

AMBIPOSITSIOONIDE MORFOSÜNTAKTILISE VARIEERUMISE KIRJELDUSI KVANTITATIIVSETE PROFIILIDE ABIL

MIRJAM RUUTMA, AKI-JUHANI KYRÖLÄINEN,
MAARJA-LIISA PILVIK, KRISTEL UIBOAED

Eesti keeles on mitmeid kaassõnu, mis võivad esineda nii ees- kui ka tagasõnadena. See on tüpoloogiliselt üsnagi haruldane, kuid eesti keeles (Ehala 1995) ja ka teistes läänemeresoome keeltes (Grünthal 2003; Huumo 2013; Lehis mets 2014) ja saami keeltes (Janda jt 2014) siiski väga tavaline, ning pole ilmset seletust, miks mõned adpositsioonid on selliseks kujunenud ja millest oleneb selliste „kahepaiksete” kaassõnade esinemine kord pre-, kord postpositsioonina. Artiklis avame nähtuse tagamaid viie sellise adpositsiooni (*läbi, mööda, vastu, üle ja ümber*) kasutuskontekstide abil, kasutades eesti murrete materjali (st mittestandardset keelt). Lähenemisviis võimaldab liita morfosüntaktilist informatsiooni sõna kirjeldusse ja esitada sõna kohta mitmekülgsemat empiirilist infot ning on sellisena keelekirjelduses rakendatav mujalgi, näiteks sõnaraamatute ja grammatikate koostamisel, tõlkimisel või lihtsalt näitelausete valikul, samuti keeleõppes.

Kõikide kaassõnade n-õ kahepaiksust on üheselt raske seletada. Janda jt (2014) järgi võivad ambipositsioonide¹ paiknemist mõjutada näiteks keelekonfliktid, kaassõnaga väljendatav suhe, keele tüpoloogilised eripärad ning register. Eesti murrete uurimus (Ruutma 2014) osutab, et vaadeldava viie kaassõna esinemine ees- või tagasõnana sõltub kõige rohkem konkreetsest kaassõnast ning alles seejärel tulevad positsiooni ennustamisel mängu teised keelesised ja -välised tegurid. Sellele toetudes vaatlemegi artiklis viit kaassõna eraldi ning tuvastame tegureid, mis vastava sõna paiknemist võiksid mõjutada ning mõõta ka nende tegurite olulisust. Töö lähtub tänapäevase murdesüntaksi uurimise põhimõtetest (Szmrecsanyi, Wolk 2011; Szmrecsanyi 2014a; 2014b), mille järgi sobib süntaksi põhi- ja erijoonte tuvastamiseks kõige paremini suur korpusandmestik, mis omakorda võimaldab vaadelda varieerumist kvantitatiivselt. Viimane on oluline, sest süntaktilist varieerumist peetakse mittediskreetseks ning piirkondlikud erinevused avalduvad pigem sagedusmustrites kui lihtsalt ühe või teise süntaktilise joone ilmnemises või mitteilmnemises ühel kindlal alal (Kortmann 2010: 842). Tänapäevases murdesüntaksis seatakse enamasti eesmärgiks sagedaste ja tavaliste keelenähtuste kirjeldamine, vastandudes sellega n-õ vana kooli dialektoloogiale, kus tähelepanu fookus on traditsiooniliselt olnud eripärastel ja harva esinevatel, enamasti leksikaalsetel ja fonoloogilistel nähtustel (Szmrecsanyi 2014b). Eesti keeles võiksid sel-

¹ Terminit *ambipositsioon* (ingl *ambiposition*) on soovitanud kasutada Claude Hagège (2010: 114). Inglisekeelses kirjanduses esineb ka nimetus *biposition* (nt Ehala 1995; Grünthal 2003; Huumo 2013). Eestikeelses kirjanduses ei ole nähtuse tähistamiseks kindlat terminit välja kujunenud.

lised sagedased ja tavalised uurimisobjektid olla näiteks mineviku liitaegade, kaassõnade, erinevate käänete kasutus vmt. Sagedus- ja korpuspõhine murdeuurimine on traditsioonilises dialektoloogias olnud tagaplaanil ning seetõttu on sagedasti esinevate konstruktsioonide ja nende varieerumismustrite kohta tegelikult kõige vähem informatsiooni.

Järgnevalt tutvustame lühidalt, mida peetakse traditsiooniliselt kõnealuste sünonüümsete grammatiliste konstruktsioonide vaheldumist mõjutavateks teguriteks, ning anname lühikese ülevaate viiest uuritavast adpositsioonist eelkõige nende abil väljendatavate funktsioonide ja nende komplemendi käänete kaudu. Seejärel selgitame üht kvantitatiivset uurimise meetodit (klassifitseerimispuud) kasutades, kas ja kui palju murdeandmestikus kodeeritud keelesisesed ja/või -välised tegurid vaadeldava ambipositsiooni paiknemist mõjutavad, ning täiendame seeläbi iga uuritava kaassõna profiili empiirilise infoga. Mõiste *profiil* all ei pea me siinses artiklis silmas ühtki konkreetset terminit (nt Gries, Divjak 2009) ega meetodit (Neset jt 2011; Janda, Lyashevskaya 2013), vaid selle üldkeelset tähendust nii, nagu see on esitatud EKSS-is: **profiil** [---] 4. tüüpiliste põhijoonte kogum; iseloomulik suund, tegevusala, kallak; omapära.

1. Ambipositsioonidest eesti keeles

Eesti keeles võib kaassõnad jagada paiknemise järgi kolme rühma: prepositsioonid, postpositsioonid ja mõlema positsioonivõimalusega adpositsioonid ehk ambipositsioonid (EKG I: 37; Palmeos 1982; Grünthal 2005: 31). Ambipositsioonide osakaal eesti adpositsioonide hulgas pole suur, vaid umbes 10 %: eesti keele seletava sõnaraamatu veebiversioonis (vaadatud 27. XII 2014) on kokku 267 adpositsiooni, millest postpositsioone on 192, prepositsioone 49 ning ambipositsioone 26. Umbes samas järgus (13 %) on ambipositsioone ka soome keeles, põhjasaami keeles on aga süstemaatiliselt nii ees- kui ka tagasõnadena kasutatavate kaassõnade osakaal suisa 22 % kõigist kaassõnadest (Janda jt 2014: 91–92).

Pre- ja postpositsioonide kujunemisele on esitatud mitmeid selgitusi (nt Grünthal 2003, 2005; Palmeos 1982), kuid sellele, miks mõned adpositsioonid on eesti keeles kujunenud ambipositsioonilisteks ja eiravad muidu üsna järjepidevat adpositsioonifraasi süntaktilist struktuuri, pole ilmset ning vastuoludeta selgitust leitud.

Ühe võimaliku seletuse ambipositsioonide ilmnemisele on välja pakkunud Martin Ehala (1995), kelle järgi võib adpositsioonidel olla kaks paiknemisvõimalust vaid sel juhul, kui kaassõna nõuab pre- ja postpositsioonina komplemendilt erinevat käänet (nt *vastu katust, katuse vastu*). Vastasel juhul tekiks adpositsiooni kahe võimaliku paiknemisega liigset mitmetähenduslikkust (Ehala 1995: 328–338). Mõnede eesti keele ambipositsioonide (*pärast* ja *vastu*) paiknemine ongi selgelt seotud nende komplemendi käändega: prepositsioonina esinedes nõuab ambipositsioon oma komplemendilt partitiivi (*pärast tööd, vastu puud*) ning postpositsioonina esinedes genitiivi (*töö pärast, puu vastu*). Ka pre- või postpositsioonilised adpositsioonifraasid on oma morfosüntaktilise ülesehituse poolest samasugused (Grünthal 2003: 84; EKG I: 37–38). Adpositsioonifraasi kaks põhilist struktuuri on seega järgmised:

- 1) prepositsioonifraas [pre + N[par]]
- 2) postpositsioonifraas [N[gen] + post]

Neid tüüpilisi adpositsioonifraasi mustreid järgivad enamasti ka ambipositsioonid *alla*, *kaudu* ja *ligi*, ent nende puhul on võimalikud ka vastupidised kombinatsioonid, nt *alla selle*, *maanteed kaudu*, *maad ligi*. Ambipositsioonid *ümbert*, *pealt* ja *üle* omakorda nõuavad oma komplemendilt mõlemas positsioonis genitiivi, nt *ümbert käe* ja *käe ümbert*, ambipositsioon *mööda* jällegi partitiivi, nt *mööda teed*, *teed mööda* (Palmeos 1982: 50, 55–59; EKG I: 37). Alati ei ole ambipositsioonide jagunemine rektsiooni järgi pre- ja postpositsioonideks niisiis sugugi reeglipärane. Riho Grünthal (2003: 84) ongi väitnud, et kaassõnafraasi sõnajärg on olnud algselt vaba ning adpositsiooni paiknemisega sobituv komplemendi käänne on keeles hiljem kujunenud. Seega võiksid n-õ tavalisest kaassõnafraasi struktuurist erinevalt käituvad kaassõnad olla jäänud vanadest kaassõnafraasi mustritest. Samas on Grünthal (2005: 32–33) täheldanud, et kui kaassõna ei nõua komplemendilt pre- ja postpositsioonina varieerudes erinevat käännet, peaks genitiiv lisama postpositsioonilise sõnajärje eelistust ja partitiiv prepositsioonilise sõnajärje eelistust. Sellise sõnajärje eelistuse põhjuseks võib olla soome-ugri keeltele omane possessiivse konstruktsiooni sõnajärg, kus genitiivne täiend peabki eelnema oma põhisõnale (Grünthal 2003: 36). Jääb aga siiski üsna ebaselgeks, miks on mõned kaassõnad oma komplemendi käännet süstemaatilisemalt „vorminud”, kui jälle teised näivad kombineeruvat endiselt üsna vabalt nii mõlema käändega samas positsioonis (nt *tänavat kaudu*, *tänavat kaudu*) kui ka sama käändega mõlemas positsioonis (nt *ümbert käe*, *käe ümbert*).

