

Komplementausega predikaatidel põhinevate diskursusemarkerite kasutus eri registrites

TIIT HENNOSTE, KÜLLI PRILLOP, KÜLLI HABICHT, HELLE METSLANG, KIRSI LAANESOO, LIINA PÄRISMAA, ELEN PÄRT, ANDRA RUMM, ANDRIELA RÄÄBIS, CARL ERIC SIMMUL

Artiklis käsitleme diskursusemarkereid, mis on kujunenud kõnelejat tähistava tegevussubjektiga pealausest, millele järgneb komplementause (*Ma arvan, et see kana on surnud.*). Selliseid markereid nimetatakse ingliskeelse terminiga *complement-taking predicate markers* (CTP-markers), meie kasutame eestikeelset terminit *komplementausega predikaatidel põhinevad markerid*, lühidalt KP-markerid.

KP-markerid esinevad eri keeltes (nt ingl *I think, I guess* 'ma arvan', sks *ich denke* 'ma arvan', *ich glaube* 'ma usun'; pr *je pense* 'ma arvan', *je crois* 'ma usun/arvan'; rts *jag tycker* 'ma arvan', *jag tror* 'ma usun' jm). Markereid on pragmaatikas uuritud palju, kuid ennekõike nende kujunemist, pragmatiseerumist (nt Aijmer 1997; Brinton 2017; Heine jt 2021). Markerite kasutust on analüüsitud valdavalt suulises suhtluses (nt Mullan, Karlsson 2012; Breitung-Siepmann 2012; O'Grady 2017; eesti keele kohta nt Hennoste 2004, 2006; Keevallik 2010).

Markerid on leitud olevat polüfunktsionaalsed. Näiteks enim uuritud *I think* toob ennekõike esile kõneleja ebakindluse (täheenduses 'arvatavasti') või subjektiivse arvamuse (täheenduses 'minu arvates'). Selle kõrval on sel ka interpersonaalsed rollid: see pehmendab väidet, markeerib eelkõnelejast erinevat seisukohta, ühtlasi kaasates teda enam suhtlusesse, avardades dialoogilist ruumi. Peale selle võib marker esineda vestluse või teksti organiseerijana, osutades nt kokkuvõttele, uuele teemale või teksti planeerimisele (vt nt Kärkkäinen 2003; Mullan 2012; Mortensen 2012; Keevallik 2010; Hennoste jt 2020; siinses numbris Laanesoo 2022).

KP-markerite täpsemaid seoseid erinevate registritega on vaadeldud statistilistes registrianalüüsides. Siin on meile olulised Douglas Biberi (1988) ja Ewa Jonssoni (2015) uurimused, kus uuritavad ei ole otseselt markerid, vaid nende tuumaks olevad verbid (ingl *private verbs*). Lisaks on kasutusel kategooria *THAT deletion*, kuhu on arvatud sidendi *that* 'et' puudumine, mis seostub inglise keeles nende konstruktsioonide pragmatiseerumisega.

Biber võrdleb oma mitmemõõtmelises analüüsis (ingl *multidimensional analysis*) suulisi ja trükitekste ning näitab, et mõlema mainitud keelejoone sagedus on tugevas korrelatsioonis mõõtmega, mida ta nimetab *involved vs. informational production*. See mõõde moodustab skaala, mille ühes otsas on afektiivne, suhtluskeskne ja üldistatud (ingl *generalized*) sisu, teises otsas kõrge infoväärtusega täpne infosisu. Skaala taustal on 1) tingimused, mis lubavad toimetamist või piiravad teksti tootmist reaalses, ja 2) keelekasutaja keskne eesmärk olla informatiivne või kaasatud (ingl

involved). Reaalajas teksti tootmine nõuab kiiret reageerimist partneri jutule ja see ei võimalda alati leida täpset väljendust (nt Biber, Conrad 2009: 236–241). Kaasatus tähendab, et tekstid jagavad personaalset infot ning väljendavad palju suhtlejate arvamusi, hoiakuid, hinnanguid, millega kaasneb vajadus modaliseerida oma sõnavalikut (Jonsson 2015: 154).

Jonsson analüüsib Biberi skaala abil netitekste ja näitab, et argised netivestlused (*split-window ICQ Chat*) on KP-markerite kasutuselt analoogilised suuliste argivestlustega ja avalikud *chat*'i lühivestlused kasutavad neid vähem. Erinevus on Jonssoni järgi seotud asjaoluga, et avalikus suhtluses ei tunne suhtlejad üksteist piisavalt, et oma eelistusi, hinnanguid jms eksplitsiitselt väljendada (Jonsson 2015: 161).

Need uuringud vaatlevad erinevaid verbe/markereid koos, ühe rühmana. Oleme varem uurinud eraldi KP-markerit *ma arvan, et* (Hennoste jt 2020). Selle analüüsi arvandmed kinnitasid teiste keelte põhjal saadud markerirühmi: marker esines eriti sageli neti- ja argivestlustes, selle kasutus vähenes institutsionaalses suulises suhtluses ja avalikes netitekstides ning oli eriti vähene trükitekstides.

Käesoleva artikli eesmärk on tuua sisse uus aspekt: uurida võrdlevalt kõige sagedamini esinenud KP-markerite kasutust eri registrites. Nende KP-markerite aluseks on verbivormid *arvan, usun, loodan, tundub, näib ja paistab* ning neid ühendab asjaolu, et nad väljendavad ütleja ebakindlust öeldu tõeväärtuse suhtes. Uurimiseks valisime samad tekstide rühmad, mida kasutasime (inter)subjektiivsuspartiklite võrdlemisel: suulised tekstid (institutsionaalne suhtlus ja argivestlus), trükitekstid (ilukirjanduslik proosa, ajakirjandus ja teadus) ning netitekstid (netivestlused ja -kommentaariid) (Hennoste jt 2021).

Meie uurimisküsimused on: 1) milline on eri markerite esinemus eri registrites, kas sellel on süsteemne seos registritega ning millised konsituatiivsed tegurid on kasutussagedusega seotud?; 2) missugune on markerite vormivariantide esinemus eri süntaktilistes positsioonides ja millised tegurid on sellega seotud?

1. Konsituatiivsed ja tekstiehituslikud tegurid

Uurijad on leidnud rea konsituatiivseid ja tekstiehituslikke tegureid, mis mõjutavad episteemiliste ja hoiakumarkerite, samuti partiklite valikut ning kasutussagedust eri registrites (vt Chafe 1982; Biber 1988; Hennoste 2000; Rühlemann 2007; Leech, Short 2007; Biber, Conrad 2009; Auer 2009; Hennoste, Pajusalu 2013; Jonsson 2015). Omalt poolt oleme lisanud jooni, mida leidsime partiklite võrdlevas analüüsis (Hennoste jt 2021). Sõnastame siin tegurid valdavalt tunnuste paaridena. Oluline on meeles pidada, et tegu ei ole vastanditega, vaid ennekõike skaala otstega. Esmalt toome välja konsituatiivsed tegurid (me ei korda iga teguri juures viiteid allikatele).

Dialoogiline vs. monoloogiline tekstitootmine. Kõige üldisem eristus on suuliste ja kirjalike tekstide vahel, mille tootmise konsituatiivsed tegurid on põhimõtteliselt erinevad. Suulist tekstitootmist iseloomustab dialoogilisus, kus suhtlejad ehitavad teksti koos. Partneri vahetud reaktsioonid (vastused küsimusele, tagasiside, parandusalgatused jms) juhivad tekstitootmist ning vajadusel loovad suhtlejad ka

lausungeid koos. Kirjalik tekstitootmine on monoloogiline, ilma tagasisideta. Teatud mõttes piirile jääb netivestlus, kus on võimalik partnerite koostöö nt lausete tootmisel.¹

Suulist keekekasutust eristab kirjalikust ka teksti kustutamatus. Sõnad ilmuvad kõnes lineaarselt üksteise järel ning neid ei ole võimalik peale väljaütlemist kustutada (nt parandamiseks). Kirjalikud tekstid on kustutatavad.

