetised* 6. *Nonaginta: Johannes Voldemar Veski* 90. sünnipäevaks 27. juunil 1963. 217–223.

- Neetar, Helmi. 1964. Aluse ja öeldise ühildumist mõjutavatest teguritest eesti murretes. *Emakeele Seltsi aastaraamat* 10. 151–166.
- Neetar, Helmi. 1965a. Aluse ja öeldise ühildumise seaduspärasustest eesti murretes. *Keel ja Kirjandus* 1. 25–29.
- Neetar, Helmi. 1965b. Aluse (kollektiivsubstantiivi) ja öeldise ühildumisest eesti murretes. *Emakeele Seltsi Aastaraamat* 11. 185–193.
- Neetar, Helmi. 1966. eks-laused. *Emakeele Seltsi aastaraamat* 12. 119–124.
- Neetar, Helmi. 1970. Määrsõnalisest täiendist eesti murretes. *Emakeele Seltsi aastaraamat* 16. 195–206.
- Neetar, Helmi. 1988. Mõnest teonimetarindist eesti murretes. *Emakeele Seltsi aastaraamat* 1986 (32). 36–45.
- Nenadic, Oleg & Michael Greenacre. 2007. Correspondence Analysis in R, with Two- and Three-dimensional Graphics: The ca Package. *Journal of Statistical Software* 20(3). 2–13.
- Nerbonne, John. 2009. Data-driven dialectology. *Language and Linguistics Compass* 3(1). 175–198.
- Nerbonne, John, Rinke Colen, Charlotte Gooskens, Peter Kleiweg & Therese Leinonen. 2011. Gabmap – A Web Application for Dialectology. *Dialectologia. Special Issue II*. 65–89.
- Nerbonne, John & Peter Kleiweg. 2007. Toward a Dialectological Yardstick. *Journal of Quantitative Linguistics* 14(2). 148–167.
- Nerbonne, John, Peter Kleiweg, Wilbert Heeringa & Franz Manni. 2008. Projecting Dialect Differences to Geography: Bootstrap Clustering vs. Noisy Clustering. Christine Preisach, Lars Schmidt-Thieme, Hans Burkhardt & Reinhold Decker (eds.), *Data Analysis, Machine Learning, and Applications. Proceedings of the 31st Annual Meeting of the German Classification Society*, 647–654. (Studies in Classification, Data Analysis, and Knowledge Organization). Berlin: Springer.
- Nikolaeva, Irina. 2010. Typology of Finiteness. *Language and Linguistics Compass* 4(12). 1176–1189.
- Nunnally, Jum C. 1978. *Psychometric theory*. 2nd ed. New York: McGraw-Hill.
- Pajusalu, Karl. 1989. Suhtlustaotlused inimkeelt muutmas. Tähelepanekuid eesti murrete verbimorfoloogiast. *Keel ja Kirjandus* 3. 142–149.
- Pajusalu, Karl. 1996. *Multiple linguistic contacts in South Estonian variation of verb inflection in Karksi*. Turku: Turun yliopisto.
- Pajusalu, Karl, Tiit Hennoste, Ellen Niit, Peeter Päll & Jüri Viikberg. 2009. *Eesti murded ja kohanimed*. 2., täiendatud trükk. Tallinn: Eesti Keele Sihtasutus.
- Pajusalu, Renate. 1994. Muutumisverbid eesti keeles ja nende vasted soome keeles. *Lähivertailuja* 40(8). Oulun Yliopiston tutkimusraporteja. 83–97.
- Pajusalu, Renate & Heili Orav. 2008. Supiinid koha väljendajana: Liikumissündmuse keelendamise asümmeetriast. *Emakeele Seltsi aastaraamat* 2007 (53). 104–121.
- Pajusalu, Renate, Ilona Tragel, Ann Veismann & Maigi Vija. 2004. *Tuumsõnade semantikat ja pragmaatikat*. Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 5. Tartu: Tartu Ülikool.
- Pall, Valdek. 1982. Idamurde vahekorraast naabermurretega ja vadjä keelega. *Keel ja Kirjandus* 5. 246–251.
- Pecina, Pavel. 2008. *Lexical Association Measures: Collocation Extraction*. Prague, Czech Republic: Faculty of Mathematics and Physics, Charles University in Prague.
- Pedersen, Ted. 1996. Fishing for Exactness. *Proceedings of the South Central SAS User's Group (SCSUG-96) Conference*, 188–200.

- Pedersen, Ted & Rebecca Bruce. 1996. What to Infer from a Description. *Technical Report 96-CSE-04*.
- Penjam, Pille. 2005. *tulema*-verbi leksikaalsest ja grammatilisest kasutusest ning modaalse funktsiooni kujunemisest Eesti kirjakeeles 17.–20. sajandil. Tartu: Tartu Ülikooli eesti keele õppetool. Käsikirjaline magistritöö.
- Penjam, Pille. 2006a. *tulema*-verbi grammatilised funktsioonid eesti kirjakeeles. *Keel ja Kirjandus* 1. 33–41.
- Penjam, Pille. 2006b. *tulema*-verbi modaalse funktsiooni kujunemine eesti kirjakeeles. *Lause argumentstruktuur*. Pille Penjam (toim.), 71–98. Tartu Ülikooli eesti keele õppetooli preprintid 2. Tartu: Tartu Ülikool.
- Penjam, Pille. 2008. *Eesti kirjakeele da- ja ma-infinitiiviga konstruktsioonid*. Tartu: Tartu Ülikooli Kirjastus.
- Piiraja, Piret. 2010. *Verb hakkama situatsiooni alguse väljendajana eesti keele ahelverbides*. Tartu: Tartu Ülikooli eesti ja üldkeeleteaduse instituut. Käsikirjaline magistritöö.
- Potseps, Vaike. 1956. *da-infiniitiv ja selle tarvitamine eesti murdeis*. Tartu: Tartu Riiklik Ülikool. Käsikirjaline diplomitöö.
- R Development CoreTeam. 2011. *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*.
- Rajandi, Henno. 1999. *Eesti impersonaali ja passiivi süntaks*. Tallinn: Eesti keele Instituut.
- Rätsep, Huno. 1957. Aspektikategooriast eesti keeles. *Emakeele Seltsi aastaraamat* 3. 72–81.
- Rätsep, Huno. 1978. *Eesti keele lihtlausete tüübid*. Tallinn: Valgus.
- Saareste, Andrus. 1938a. *Eesti murdeatlas: tutvustusvihk = Atlas des parlers estoniens: enchantillon*. Tartu: Eesti Kirjanduse Selts.
- Saareste, Andrus. 1938b. *Eesti murdeatlas: Atlas des parlers estoniens*. Tartu: Eesti Kirjanduse Selts.
- Sahkai, Heete. 2011. *Teine grammatika. Eesti keele teonimede süntaks konstruktsiooni-põhises perspektiivis*. Tallinn: Tallinna Ülikool.
- Sang, Joel. 1974. *Eitavad laused Kihnu murrakus*. Tartu: Tartu Riikliku Ülikooli eesti keele kateeder. Käsikirjaline diplomitöö.
- Sang, Joel. 1975. Eitus Kihnu murrakus. *Keel ja Kirjandus* 3. 155–162.
- Saukkonen, Pauli. 1965. *Itämerensuomalaisten kielten tulosijainfinitiivirakenteiden historiaa I. Johdanto. Adverbiaali infinitiivi*. Helsinki: Suomalais-Ugrilainen Seura.
- Saukkonen, Pauli. 1966. *Itämerensuomalaisten kielten tulosijainfinitiivirakenteiden historiaa II. Adverbiaali infinitiivi (täydennys). Adnominaali infinitiivi. Absoluutti infinitiivi. Elliptinen infinitiivi*. Helsinki: Suomalais-Ugrilainen Seura.
- Schulz, Monika. 2010. *Morphosyntactic variation in British English Dialects. Evidence from Possession, Obligation and Past Habituality*. Freiburg: Albert-Ludwigs-Universität. PhD Dissertation.
- Schutter, Georges de. 2010. Dialectology. Mirjam Fried, Jan-Ola Östman & Jef Verschueren (eds.), *Variation and Change: Pragmatic Perspectives*, 73–80. Amsterdam, Netherlands: John Benjamins Publishing Company.
- Séguy, Jean. 1971. La relation entre la distance spatiale et la distance lexicale. *Revue de Linguistique Romane* 35(138). 335–357.
- Séguy, Jean. 1973. La dialectometrie dans l'Atlas linguistique de la Gascogne. *Revue de Linguistique Romane*, 37(145). 1–24.
- Seiler, Guido. 2004. On three types of dialect variation and their implications for linguistic theory. Evidence from verb clusters in Swiss German dialects. Bernd