Kui jätta kaassõnafraasi sõnajärje ennustamisel kõrvale käändeline ühildumine vastavalt kaassõna positsioonile, on ambipositsioonifraasi sõnajärg seotud ka komplemendi pikkusega. Kui komplement on pika struktuuriga (nt sisaldab lisaks noomenile ka täiendeid, kusjuures täiendiks võib olla ka relatiivlause), on adpositsioonidel kalduvus esineda prepositsioonidena (Huumo 2013). Selle väite kontrollimiseks murdematerjali põhjal oleme komplemendi pikkust analüüsinud ka käesolevas töös (täpsemalt vt ptk 4).

Ambipositsioonide paiknemise oluliseks ennustajaks on sageli ka adpositsiooni tähendus ehk adpositsiooni ja selle komplemendi vahelise suhte tüüp (Huumo 2013; Grünthal 2003). Ambipositsioonil võib pre- ja postpositsioonina esinedes olla erinevaid tähendusi (nt väljendab prepositsioonifraas *pärast abiellumist* ajalist, ent postpositsioonifraas *abiellumise pärast* põhjuslikku suhet), kuid sageli püsib see mõlemas positsioonis ka samana, nt väljendavad *ümbert maja* ja *maja ümbert* mõlemad ruumilist suhet (Palmeos 1982; EKG I: 37; vt ka ptk 2). Kaassõna positsiooni sidumisel selle väljendatud suhte tüübiga on läänemeresoome ja saami keelte puhul varasemalt täheldatud, et kui ambipositsiooniga võib märkida nii ajalist kui ka ruumilist suhet, on prepositsioonide kasutus seotud rohkem ajasuhte väljendamisega ning postpositsioonid on tavalisemad ruumilistes tähendustes (Grünthal 2005: 33; Huumo 2013: 320; Janda jt 2014: 91, 100). Kui arvestada, et postpositsioonid on tekkelt vanemad kui prepositsioonid ning on esmalt väljendatud vaid ruumisuhet (Palmeos 1982: 69–70; Grünthal 2003, 2005), võikski ruumilise tähenduse väljendamine olla seotud vanema postpositsioonilise struktuuriga. Ometi ei ole eesti keeles adpositsioonide ning nende väljendatavate suhete seost täpsemalt vaadeldud.

Käesolevas töös oleme kaasanud analüüsi adpositsiooni tähenduse kui ühe võimaliku teguri, mis ambipositsiooni paiknemist mõjutab (täpsemalt vt ptk 4 ja 5).

Kaassõnade esinemist mingis keelele ebatüüpilisemas positsioonis (eesti keeles nt prepositsioonidena) on selgitatud ka võõrmõjude ja keelekontaktidega (nt Grünthal 2005: 30; Janda jt 2014), teisalt on väidetud, et üldiselt pole kaassõnad võõrmõjutustele sugugi vastuvõtlikud (Lass 1997: 190; Grünthal 2003). Näitena on Grünthal (2003: 53) toonud liivi keele, kus läti keele mõjud ilmnevad nii leksikaalsel, foneetilisel kui ka süntaktilisel tasandil, kuid kaassõnafraasi struktuur näib olevat märkimisväärse kontaktkeele mõjuta. Eesti keele kohta on Mati Hint (1990) väitnud, et prepositsioonirohke vene keel on suurendanud eesti keele mõlema paiknemisvõimalusega kaassõnade prepositsioonilise kasutuse osakaalu, kuid selle on kahtluse alla seadnud Ehala (1994: 177–178), kes toob välja, et näiteks hoolimata pikemast ja varasemast kontaktist samuti põhiliselt prepositsioone kasutava saksa keelega, pole tolle perioodi eesti keeles kaassõnade kasutusel sarnaseid tendentse täheldatud. Seega võiks eeldada, et sama kaassõna kasutus pre- või postpositsioonina võiks olla tingitud pigem keelesisestest teguritest ning vähem mõjutatud kontaktkeeltest.

Järgnevas peatükis kirjeldame lähemalt uuritavaid ambipositsioone ja nende tähendusi.

2. läbi, mööda, vastu, üle ja ümber

Kaassõnu on eesti keeles sageli uuritud ühe kaupa, erinevatest teoreetilistest lähtepunktidest või erineva rõhuasetusega (nt tähenduse prototüüpsus, grammatiseerumine, vt Ehala 1994, 1995; Ehasalu jt 1997; Habicht 2000; Habicht, Penjam 2007; Veismann 2009 jt). Ülevaatlikumaid käsitlusi leiab mitmetest grammatikatest (nt EKG I: 33–39; EKG II: 137–139; Palmeos 1982) ja mõnest monograafiast (Grünthal 2003). Artiklis kasutame uuritavate ambipositsioonide sagedasematest funktsioonidest ja fraasistruktuurist ülevaate andmiseks põhiliselt Paula Palmeose (1982) süsteemset ja põhjalikku kaassõnade käsitlust, milles erinevalt teistest allikatest on tehtud kaassõnade kirjeldamisel tähelepanekuid ka murrete kohta, ning eesti keele akadeemilist grammatikat (EKG I; EKG II). Peatükis esitatud näited on autorite omad. Kokkuvõtlikult on info esitatud tabelis 1.

Ambipositsiooni *läbi* komplement esineb nii pre- kui ka postpositsiooniga genitiivis, ainsa erandina ajasuhet tähistades, kui postpositsioonilise *läbi* komplement on nominatiivis (*suvi läbi*).² Prepositsioonina väljendab *läbi* aja- (*läbi suve*), ruumi- (*läbi akna*) ja viisisuhet (*läbi häda ja vaeva*). Samad tähendused on sõnal ka postpositsioonina esinedes, kuid lisaks sellele väljendab *läbi* postpositsioonina ka vahendi (*käe läbi*), abinõu (*õnne läbi*) ja põhjuse (*ülekohtu läbi*) suhteid (Palmeos 1982: 25–39, 60–61; EKG I: 35, 37; EKG II: 138–139).

² EKG (II: 139) on nimetanud ka aja- või kohatähendust märkiva postpositsioonilise *läbi* komplementi elatiivis esinemise võimalust (nt *jooksis koridorist läbi*). Sellistel juhtudel pole aga sugugi selge, kas tegu on postpositsiooni või afiksaaladverbiga, samuti pole elatiivse komplementiga *läbi*-sõnu postpositsioonideks märgitud murdekorpuses, mistõttu selles artiklis me neid ühendeid kaassõnafraasideks ei loe.

Ambipositsioonide funktsioonid

	prepositsioon + N[gen]	prepositsioon + N[par]	prepositsioon + N[par]	N[par] + postpositsioon	N[nom] + postpositsioon
<i>läbi</i>	AEG VIIS KOHT	–	VAHEND VIIS SEISUND KOHT ABINÕU PÕHJUS	–	AEG
<i>mööda</i>	–	KOHT	–	KOHT viis	–
<i>vastu</i>	–	KOHT AEG viis	KOHT VIIS SUHTUMINE SUHE ASENDUS	–	
<i>üle</i>	KOHT AEG kvantiteet VIIS seisund	–	KOHT PÕHJUS KVANTITEET	–	–
<i>ümbes</i>	KOHT	–	KOHT LIGIKAUDSUS	–	–

mööda nõuab nii pre- kui ka postpositsioonina oma komplementilt partitiivi. Mõlemas positsioonis võib *mööda* väljendada ruumisuhet, kusjuures kaasõna erinev paiknemine ei muuda selle tähendust (*mööda teed ~ teed mööda*). Postpositsioonina võib *mööda* tähistada ka viisisuhet (*enese heaksarvamist mööda*) (Palmeos 1982: 26, 53–54; EKG I: 35, 37; EKG II: 139).

Adpositsioonifraas, mis sisaldab ambipositsiooni *vastu*, järgib tüüpilist pre- või postpositsioonifraasi ülesehitust. Prepositsioonina on *vastu* komplement partitiivis ning postpositsioonina genitiivis. Prepositsioonina tähistab

vastu ruumi- (*vastu akent*), aja- (*vastu õhtut*) ja viisisuhet (*vastu tahtmist*), postpositsioonina annab aga lisaks ruumi- (*akna vastu*) ja viisisuhte (silm silma vastu) edasi ka suhtumist (*minu vastu ollakse kuri*), suhteid (*ei tunne tema vastu huvi*) või asendust (*vahetab kinga teise vastu*) (Palmeos 1982: 23–47, 58–59; EKG I: 39; EKG II: 138–139).