Reaalajas teksti tootmine vs. toimetamise võimalus. Selle skaala keskne mõõdik on tekstiüksuse koostamiseks kasutada olev aeg. Skaala ühes otsas paiknevad reaalajas loodud tekstid, milles kõneleja ei mõtle lausungit enne väljaütlemist peas valmis, vaid loob selle paralleelselt väljaütlemise või kirjutamisega. Siia kuuluvad suulised tekstid, aga ka netivestlused. Skaala teises otsas on tekstid, mille loomisel ajasurve puudub ja mis lubavad toimetamist (nt ilukirjandus- ja teadustekstid). Teisalt ei tähenda toimetamise võimalus kohustust. Nt kirjalikes netivestlustes ja -kommentaaries annab aeglasem tempo enam redigeerimisvõimalusi kui suulises vestluses, kuid netivestluse tekstid on pea alati reaalajas toodetud ja ka kommentaaride tekste tihti ei toimetata. Redigeerimine tähendab ka seda, et järgitakse kindlaid malle, reegleid jms, millega tekst püütakse vastavusse viia. See seostub erinevate sotsiaalsete piirangutega ja on oluline ennekõike trükitekstides. Omavahel erinevad ilukirjandustekstid, milles valitseb *licentia poetica*, ja ametlikud tekstid, mida juhivad toimetamismõõdikud.

Silmast silma vs. vahendatud suhtlus ja multimodaalsus. Vahendatud suhtluse all mõeldakse igasugust suhtlust, mis ei toimu silmast silma (nt telefoni- ja arvutisuhtlus, raamatu lugemine jms). Kirjalik suhtlus on alati vahendatud, suuline suhtlus võib olla nii silmast silma kui ka vahendatud. Vahetuse/vahendatusega seostub multimodaalsus ehk mitteverbaalsete ja kehakeeleliste vahendite kasutamine. Multimodaalne komplekt esineb silmast silma suulises suhtluses ja suurel määral ka videosuhtluses (nt Zoomis). Vahendatud suulises suhtluses võib kehakeele komponent puududa (nt telefonis) ning netis on kasutusel ka eripärased mitteverbaalsed vahendid (nt emotikonid).

Argine vs. avalik vs. ametlik suhtlussfäär. Keskne argisuhtluse variant on suuline argivestlus, aga selle alla kuulub ka netivestlus. Suulise keele uurimises vastandub argisfäärile institutsionaalne ehk ametlik suhtlus (ingl *institutional interaction*), milles vähemalt üks suhtlejatest on institutsiooni esindaja ning keekekasutust mõjutavad vastava institutsiooni kujundatud keelelised ja tekstiehituslikud normid. Kuigi ametlikku suhtlussfääri kuulub erinevaid variante (nt ostmise-müümise kaupluses, koolitunnid, koosolekud, arsti ja patsiendi suhtlus jm), vaadatakse seda suulise keekekasutuse analüüsis ühe tervikuna.

Trüki- ja netitekstide korral institutsionaalsuse terminit tavaliselt ei kasutata. Nende ühe osa moodustab samuti ametlik suhtlus, kus kirjutaja esindab kindlat institutsiooni ja järgib seal kehtivaid teksti- ja keelenorme (nt ajakirjandus ja teadus). Selle kõrval esineb ka suhtlus, milles osalejad ei esinda institutsioone ja keelekasutusnormid on vabamad (nt netikommentaaris ja ilukirjanduslik proosa). Selle

¹ Spetsiifilistes olukordades on võimalik ka suuline monoloogiline tekstitootmine, milles tagasiside puudub (nt üksi tehtav raadiosaade). Selliseid tekste me ei vaatle.

kaksikjaotuse tarvis eraldi termineid välja kujunenud ei ole. Esimest nimetame ametlikuks ja teist avalikuks suhtlussfääriks ning tekste vastavalt ametlikeks ja avalikeks tekstideks.

Suhtlussfäärilise jaotusega seostub keelekasutuse sotsiaalne kontroll. Argisuhtluses on sotsiaalne kontroll keele- ja suhtlusnormide järgimise üle väike ja kõnelejad pööravad oma sõnumi keelelisele vormile vähe tähelepanu. Suulises institutsionaalses suhtluses on kontroll tugevam. Kontroll kirjaliku suhtluse üle on ametlikus sfääris range, kuid avalikus oluliselt lödvem.

Järgnevalt esitame funktsioonide ja ehitusega seotud tegurid, mis mõjutavad KP-markerite ja partiklite kasutust tekstis.

Vestlus vs. narratiiv vs. seletus. Suhtlus jaguneb monoloogiks ja dialoogiks ehk vestluseks. Vestlust eristab muudest tekstidest kaks asjaolu. Esiteks, vestluses vaheldavad kaks või enam suhtlejat lühikesi kõnevoore ja partner peab reageerima eelnevale voorule kiiresti (normist pikemad pausid kõnevoorude vahel viitavad üldjuhul mingile probleemile). Teiseks, kasutada tuleb ka vestlust juhtivaid keelevahendeid: suhtlejad peavad ehitama ja vaheldama kõnevoore, osutama voo üleandmise kohta jms. Vestlus esineb põhiliselt suulises suhtluses, aga veidi teisel kujul ka saadetisi (ingl *post, transmission unit*) vaheldava netivestlusena.

Monoloogi all mõeldakse sama kõneleja või kirjutaja esitatud pikemat teksti. Selles on kaks kesket retoorilist varianti: jutustav ja seletav-analüüsiv tekst. Jutustavate tekstide keskne esindaja on ilukirjanduslik proosa. Seletavate tekstide tuumaks on analüüs, seletamine, tõestamine ja iseloomulikuks esindajaks teadustekstid. Nende vahele jäävad jutustamist ja seletamist kombineerivad tekstid, sh mitmed ajakirjandusžanrid. Siiski esinevad nii jutustavad kui ka seletavad tekstid ka suulises suhtluses ja netivestluses. Jutustavad tekstid tulevad rohkem esile argisuhtluses ja seletavad tekstid institutsionaalses suhtluses (nt teadusliku dispuudina).

Dialoogi kasutamine kirjalikus tekstis ja suulise dialoogi imitatsioon. Dialoogi retoorilise tekstitootmise võttena kasutab ennekõike ilukirjandus, kus eristub autoritekst ja tegelaskõne. Samuti esineb dialoogi ajakirjanduses ja netikommentaarides. Ilukirjanduses on tüüpiliselt tegu suulise dialoogi imitatsiooniga, mille olulisi eesmärke on suulise keelekasutuse illusiooni loomine. Ajakirjanduses ja netikommentaarides imiteerimist ei ole. Ajakirjandusdialoogi valitsev vorm on küsimuste ja vastuste vaheldamine intervjuudes, netikommentaarides seisneb dialoogikommenteerija reageerimises mõnele eelnevale kommentaarile.