- Kortmann (ed.), *Dialectology meets typology: dialect grammar from a cross-linguistic perspective*, vol. 153, 367–399. Berlin: Mouton de Gruyter.
- Siewierska, Anna. 2008. Introduction: impersonalization from a subject-centred vs. agent-centred perspective. *Transactions of the Philological Society* 106(2). 115–137.
- Siewierska, Anna & Willem B. Hollmann. 2007. Ditransitive clauses in English with special reference to Lancashire dialect. *Structural-functional Studies in English Grammar: In Honor of Lachlan Mackenzie*. 83–102.
- Spruit, Marco René. 2008. *Quantitative perspectives on syntactic variation in Dutch dialects*. University of Amsterdam. PhD Dissertation.
- Szmrecsanyi, Benedikt. 2008a. Forests, trees, corpora, and dialect grammars. Benedikt Szmrecsanyi & Bernhard Wälchli (eds.), *Aggregating dialectology, typology, and register analysis: Linguistic variation in text and speech*. Berlin: Mouton De Gruyter.
- Szmrecsanyi, Benedikt. 2008b. Methods and objectives in contemporary dialectology. Ilja A. Seržant & Björn Wiemer (eds.), *Современные методы в диалектологии. Ареал северных, северо-западных русских и белорусских говоров / Contemporary approaches to dialectology: The area of North, Northwest Russian and Belarusian vernaculars*.
- Szmrecsanyi, Benedikt. 2008c. About text frequencies in historical linguistics: disentangling environmental and grammatical change. Martin Hilpert & Hubert Cuyckens (eds.), *Corpus Linguistics and Linguistic Theory* (Special issue).
- Szmrecsanyi, Benedikt. 2012. Analyticity and syntheticity in the history of English. Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford Handbook of the History of English*, 654–665. Oxford: Oxford University Press.
- Szmrecsanyi, Benedikt. 2013. *Grammatical Variation in British English Dialects: A Study in Corpus-Based Dialectometry*. Cambridge: Cambridge University Press.
- Szmrecsanyi, Benedikt & Bernd Kortmann. 2009. The morphosyntax of varieties of English worldwide: A quantitative perspective. *Lingua* 119(11). 1643–1663.
- Szmrecsanyi, Benedikt & Christoph Wolk. 2011. Holistic corpus-based dialectology. Stefan Th. Gries (ed.), *Revista Brasileira de Linguística Aplicada* 11(2). 561–592.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collocations: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1). 1–43.
- Tael, Kaja. 1988. Sõnajärjemallid eesti keeles (võrrelduna soome keelega): Läänemere-soome keeleteaduse sümpoosion, Turku, 30.08–2.09. 1988. *Preprint / Eesti NSV Teaduste Akadeemia Ühiskonnateaduste osakond*, KKI-56. Tallinn: Eesti NSV Teaduste Akadeemia, Keele ja Kirjanduse Instituut.
- Tanning, Salme. 1961. *Eesti murded I. Mulgi murdetekstid*. Mari Must & Aili Univere (toim.). Tallinn: Eesti Riiklik Kirjastus.
- Tanning, Salme. 2004. *Karksi murrak*. Karl Pajusalu (toim.). Tallinn: Eesti Keele Instituut.
- Tauli, Valter. 1980. *Eesti grammatika II. Lauseõpetus*. Uppsala: Finsk-ugriska institutionen.
- Tidwell, Doug. 2002. Introduction to XML. *USA: IBM*. Retrieved December 20, 2002.
- Toomse, Mihkel. 1998. *Lõuna-Eesti murded 1–30: kaardid*. (Ed.) Karl Pajusalu. (56). Turku: Turun yliopiston suomalaisen ja yleisen kielitieteen laitoksen julkaisuja.
- Torn-Leesik, Reeli. 2007. Voice and modal verbs in Estonian. *Linguistica Uralica* 43(3). 173–186.