Pre- ja postpositsioonilise *üle* komplement esineb alati genitiivis. Prepositsioonina väljendab *üle* ruumi- (*üle tee*), aja- (*üle aasta*), kvantiteedi- (*üle kahe kilo*) ja viisi/seisundi- (*üle jõu*) tähendust. Postpositsioonina esineb see ruumi- (*maja üle ripuvad oksad*), põhjuse- (*hea meel uudise üle*) ja harva ka kvantiteedisuhte tähistajana (*viie kilo üle*) (Palmeos 1982: 24–43, 50; EKG I: 35–37; EKG II: 139).

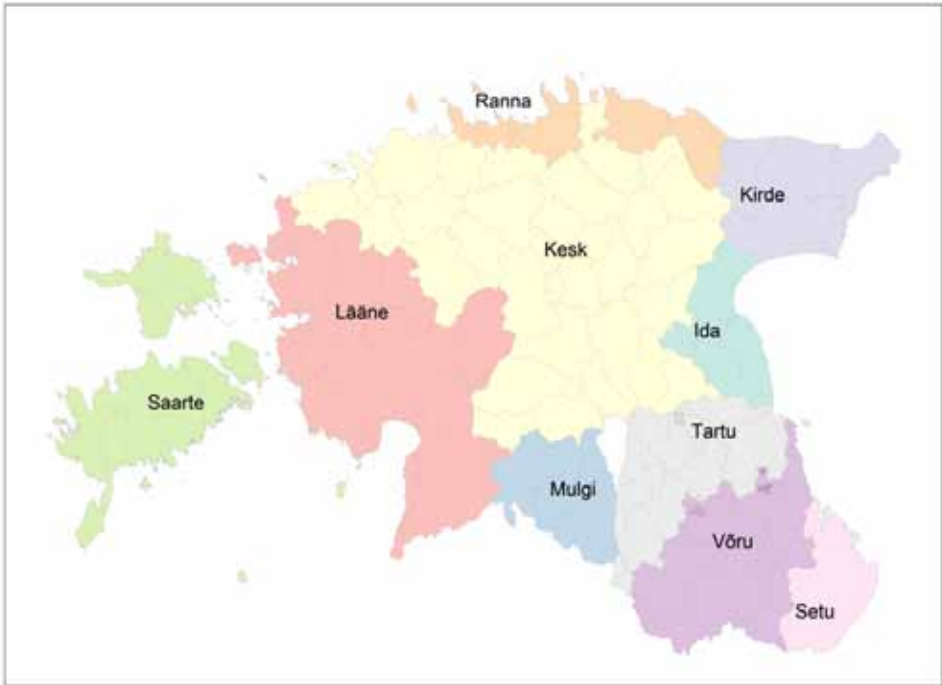
Ambipositsioon *ümb*er nõuab samuti oma komplemendilt mõlemas positsioonis alati genitiivi. Ruumisuhet väljendatakse pre- või postpositsiooniga (*ümb*er aia ~ aia *ümb*er), kuid ligikaudsust väljendatakse vaid postpositsiooniga (*kahe kilo ümb*er) (Palmeos 1982: 24–42, 50–51; EKG I: 36; EKG II: 139).

3. Andmestik

Uurimuse andmestik on pärit eesti murrete morfoloogiliselt märgendatud korpusest (EMK 2014). Murdekorpuses on esindatud kõik eesti murdealad ning korpuse koostamise üks eesmärkidest ongi luua eeldused murrete kontrastiivseks uurimiseks. Korpuse tekstid põhinevad traditsiooniliste murdeintervjuude lindistustel, kus keelejuhti küsitletakse talle tuttavalt territooriumil ning intervjuude teemad puudutavad minevikusündmusi, tööde tegemist, eluolu jne. Lindistused on peamiselt pärit 1960.–1970. aastatest. Uurimismaterjali kogumise hetkel (september 2014) oli korpuses üle 830 000 morfoloogiliselt märgendatud sõna. Joonis 1 esitab murdekorpuse aluseks oleva murdejaotuse. Traditsiooniliselt jaotatakse eesti murded lõuna- ja põhjaeesti murderühmaks. Esimesse kuuluvad Mulgi, Tartu, Võru ja Setu murre ning teise rühma moodustavad ülejäänud murded: saarte, lääne, kesk, ida-, ranna- ja kirdemurre.

Põhja- ja lõunaeesti murderühma vahelised erinevused on suurimad fonoloogilisel, morfoloogilisel ja leksikaalsel tasandil. (Morfo)süntaktilisi jooni traditsioonilised murdejaotused arvesse ei võta, mistõttu pole eriti palju infot, kas murded jaguneksid sellel tasandil kuidagi teisiti. Käesolev töö ja rakendatav meetod (vt ptk 5) annavad infot ka selle kohta, kas samasugune murdejaotumine toimub ambipositsioonide kui morfosüntaktilise nähtuse tasandil.

Uurimismaterjal on kogutud murdekorpuse morfoloogiliselt märgendatud osast automaatselt R-skriptidega (R Development Core Team). Esialgne andmestik koosnes kõigist korpusest tuvastatud pre- ja postpositsioonidest, millest omakorda valiti mõlemas positsioonis esinemise sageduse järgi viis uuritavat kaassõna. Lisaks lemma infole saime ära kasutada murdekorpuse pre- ja postpositsiooni märgendiga tähistatud osa. Iga kaassõna jaoks kaassime viiesõnalise konteksti vasakult ja paremalt. Kõik kontekstiüksused on läbi vaadatud käsitsi, selle käigus on eemaldatud märgendusvead (enamasti kaassõnadeks märgendatud afiksaaladverbid). Kokku jäi andmestikku 1062 vaatlust ehk kaassõnade kasutuskorda. Nende uuritavad tunnused on samuti kodeeritud käsitsi ning kodeerimisskeemi kirjeldabki täpsemalt järgnev peatükk.



Joonis 1. Murdekorpuses kasutatav murdejaotus.³

4. Andmestiku kodeerimine

Kõik uuritavate kaassõnade esinemised koos oma kontekstiga on läbi vaadatud ning võimalikud ambipositsioonide paiknemist mõjutavad tegurid on kodeeritud käsitsi. Järgnevalt annamegi ülevaate kõigist andmestikus kodeeritud tunnustest. Esitame iga tunnuse tasemed ning põhjused, miks oleme eeldanud, et tunnus võib kaassõna paiknemist mõjutada. Kodeeritud seletavad tunnused on valitud varasemate uurimuste ja autorite oletuste põhjal (täpsemalt iga tunnuse kirjelduse juures).

Adpositsiooni asend (ADPOSITSION). See on analüüsitav tunnus, mille varieerumist ennustame ülejäänud kodeeritud (kirjeldavate) tunnuste abil. Uuritav tunnus on binaarne: ambipositsioon esineb prepositsioonina või postpositsioonina.

Murre (MURRE). Tähistab murdeala, kust analüüsitav lause pärit on. Kasutame murdekorpuse aluseks olevat murdejaotust, mis eristab 10 murdeala: idamurre (*ida*), keskmurre (*kes*), kirdemurre (*kir*), läänemurre (*laa*), Mulgi murre (*mul*), rannamurre (*ran*), saarte murre (*saa*), Setu murre (*set*), Tartu murre (*tar*) ja Võru murre (*vor*). Tunnuse kaasamise eesmärk on uurida ambipositsiooni kasutuse geograafilist varieerumist.

³ Kaart on muudetud versioon Eesti kihelkondade kaardist (EKI 2014).

Komplemendi kääne (KOMPL_KAANE). Kuigi tüüpiliselt on prepositsioonid partitiivse ning postpositsioonid genitiivse rektsiooniga, võib ambipositsioonide puhul komplemendi kääne püsida mõlemas positsioonis ka samana (vt täpsemalt ptk 2). Kui erinev paiknemine komplemendi käännet ei muuda, peaks genitiivse rektsiooniga olema tavalisem postpositsioonilisus ja partitiivse rektsiooniga prepositsioonilisus (Grünthal 2005: 32–33). Siinses uurimismaterjalis esinevad komplemendid nominatiivis (*nom*), genitiivis (*gen*) ja partitiivis (*par*).

Adpositsiooni asend verbi suhtes (SONAJARG). Tunnus tähistab seda, kas kaassõna fraas eelneb (*vasak*) või järgneb (*parem*) öeldisverbile. Öeldisverbi puudumist on tähistatud kodeeringuga *ellipsis*.

Lausungit on peetud suulise kommunikatsiooni põhiüksuseks ja tüüpiliselt koosneb see predikaadist ja sellega seotud fraasidest (Thompson, Couper-Kuhlen 2005: 810, 2. alaviide). Keeleteaduses üldlevinud seisukoht on, et fraasi moodustavate elementide järjekord ei ole juhuslik ning keerulisematel fraasidel on kalduvus paikneda lausungi lõpus. Lisaks teatud lausetüüpide tingitud mallidele mõjutab lause moodustajate paiknemist eesti keeles ka lause infostruktuur, mille kohaselt uus info peaks esitatama üldiselt lause lõpu poole. Sellest eeldusest lähtudes võib lause moodustajate järjekord kaudselt öelda üht-teist ka moodustajasiseselt (antud juhul kaassõna fraasi) järjekorra kohta. Nõnda oleme pidanud otstarbekaks kaasata analüüsi seletava tunnuse, mis kontrollib hüpoteesi, kas uue info esitamisel eelistatakse kahe paiknemisvõimalusega kaassõnade puhul (mingitel murdealadel) pigem pre- või postpositsioone.