Kaasav vs. informeeriv sisu. Kaasavat sisu iseloomustab ennekõike autori seisukohtade, tunnete ja hoiakute väljendamine. Maksimaalselt informeerivas tekstis valitseb faktide esitamine ning kaasatuse näitamist välditakse. Skaala kaasavas otsas on argisfääri kuuluvad argi- ja netivestlused, aga kaasatus on oluline ka avalikes netikommentaarides ja ilukirjanduses. Informatiivsesse otsa jääb teadus, samal ajal on ajalehetekstide seas nii informeerivad uudised kui ka suuremat kaasatust vahendavad arvamused. Kaasatus märgib ka seotust partneriga, mis väljendub tema kaasaarvamises, osalemisvõimaluste suurendamises ning selle kaudu solidaarsuse ja kokkukuuluvuse väljendamisega. Selles mõttes saab kaasatus esineda dialoogi võimaldavates tekstides (siin analüüsitud suulised ja netitekstid, kaudselt ka ilukirjanduse

Tabel 1. Konsituatiivsed ja tekstiehituslikud tegurid.

Registrid	Konsituatiivsed tegurid			Tekstiehituslikud tegurid				
	Dialoogiline / monoloogiline loomus	Reaalajas / toimetatav	Silmast silma / vahendatud	Argi / avalik / ametlik	Vestlus / narratiiv / seletus	Dialoog tekstis	Kaasav / informeeriv	Seos tõega
Argivestlus	D	R	SS/V	AR	V	ei	K	ei
Institutsionaalne suhtlus				AM			I/K	jah
Netivestlus	M(D)		V	AR			K	ei
Netikommentaariid	M	(R/T)		AV	?	(jah)		
Ilukirjandus		T			J	jah	K/I	
Ajakirjandus				AM	S/J	(jah)	I	jah
Teadus					S	ei		

Lühendid: AM = ametlik; AR = argine; AV = avalik; D = dialoogiline; I = informeeriv; J = jutustav; K = kaasav; M = monoloogiline; R = reaalaajas; S = seletav; SS = silmast silma; T = toimetatud; V = vahendatud.

dialoogid). Samal ajal ei tähenda kaasahaaramise võimalus alati selle kasutust. Nii on kaasatus tavaline argisuhtluses, kuid võib puududa institutsionaalses suhtluses, milles valitseb enamasti ühe suhtluspoole võim.

Teksti seotus tõega iseloomustab tekste, mille tootja peab oluliseks ja tõstab esile seda, et öeldu on maksimaalselt tõsikindel. Selline hoiak iseloomustab eriti teadust, aga ka ajakirjandust ja suulist institutsionaalset suhtlust. Muudes siin vaadatud registrites see kuigi oluline ei ole.

Kui eeltoodud tegureid koos vaadata, siis on näha, et teravate piiridega registrikooslusi nende alusel ei kujune (tabel 1). Üldistavalt saab öelda, et ühele poole jäävad argivestlus, suuline institutsionaalne suhtlus ja netivestlus. Neid ühendab dialoogilisus, reaajas tekstitootmine ja vestlus. Nende sees eristub ühelt poolt institutsionaalne suhtlus, mida iseloomustab ametlikkus, infokeskus ja tõega seotus. Teisalt erinevad suulised registrid ja netivestlus. Netivestlust ehitatakse ennekõike monoloogiliselt, see on alati vahendatud ja selle multimodaalsed võimalused erinevad suulise suhtluse omadest. Ka on selles tekstitootmise tempo kirjaliku vormi tõttu aeglasem kui kõnes.

Teise poole moodustavad teadus, ajakirjandus, kirjandus ja ka netikommentaarid. Neid ühendab monoloogilisus, vahendatus, kehakeele puudumine, toimetamise võimalus, mitteargisus, vestluse puudumine. Selles rühmas on tugev sisepiir meedia ja teaduse ning kirjanduse ja kommentaaride rühma vahel, mida eristavad kaasatus/informeeritus, redigeerimise eripärad, seostumatus/seos tõega ning avalikkus/ametlikkus.

2. Uuritavad markerid ja registrid

KP-markerite ja analüüsitavaate üksuste piiritlemisel oleme lähtunud samadest tunnustest nagu *ma arvan* puhul (vt Hennoste jt 2020: 66). Alati on tegu markeriga, kui:

- selles puudub sidend *et* (nt *ma arvan / arvan*);
- üksus esineb (osa)lause algusest tagapool (lause sees, lõpus või selle järel jätkuna);
- *et* ei ole suulises tekstis prosoodilise piiriga verbist eraldatud.

Lause alguses olevad ja sidendit *et* sisaldavad üksused võivad olla nii markerid kui ka pealauseid. Eristuse teeme vastavalt sellele, kas dialoogi või teksti pealiin liigub mööda formaalset pea- või kõrvallauset (Thompson 2002). Markerite hulka arvatakse juhtumid, milles teksti või dialoogi pealiin liikus mööda formaalset kõrvalauset. Juhul kui pealause ei seostunud üheselt suhtluse pea- või kõrvalliiniga, kasutasime markeri eristamiseks teisi tunnuseid.

Analüüsime siin ainult markeri ainest sisaldavaid üksusi, milles *ma/mulle* ja *et* esinevad või mitte. Iga marker võib esineda vormiliselt neljal kujul (nt *ma arvan*, *et*; *ma arvan*; *arvan*, *et*; *arvan*; *mulle tundub*, *et*; *mulle tundub*; *tundub*, *et*; *tundub*).² Kõrvalle on jäetud konstruktsioonid, milles on lisaks muid lauseliikmeid (sh kõrval-

² Tähistame markeriteks olevaid konstruktsioone kujuga (*ma*) *arvan* (*et*), (*mulle*) *tundub* (*et*), lühivariandina ka ainult konstruktsiooni esindava verbiga: *arvan*, *tundub*.

lauseid) või mittelauselikelemente (*tegelikult, aga, no, ka, küll, kuigi* jms). Samuti oleme jätnud kõrvale üksused, milles esines aluse ja öeldise pöördõnajakärg (*arvan ma*) või esiletõstetud ja üldjuhul kontrastiivsusele viitav pikk asesõnavariant *mina/minule* (seletusi vt Hennoste jt 2020: 66–74).

Uuritavate markerite aluseks olevad verbid väljendavad kõik ebakindlust öeldava suhtes ja on teatud määral kasutatavad sünonüümidenä. Samal ajal on neil oma erijooned. *Arvan* on leksikaalse tähenduse poolest kõige abstraktsem, väljendades oletust millegi suhtes, mis pole päris kindel, st selle isiklikul hinnangul tõenäoseks pidamist ning milleski mingil seisukohal olemist (vt Sõnaveeb). *Arvan* kannab kahte raskesti eristatavat põhitähendust. Ühel juhul tõuseb esile isikliku arvamuse avaldamine ehk 'minu arvates' (näide 1) ja teisel juhul üldisem ebakindluse väljendamine propositsiooni tõeväärtuse suhtes ehk 'arvatavasti' (näide 2).³

- (1) **Arvan**, sa oled lihtsalt rumal inimene. (KOMM)
- (2) **Arvan, et** igauks on kokku puutunud oma elus inimestega, kellele väga meeldib igasugu õpetusi jagada, kes justkui teavad „kuidas asjad tegelikult käima peaksid”. (TEA)

Usun esineb kontekstides, kuhu sobiks ka *arvan*, ning väljendab sel juhul tugevamat kindlust propositsiooni tõeväärtuse suhtes (näide 3). *Loodan* lisab tugevama subjektiivse hoiaku kui *usun* ja *arvan*, kasutaja väljendab, et öeldav meeldiks talle, kui öeldu oleks tõsi, kuid tal on selles kahtlused (näide 4).