- Torn-Leesik, Reeli. 2009. The voice system of Estonian. *STUF: Sprachtypologie und Universalienforschung* 62(1–2). 72–90.
- Torn-Leesik, Reeli & Virve-Anneli Vihman. 2010. The Uses of Impersonals in Spoken Estonian. *SKY Journal of Linguistics* 23. 301–343.
- Tragel, Ilona. 2003. *Eesti keele tuumverbid*. Tartu: Tartu Ülikooli Kirjastus.
- Tragel, Ilona & Külli Habicht. 2012. Grammaticalization of Estonian *saama* "to get". *Linguistics* 50(6). 1371–1412.
- Uibo, Udo. 2010. Etümoloogilisi märkmeid (VIII). *Keel ja Kirjandus* 1. 45–52.
- Uiboed, Kristel. 2010a. Ühendverbid eesti murretes. *Keel ja Kirjandus* 1. 17–36.
- Uiboed, Kristel. 2010b. Statistilised meetodid murdekorpuse ühendverbide tuvastamiseks. *Eesti Rakenduslingvistika Ühingu aastaraamat* 6. 307–326.
- Uiboed, Kristel, Cornelius Hasselblatt, Liina Lindström, Kadri Muischnek & John Nerbonne. 2013. Variation of verbal constructions in Estonian dialects. *Literary & Linguistic Computing* 28(1). 42–62.
- Univere, Aili. 1966. Põhja-Eesti *hakkama* ja Lõuna-Eesti *nakkama*. *Emakeele Seltsi aastaraamat* 12. 109–117.
- Univere, Aili. 1988. Idamurre. *Emakeele Seltsi aastaraamat* 1986 (32). 59–93.
- Univere, Aili. 1996. *Eesti murded IV. Idamurde tekstid*. Mari Must & Valdek Pall (toim.). Tallinn: Eesti Teaduste Akadeemia. Eesti Keele Instituut.
- Uuspõld, Ellen. 1989. Modaalasusest ja modaalset predikaadist eesti keeles. *Keel ja Kirjandus* 8. 468–477.
- Vangsnes, Øystein Alexander. 2007. Scandinavian dialect syntax (before and after) 2005. *Nordlyd* 34(1). 7–24.
- Wiechmann, Daniel. 2008. On the computation of collostruction strength: Testing measures of association as expressions of lexical bias. *Corpus Linguistics & Linguistic Theory* 4(2). 253–290.
- Wiedemann, Ferdinand Johann. 1875. *Grammatik der ehstnischen Sprache: zunächst wie sie in Mittelehstland gesprochen wird, mit Berücksichtigung der anderen Dialekte*. St.-Petersbourg: Imprimerie de l'Académie Impériale des sciences.
- Wiedemann, Ferdinand Johann. 2011. *Eesti keele grammatika*. Ellen Niit & Heli Laanekask (toim.). Tallinn: Eesti Teaduste Akadeemia Emakeele Selts.
- Viitso, Tiit-Rein. 2003. Structure of the Estonian language. Phonology, morphology and word formation. Mati Ereht (ed.), *Estonian Language*, 9–92.
- Wilcock, Graham. 2009. *Introduction to linguistic annotation and text analytics*. San Rafael, California: Morgan & Claypool.
- Wilson, Andrew. 2005. Modal verbs in written Indian English: A quantitative and comparative analysis of the Kolhapur corpus using correspondence analysis. *ICAME JOURNAL Computers in English Linguistics* 29. 151–170.
- Vitsberg, Eduard. 1958. *Eitus eesti murretes*. Tartu Riikliku Ülikooli eesti keele kateeder. Käsikirjaline diplomitöö.
- Vogelaar, Gunther de & Guido Seiler. 2012. Introductory remarks. Gunther de Vogelaar & Guido Seiler (eds.). *The Dialect Laboratory: Dialects as a Testing Ground for Theories of Language Change*, 1–31. Amsterdam, Netherlands: John Benjamins Publishing Company.
- Östman, Jan-Ola & Graeme Trousdale. 2013. Dialects, Discourse, and Construction Grammar. Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford Handbook of Construction Grammar*, 476–490. Oxford: Oxford University Press.
- Ylikoski, Jussi. 2003. Defining Non-Finites: Action Nominals, Converbs and Infinitives. *SKY Journal of Linguistics* 16. 185–237.

SUMMARY IN ENGLISH

Verb Constructions in Estonian Dialects

The present thesis examines the variation of verb constructions in Estonian dialects. The study utilizes an aggregate dialectal approach and quantitative methods to explore the variation of the syntactic phenomenon – finite and non-finite verb constructions – under investigation.

There are three main objectives in this study: linguistic, computational, and methodological. The most important linguistic objective is to explore the syntactic variation in Estonian dialects because to date we have very limited knowledge about the variation on a syntactic level. Additionally, this study examines the syntactic variation by combining traditional descriptive dialectology with contemporary linguistic theory, specifically Construction Grammar. This study assumes that dialects should not be studied in isolation from other linguistic fields and can be described within contemporary linguistic frameworks. In addition, modern linguistic theories can also benefit from dialectal data and research findings. The major computational objective of this thesis is to create a common technical basis for doing quantitative syntactic research using the Estonian Dialect Corpus (EMK 2011). Finally, the methodological objective is to test contemporary statistical methods and to compare them with simpler frequency-based co-occurrence findings.

The main hypotheses of the present research are the following.

- 1) Finite and non-finite verb constructions vary considerably across dialects. Based on previous studies, I expect that greater differences emerge between western and eastern dialects as opposed to the traditional dialect classifications which state greater differences between northern and southern dialects. One reason for that might be that traditional Estonian dialect classifications do not take syntax into account and rely only on phonetic, morphological, and lexical features.
- 2) Based on the previous statement, I expect that western dialects use more analytical ways of expression, i.e. use more finite and non-finite verb constructions. These assumptions follow from previous studies (Uibo 2010a; Uibo *et al.* 2013).
- 3) Methodologically, I expect that simple co-occurrences of certain linguistic units have explanatory power when studying syntactic constructions and that they do not perform considerably worse than more sophisticated statistical methods.

To date, syntactic phenomena in dialects have been little studied; however, in recent decades, dialect syntax has developed into a lively and interesting separate field of research (Fried 2010: 9; Macaulay 2010: 67; Schutter 2010: 74). Although dialect and spoken syntax share common properties, dialect syntax has unique characteristics that cannot simply be reduced to spoken register; specifically, spatial dimension is of central interest here (Auer 2004: 71–72).

Syntactic change, compared to phonological and lexical change for instance, is less perceivable (Labov 1991: 277), but there is no reason to assume that syntactic phenomena do not vary geographically (Szmrecsanyi (ilmumas a) [*to appear*]). Syntactic phenomena usually have wider geographical coverage and are seldom restricted to smaller areas. Thus, syntactic variation is rather described by preferential tendencies and frequency patterns than via categorical observations (Kortmann 2010: 842). For these reasons, this study contributes to the new field of dialect syntax by exploring the nature of syntactic variation and by adding new language data to the field.

Frequently occurring grammatical verbs and forms are of central interest, because this study emphasizes the importance of frequency (Bybee 2007; Divjak & Gries 2012 among many others). The frequency data has been considered important especially when studying language variation and it has been stated to be an important source of information when exploring linguistic variation (Bod *et al.* 2003: 3; Szmrecsanyi 2013).

This study approaches data and analyses in an aggregational manner. Aggregate dialectology deals with many different linguistic features at once instead of observing a single carefully selected idiosyncratic phenomenon (Nerbonne 2009). Large scale aggregate syntactic studies have been performed previously for instance on English dialects (Grieve 2009; Szmrecsanyi 2013) and Dutch dialects (Spruit 2008). The current study explores only certain verbal units, but is still aggregational in nature. For instance, no assumptions or restrictions have been made regarding different non-finite forms and finite verbs when gathering the data. The analysis comprises all the verbal units detected from the corpus and they are analyzed together in an aggregational manner. Generalizations about dialectal differences are drawn mostly based on the results of aggregate. As the results reveal, the constructions in the focus are considerably more heterogeneous group of items than one would expect at first glance.

This study examines finite and non-finite verb constructions which can be typologically defined as auxiliary verb constructions (Anderson 2006; Heine 1993; Kuteva 2001). This contrasts the traditional description of Estonian standard language as finite verbs in these constructions are not analyzed as auxiliary verbs. However, many finite verbs in Estonian tend to function as auxiliaries when they co-occur with certain non-finite forms. These constructions form a continuum in terms of auxiliary-hood. It is not the case with all such units, but all the constructions observed in this study have the potential to develop into auxiliaries, based on the statement that Bolinger (1980: 297) makes: the moment a verb takes a non-finite complement, that verb also has the potential to develop into an auxiliary, some verbs of this kind may just stay in the beginning of that stage of development.

This thesis uses the Construction Grammar framework (Croft 2001; Fillmore 1989; Fried & Östman 2004; Goldberg 1995), which assumes that atomic structures of language are form-meaning pairings – constructions, hence the present study makes an assumption that finite and non-finite verb pairs are holistic units that convey grammatical and/or semantic meanings. The study follows an

usage-based approach (Bybee 2010; Langacker 1987; Kemmer & Barlow 2000), which states that the object of investigation should be the real language usage data which is the best evidence for making conclusions about language systems in general. The data analysed in the present study is comprised of traditional recorded dialect interviews, i.e. spoken spontaneous language.