Komplemendi referendi elusus (ELUSUS). Grammatiliste nähtuste kirjeldamisel on elusust peetud fundamentaalseks kognitiivseks kategooriaks ning paljud uurimused on osutanud selle olulisele mõjule just grammatilistes alternatsioonides, näiteks *s*-genitiivi ja *of*-genitiivi vaheldus inglise keeles (Hinrichs, Szmrecsanyi 2007) või daativi alternatsioon (Bresnan, Ford 2010). Eesti keeles on kaassõna ja käändevaheldust põhjalikult uurinud Jane Klavan (2012), kuid tema uurimuses kaassõna ja käände vaheldumisel elusus statistiliselt oluliseks teguriks ei osutunud (Klavan 2012: 155). Siinses uurimuses testime, kas elusus võiks mõjutada ambipositsiooni paiknemist. Adpositsiooni komplement on liigitatud oma referendi elususe järgi elusaks või elutuks. Liigituse aluseks on võetud Bresnani jt (2007) lihtsustatud binaarne klassifikatsioon: elusatele entiteetidele viitavad sellised noomenid nagu pärisnimed, üldnimed ja asesõnad, lisaks ka noomenid, mis tähistavad organisatsioone, institutsioone ja muid kollektiivseid entiteete, mis võimaldavad teatavaid agendivõimalusi; noomenid, mis ei kuulu eelnevalt nimetatute hulka, on kodeeritud elututele entiteetidele viitavateks.

Komplemendi arv (KOMPL_ARV). Tuomas Huumo (2013) on leidnud soome keele ambipositsioonide puhul, et mida keerukam on komplemendi struktuur, seda tavalisem peaks olema ambipositsiooni prepositsiooniline kasutus. Siinses töös vaatlemegi, kas see vastab ka eesti keele murdematerjalis tõele. Oleme kodeerinud kaassõna fraasi komplemendi arvu ainsuseks (*sg*) ja mitmuseks (*pl*) ning viimast peame eesti keeles grammatiliselt komplekssemaks üksuseks. Oletame, et mitmusevormidega eelistatakse pigem prepositsioone ja ainsusevormidega postpositsioone.

Aeg (AEG). Tunnus tähistab verbi aega. Oleme eristanud oleviku ning mis tahes minevikuvormid. Verbi puudumist tähistab tunnus *ellips*. Sarnaselt pluurali markeeritusele võib ka minevikuvormi käsitleda morfosemantiliselt keerukamana (Dressler 2000: 288). Ehkki verb ei ole ambipositsiooni komplement ning verbivorm ei mõjuta seetõttu pre- või postpositsiooni kasutust nii nagu komplementi struktuur, võib seda pidada üldisemalt konteksti kompleksuse indikaatoriks, mis võib mingil määral mõjutada ka ambipositsiooni paiknemist.

Fraasi pikkus (FRAASI_PIKKUS). Fraasi pikkust (või n-õ raskust) on peetud oluliseks mõjutajaks paljudes eri süntaktilisi nähtusi käsitlevates empiirilistes uurimustes (nt Bresnan jt 2007; Hinrichs, Szmrecsanyi 2007; Klavan 2012; Tagliamonte, Baayen 2012 jne). Seejuures on pikkuse või raskuse all mõeldud fraasi süntaktilist (nt märkide, sõnade arv), fonoloogilist (nt silpide, häälikute arv), semantilist või grammatilist kompleksust (nt markeeritus). Selles töös kasutame fraasi pikkust sõnades. On oletatud, et prepositsioonilisus on tavalisem pikkade fraaside puhul (nt Huumo 2013), seega võiks fraasi kompleksus mõjutada ka uuritavate ambipositsioonide paiknemist.

Adpositsiooni väljendatav suhe (ADP_SUHE). Nagu 3. peatükis näidatud, väljendavad ambipositsioonid lauseliikmete erinevaid suhteid, kusjuures vastavalt ambipositsiooni asendile fraasis võib selle funktsioon muutuda. Seega on adpositsiooni paiknemise oluliseks mõjutajaks just nimelt adpositsiooni väljendatav suhe. Siinsesse töösse on ambipositsioonide erinevad funktsioonid kaasatud põhjusel, et selgitada empiirilisel, kuidas täpselt ambipositsioonide paiknemine ning väljendatav tähendus omavahel seotud on: kas teatud tähendusega eelistatakse üht paiknemisvõimalust teisele isegi sel juhul, kui tähenduse väljendamiseks on võimalikud mõlemad positsioonid, või kas funktsioon on mõne kaassõna puhul oluline paiknemise ennustaja vaid koosmõjus mõne teise seletava muutujaga. Oleme eristanud järgnevad suhted: ABSTRAKTNE (*abs*), AEG (*aeg*), MÕÖDE (*moode*) ja RUUM (*ruum*). Jaotus põhineb Ehala (1994: 184) kaassõnade uurimuse semantilisel liigitusel, kus kaassõnad on jagatud nelja kategooriasse: ruumiline, ajaline, mõõtmeline ja abstraktne. Selline liigitus lihtsustab 3. peatükis kirjeldatud EKG (I: 34–36; II: 137–139) ja Palmeose (1982) kaassõnade funktsioonide detailsemad kirjeldusi ning sobib paremini statistilise andmeanalüüsi vajadustega: kuna mitmed kaassõnade väljendatavad spetsiifilised suhted tulevad andmestikus esile üksikjuhtudel, oleks nende kaasamine kvantitatiivsesse analüüsi keeruline.

RUUM tähistab kohasuhet, AEG ajasuhet. Kategooriasse ABSTRAKTNE kuuluvad põhjust, viisi ja seisundit, vahendit või abinõu, suhtumist või suhet ning asendust väljendavad kaassõnad. MÕÖDE aga on klass, mis koondab kvantiteeti (hulka, mõõtu ja määra) ja ligikaudsust väljendavaid funktsioone.

Tabel 2 esitab kokkuvõtlikult kõigi kodeeritud muutujate ja nende tasemete sagedusjaotused analüüsitud murdeandmestikus. Fraasi pikkuseks on valitud 3–5 sõna ($M = 3,89$ ja $SD = 0,51$).

Tabel 2.

Analüüsitud andmed kodeeritud muutujate kaupa

Muutujad	Sagedusjaotus
Adpositsioon	
postpositsioon (Post)	380
prepositsioon (Pre)	682
Murre	
ida	51
kesk (kes)	182
kirde (kir)	52
lääne (laa)	256
Mulgi (mul)	82
ranna (ran)	85
saarte (saa)	161
Setu (set)	41
Tartu (tar)	85
Võru (vor)	67
Komplemendi kääne	
genitiiv (gen)	760
nominatiiv (nom)	20
partitiiv (par)	282
Sõnajärg	
ellips	49
parem	760
vasak	253
Komplemendi elusus	
elus	57
elutu	1005
Komplemendi arv	
mitmus (pl)	111
ainsus (sg)	951
Aeg	
ellips	49
minevik	761
olevik	252
Adpositsiooni väljendatav suhe	
ABSTRAKTNE (abs)	130
AEG	75
MÕÕDE (moode)	84
RUUM	773

5. Meetod

Keeleteaduslikud, eriti just dialektoloogilised ja sotsiolingvistikud on andmeanalüüsi seisukohalt problemaatilised, reeglina on sageduserinevused väga suured, valimid eri suurusega ning andmepunktide jaotus ebahütlane.⁴ Siin pole erand ka käesoleva uurimuse materjal. Leidmaks vastust küsimustele, millised uuritavatest tunnustest koos või eraldi mõjutavad ambipositsiooni esinemispositsiooni kõige rohkem suulises spontaanses murdematerjalis ning kas eri kaassõnade ja kodeeritud tunnuste lõikes esineb erinevusi ka murretes, rakendasime klassifitseerimispuude meetodit (Breiman jt 1984; Hothorn jt 2006), mis suudab nimetatud probleemidega toime tulla. Meetod põhineb vaatluste n-ö korduval ositamisel: vaatlused jagatakse binaarsetesse klassidesse vastavalt sellele, kui tugevalt mingi kodeeritud seletav tunnus (nt kaassõna funktsioon) uuritava tunnusega (kaassõna paiknemine) seotud on. Esimese jagunemise annab enamasti üks kõige olulisematest seletavatest tunnustest ning selle tunnuse osas sarnased vaatlused grupeeritakse. Seejärel jätkub ositamine juba kahes tekkinud klassis vastavalt järgmistele olulistele seletavatele muutujatele nendes klassides, seejärel nende muutujate põhjal jaotatud klassides jne. Jagunemine kestab seni, kuni edasine jagamine pole enam põhjendatud, st statistiliselt pole järgnevad tunnused enam uuritava muutuja seisukohalt olulised. Protsessi lõpptulemusena peaks kõik vaatlused olema jagatud uuritava tunnuse suhtes optimaalselt rühmadesse (Strobl jt 2009: 325–327). Ehkki klassifitseerimispuid on keeleteaduses kasutatud senini võrdlemisi vähe, on need andnud väga häid tulemusi, võimaldades lihtsalt visualiseerida ka keerulisi interaktsioone (nt Tagliamonte, Baayen 2012; Baayen jt 2013; Kyröläinen 2013).