- (3) **Ma usun**, see oli Martha Eggert, kes laulis. (ILU)
- (4) **Ma loodan, et** endise projekti järgi ei lubata hoonet üles ehitada. (AJA)

Süntaktiliselt on nende kolme markeri materjaliks normaallause, kus subjekt (*ma*) on nominatiivis ja verb ühildub sellega. Viide hinnanguandjale, 'minule', säilib markeris verbi isikuvormi kaudu ka siis, kui *ma* puudub.

Tundub, näib ja paistab väljendavad kasutus kontekstide põhjal otsustades nõrgemat tõsikindlust kui *arvan* (näide 5). Nende materjaliks on kogejalause, kus tegevussubjekt (*mulle*) on allatiivis, verb on kolmanda pöörde vormis ega viita hinnanguandja isikule. Kui *mulle* puudub, pole markeri vormis mingit viidet sellele, kellelt hinnang pärineb. See avab võimaluse hinnangu ebaisikuliseks mõtestamiseks.

- (5) **Tundub, et** oled mõistlik inimene. (KOMM)

Uuritavad registrid jagunevad kolmeks (tabel 2): suulised tekstid (institutsionaalne suhtlus ja argivestlus), trükitekstid (ilukirjanduslik proosa, ajakirjandus ja teadus) ning netitekstid (netivestlused ja -kommentaariid) (vt ka Hennoste jt 2021; Prillop jt 2021). Suulised ja trükitekstid esindavad traditsioonilisi registreid, mida on kasutatud paljudes registrivõrdlustes. Nende kõrvale oleme valinud kaks netitekstide rühma. Netivestlus on suulise argivestluse netianaloog, mis on käsitletav omaette

³ Näited pärinevad artiklist Hennoste jt 2020. Lühendite seletusi vt artikli lõpust.

(alam)registrina. Veebiportaalides meediatekstidele lisatavad kommentaarid jäävad pigem registri piirile, siin käsitleme neid netiregistrina.

Tabel 2. Uuritavad registrid (Hennoste jt 2021).

Suhtlussfäär	Suuline register	Netiregister	Trükiregister
Argine	argivestlus	netivestlus	–
Avalik	–	netikommentaariid	ilukirjandus
Ametlik	institutsionaalne suhtlus	–	ajakirjandus, teadus

3. Materjal ja meetod

Uuritav materjal on võetud portaali Keeleveeb korpustest, „Suulise eesti keele korpusest” ja ühest projekti PRG341 töörühma korpusest: netivestluste korpusest, mis sisaldab aastatel 2010–2020 kogutud Skype’i, Messengeri jms vestlusi. Kokku valisime uurimiseks seitse korpust (vt ka Prillop jt 2021; korpuste suurus on antud sulgudes tekstisõnades).

I. Trükitekstid

- a) ilukirjanduslik proosa (5,8 miljonit)
- b) ajakirjandustekstid (32,9 miljonit)
- c) teadustekstid (1,3 miljonit)

II. Netitekstid

- a) netivestlused (125 000)
- b) netikommentaariid (1,8 miljonit)

III. Suulised tekstid

- a) institutsionaalne suhtlus (312 000)
- b) argivestlus (240 000)

Analüüsis kombineerisime kvantitatiivseid meetodeid kvalitatiivse semantilise ja pragmaatilise analüüsiga. Kõigepealt tegime korpustes esindatud kasutusjuhtude kvalitatiivse analüüsi, mille käigus eraldasime uuritavate verbidega konstruktsioonide seast KP-markerid. Seejärel analüüsisime markerite kasutust statistiliselt. Lõpuks analüüsisime kvalitatiivselt konsituatiivsete tegurite ning markerite vahelisi seoseid.

4. Tulemused

Järgnevalt esitame tulemused kahes osas. Esimeses osas vaatame markerite esinemisagedusi, teises eri vormiga markerite positsioonide varieerumist. Mõlemas osas esitame ka varieerumistega seotud konsituatiivsed tegurid.

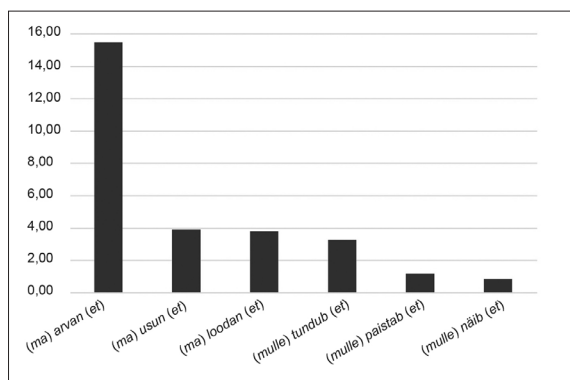
4.1. Markerite esinemissagedus ja sagedusega seotud konsituatiivsed tegurid

Meie esimene küsimus on, kas markerite kasutussagedus erineb registriti ning kas registrid koonduvad sageduse alusel suuremateks rühmadeks. Analüüsiks võtsime igast korpusest juhuvalikuga 200 lauset iga uuritava verbi kohta (või kõik laused, kui neid oli kokku vähem kui 200) ja valisime välja markeristunud kasutused. Kokku jäi valimisse 1842 markerit (tabel 3).

Tabel 3. Markerite absoluutsagedused valimis.

<i>(ma) arvan (et)</i>	542	<i>(mulle) tundub (et)</i>	210
<i>(ma) usun (et)</i>	386	<i>(mulle) paistab (et)</i>	142
<i>(ma) loodan (et)</i>	404	<i>(mulle) näib (et)</i>	158

Selleks et võrrelda markerite levikut eri registrites, leidsime markeristunud kasutuste osakaalud iga verbi jaoks ja arvutasime markeri hinnangulised normaliseeritud sagedused 10 000 tekstisõna kohta, arvestades iga uuritava verbi kogusagedust korpuses.



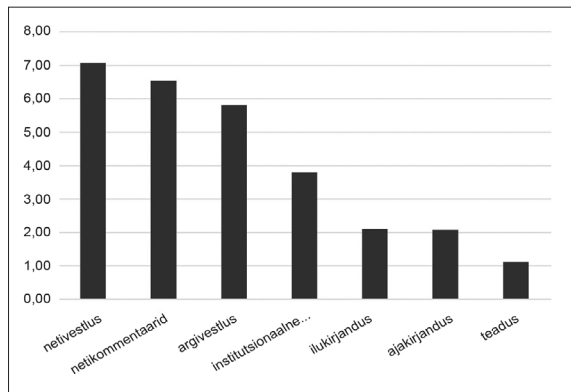
Joonis 1. Eri KP-markerite hinnanguline esinemiskordade arv 10 000 tekstisõna kohta.

Analüüs näitab, et konkreetsete markerite normaliseeritud esinemissagedus erineb (joonis 1): kõige sagedasem marker on *(ma) arvan (et)*, mida kasutatakse umbes sama palju kui kõiki muid markereid kokku. Teise rühma kuuluvad markerid verbi-dest *usun*, *loodan* ja *tundub* ning eriti harvad on *paistab* ja *näib*.