The data were obtained from the morphologically annotated Corpus of Estonian Dialects. The corpus contains 613 124 running words of text from all ten dialects of Estonian. The map of Estonian dialects is presented on Figure 2. Traditional dialect classifications rely on phonological, morphological, and lexical differences. Based on these linguistic levels, North-Estonian (the Coastal, Eastern, Insular, Mid, Northeastern, and Western dialects) and South-Estonian (Mulgi, Tartu, Võru, and Seto) dialect groups are considered to be the most distinctive. Sometimes Coastal and Northeastern dialects are described as a separate subgroup in the North-Estonian dialect group.

The corpus texts were first manipulated automatically. Clause boundaries for construction extraction were set automatically by utilizing a syntactic parser of Estonian (Müürisep 2000), which has been adapted for dialectal text parsing (Lindström & Müürisep 2009). After setting the clause boundaries, all the combinations with non-finite form and finite verb were composed automatically. Combinations that occurred less than three times in the data were excluded from further analysis. The dataset containing all pairs of non-finite form and finite verb in all ten dialects was the basis for further analysis. The precision of the detection process was 92%.

The present study utilizes several quantitative methods. Collostructional methods (Stefanowitsch & Gries 2003) are used for construction extraction by calculating statistical association scores for two verbal units: the finite verb and non-finite form. Collostructional findings are compared with simple co-occurrence frequency data, while quantitative analyses are separately conducted for constructions extracted with collostructional methods and for verbal units based on their clausal co-occurrence frequencies. For dialect classification purposes, Correspondence Analysis (Greenacre 2007) and fuzzy clustering (Everitt *et al.* 2011) are applied. Different possibilities available in R (R Development CoreTeam 2011) and Gabmap (Kleiweg *et al.* 2011) were utilized for visualization purposes.

Collostructional methods appeared to be quite successful for construction extraction and contributed to eliminating random noise from the quantitative data, although these benefits came with some loss of information. Simple co-occurrence frequencies turned out to be quite successful for detecting variational patterns and the findings correlated strongly with the collostructional findings.

Quantitative aggregational results revealed interesting tendencies. Frequencies of verb constructions vary considerably across dialects. The most distinctive groups were formed between eastern and western dialects, instead of the traditional North-South distinction. The eastern group consists of North-eastern, Eastern, Mulgi, Seto, Tartu, Võru dialects and the western group of Insular,

Mid, Coastal, and Western dialects (see figures 9 and 10). This is an interesting result in light of previous dialect classifications which do not take syntactic features into account. Southern and northern dialect groups are traditionally considered to be the most different. The present study does not confirm this and provides a completely different dialect classification.

Constructions extracted with collostructional methods and instances collected based on clausal co-occurrence frequencies are analysed qualitatively in order to explore which meanings certain verb pairings carry in dialects and how these form-meaning pairings vary across dialects. The qualitative analysis included only verb constructions that emerged from the quantitative analysis and/or had frequency (normalized or absolute) of 100 instances in the data. Additional motivation for performing the qualitative analysis was the need to evaluate the outcomes of the quantitative analysis, i.e. to investigate whether the automatically extracted constructions were really existent in the corpus or just statistically co-occurring items.

First, all the constructions with the first infinitive (*da*-infinitive) were analysed. With the verb *jõudma* ‘to manage, to reach’ the *da*-infinitive can express either physical or mental managing. With the verb *laskma* ‘to let’ the *da*-infinitive can express instrumental causative or different modal meanings. The *da*-infinitive and verb *olema* ‘to be’ form detachment, distance, mental, possessive, property, predicative, idiomatic, and modal constructions. The *da*-infinitive and finite verb *saama* ‘to get, become’ can form impersonal, mental, modal, passive, and idiomatic constructions. The verb *tahtma* ‘to want’ expresses modality, proximativity, desire/wish, and resistance with the *da*-infinitive. The verb *tohtima* ‘may’ and *võima* ‘can, may’ only form modal constructions with the *da*-infinitive. The motion verb *tulema* ‘to come’ forms attributive, finale, and modal constructions with *da*-infinitive.

The *ma*-infinitive (supine) forms inchoative constructions and idiomatic expressions with the finite verb *hakkama* ‘to start’. With the verb *jääma* ‘to stay, remain’ supine expresses inchoative meaning. The motion verb *minema* ‘to go’ forms verbal particle, inchoative, locative constructions and idiomatic expressions with the supine. The finite verb *olema* ‘to be’ expresses modality, property, and proximativity with the *ma*-infinitive form. The verb *panema* ‘to put’ and the supine form inchoative or causative constructions, and also idiomatic expressions. The modal verb *pidama* ‘must’ and the supine can form two kinds of quotative constructions and also avertive, proximative, and modal constructions. The supine and finite verb *saama* ‘to get, become’ express future, causativity, managing, but can also form idiomatic expressions. The motion verb *tulema* ‘to come’ expresses inchoativity and locativity, when it appears with the supine form, but can also form verb particle constructions. The inessive form of the supine can express habituality and locativity when it appears with the finite verb *käima* ‘to go, walk’. Different constructions form with the supine inessive and the finite *olema* ‘to be’: locative, property, progressive, proximative, and existential constructions.

Passive past participle form impersonal, passive, and possessive perfect constructions when it appears with finite *olema* ‘to be’, additionally to these it can form resultative constructions when it appears with finite *saama* ‘to get, become’.

Aggregational quantitative analysis was applied also to qualitative results. First, separate analyses were made for different non-finite verb constructions (*da*-, *ma*-infinitive and passive past participle). Then, all the non-finite verb constructions were aggregated into one analysis. The quantitative analysis with *da*-infinitive constructions reveals a continuum rather than distinctive groups of dialects. One side of the continuum are the North-eastern, Western, Insular, and Mid dialects. The Coastal, Eastern, and Tartu dialects are in the centre of that continuum. The Võru and Mulgi dialects are more on the other side and Seto is even further on the other side. Võru and Mulgi dialects are almost identical in respect to *da*-infinitive constructions. The Seto dialect is very different from all the others and the North-eastern dialect is also quite distinctive.

The quantitative analysis of supine constructions revealed more distinctive groups than the *da*-infinitive construction analysis. The Mulgi and Seto dialects are very different from all other dialects, but also from each other. The first more homogeneous group includes Mulgi, North-eastern, Western, Insular, and Mid dialects. The second more coherent group consists of Coastal, Eastern, Võru, Tartu, and Seto dialects. In this analysis Võru and Tartu dialects are almost identical, but surprisingly very different from Seto and even more different from the Mulgi dialect. Eastern dialect seems to be similar to Coastal dialect and is more similar to Võru and Tartu than other traditional North-Estonian dialects.

Quantitative analysis of passive past participle constructions reveals very differences between western (Coastal, Mid, Western, Insular) and eastern (North-eastern, Eastern, Tartu, Mulgi, Võru, Seto) dialects. Here the Eastern and North-eastern dialects are almost identical. The Võru, Mulgi, and Seto dialects are very similar, but here Tartu seems to be a bit more different from other South-Estonian dialects.

More general results confirm some statements made earlier about syntactic variation. Kortmann (2010: 842) claims syntactic variation to be less restricted to certain areas and the spread of syntactic phenomena can be observed rather as patterns of preferences. Results presented in this work confirm these statements. There were very few constructions which did not occur in all the areas, which might be also the result of small dataset. Differences emerged from the frequencies of use.