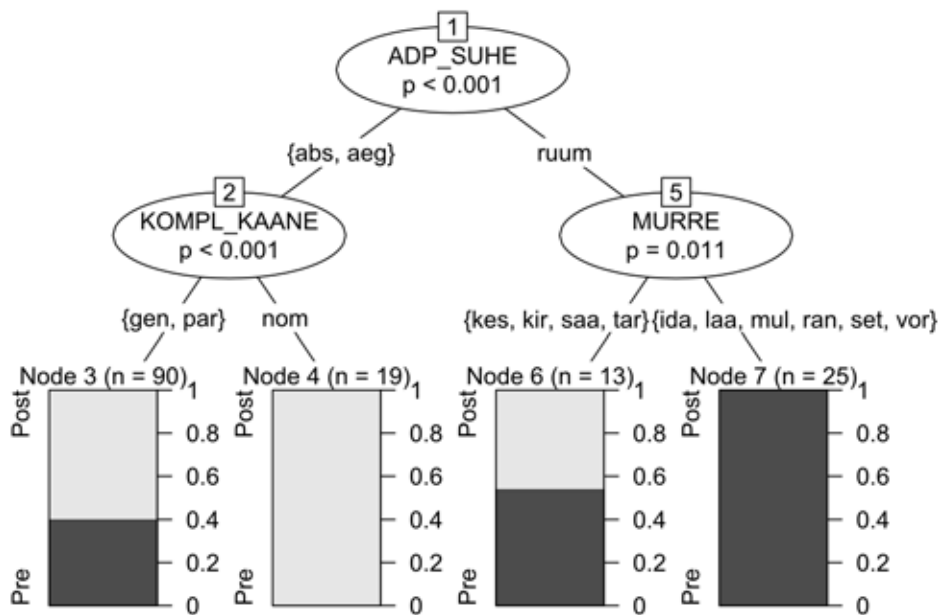
6. Ambipositsioonide klassifitseerimispuud

Käesolevas peatükis kirjeldame, millised olulisemad tegurid ja mil moel üksikute kaassõnade kohta moodustatud klassifitseerimispuudes esile tulid, ning illustreerime tulemusi näitelausetega. Näitelaused on valitud puu hargnemiste illustreerimiseks, ilmestades erinevaid tunnustekombinatsioone.

6.1. läbi

Nagu selgus ka 3. peatükis, on ambipositsioonil *läbi* rohkelt tähendusi ning väljendatavast suhtest sõltub paljuti ka selle adpositsiooni paiknemine. Siinse materjali põhjal on abstraktset ja ajasuhet tähistavate ambipositsioonide paiknemine erinev ruumisuhet tähistavate ambipositsioonide paiknemisest. Esimesel juhul on nominatiivse komplemendiga *läbi* (1) alati postpositsioon, genitiivse ja partitiivse komplemendiga aga vaid veidi üle poolte (60 %) juhtudest postpositsioon (2) ja umbes 40 %-lise osakaaluga prepositsioon (3). Ouline on märkida, et erinevalt murdekorpusest ei esine ei Palmeose (1982) ega EKG käsitluse järgi *läbi* komplement partitiivis.

⁴Neid probleeme on kompaktsetl kirjeldanud Tagliamonte ja Baayen (2012).



Joonis 2. Kaassõna *läbi* paiknemist mõjutavad olulised tegurid klassifitseerimispuus.

- (1) ja nii (.) nii kõik *talv läbi* egä päev üits (.) ütspää+dükk (Tartu)
- (2) nohh eks = ta *sell suure`kurjuse läbi* lolliks ta läks (lääne)
- (3) `õhtu tulime *läbi üüd* tagasi `jälle (saarte)

Ruumisuhet tähistades tulevad ambipositsiooni *läbi* paiknemisel esile murdelised erinevused, ehkki murretest moodustuvad rühmad ei ole traditsiooniliselt ootuspärased põhja- ja lõuna- ega ka ida- ja läänerühm. Kesk-, kirde-, saarte ja Tartu murdes esineb *läbi* pre- ja postpositsioonina võrdselt (4). Ida-, lääne-, Mulgi, ranna-, Setu ja Võru murdes aga on ruumisuhet tähistades *läbi* alati prepositsioon (5).

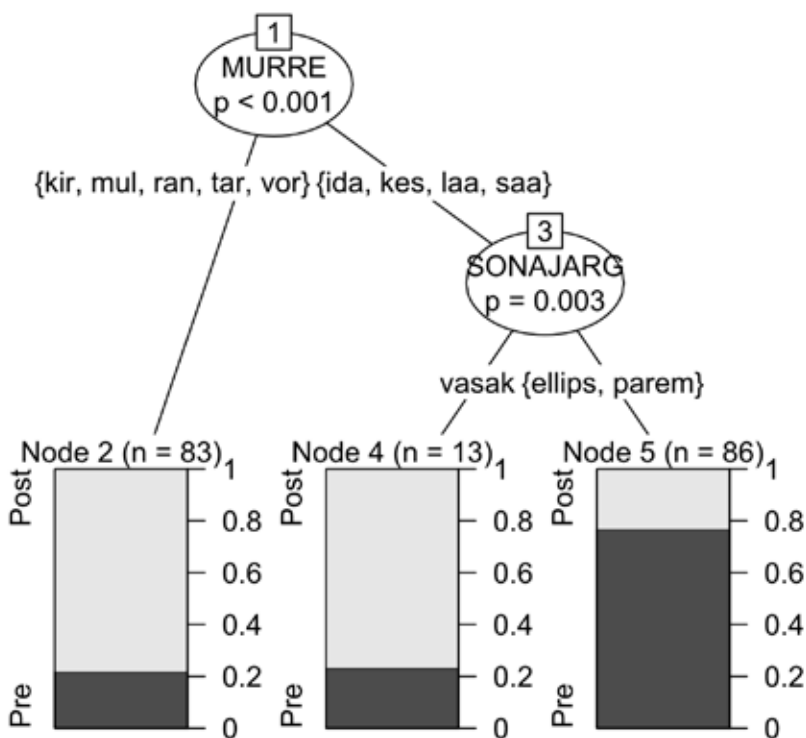
- (4) *tulli sinnä võttap mull käest`kinni ja (.) viib läbi lume* (Tartu)
- (5) *kattõgese poisiq lääväq läbi mõtsa* (Setu)

Niisiis selgub *läbi* puhul, et ruumisuhet tähistades on ambipositsioon enamasti prepositsiooniline ning abstraktses ja aja dimensioonis enamjaolt postpositsiooniline. Nõnda võib täheldada eesti murretes *läbi* paiknemise juures vastupidiseid tendentse kui teistes lähisugulaskeeltes, kus mõlema funktsiooni võimalikkuse korral väljendavad ambipositsioonid ajasuhet peamiselt prepositsioonidena ja ruumisuhet postpositsioonidena (nt Grünthal 2005: 33; Huumo 2013: 320; Leino 1993: 217–218). Samuti ei tundu siin seoseid olevat

genitiivi ja partitiivi ning adpositsiooni paiknemise vahel: genitiivse ja partitiivse komplemendiga ambipositsioonid näivad ühtmoodi paiknevat vähemalt abstraktset ja ajasuhet väljendades pea võrdselt nii pre- kui ka postpositsioonidena.

6.2. *mööda*

Klassifitseerimispuu põhjal on ambipositsiooni *mööda* paiknemine murde- ti erinev. Moodustub kaks rühma: esimesse kuuluvad kirde-, Mulgi, ranna-, Tartu ja Võru murre ning teise ida-, kesk-, lääne- ja saarte murre. Lõunaees- ti murretes peaks küll ambipositsiooni *mööda* kasutatama harva, sest seda asendab enamasti postpositsioon *pidi* (Ploom 1955: 155–158). Setu murre po- legi siin esindatud, sest analüüsitud materjalis puudusid juhud, kus *mööda* Setu murdes esineks.



Joonis 3. Kaassõna *mööda* paiknemist mõjutavad olulised tegurid klassifitseerimispuus.

Kaassõna *mööda* kasutuste järgi eristuvad niisiis pigem geograafiliste äärealade rühm ning omalaadne keskalade vöö (saarte, lääne-, kesk- ja idamurre), mis kulgeb läänest itta. Esimeses esineb *mööda* peamiselt postpositsioonina (osakaal on 80 %) (6).

(6) *tiid =müüdä* *ess jole`julgenu tulla* (Mulgi)

Teises moodustunud murderühmas mõjutab ambipositsiooni *mööda* paiknemist ka sõnajärg. Kui kaassõnafraas eelneb öeldisverbile (mida esineb andmestikus küll võrdlemisi harva), paikneb *mööda* sagedamini oma komplemendi järel (7). Öeldisverbile järgnedes või finitiverbi ellipsiga kontekstides on *mööda* enamjaolt prepositsioon (8). Kuna lause lõpu poole esitatud kaassõnafraas peaks eesti keele lause infostruktuuri järgi olema uus info, annab 80 %-line prepositsioonide osakaal öeldisverbile järgnevas kontekstis alust järeldada, et uue informatsiooni esitamisel eelistatakse ambipositsiooni *mööda* ida-, kesk-, lääne- ja saarte murdes kasutada prepositsioonina.