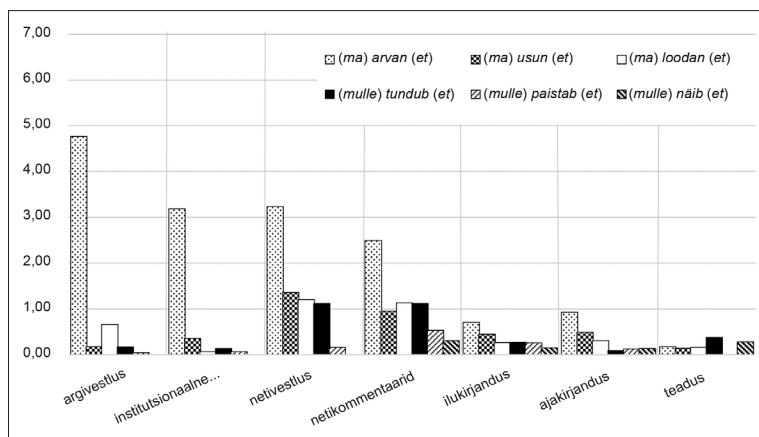
Teisalt erineb markerite kogusagedus valimites väga palju (joonis 2). Selle alusel saab registrid koondada viide rühma:

- 1) netivestlus ja netikommentaarid,
- 2) suulised argivestlused,
- 3) institutsionaalne suhtlus,
- 4) ilukirjanduslik proosa ja ajakirjandus,
- 5) teadus.

Lisaks erinevad registrid markerite mitmekesisuse ja jaotuse poolest (joonis 3). Ühele poole jäävad mitmekesise markerikasutusega kirjalikud tekstid, teisele poole tugevalt ühe markeri (*ma arvan (et)*) kesksed suulised tekstid. Kirjalikud tekstid omakorda jagunevad kaheks. Netitekstides on suur (*ma arvan (et)*) markeri ülekaal, trükitekstides on eri markerite levik võrdlemisi sarnane.



Joonis 2. KP-markerite hinnanguline kogusagedus registrites 10 000 tekstisõna kohta.



Joonis 3. Eri KP-markerite hinnanguline esinemiskordade arv 10 000 tekstisõna kohta registrites.

Järgmine küsimus on, millised konsituatiivsed tegurid seostuvad markerite sageduserinevustega. Joonistelt 1 ja 3 on näha, et suuresti määrab markerite kogusageduse (*ma arvan (et)*). Seetõttu vaatame seda eraldi. (*ma arvan (et)*) kasutuse poolest moodustavad registrid kontiinumi: argivestlus > institutsionaalne suhtlus ja netivestlus > netikommentaariid > ilukirjandus ja ajakirjandus > teadus (vt joonist 3).⁴

⁴ Võrreldes meie varasema (*ma arvan (et)*) uuringuga (Hennoste jt 2020) on siinse uuringu aluseks olnud netivestluste korpuses selle markeri hinnanguline sagedus 10 000 tekstisõna kohta oluliselt väiksem ega ületa enam sagedust argivestlustes. Praegune materjal on mahukam ning kogutud rohkematelt vestlejatelt, nii et üksikisiku keelekasutuse mõju on väiksem.

Skaala ühes otsas on argivestlus. Tegu on reaajas toimuva kaasatusekeskse vestlusega. Kaasatuse vähenemise mõju markerite kasutusele paistab silma, kui võrdleme argivestlust infokeskse ja sageli ühe suhtluspoole võimukeskse institutsionaalse suhtlusega. Teisalt on markerit vähem ka netivestlustes ja netikommentaaries, mis on kaasatusekesksed. Netivestlus toimub samuti reaajas, kuid kirjaliku tekstina on seal rohkem aega tekstitootmiseks ja sõnumi täpsemaks vormistamiseks kui suulises kõnes. Markerit sagedus väheneb veelgi netikommentaaries, mis ei nõua reaajas tekstitootmist.

Skaala teise otsa moodustab markerit väga vähe kasutatav teaduskirjandus. Siin on piiravaks teguriks (*ma*) *arvan* (*et*) semantika, mis ei sobi teaduse tõekesksusega ja kaldega impersonaalsusele (Hennoste jt 2020). Sama seletus sobib suurel määral ka ajakirjanduse tarvis. Üllatav on aga markeri vähesus ilukirjanduses, mis on olemuselt kaasav tekst ning sisaldab sageli suulist vestlust imiteerivat dialoogi ja selles väljenduvaid hoiakuid jms. Üldistavalt on (*ma*) *arvan* (*et*) kasutus siiski kooskõlas varem KP-markerite kohta sõnastatud konsituatiivsete teguritega (Biber 1988; Jonsen 2015).

Kui võrdleme (*ma*) *arvan* (*et*) sagedusi teiste markerite sagedustega, siis näeme kahte olulist erinevust (vt joonist 3). Esiteks, suulises suhtluses valitseb tugevalt (*ma*) *arvan* (*et*), kõik teised markerid esinevad seal väga harva. (*ma*) *usun* (*et*) ja (*ma*) *loodan* (*et*) on suulistes tekstides samas sagedusjärgus trükitekstidega. Väga sarnase funktsiooniga (*mulle*) *tundub/paistab/näib* (*et*) esinevad samuti harva, sealjuures *näib*-verbiga marker puudub suulise keele materjalis hoopis. Ennekõike suulise tekstitootmise kiire tempo ei luba markereid täpsemalt valida ja eelistab kordamisi. Aeglasem tempo võimaldab valida markereid hoolikamalt ja suurendada nende mitmekesisust. Seda näitab ka netisuhtluse markerite suurem mitmekesisus. Samuti on kirjas, sealhulgas netis võimalik markeri varjatud vahetamine enne sõnumi ära saatmist, mis suulises suhtluses võimalik ei ole (suulise teksti kustutamatus tõttu oleks seal asendatud markerid tekstis näha, kuid selliseid näiteid meie materjalis ei esinenud).

Teiseks, netitekstid kasutavad kokku enam markereid kui suulised tekstid (vt joonist 2). Suulise institutsionaalse suhtluse puhul tuleb siin kõne alla infokesksuse mõju, mis vähendab markerite hulka. Ülejäänud suulised ja netiregistrid on aga kaasatusekesksed ning seega peaks nendega seostuma tegureid, mis suurendavad kaasatuse väljendamist just netis. Selliseid tegureid välja toodute hulgas ei leidu.

Trükitekstid kasutavad netiga analoogiliselt mitmekesiseid markereid, samal ajal esineb seal markereid võrdlemisi harva. Üks võimalik seletus on trükitekstide suurem toimetatus, mis võimaldab täpsemat teksti. Ajakirjanduses ja eriti teaduses on oluline ka seos tõega ja infokesksus, mis eriti ei eelda ebakindluse markereid ega isiklike hoiakute väljendamist. Probleemse rühma moodustab jällegi ilukirjandus, mis kasutab kokkuvõttes sama palju markereid kui trükiajakirjandus, kuigi on toimetatud ja kaasav tekst, mis ei tegele tõsikindla info vahendamisega. Selle seletuseks võib olla asjaolu, et KP-markerid ei sobi autoriteksti, vaid ennekõike otsesesse kõnesse. Seda toetab ka fakt, et dialoogivabas teaduses on markereid vähe, tsitaatidena otsest kõnet sisaldavates ajakirjandustekstides rohkem.

Eraldi teema on KP-markerite vormivariantide kasutus. Markerid võivad esineda neljal kujul (nt *ma usun, et; ma usun; usun, et; usun*). Valdavalt esinevad markerid sidendiga *et*, sealjuures on sidendiga *arvan-* ja *loodan-*markereist 78% (näited 6, 7), *usun-*markereist 90% (näide 8), *näib* kasutusest 94% (näide 9), *paistab* 96% ja *tundub* 97%.

Asesõna kasutuse poolest jagunevad markerid kolmeks. *Arvan* esineb valdavalt asesõnaga (75%) (näide 6), *loodan* ja *usun* asesõnaga ja ilma selleta variandid jagunevad umbes pooleks (*loodan* 44%, *usun* 48%) (näited 7, 8) ning kolmikus *tundub, näib* ja *paistab* valitseb ilma asesõnata variant (*tundub* esineb asesõnaga 23% juhtudest, *näib* 18%, *paistab* 1%) (näide 9). Üksnes verbist koosnevad markerid esinesid väga harva, olenemata verbist.