The thesis makes a considerable contribution to the knowledge of Estonian dialects and to the field of dialect syntax in general. The results provide new information about Estonian dialects and their syntactic variation. The current work suggests that understudied syntactic phenomena in dialects provide new knowledge about dialects and syntactic variation in general. Findings demonstrate how including a new linguistic level in the analysis can shed light on variational knowledge from a totally different perspective. The results emphasize the importance of syntax when creating dialect classifications.

LISA I. MURDEKORPUSE MORFOLOOGILISELT MÄRGENDATAVAD KATEGOORIAD

Morfoloogilisel märgendamisel eristatud sõnaklassid ja nende lühendid

Substantiiv (nimisõna)	S	nt <i>kas 's</i>
Pärisnimi	H	<i>Jüri, Pärnumaa</i>
Verb (tegusõna)	V	nt <i>ostma</i> , v.a <i>olema</i> abiverbina
Abiverb	Va	ainult <i>olema</i> -verbi vormid
Adverb (määrsõna)	Adv	nt <i>täna</i>
Proadverb	ProAdv	<i>siin, seal, kunagi, kõikjal</i>
Abimäärsõna	Adva	verbi osad, nt <i>välja (mõtlemas)</i>
Numeraal (arvsõna)		
põhiarvsõnad	Nump	nt <i>kaks</i>
järgarvsõnad	Numj	nt <i>teine</i>
Adjektiiv (omadussõna)	A	<i>vana</i>
Adjektiiv komparatiivis	Ak	<i>vanem</i>
Adjektiiv superlatiivis	As	<i>vanim</i>
Pro-sõnad		
Prosubstantiiv	ProS	Nt <i>see, too, tema, mina</i>
Proadjektiiv	ProA	Nt <i>niisuke, sihuke</i>
Proadverb	ProAdv	Nt <i>siin, seal</i>
Kaassõnad		
Postpositsioon	Post	nt <i>maja taga</i>
Prepositsioon	Pre	nt <i>pärast sööki</i>
Diskursusepartiklid	Par	
Suhtlussõnad	Suht	<i>aitäh, palun, tere</i>
Onomatopoeetilised sõnad	Ono	<i>mürts</i>
Küsisõna	Intr	nt <i>kas</i> , ka relatiiv-interrogatiivpronoomen (<i>kes, kelle, keda</i> jne), küsivad-siduvad adverbid (<i>miks, millal, kus</i> jne)
Konjunktsioon (sidesõna)	Konj	nt <i>ja</i> , sh piiripartiklid
Eitussõna liitvormides	Mn	<i>ei, mitte</i>
Võrdlussõna liitsuperlatiivis	Ms	<i>kõige</i>
Muud	Muu	ei oska määrata või sobiv kategooria puudub
Määramata	X	ei ole võimalik määrata (poolikud sõnad)

Käändsõnadele (S, ProS, A, Ak, As, ProA, H, Nump, Numj, Intr) märgitud morfoloogilised tunnused ja lõpud

ainsus	sg
mitmus	pl
nominatiiv (nimetav)	n
genitiiv (omastav)	g
partitiiv (osastav)	p
illatiiv (sisseütlev)	ill
additiiv (lühike sisseütlev)	add
inessiiv (seesütlev)	in
elatiiv (seestütlev)	el
allatiiv (alaleütlev)	all
adessiiv (alalütlev)	ad
ablatiiv (alaltütlev)	abl
translatiiv (saav)	tr
terminatiiv (rajav)	ter
essiiv (olev)	es
abessiiv (ilmaütlev)	ab
komitatiiv (kaasaütlev)	kom
instruktiiv (viisiütlev)	inst
possessiivsufiks	poss

Pöörd sõnadele (V, Va) märgitud morfoloogilised tunnused ja lõpud

<i>da</i> -infinitiiv	inf
<i>des</i> -gerund	ger
<i>ma</i> -supiin	sup
<i>tav</i> -partitsiip	tav
<i>nud</i> -partitsiip	nud
<i>tud</i> -partitsiip	tud
<i>v</i> -partitsiip	v
Personaal	ps
Impersonaalne passiiv	ips
Personaalne passiiv	pas
Indikatiiv (kindel)	ind
Konditsionaal (tingiv)	knd
Imperatiiv (käskiv)	imp
Jussiiv (möönev)	jus
Kvotatiiv (kaudne)	kvt
Potentsiaal	pot
Preesens (olevik)	pr
Imperfekt (lihtminevik)	ipf
Ainsus	sg
Mitmus	pl
Esimene isik (mina, meie)	1
Teine isik (sina, teie)	2
Kolmas isik (tema, nemad)	3
eituse märgend verbivormi lõpus	neg
eitussõna	Mn

LISA 2. VERBIÜHENDITE NORMALISEERITUD SAGEDUSTE ANALÜÜSI ALGANDMED

Inf+V _{fin}	RAN	IDA	SAA	KES	MUL	KIR	SET	TAR	LÄÄ	VÕR
inf_andma	0	0	0	4	0	0	0	0	0	0
inf_julgema	0	0	0	2	0	0	0	0	0	0
inf_jõudma	0	4	0	0	0	0	0	7	0	0
inf_käskima	0	0	2	0	0	0	0	0	0	0
inf_laskma	0	8	5	4	0	0	0	0	2	0
inf_lubama	0	0	2	0	0	0	0	0	0	0
inf_olema	19	20	16	25	5	18	9	7	26	5
inf_oskama	0	0	1	0	0	0	0	0	2	0
inf_pidama	0	0	0	0	0	0	0	0	2	0
inf_saama	19	14	13	20	0	47	9	11	25	6
inf_tahtma	10	7	9	9	0	0	0	0	12	0
inf_tohtima	0	0	5	3	0	0	0	0	0	0
inf_tulema	6	6	4	9	7	5	0	0	4	5
inf_täidima	6	0	0	0	0	0	0	0	0	0
inf_viitsima	0	0	1	0	0	0	0	0	0	0
inf_võima	0	8	5	9	0	0	0	0	2	0
ma_ajama	0	0	5	9	0	5	0	0	6	0
ma_hakkama	173	124	215	165	34	134	108	82	140	108
ma_heitma	0	0	0	0	0	0	0	4	0	0
ma_hüüdma	0	0	0	2	0	0	0	0	0	0
ma_juhtuma	0	0	4	5	0	0	0	0	8	0
ma_jätma	0	0	0	2	0	0	0	0	0	0
ma_jääma	12	10	7	12	0	12	0	13	9	7
ma_kippuma	0	0	0	0	0	0	0	0	2	0
ma_kukkuma	0	0	2	0	0	0	0	0	0	0
ma_laskma	0	0	2	0	0	0	0	0	0	6
ma_lööma	0	0	0	0	0	5	0	0	0	0
ma_minema	52	54	40	51	21	35	114	51	43	50
ma_nägema	0	0	0	2	0	0	0	0	2	0
ma_olema	61	40	32	24	50	32	11	14	32	30
ma_panema	19	24	30	39	12	12	40	23	47	24
ma_pidama	96	84	176	167	162	184	65	61	153	80
ma_saama	22	7	16	24	0	15	0	13	25	9
ma_saatma	0	0	1	0	0	0	0	0	2	0