(7) *suurt tied = mööda te lähte Mustve = puole* (ida)

(8) *ja näh vere+burikkad mööda kasukka kraed`alla* (lääne)

Põhilised *mööda* kasutuserinevused johtuvad klassifikatsioonipuu põhjal siiski murdealadest: kirde-, Mulgi, ranna-, Tartu ja Võru murdes on *mööda* valdavalt postpositsioon, ida-, kesk-, lääne- ja saarte murdes aga prepositsioon.

6.3. *vastu*

Kaassõna *vastu* kõige olulisemaks paiknemise mõjutajaks osutub klassifitseerimispuus mitte kuigi üllatuslikult komplemendi kääne: genitiivse komplemendiga on *vastu* valdavalt postpositsioon ning partitiivse komplemendiga prepositsioon. Kuigi genitiiviga peaks ambipositsioon *vastu* alati järgnema oma komplemendile (9), siis mõnikord esineb ida- ja rannamurdes *vastu* ka prepositsioonina (10).

(9) *e siiss painuttad põlve`vastu* (saarte)

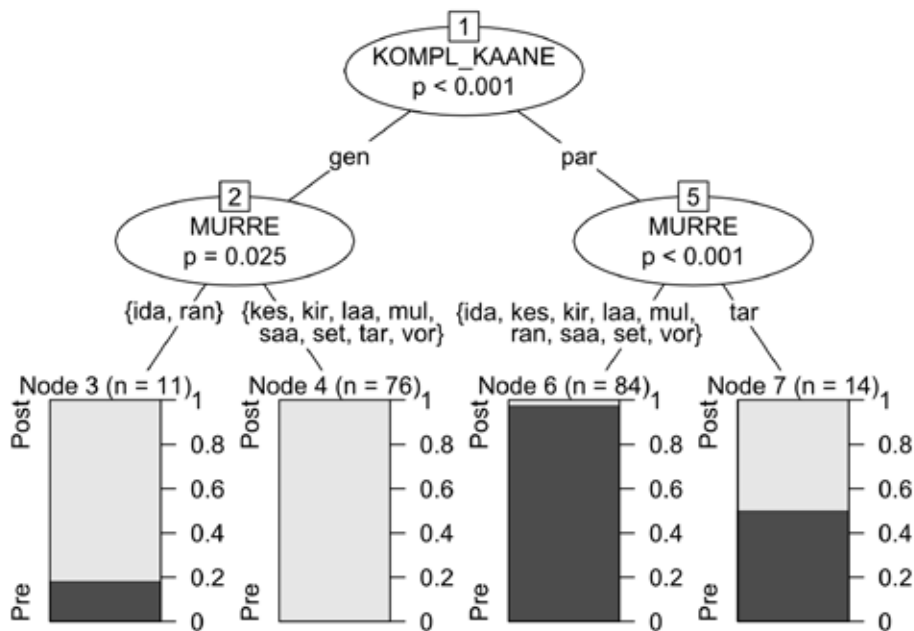
(10) *ja siis sie (.) oli siel nii (.) *vasta *tuule *mennä* (ranna)

Partitiivse komplemendiga *vastu* on murdematerjali põhjal pea alati prepositsioon (11), kuid Tartu murdes on ligi pooled juhtumitest postpositsioonid (12).

(11) *panin nina vasta akkent* (lääne)

(12) *tõmmass nigu miu (.)`endä`rinda`vasta* (Tartu)

Üldjoontes on ambipositsiooni *vastu* olulisim paiknemise mõjutaja ka murdematerjalis komplemendi kääne. Seega järgib *vastu* tüüpilist pre- ja postpositsioonifraasi morfosüntaktilist struktuuri: genitiiviga on eelistatud postpositsiooniline ning partitiiviga prepositsiooniline sõnajärg. Komplemendi käände kõrval on oluline tunnus ka murre, kus ida- ja rannamurre eristuvad ülejäänud murretest paari prepositsioonide kasutusjuhuga genitiivi rühmas;



Joonis 4. Kaassõna *vastu* paiknemist mõjutavad olulised tegurid klassifitseerimispuus.

partitiivirühmas, kus selgelt prevaleerib prepositsioonide kasutus, erineb aga teistest Tartu murre, kus pre- ja postpositsiooniline kasutus on peaaegu võrdne.

6.4. üle

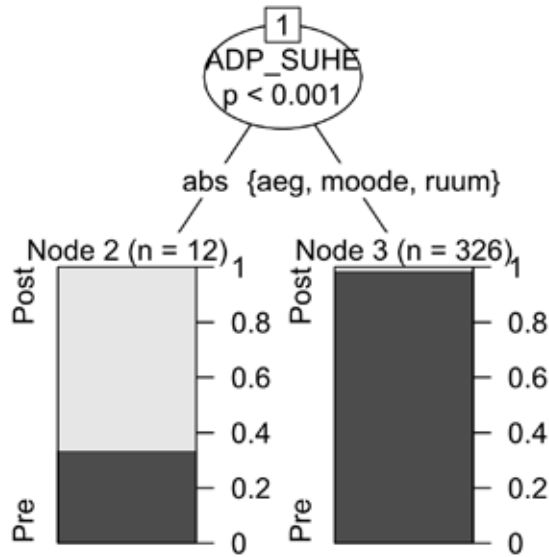
Ehkki ambipositsiooni *üle* tähendused on nii pre- kui ka postpositsioonina esinedes sarnased (vt ptk 2), osutab klassifitseerimispuu, et ambipositsiooni paiknemine on eelkõige seotud just tähistatava suhtega. Abstraktset suhet väljendades on *üle* tüüpiliselt postpositsioon ja vaid umbes 30 %-lise osakaaluga prepositsioon (13). Aja-, mõõtme- ja ruumisuhet tähistades esineb *üle* aga peaaegu alati prepositsioonina (14).

(13) *nohh aga see oli moo jaoks kõik (.) üle jõu* (lääne)

(14) *kuus (.) ritta oll' üle uibu+aija pikkält nigu* (Võru)

Tuleb märkida, et võrreldes muude tähendustega väljendab *üle* abstraktset suhet murdematerjalis siiski harva ja esineb prepositsioonina peamiselt väljendis *selle üle*, mis sageli on seotud ka verbi rektatsiooniga (*mis ma selle üle pahandan*).⁵

⁵ Täname retsensenti sellele tähelepanu juhtimast.



Joonis 5. Kaassõna *üle* paiknemist mõjutavad olulised tegurid klassifitseerimispuus.

6.5. *üumber*

Ambipositsioon *üumber* on mõõtmesuhet väljendades harva prepositsioon – ka Palmeose järgi järgneb see ambipositsioon ligikaudsust tähistades oma komplemendile (15). Ruumisuhteid tähistades on aga *üumber* kirjakeeles nii pre- kui ka postpositsioon. Murdematerjaliski on see nõnda, kuid komplemendi elusus näib siin paiknemist mõjutavat. Elusa komplemendiga on ruumisuhet tähistav *üumber* enamasti postpositsioon (16). Prepositsioonina esineb *üumber* vaid alla 20 %-l juhtudest (18). Elutu komplemendiga aga on sama ambipositsioon harvem postpositsioon (17) ning sagedamini (60 %-lise osakaaluga) just pigem prepositsioon (19).

(15) *temä olli söss joba kah (.)`seitsme+`gümne`aaste ümmer* (Mulgi)

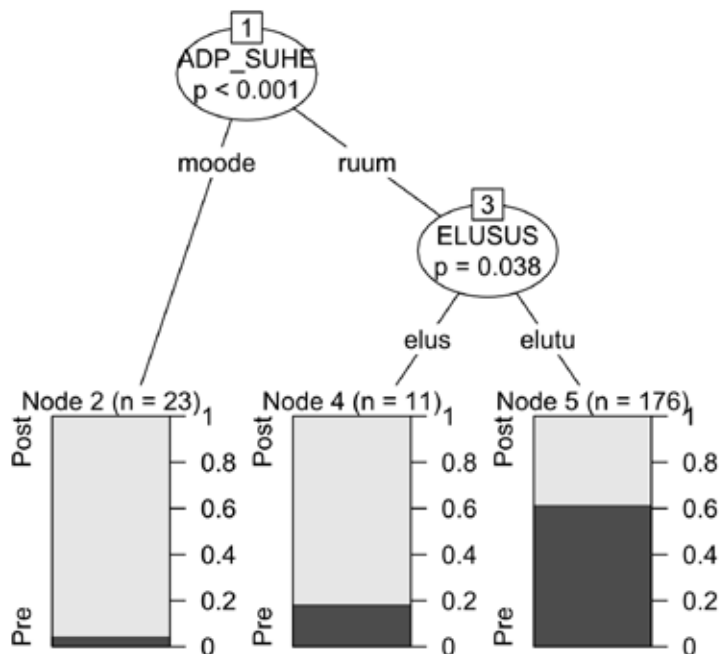
(16) *loomad`jooksid muu ümmere* (lääne)

(17) *`surnu+aed oli`enne just vana kirikku`ümmer* (lääne)

(18) *tuu syss täñittäss käu`ümbre latsõ* (Setu)

(19) *niid`pañtti seasi`ümmer`vankre kõik igale`poole* (kesk)

Seda, miks just elusus kaassõna *üumber* (ja ei ühegi teise kaassõna) paiknemise ennustamisel oluliseks osutub, on aga raske seletada. Erinevused lausete 16 ja 17 ning 18 ja 19 vahel ei tulene tõenäoliselt otseselt mitte entiteetide elususest, vaid on pigem seotud kirjeldatud situatsioonide aspektuaalsete omadustega, mida siinses analüüsis kodeeritud ei ole, kuid mida kindlasti peab põhjalikumalt uurima.