- (6) mhmh .hh **ma arvan et** ma võibolla=sis `meilin et=sis=on `mõtlemisaega `ka.
(SUUL IN)
- (7) **Ma loodan, et** mul ei lähe enam kaua aega. (ILU)
- (8) „[---] **Usun, et** Eestil ei ole sellist suurt ohtu, et võidakse vastu võtta otsus hakata ELiga ühinemist grupiviisiliselt käsitlema,” arvas Vähi. (AJA)
- (9) **Näib, et** meediatarbijad janunevad eelkõige elamuste järele. (TEA)

Vormivariantide kasutus erineb registriti. Ilukirjanduses, netivestluses, institutsionaalses suhtluses ja vähemal määral ka argivestluses valitsevad asesõnalised variandid, mille seas on keskne *ma+V+et* (näited 6 ja 7). Ajakirjanduses, netikommendaarides ja vähemal määral ka teaduses valitsevad ilma asesõnata variandid, mille seas on keskne *V+et* mall (näited 8 ja 9). Seega valitseb asesõnaline mall vestlustes ja üldisemalt dialoogi sisaldavates registrites.

Teisalt on erinevus sidendiga ja ilma sidendita variantide kasutuses. Ilma sidendita variant esineb sageli suulistes registrites, argi- ja institutsionaalses vestluses (vastavalt 40%-l ja 31%-l juhtudest) (näide 10). Mujal valitsevad sidendiga variandid (näited 7–9). Seega paistab siin välja suulise ja kirjaliku keele vaheline erinevus.

- (10) H: `põhiline on sis `see=et nagu det `sembri alguses tegelikult `selguvad need (.)
põhiasjad või.
T: **ma arvan** pisut `varem. (0.5) aga `siiski jah. (SUUL ARG1)

4.2. Markerite positsioonid ja nende seosed konsituatiivsete teguritega

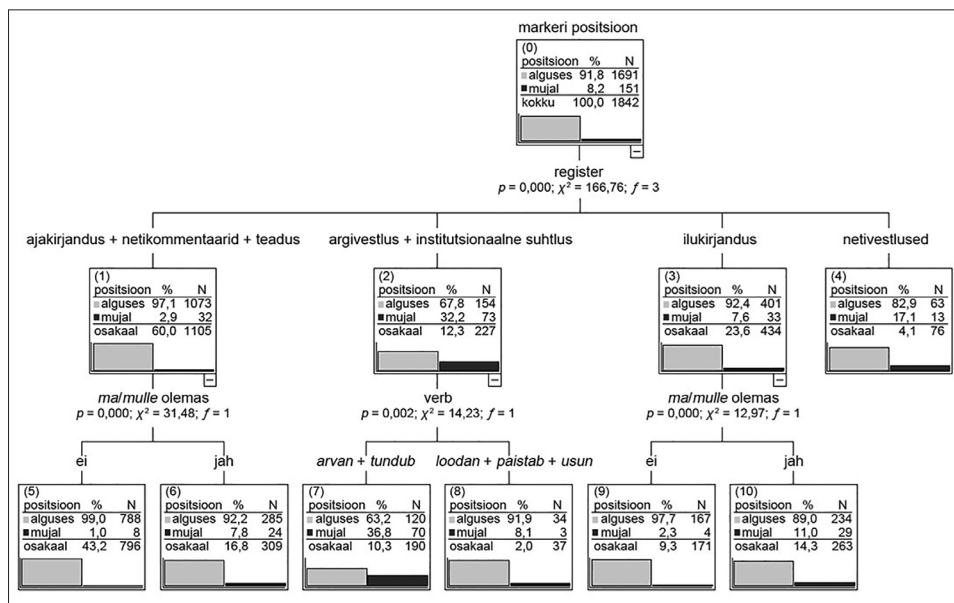
Meie teine uurimisküsimus oli, milline on markerite vormivariantide esinemus eri süntaktilistes positsioonides ja millised tegurid on positsioonidega seotud. Analüüsitud tekstides valitseb markerite alguspositsioon (lauses või osalauses), muud positsioonid esinevad harva (8,2%-l juhtudest). Seetõttu eristame analüüsis ainult alguspositsiooni ja muid positsioone kokku.

Kõigepealt uurime, kas markeri nihkumine lause algusest tahapoole on mõnes registris või registrigrupis sagedasem kui teises või seostub positsioon pigem markeri struktuuri (*ma/mulle* olemasolu, verb) kui registriga. Seejärel selgitame, kas

leitud seaduspärad on seletatavad konsituatiivsete tegurite abil või/ja mis osas tulenevad need muudest teguritest (verbi semantikast, markeristumise määrast vms).

Esmaseks kvantitatiivseks analüüsiks kasutasime otsustuspuude meetodit CHAID (*Chi-square Automatic Interaction Detector*),⁵ mille tulemuse esitab joonis 4.

Arvud näitavad, et markeri positsioon on tugevalt seotud registriga (joonised 4 ja 5). Suulistes registrites on marker algasendist tahapoole nihkunud ca 30%-l juhtudest (grupp 2 joonisel 4; näide 11), netivestlustes 17% (grupp 4; näide 12), ilukirjanduslikus proosas 8% (grupp 3), ülejäänud kirjalikes registrites alla 3% (grupp 1).



Joonis 4. KP-markeri süntaktilise positsiooniga seotud mõjutegurid.

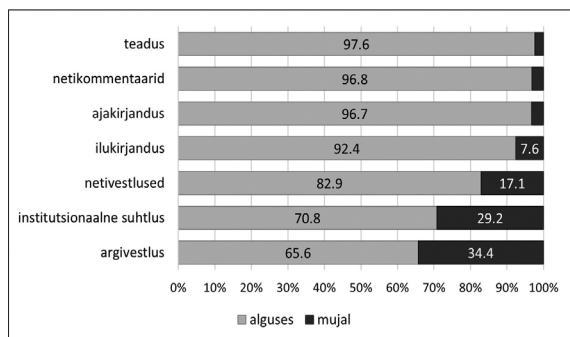
(11) N: a hinna- `hinnad nagu=on `samad=vä.

M: `hinnad on niuke **ma=arvan** siuke `keskmine siuke `hinnatase on `sama mis Eesti `möni asi `kallim, `möni asi `odavam. (SUUL ARGJ)

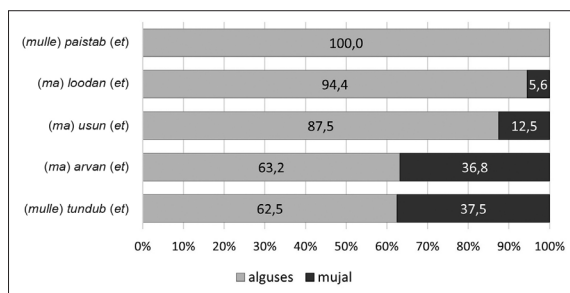
(12) eraldi mooduli peab laadima selle jaoks alla, **tundub** (NV)

Kesksed mõjurid on sealjuures reaajas tekstitootmine ja dialoog (vestlus ja dialoogi kasutus ilukirjanduses). Suulisest suhtlusest aeglasema tempoga netivestluses on marker tahapoole nihkunud oluliselt vähem. Kustutamatusse tõttu pole suulisel keeles võimalik ka markeri asukoha hilisem korrigeerimine, kuna markeri algusesse viimine nõuaks kogu lause uuesti ütlemist.