Inf+V _{fin}	RAN	IDA	SAA	KES	MUL	KIR	SET	TAR	LÄÄ	VÕR
ma_tahdma	0	0	3	0	0	0	0	0	0	0
ma_teadma	9	0	0	2	0	0	0	0	0	0
ma_tegema	0	0	0	0	0	0	0	0	3	0
ma_tooma	0	0	1	0	0	0	0	0	0	0
ma_tulema	20	26	12	15	14	12	28	20	8	17
ma_viima	0	0	2	2	0	0	0	0	3	0
ma_võima	0	0	3	0	0	0	0	0	2	0
ma_võtma	0	0	2	3	0	0	0	0	3	5
ma_õpetama	0	0	0	0	0	0	0	0	2	0
ma_õppima	0	0	0	2	0	0	0	0	0	0
mas_hakkama	0	0	3	0	0	0	0	0	0	0
mas_käima	26	33	44	23	14	27	0	4	34	7
mas_olema	28	18	21	16	11	24	11	10	18	28
mas_panema	0	0	1	0	0	0	0	0	0	0
mas_pidama	0	0	5	0	0	0	0	0	0	0
mas_saama	0	0	0	4	0	0	0	0	2	0
mast_tulema	0	0	0	0	0	6	0	0	0	0
mata_jääma	6	7	3	0	0	0	0	6	3	8
mata_olema	7	7	3	5	0	8	0	6	6	0
nud_ajama	0	0	0	2	0	0	0	0	3	0
nud_andma	0	4	2	3	7	0	0	0	4	0
nud_elama	0	6	0	0	0	0	0	0	2	0
nud_hakkama	0	0	4	0	0	5	0	0	0	0
nud_jutustama	0	0	0	0	5	0	0	0	0	0
nud_jääma	4	0	3	2	0	0	0	4	4	0
nud_kaema	0	0	0	0	0	0	0	0	0	5
nud_keerama	0	0	0	2	0	0	0	0	0	0
nud_ketrama	0	0	0	0	7	0	0	0	0	0
nud_kurtma	0	0	0	0	5	0	0	0	0	0
nud_kustuma	0	0	0	0	0	0	0	0	2	0
nud_kutsuma	0	0	2	0	7	0	0	0	0	0
nud_kuulma	0	0	0	2	0	0	0	0	2	0
nud_kõnelema	0	0	0	0	9	0	0	4	0	0
nud_käima	16	0	2	2	9	0	0	0	4	0
nud_laskma	0	0	0	0	0	0	0	0	5	0
nud_lugema	0	0	0	0	0	0	0	0	4	0

Inf+V _{fin}	RAN	IDA	SAA	KES	MUL	KIR	SET	TAR	LÄÄ	VÕR
nud_magama	0	0	3	2	0	0	0	0	0	0
nud_minema	20	10	7	9	29	5	0	10	17	7
nud_naerma	0	0	0	4	0	0	0	0	0	0
nud_nägema	4	0	2	4	5	0	0	6	2	6
nud_olema	515	512	622	487	1030	291	162	424	555	431
nud_pagema	0	0	0	0	5	0	0	0	0	0
nud_panema	4	0	8	4	0	0	0	0	7	0
nud_pidama	9	0	2	6	0	0	0	0	2	0
nud_rääkima	0	10	7	9	0	6	0	0	11	0
nud_saama	25	10	12	9	29	5	0	4	18	0
nud_seletama	0	0	0	0	0	0	0	0	0	3
nud_surema	6	0	1	0	0	0	0	0	2	0
nud_tahtma	0	0	4	4	0	0	0	0	3	0
nud_taidma	6	0	0	0	0	0	0	0	0	0
nud_teadma	6	7	2	4	5	9	0	7	9	6
nud_tegema	7	0	6	9	29	0	0	0	4	0
nud_tohtima	0	0	0	2	0	0	0	0	2	0
nud_tooma	0	0	0	4	9	0	0	0	4	0
nud_tulema	7	21	7	7	25	0	0	11	9	3
nud_tõmbama	0	0	0	0	9	0	0	0	0	0
nud_vaatama	6	0	0	4	5	0	0	0	6	0
nud_valama	0	0	0	2	0	0	0	4	0	0
nud_viima	0	0	0	2	0	0	0	0	4	0
nud_võima	4	0	0	2	0	0	0	0	0	0
nud_võtma	0	0	2	6	7	0	0	0	3	0
nud_ütleva	4	0	1	0	0	6	0	4	8	5
tud_andma	0	0	1	0	0	0	0	0	0	0
tud_hakkama	0	0	0	0	0	0	0	0	2	0
tud_heitma	0	0	1	0	0	0	0	0	0	0
tud_hoidma	0	0	1	0	0	0	0	0	0	0
tud_katkuma	0	0	1	0	0	0	0	0	0	0
tud_kutsuma	0	0	2	0	0	0	0	0	0	0
tud_kuulma	0	0	0	0	0	0	0	0	0	3
tud_käima	0	0	2	0	0	0	0	0	0	0
tud_minema	0	4	2	0	0	0	0	0	4	3
tud_olema	131	167	210	149	191	115	105	145	157	162
tud_panema	0	0	5	2	0	0	0	0	4	0

Inf+V _{fin}	RAN	IDA	SAA	KES	MUL	KIR	SET	TAR	LÄÄ	VÕR
tud_pidama	0	4	8	8	5	0	0	0	6	0
tud_rääkima	0	0	3	0	0	0	0	0	2	0
tud_saama	181	28	192	212	21	21	11	16	186	19
tud_teadma	4	7	2	0	0	0	0	0	0	0
tud_tegema	4	0	5	0	0	0	0	0	6	0
tud_tooma	0	0	0	0	0	0	0	0	2	0
tud_tulema	0	0	1	2	0	0	0	0	0	0
tud_tõmbama	0	0	1	0	0	0	0	0	0	0
tud_võtma	0	0	1	0	0	0	0	0	2	0
tud_ütleva	0	0	0	0	0	5	0	0	0	0
v_ajama	0	0	0	0	0	0	9	0	0	0
v_olema	4	0	4	2	0	0	0	0	0	6
v_panema	0	0	2	0	0	0	0	0	2	0