Joonis 6. Kaassõna *ümb*er paiknemist mõjutavad olulised tegurid klassifitseerimispuus.

7. Kokkuvõte

Artiklis esitasime eesti murrete materjali põhjal viie nn kahepaikse kaassõna ehk ambipositsiooni (*läbi*, *mööda*, *vastu*, *üle* ja *ümb*er) morfosüntaktilis-geograafilise kirjelduse, näidates, millised keelesisesed ja ka -välised tegurid kaassõna esinemist kas ees- või tagasõnana enim mõjutavad. Kasutasime selleks klassifitseerimispuid, mis on rakendatavad ka muude keelenähtuste uurimisel. Lisaks näitasime, kuidas klassifitseerimispuu abil on võimalik tuvastada vaadeldava konstruktsiooni kirjeldamiseks kõige olulisemaid tegureid.

Kaassõna *läbi* paiknemise määrab eelkõige väljendatava suhte tüüp: ruumisuhet tähistades kasutatakse seda enamasti prepositsioonina, abstraktset ja ajasuhet väljendades aga pigem postpositsioonina. Seejuures on huvitav, et varasemates uurimustes lähisugulaskeelte (suuresti küll soome keele) ambipositsioonide kohta on esile toodud just nimelt vastupidine tendents. Niisiis analüüsitud eesti murdeandmestiku põhjal ei ole vähemalt *läbi* näitel eesti keeles kaassõna paiknemine väljendatava suhtega samal viisil seotud.

Funktsionaalselt homogeense ehk uurimismaterjalis alati ruumisuhet väljendanud kaassõna *mööda* analüüsis on hästi näha murdeala mõju kaassõna pre- või postpositsioonilisele kasutusele. Eristuvad äärealad kirdes ja kagus, kus eelistatakse valdavalt postpositsiooni, ning keskala, kus eelistatakse pre-

positsiooni. Erandiks viimases rühmas on kontekstid, kus kaassõnafaas eelneb verbile. Selliseid paiknemise eelistusi on paraku raske seletada keelekontaktidega, kuna vene keele mõju idas ning läti keele mõju lõunas soodustaks kirdes ja kagus postpositsioonide asemel pigem hoopis prepositsioonide kasutust. Samuti ei peaks kaassõnade süsteem olema võõrmõjutustele kuigi aldis. Ei saa ka päriselt väita, et erinevused tuleksid sellest, et nimetatud äärealadel kasutatakse *mööda* asemel mingeid teisi lekseeme (nt *pidi*), kuna esinemisjuhtusid on mõlemas murrete rühmas üsna võrdselt.

Kui komplementi kääne võib vastavalt ambipositsiooni paiknemisele varieeruda (*läbi* ja *vastu*), osutub see tegur ka oluliseks kaassõna paiknemise ennustajaks. Erinevalt kirjakeelest võib *vastu* asend murretes nii genitiivse kui ka partitiivse komplementi suhtes siiski varieeruda.

Kaassõna *üle* on murdekorpuses valdavalt prepositsioon ning väljendab enamasti ruumi- ja ajasuhet või ka spetsiifilisemat mõõtmesuhet (*üle kahekümne*). Postpositsioonina esineb *üle* harva. Kuigi genitiivis komplementiga peaks olema eelistatud kasutus postpositsioonina, siis *üle* (ning mingil määral ka teised uuritud kaassõnad) oma genitiivsest rektsioonist hoolimata enamasti seda postpositsioonilist struktuuri ei järgi.

Kaassõna *ümber* olulisimaks paiknemise mõjutajaks on väljendatava suhte tüüp. Mõõtmesuhet väljendades, mis on sarnane abstraktse dimensiooniga, ja elusate referentidega eelistatakse pigem postpositsioone, elutute referentidega prepositsioone. Elususe esiletuleku taustal võivad olla tegurid, mida siinse uurimusse kaasatud ei olnud, nagu näiteks kirjeldatud situatsioonide aspektuaalsed omadused.

Verbi aeg ja kaassõnafaasi pikkus ühegi uuritava ambipositsiooni paiknemise mõjutajana statistiliselt oluliseks ei osutunud.

Edaspidi tasub kindlasti keskenduda teguritele, mis käesolevast uurimisest on välja jäänud, ent mis võiksid ambipositsioonide paiknemist mõjutada. Üheks selliseks potentsiaalselt oluliseks mõjutajaks võiks olla predikaatverbi semantiline klass. Kindlasti võiks kaaluda kaassõnafaasi raskuse erinevat operatsionaliseerimist, näiteks fraasi pikkust lugeda silpide arvu põhjal mitte ainult transkriptsioonimärkidena.

Klassifitseerimispuude meetodit saab kokkuvõttes üsna hästi rakendada grammatiliste (ent miks mitte ka leksikaalsete või fonoloogiliste) alternatsioonide uurimisel, kasutades empiirilisi keeleandmeid. Viimane on ühtlasi nii meetodi suureks plussiks kui mõneti ka miinuseks, kuna empiirilise andmeanalüüsi tulemused (ka mis tahes muu meetodi rakendamisel) on paratamatult sõltuvad muuhulgas sisendmaterjali mahust, keeleteaduse puhul materjali diskursuslikest omadustest ning ka kodeeritavate tunnuste valikust, kodeerimise detailsusastmest jne. Seda kõike peab tulemuste tõlgendamisel arvesse võtma. Teisalt annab empiiriline materjal ja kvantitatiivsete meetodite rakendamine (siinse töös klassifitseerimispuu) objektiivset infot reaalse keelekasutuse (*pro* teadmise) kohta, võimaldab arvesse võtta ka kasutusagedusi ning on sellisena oluline keelekirjelduse osa ning üks täiendava info allikaid.

Artikli valmimist on toetanud Eesti Teadusagentuuri projekt PUT90 „Eesti murrete süntaks” ja Euroopa Liidu Euroopa Sotsiaalfondi programm DoRa, mida viib ellu Sihtasutus Archimedes.

Kirjandus

- Baayen, R. Harald, Endresen, Anna, Janda, Laura A., Makarova, Anastasia, Nessel, Tore 2013. Making choices in Russian: Pros and cons of statistical methods for rival forms. – *Russian Linguistics*, kd 37, nr 3, lk 253–291.
- Breiman, Leo, Friedman, Jerome, Stone, Charles J., Olshen, Richard A. 1984. *Classification and Regression Trees*. New York: Chapman and Hall.
- Bresnan, Joan, Cueni, Anna, Nikitina, Tatiana, Baayen, R. Harald 2007. Predicting the dative alternation. – *Cognitive Foundations of Interpretations*. Toim Gerlof Bouma, Irene Krämer, Joost Zwarts. Amsterdam: Royal Netherlands Academy of Arts and Sciences, lk 69–94.
- Bresnan, Joan, Ford, Marilyn 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. – *Language: Journal of the Linguistic Society of America*, kd 86, nr 1, lk 168–213.
- Dressler, Wolfgang U. 2000. Naturalness. – *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*. Halbbd. 1. = *Morphology: An International Handbook on Inflection and Word-Formation*. Vol. 1. Toim Geert Booij, Christian Lehmann, Joachim Mugdan, Wolfgang Kesselheim, Stavros Skopeteas. Berlin–New York: de Gruyter, lk 288–296.
- Ehala, Martin 1994. Russian influence and the change in progress in the Estonian adpositional system. – *Linguistica Uralica*, kd 30, nr 3, lk 177–193.
- Ehala, Martin 1995. Explaining the bipositional head: Complement order in adpositional systems. – *Proceedings of LP'94: Item Order in Natural Languages*, Prague, August 16–18, 1994, *Proceedings of the Conference*. Toim Bohumil Palek, František Dane, Osamu Fujimura, Jiri V. Neustupný. Prague: Charles University Press, lk 325–339.
- Ehasalu, Epp, Habicht, Külli, Kingisepp, Valve-Liivi, Peebo, Jaak 1997. *Eesti keele vanimad tekstid ja sõnastik*. (Tartu Ülikooli eesti keele õppetooli toimetised 6.) Tartu: Tartu Ülikool.
- EKI 2014 = Eesti Keele Instituudi kohanimeandmebaasi kihelkonnapiiride andmestik.
- EKG I = Mati Ereht, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare, *Eesti keele grammatika I. Morfoloogia*. Sõnamoodustus. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut, 1995.
- EKG II = Mati Ereht, Reet Kasik, Helle Metslang, Henno Rajandi, Kristiina Ross, Henn Saari, Kaja Tael, Silvi Vare, *Eesti keele grammatika II. Süntaks*. Tallinn: Eesti Teaduste Akadeemia Eesti Keele Instituut, 1993.
- EKSS = Eesti keele seletav sõnaraamat. <http://www.eki.ee/dict/ekss/> (27. XII 2014).
- EMK 2014 = Murdekorpus. <http://www.murre.ut.ee/mkweb/> (13. I 2016).
- Gries, Stefan Th., Divjak, Dagmar 2009. Behavioral profiles: A corpus-based approach to cognitive semantic analysis. – *New Directions in Cognitive Linguistics*. Toim Vyvyan Evans, Stephanie S. Pourcel. Amsterdam–Philadelphia: John Benjamins, lk 57–75.
- Grünthal, Riho 2003. Finnic adpositions and cases in change. (Suomalais-Ugri-laisen Seuran toimituksia 244.) Helsinki: Société Finno-Ougrienne.
- Grünthal, Riho 2005. Miksi itämerensuomessa on prepositioita? – *Virittäjä*, nr 1, lk 28–51.
- Habicht, Külli 2000. Grammaticalization of adpositions in old literary Estonian. – *Estonian: Typological Studies IV*. (Publications of the Department of Estonian of the University of Tartu 14.) Tartu, lk 19–58.