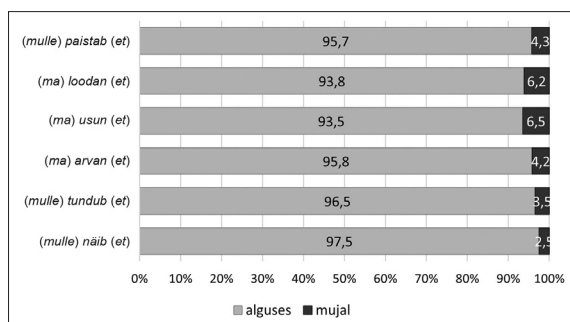
⁵ Tegü on masinõppemeetodiga, mis jagab vaatlused (siinses uuringus näitelauseid) etteantud mõjutegurite (siin register, verb, ma/mulle olemasolu) alusel iteratiivselt klassidesse, kasutades klassidevahelise erinevuse hindamiseks χ^2 statistikut. Iga hargnemine puus toimub antud tasan-dil statistiliselt kõige olulisema teguri alusel. (Kass 1980) Otsustuspuu genereerimisel kasutasime tarkvarapaketti IBM SPSS Statistics. p-väärtused on Bonferroni korrigeerimisele. Usaldusnivoo nii klasside ühendamiseks kui ka eraldamiseks on 95%.



Joonis 5. KP-markeri positsioon (osa)lauses.



Joonis 6. KP-markeri positsioon olenevalt verbist argivestlustes ja institutsionaalses suhtluses (%). Jooniselt puudub (mulle) näib (et), mida siinse materjali suulistes registrites ei esinenud.



Joonis 7. KP-markeri positsioon (%) kirjalikes registrites verbist olenevalt.

Teisalt selgus, et eri markerid on nihkunud tahapoole eri määral. Kõigepealt on erinevus kirjalike ja suuliste registrite vahel: suulises registris on (ma) arvan (et) ja (mulle) tundub (et) oluliselt sagedamini tahapoole nihkunud kui ülejäänud markerid (joonis 6; grupid 7 ja 8 joonisel 4). Seda ei saa seletada konsituatiivsete teguritega. Kirjalikes tekstides ei ole erinevused markerite vahel statistiliselt olulised (joonis 7).

Lisaks näitab analüüs, et statistiliselt on seotud ka KP-markeri asukoht ja ase-sõna kasutus. Kõigis registrites ja kõigi markerite puhul kehtib põhimõte, et lause

sees/lõpus esineb *ma/mulle* sagedamini kui lause alguses. Nt kõige kindlamini paikneb marker lause alguses ajakirjandus- ja teadustekstides ning netikommentaari-des, kusjuures 99%-l juhtudest on neis registrites lausealguline selline marker, mis ei sisalda asesõna (grupp 5 joonisel 4). Asesõna kasutuse seos registri ja markeri nihkumisega vajab eraldi uuringut, sh võrdlust markeristumata verbidega, ja sellesse me siin ei süvene.

5. Arutelu ja kokkuvõte

Analüüsisime KP-markereid (*ma arvan (et)*, (*ma*) *usun (et)*, (*ma*) *loodan (et)*, (*mulle*) *näib (et)*, (*mulle*) *tundub (et)*, (*mulle*) *paistab (et)*) ja nende kasutust suulises argivestluses, suulises institutsionaalses suhtluses, ilukirjanduslikus proosas, ajakirjanduses, teaduses, netivestlustes ja netikommentaari-des.

Markerite kasutussageduse analüüs näitas, et selle alusel on registrid koondatavad viide rühma, milles markerid esinevad alaneva sagedusega: netivestlused ja netikommentaari-d, suulised argivestlused, institutsionaalne suhtlus, ilukirjanduslik proosa ja ajakirjandus, teadus (vt joonist 2). Samal ajal selgus, et markeri (*ma arvan (et)*) hinnanguline sagedus oli materjalis suurem kui muudel markeritel kokku ja määras sellega ka markerite kogusageduse (vt joonist 1).

Täpsem analüüs näitas, et just (*ma arvan (et)*) sagedus vastab hästi varasemates uurimustes leitud KP-markerite üldisele sagedusjaotusele (vrd Jonsson 2015: 214). Kõiki markereid arvestades tulevad välja olulised erinevused eri registrite vahel (vt joonist 3).

Esiteks, suulises suhtluses valitseb (*ma arvan (et)*), muid markereid esineb seal väga vähe. See on sarnane episteemiliste partiklite võrdleva analüüsi tulemusega, kus samal viisil toimib partikkel *vist* (Hennoste jt 2021). Seega valitseb suulises suhtluses erinevalt kirjalikust üks KP-marker või ebakindluse partikkel. See on kooskõlas suulise suhtluse ökonoomiapõhimõttega, mis soosib varieerumise asemel kordusi (vt nt Rühlemann 2007).

Teiseks, kirjalikud registrid kasutavad mitmekesisest markerite valikut, sealjuures nii toimetamist võimaldavates trükitekstides ja netikommentaari-des kui ka argivestlusega analoogilises reaajas netivestluses (vt joonist 3). Siin on keskeks mõju-riks kirjaliku teksti tootmise aeglasem tempo, mis võimaldab valida markereid täpsemalt ja suurendada nende mitmekesisust.

Kolmandaks, trükitekstides on markereid kokku vähem kui netitekstides (vt jooniseid 2 ja 3). Üks võimalik seletus on, et tänu toimetamisele on trükitekstid täpsemad ega vaja ebakindlusele osutamist. Ajakirjanduses ja eriti teaduses on oluline ka seos tööga ja infokesksus, mis ei eelda ebakindluse markereid ega isiklike hoiakute väljendamist. Samuti on trükitekst suunatud korruga paljudele adressaatidele väljaspool reaalaega ning sellisena ei ava dialoogilist ruumi.

Neljandaks, suulises ja netisuhtluses esineb (*ma arvan (et)*) muudest markeritest oluliselt sagedamini, trükitekstides on eri markerite hinnangulised sagedused võrdlemisi sarnased (vt joonist 3). Siin on keskeks eristajaks interpersonaalsus. Suuline

suhtlus ja netivestlus on reaajas vestlused, trükitekstid toimetatud narratiivid ja seletused. Vestlustes tõuseb esile interpersonaalne suhtlusruumi avardamine. See on markeri (*ma arvan (et)*) oluline roll ja võimaldab seletada selle markeri suuremat sagedust vestlustes (vt ka siinses numbris Laanesoo 2022). Vaherühmaks on siin netikommentaariid, mis moodustatakse küll valdavalt reaajas, kuid on vaid osaliselt vestlused. Siiski on see interpersonaalsust võimaldav register.

Uurimus andis kaks üllatavat tulemust. Esiteks, netitekstid kasutavad kokkuvõttes enam markereid kui suulised tekstid (vt joonist 2). See joon ei ole seletatav eelnevalt leitud konsituatiivsete teguritega. Üks võimalik seletus on, et internetis väljendavad suhtlejad oma ebakindlust rohkem verbaalselt, kuna puuduvad suulise keelega analoogilised mitteverbaalsed vahendid (nt intonatsioon, miimika jm). Ka on võimalik, et vahendatud netisuhtluses väljendavad inimesed oma hoiakuid vabamalt kui silmast silma suulises suhtluses või kõnelevad enam suhtumistest, tunnetest jms kui suulises suhtluses ning seetõttu esineb seal ka rohkem KP-markereid.