LISA 3. KOLLOSTRUKTSIOONILIST ANALÜÜSI ARVESTAVA NORMALISEERITUD SAGEDUSTE ANALÜÜSI ALGANDMED

Inf+V _{fin}	RAN	IDA	SAA	KES	MUL	KIR	SET	TAR	LÄÄ	VÕR
inf_andma	0	0	0	4	0	0	0	0	0	0
inf_julgema	0	0	0	2	0	0	0	0	0	0
inf_jõudma	0	4	0	0	0	0	0	7	0	0
inf_käskima	0	0	2	0	0	0	0	0	0	0
inf_laskma	0	8	5	4	0	0	0	0	2	0
inf_lubama	0	0	2	0	0	0	0	0	0	0
inf_oskama	0	0	1	0	0	0	0	0	2	0
inf_saama	19	14	13	0	0	47	9	11	25	6
inf_tahtma	10	7	9	9	0	0	0	0	12	0
inf_tohtima	0	0	5	3	0	0	0	0	0	0
inf_tulema	6	0	4	9	7	0	0	0	4	5
inf_täidima	6	0	0	0	0	0	0	0	0	0
inf_viitsima	0	0	1	0	0	0	0	0	0	0
inf_võima	0	8	5	9	0	0	0	0	2	0
ma_ajama	0	0	5	9	0	0	0	0	6	0
ma_hakkama	173	124	215	165	34	134	108	82	140	108
ma_heitma	0	0	0	0	0	0	0	4	0	0
ma_juhtuma	0	0	4	5	0	0	0	0	8	0
ma_jääma	12	0	7	12	0	12	0	13	9	0
ma_kippuma	0	0	0	0	0	0	0	0	2	0
ma_kukkuma	0	0	2	0	0	0	0	0	0	0
ma_minema	52	54	40	51	21	35	114	51	43	50
ma_panema	19	24	30	39	12	12	40	23	47	24
ma_pidama	96	84	176	167	162	184	65	61	153	80
ma_saatma	0	0	1	0	0	0	0	0	0	0
ma_tulema	20	26	12	15	0	0	28	20	0	17
ma_õpetama	0	0	0	0	0	0	0	0	2	0
ma_õppima	0	0	0	2	0	0	0	0	0	0
mas_käima	26	33	44	23	14	27	0	4	34	7
mas_olema	0	0	0	0	0	0	0	0	0	28
mast_tulema	0	0	0	0	0	6	0	0	0	0
mata_jääma	6	7	3	0	0	0	0	6	3	8

Inf+V _{fin}	RAN	IDA	SAA	KES	MUL	KIR	SET	TAR	LÄÄ	VÕR
nud_kaema	0	0	0	0	0	0	0	0	0	5
nud_kustuma	0	0	0	0	0	0	0	0	2	0
nud_kuulma	0	0	0	0	0	0	0	0	2	0
nud_magama	0	0	3	2	0	0	0	0	0	0
nud_naerma	0	0	0	4	0	0	0	0	0	0
nud_olema	515	512	622	487	1030	291	162	424	555	431
nud_rääkima	0	10	7	9	0	0	0	0	11	0
nud_surema	6	0	0	0	0	0	0	0	0	0
nud_taidma	6	0	0	0	0	0	0	0	0	0
nud_teadma	0	0	0	0	0	9	0	0	9	0
nud_tegema	0	0	0	9	29	0	0	0	0	0
nud_tooma	0	0	0	4	0	0	0	0	0	0
nud_vaatama	6	0	0	4	5	0	0	0	6	0
nud_valama	0	0	0	2	0	0	0	0	0	0
nud_võtma	0	0	0	6	0	0	0	0	0	0
nud_ütleva	0	0	0	0	0	0	0	0	8	0
tud_heitma	0	0	1	0	0	0	0	0	0	0
tud_hoidma	0	0	1	0	0	0	0	0	0	0
tud_katkuma	0	0	1	0	0	0	0	0	0	0
tud_olema	0	167	0	0	191	115	105	145	0	162
tud_saama	181	28	192	212	21	21	0	16	186	19
tud_teadma	0	7	0	0	0	0	0	0	0	0
tud_tegema	0	0	5	0	0	0	0	0	6	0
v_ajama	0	0	0	0	0	0	9	0	0	0
v_panema	0	0	2	0	0	0	0	0	2	0

ELULOOKIRJELDUS

Nimi: Kristel Uiboaed
Sünniaeg: 9. märts 1984
Kodakondsus: Eesti
Aadress: Tartu Ülikool, eesti ja üldkeeleteaduse instituut,
Jakobi 2-430, 51004 Tartu
Telefon: +372 737 6127
E-post: kristel.uiboaed@ut.ee

Hariduskäik

2008–2013 Tartu Ülikool, Filosoofiateaduskond, Eesti ja üldkeeleteaduse instituut, eesti keele doktorantuur
2011 kevad Groningeni Ülikool, külalisdoktorant
2005–2008 Tartu Ülikool, Filosoofiateaduskond, Eesti ja üldkeeleteaduse instituut, humanitaarteaduse magister (arvutilingvistika)
2007 Jyväskylä Ülikool, Humanitaarteaduskond, vahetusüliõpilane (soome keel)
2006–2007 Tampere Ülikool, Infoteaduskond, vahetusüliõpilane (infoteadus)
2002–2005 Tartu Ülikool, Eesti ja soome-ugri keeleteaduse osakond humanitaarteaduse bakalaureus (arvutilingvistika, kõrvalerialad: soome keel ja kultuur, eesti keel)
1999–2002 Tartu Hugo Treffneri Gümnaasium

Teenistuskäik

2013 Tartu Ülikool, Eesti ja üldkeeleteaduse instituut, nooremteadur
2008–2013 Tartu Ülikool, Eesti ja üldkeeleteaduse instituut, arvutilingvistika ja korpuskeele spetsialist

Publikatsioonid

Uiboaed, Kristel. 2013. Kollostruktsioonilised meetodid ja konstruktsioonilise varieerumise tuvastamine. *Eesti ja soome-ugri keeleteaduse ajakiri ESUKA / Journal of Estonian and Finno-Ugric Linguistics JEF*, 4(1), 185–204.
Uiboaed, Kristel, Cornelius Hasselblatt, Liina Lindström, Kadri Muischnek, John Nerbonne. 2013. Variation of Verbal Constructions in Estonian Dialects. *Literary and Linguistic Computing*, 28(1), 42–62.
Kaalep, Heiki-Jaan, Kadri Muischnek, Kristel Uiboaed, Kaarel Veski. 2010. The Estonian Reference Corpus: its composition and morphology-aware user interface. *Frontiers in Artificial Intelligence and Applications: The Fourth International Conference Human Language Technologies: The Baltic Perspective*, Riga, Latvia, October 7–8, 2010. Skadiņa, Inguna, Vasiljevs, Andrejs (eds.). IOS Press (Human Language Technologies – The Baltic

- Perspective – Proceedings of the Fourth International Conference Baltic HLT 2010), 143–146.
- Uiboed, Kristel. 2010. Statistilised meetodid murdekorpuse ühendverbide tuvastamisel. *Eesti Rakenduslingvistika Ühingu aastaraamat 6*. Eesti Keele Sihtasutus, 307–326.
- Uiboed, Kristel. 2010. Ühendverbid eesti murretes. *Keel ja Kirjandus*, 1, 17–36.
- Lindström, Liina, Liisi Bakhoff, Mari-Liis Kalvik, Anneliis Klaus, Rutt Läänemets, Mari Mets, Ellen Niit, Karl Pajusalu, Pire Teras, Kristel Uiboed, Ann Veismann, Eva Velsker. 2006. Sõnaliigituse küsimusi eesti murrete korpuse põhjal – E. Niit (toim.). *Keele ehe*. Tartu: Tartu Ülikooli eesti keele õppe-
tooli toimetised 30, 154–167.