- Habicht, Külli, Penjam, Pille 2007. Kaassõna keeleuurija ja -kasutaja käsitluses. – Emakeele Seltsi aastaraamat 2006 (52). Tallinn: Teaduste Akadeemia Kirjastus, lk 51–65.
- Hagège, Claude 2010. Adpositions. Oxford: Oxford University Press.
- Hinrichs, Lars, Szendrői, Benedikt 2007. Recent changes in the function and frequency of Standard English genitive constructions: A multivariate analysis of tagged corpora. – *English Language and Linguistics*, kd 11, nr 3, lk 437–474.
- Hint, Mati 1990. Vene keele mõjud eesti keelele. – *Akadeemia*, nr 2, lk 1383–1404.
- Hothorn, Torsten, Hornik, Kurt, Zeileis, Achim 2006. Unbiased recursive partitioning: A conditional inference framework. – *Journal of Computational and Graphical Statistics*, kd 15, nr 3, lk 651–674.
- Huumo, Tuomas 2013. Many ways of moving along a path: What distinguishes prepositional and postpositional uses of Finnish path adpositions? – *Lingua: International Review of General Linguistics*, nr 133, lk 319–335.
- Janda, Laura A., Antonsen, Lene, Baal, Berit Anne Bals 2014. A radial category profiling analysis of North Sámi ambipositions. – *High Desert Linguistics Society Proceedings*, nr 10, lk 91–102.
- Janda, Laura A., Lyashenskaya, Olga 2013. Semantic profiles of five Russian prefixes: *po-*, *s-*, *za-*, *na-*, *pro-*. – *Journal of Slavic Linguistics*, kd 21, nr 2, lk 211–258.
- Klavan, Jane 2012. Evidence in Linguistics: Corpus-Linguistic and Experimental Methods for Studying Grammatical Synonymy. (Dissertationes linguisticae Universitatis Tartuensis 15.) Tartu: University of Tartu.
- Kortmann, Bernd 2010. Areal variation in syntax. – *Language and Space: Theories and Methods: An International Handbook of Linguistic Variation*. Toim Peter Auer, Jürgen E. Schmidt. Berlin–New York: Mouton de Gruyter, lk 837–864.
- Kyröläinen, Aki-Juhani 2013. Reflexive Space: A Constructionist Model of the Russian Reflexive Marker. (Annales Universitatis Turkuensis B 363.) Turku: University of Turku. <http://urn.fi/URN:ISBN:978-951-29-5394-3>
- Lass, Roger 1997. Historical Linguistics and Language Change. Cambridge: Cambridge University Press.
- Lehismets, Kersten 2014. Suomen kielen väylää ilmaisevien adpositioiden *yli*, *läpi*, *kautta* ja *pitkin* kognitiivista semantiikkaa. (Dissertationes linguisticae Universitatis Tartuensis 21.) Tartu: Tartu Ülikooli Kirjastus.
- Leino, Pentti 1993. Polysemia – kielen moniselitteisyys. – P. Leino, Suomen kielen kognitiivista kielioppia 1. (Kieli 7.) Helsinki: Helsingin yliopiston suomen kielen laitos, lk 207–220.
- Neset, Tore, Endresen, Anna, Janda, Laura A. 2011. Two ways to get out: Radial category profiling and the Russian prefixes *vy-* and *iz-*. – *Zeitschrift für Slawistik*, kd 56, nr 4, lk 377–402.
- Palmeos, Paula 1982. Eesti keele grammatika. II osa. Neljas vihik. Kaassõna. Teine trükk. Tartu: Tartu Riiklik Ülikool.
- Ploom, Harald 1955. Kaassõnad eesti murdeis. Diplomitöö. Käsikiri Tartu Ülikooli eesti keele õppetoolis.
- R Development Core Team 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. <http://www.R-project.org/> (13. I 2016).

- Ruutma, Mirjam 2014. Kaassõnade läbi, mööda, vastu, üle ja ümber pre- ja postpositsioonilisest varieerumisest eesti murretes. Bakalaureusetöö. Käsikiri Tartu Ülikooli eesti ja üldkeeleteaduse instituudis.
- Szmrecsanyi, Benedikt 2014a. Forests, trees, corpora, and dialect grammars. – Aggregating Dialectology, Typology, and Register Analysis. Linguistic Variation in Text and Speech. Toim Benedikt Szmrecsanyi, Bernhard Wälchli. Berlin–Boston: Walter De Gruyter, lk 89–112.
- Szmrecsanyi, Benedikt 2014b. Methods and objectives in contemporary dialectology. – 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. Toim Ilja A. Seržant, Björn Wiemer. Bergen: Department of Foreign Languages, University of Bergen, lk 81–92.
- Szmrecsanyi, Benedikt, Wolk, Christoph 2011. Holistic corpus-based dialectology. – Brazilian Journal of Applied Linguistics / Revista Brasileira de Linguística Aplicada, kd 11, nr 2, lk 561–592.
- Strobl, Carolin, Malley, James, Tutz, Gerhard 2009. An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. – Psychological Methods, kd 14, nr 4, lk 323–348.
- Tagliamonte, Sali A., Baayen, Harald R. 2012. Models, forests, and trees of York English: Was/were variation as a case study for statistical practice. – Language Variation and Change, kd 24, nr 2, lk 135–178.
- Thompson, Sandra A., Couper-Kuhlen, Elizabeth 2005. The clause as a locus of grammar and interaction. – Language and Linguistics, kd 6, nr 4, lk 807–837.
- Veismann, Ann 2009. Eesti keele kaas- ja mäarsõnade semantika võimalusi. (Dissertationes linguisticae Universitatis Tartuensis 11.) Tartu: Tartu Ülikooli Kirjastus.

Descriptions of the morphosyntactic variation of ambipositions by means of quantitative profiles

Keywords: adpositions, classification trees, Estonian dialects, dialectology, corpus linguistics, Estonian language

Certain Estonian adpositions can occur either as prepositions or as postpositions (e.g. *üumber maja*, *maja ümber* 'round the house'). In this study, we refer to these as ambipositions and the goal of this study is to find the motivation for this positional variation in usage. The data come from the Estonian Dialect Corpus, which contains natural dialectal speech, allowing us to simultaneously include both linguistic as well as dialectal factors as part of the analysis. Here, we have included five ambipositions, *läbi* 'through', *mööda* 'along', *vastu* 'against', *üle* 'over' and *ümber* 'around', that occurred with sufficient frequency in both positions. All utterances of these ambipositions in the corpus were included as part of this study and the results of the quantitative analysis, based on classification trees, showed that out of eight factors considered for the analyses, six influenced the realization of these ambipositions in usage, either as preposition or as postposition. The following factors significantly influenced their realization: dialect, case of the complement,

word order, animacy of the complement's referent, number, semantic function of the adposition. Verb tense and length of the adpositional phrase were not significant for any of the analyses. In addition to identifying the abovementioned factors that influenced their realization globally, we provide an analysis for each of the ambipositions displaying their morpho-syntactic and geographical profile. In general, these profiles can be used, for example, in lexicography or in language teaching. Finally, the results indicate that the usage profiles of these ambipositions are rather specific and their realization appears to be more strongly influenced by linguistic rather than dialectal factors.

*Mirjam Ruutma (b. 1990), BA, University of Tartu, MA Student,
mirjam.ruutma@gmail.com*

Aki-Juhani Kyröläinen (b. 1977), PhD, University of Tartu, Department of Estonian, Visiting Lecturer, akkyro@gmail.com

*Maarja-Liisa Pilvik (b. 1989), MA, University of Tartu, PhD Student,
maarja-liisa.pilvik@ut.ee*

Kristel Uiboaed (b. 1984), PhD, University of Tartu, Department of Estonian, Research Fellow in Estonian Dialects, kristel.uiboaed@ut.ee