CURRICULUM VITAE

Name: Kristel Uiboaed
Date of Birth: 9th of March 1984
Citizenship: Estonian
Address: University of Tartu, Institute of Estonian and General Linguistics, Jakobi 2-430, 51004 Tartu
Telephone: +372 737 6127
E-mail: kristel.uiboaed@ut.ee

Education

2008–2013 University of Tartu, Faculty of Philosophy, Institute of Estonian and General Linguistics, PhD student (Estonian Language)
2011 spring University of Groningen, visiting scholar
2005–2008 University of Tartu, Faculty of Philosophy, Institute of Estonian and General Linguistics Master of Arts in Humanities (Computational Linguistics)
2007 University of Jyväskylä, exchange student (Finnish language)
2006 – 2007 University of Tampere, exchange student (information studies)
2002–2005 University of Tartu, Faculty of Philosophy, Department of Estonian and Finno-Ugric linguistics, Bachelor of Humanities (major Computational Linguistics, minors Finnish language and Culture; Estonian language)
1999–2002 Tartu Hugo Treffner Gymnasium

Professional Employment

2013 University of Tartu, Department of Estonian and General Linguistics, Junior Research Fellow
2008–2013 University of Tartu, Department of Estonian and General Linguistics, specialist of computational linguistics and corpora

Publications

Uiboaed, Kristel. 2013. Kollostruktsioonilised meetodid ja konstruktsioonilise varieerumise tuvastamine. *Eesti ja soome-ugri keeleteaduse ajakiri ESUKA / Journal of Estonian and Finno-Ugric Linguistics JEF*, 4(1), 185–204.
Uiboaed, Kristel, Cornelius Hasselblatt, Liina Lindström, Kadri Muischnek, John Nerbonne. 2013. Variation of Verbal Constructions in Estonian Dialects. *Literary and Linguistic Computing*, 28(1), 42–62.
Kaalep, Heiki-Jaan, Kadri Muischnek, Kristel Uiboaed, Kaarel Veskis. 2010. The Estonian Reference Corpus: its composition and morphology-aware user interface. *Frontiers in Artificial Intelligence and Applications: The Fourth International Conference Human Language Technologies: The Baltic Perspective*, Riga, Latvia, October 7–8, 2010. Skadiņa, Inguna, Vasiljevs, Andrejs (eds.). IOS Press (Human Language Technologies – The Baltic

- Perspective – Proceedings of the Fourth International Conference Baltic HLT 2010), 143–146.
- Uiboed, Kristel. 2010. Statistilised meetodid murdekorpuse ühendverbide tuvastamisel. *Eesti Rakenduslingvistika Ühingu aastaraamat 6*. Eesti Keele Sihtasutus, 307–326.
- Uiboed, Kristel. 2010. Ühendverbid eesti murretes. *Keel ja Kirjandus*, 1, 17–36.
- Lindström, Liina, Liisi Bakhoff, Mari-Liis Kalvik, Anneliis Klaus, Rutt Läänemets, Mari Mets, Ellen Niit, Karl Pajusalu, Pire Teras, Kristel Uiboed, Ann Veismann, Eva Velsker. 2006. Sõnaliigituse küsimusi eesti murrete korpuse põhjal – E. Niit (toim.). *Keele ehe*. Tartu: Tartu Ülikooli eesti keele õppe-
tooli toimetised 30, 154–167.

DISSERTATIONES PHILOLOGIAE ESTONICAE UNIVERSITATIS TARTUENSIS

1. **Ülle Viks.** Eesti keele klassifikatoorne morfoloogia. Tartu, 1994.
2. **Helmi Neetar.** Deverbaalne nominaaltuletus eesti murretes. Tartu, 1994.
3. **Ülo Valk.** Eesti rahvausu kuradi-kujutelm. Tartu, 1994.
4. **Arvo Eek.** Studies on quantity and stress in Estonian. Tartu, 1994.
5. **Reet Kasik.** Verbid ja verbaalsubstantiivid tänapäeva eesti keeles. Tartu, 1994.
6. **Silvi Vare.** Nimi- ja omadussõnatuletus tänapäeva eesti kirjakeeles. Tartu, 1994.
7. **Heiki-Jaan Kaalep.** Eesti keele ressursside loomine ja kasutamine keele- tehnoloogilises arendustöös. Tartu, 1998.
8. **Renate Pajusalu.** Deiktikud eesti keeles. Tartu, 1999.
9. **Vilja Oja.** Linguistic studies of Estonian colour terminology. Tartu, 2001.
10. **Külli Habicht.** Eesti vanema kirjakeele leksikaalsest ja morfosüntaktilisest arengust ning Heinrich Stahli keele eripärast selle taustal. Tartu, 2001.
11. **Pire Teras.** Lõunaeeesti vokaalisüsteem: Võru pikkade vokaalide kvaliteedi muutumine. Tartu, 2003.
12. **Merike Parve.** Väited lõunaeeesti murretes. Tartu, 2003.
13. **Toomas Help.** Sõnakeskne keelemudel: Eesti regulaarne ja irregulaarne verb. Tartu, 2004.
14. **Heli Laanekask.** Eesti kirjakeele kujunemine ja kujundamine 16.–19. sajandil. Tartu, 2004.
15. **Peeter Päll.** Võõrnimed eestikeelses tekstis. Tartu, 2005.
16. **Liina Lindström.** Finiitverbi asend lauses. Sõnajärg ja seda mõjutavad tegurid suulises eesti keeles. Tartu, 2005.
17. **Kadri Muischnek.** Verbi ja noomeni püsiühendid eesti keeles. Tartu, 2006.
18. **Kanni Labi.** Eesti regilaulude verbisemantika. Tartu, 2006.
19. **Raili Pool.** Eesti keele teise keelena omandamise seaduspärasusi täis- ja osasihitise näitel. Tartu, 2007.
20. **Sulev Iva.** Võru kirjakeele sõnamuutmissüsteem. Tartu, 2007.
21. **Arvi Tavast.** The translator is human too: a case for instrumentalism in multilingual specialised communication. Tartu, 2008.
22. **Evar Saar.** Võrumaa kohanimed analüüs enamlevinud nimeosade põhjal ja traditsioonilise kogukonna nimesüsteem. Tartu, 2008.
23. **Pille Penjam.** Eesti kirjakeele *da-* ja *ma-*infinitiiviga konstruktsioonid. Tartu, 2008.
24. **Kristiina Praakli.** Esimese põlvkonna Soome eestlaste kakskeelne keele- kasutus ja koodikopeerimine. Tartu, 2009.
25. **Mari Mets.** Suhtlusvõrgustikud reaajas: võru kõnekeele varieerumine kahes Võrumaa külas. Tartu, 2010.

26. **Karen Kuldnohk.** Militaarne retoorika. Argumentatsioon ja keeleline mõjutamine Eesti kaitsepoliitilises diskursuses. Tartu, 2011.
27. **Kai Tafenau.** Uue Testamendi tõlkimisest Rootsi ajal: käsikirjad, tõlkijad ja eesti kirjakeel. Tartu, 2011.
28. **Külli Prillop.** Optimaalsusteoreetiline käsitus eesti keele fonoloogilisest kujunemisest. Tartu, 2011, 261 lk.
29. **Pärtel Lippus.** The acoustic features and perception of the Estonian quantity system, Tartu, 2011, 146 p.
30. **Lya Meister.** Eesti vokaali- ja kestuskategooriad vene emakeelega keelejuhtide tajus ja häälduses. Eksperimentaalfoneetiline uurimus. Tartu, 2011, 145 lk.
31. **Kersti Lepajõe.** Kirjand kui tekstiliik. Riigieksamikirjandite tekstuaalsed, retoorilised ja diskursiivsed omadused. Tartu, 2011, 141 lk.
32. **Tiit Hennoste.** Grammatiliste vormide seoseid suhtlustegevustega eesti keelses suulises vestluses. Tartu, 2013, 236 lk.
33. **Helena Metslang.** Grammatical relations in Estonian: subject, object and beyond. Tartu, 2013, 443 p.