Teiseks, üllatav on markerite suhteline vähesus ilukirjanduses, mis on olemuselt kaasav register, kus esineb osaliste dialoog ja väljendatakse palju hoiakuid ja tundeid. Seda võib seletada asjaolu, et KP-markerid ei sobi ilukirjanduse autoriteksti, kuna nad on kasutatavad üksnes juhul, kui kasutaja esineb enda nimel. Seda toetab ka fakt, et ebakindlust väljendavaid partikleid, mis sobivad nii autoriteksti kui ka tegelaste dialoogi, on ilukirjanduses oluliselt enam kui ajakirjanduses (Hennoste jt 2021).

Markerite vormivariantide, nende eri süntaktilistes positsioonides esinemuse ja sellega seotud tegurite analüüs näitas, et eri verbid eelistavad erinevaid vormivariante: *arvan* esineb valdavalt asesõnaga, *loodan* ja *usun* jagunevad pooleks ning *tundub*, *näib* ja *paistab* puhul valitseb ilma asesõnata variant. Üksnes verbist koosnevad KP-markerid on kõigis vormivariantides väga harvad.

Kõigi markerite keskne positsioon on lausealguline, samal ajal on registrite vahel olulised erinevused. Enim on markerid liikunud tahapoole suulistes registrites (*ca* 30%), kusjuures (*ma arvan (et)*) ja (*mulle*) *tundub (et)* esinevad tagapool oluliselt sagedamini kui ülejäänud (vt jooniseid 4 ja 6). Järgnevad netivestlused ja ilukirjanduslik proosa. See näitab, et tugevaimalt mõjutab markerite nihkumist reaajas teksti tootmine ja dialoog. KP-markerite tahapoole nihkumine on suulistele registritele üldiselt omane nähtus (vt nt Auer, Lindström 2016).

Kokkuvõttes koonduvad kesksed markerite sagedust, vormivariantide valikuid ja asukohta mõjutavad konsituatiivsed tegurid dialoogi ümber. Olulisemad on siin dialoogiline ja reaajas teksti tootmine, mis suurendavad nii markerite hulka, ühe markeri keskust kui ka markerite tahapoole nihkumist. Sealjuures on eraldi vaja esile tõsta reaajas tekstide tootmise tempot, mis eristab suulisi vestlusi netivestlusest. Järgmised olulised jooned on tekstide ehitusega seostuvad tegurid vestlus ja dialoogi imitatsioon kirjanduses, mis suurendavad markerite hulka ja tahapoole nihkumist. Samal ajal oli näha, et eri markerite ja eri registrite puhul tulevad esile ka teised konsituatiivsed ja tekstikesksed tegurid. Seega ei ole KP-markerite kasutus taandatav mõne üksiku teguri mõjule, rolli mängivad eri tegurite kombinatsioonid.

Lisaks tuleb mõjuritena esile tuua kaks asjaolu. Esiteks on oluline KP-marke-rite endi semantika, mis ei sobi registritega, milles on oluline tekstide tõekesksus ja impersonaalsus. Teiseks on oluline markerite erinev pragmatiseerumine.

Vaadates markerite vormi ja positsioone, saame ka üldistada, et analüüsitud ver- bid on erineval määral pragmatiseerunud. Ühelt poolt seostub pragmatiseerumine sidendi *et* kadumisega, mis on eelduseks konstruktsiooni liikumisele lause algus- positsioonist tahapoole. Vormi poolest on enim pragmatiseerunud *arvan-*, vähem *usun-* ja *loodan-* ning minimaalselt *tundub-*, *näib-* ja *paistab-* markerid. Teisalt võib ka markeri sagedasemat tahapoole nihkumist seletada kaugemale arenenud pragma- tiseerumisega. Sellest seisukohast on enim pragmatiseerunud *arvan-* ja *tundub-* mar- kerid. See lubab järeldada, et vorm ja positsioon on teineteisest sõltumatud tegurid. Seda kinnitab ka ilukirjandus, kus markerid ei esine suurema sagedusega kui aja- kirjanduses, kuid samal ajal on oluliselt sagedamini lauses tahapoole nihkunud.

Kokkuvõttes on enim pragmatiseerunud marker (*ma*) *arvan* (*et*), mis on polü- funktsionaalne analoogiliselt inglise markeriga *I think*. Praegu puuduvad analüüsid selle kohta, millistes funktsioonides esinevad ülejäänud siin vaadeldud KP-markerid. Kuna neid kasutatakse oluliselt vähem, siis võib oletada, et neil on ka vähem funk- tsioone. Sellest lähtudes esitame hüpoteesi, et (*ma*) *arvan* (*et*) on pragmatiseerunud üldmarker, mis ühelt poolt esineb paljudes funktsioonides ja teisalt sobib kergemini teiste KP-markerite funktsionaalseks asendajaks kui vastupidi.

Saadud tulemused seavad kahtluse alla üldistavad analüüsid, milles eri partiklid ja KP-markerid võetakse kokku üheks rühmaks. Meie analüüs näitab, et oluline on eristada vähemalt kaht rühma. Ühele poole jäävad üldised ja paljudesse konteksti- desse sobivad diskursusemarkerid (nagu *vist* ja (*ma*) *arvan* (*et*)), teisele poole muud.

Artikli autorid kuuluvad projekti PRG341 „Pragmaatika grammatika kohal: subjektiivsus ja intersubjektiivsus eesti keele registrites ja tekstiliikides“ tööühma. Artikli valmimist on toetanud Eesti Teadusagentuur ja Euroopa Regionaalarengu Fond (Eesti-uuringute Tippkeskus).

LÜHENDID

AJA = ajakirjandus; **ILU** = ilukirjanduslik proosa; **KOMM** = netikommentaarid; **NV** = neti- vestlused; **SUUL ARG** = suuline argivestlus; **SUUL IN** = suuline institutsionaalne suhtlus; **TEA** = teadus.

TRANSKRIPTSIOONIMÄRGID

.	langev intonatsioon	`	rõhutatud sõna või silp
,	poollangev intonatsioon	si-	pooleli jäänud sõna
(.)	mikropaus (0.2 sekundit või lühem)	.hh	häälekas sissehingamine
(0.8)	pikem paus sekundites	=	kokkuhääldamine

VEEBIVARAD

EKSS = Eesti keele seletav sõnaraamat. <http://www.eki.ee/dict/ekss/>

ENC19 = Jelena Kallas, Kristina Koppel 2020. Eesti keele ühendkorpus 2019. Eesti Keele-ressursside Keskus. <https://doi.org/10.15155/3-00-0000-0000-0000-08489L>

Keeleveeb. <https://www.keeleveeb.ee/>

Suulise eesti keele korpus. <https://keeleressursid.ee/et/220-suulise-est-kee-kele-korpus>

Sõnaveeb = EKI ühendsõnastik 2020. Eesti Keele Instituut. <https://sonaveeb.ee>

KIRJANDUS

- Aijmer, Karin 1997.** *I think* – an English modal particle. – Modality in Germanic Languages: Historical and Comparative Perspectives. Toim Toril Swan, Olaf Jansen-Westvik. Berlin–New York: Mouton de Gruyter, lk 1–47. <https://doi.org/10.1515/9783110889932.1>
- Auer, Peter 2009.** On-line syntax: Thoughts on the temporality of spoken language. – Language Sciences, kd 31, lk 1–13. <https://doi.org/10.1016/j.langsci.2007.10.004>
- Auer, Peter; Lindström, Jan 2016.** Left/right asymmetries and the grammar of pre- vs. post-positioning in German and Swedish talk-in-interaction. – Language Sciences, kd 56, lk 68–92. <https://doi.org/10.1016/j.langsci.2016.03.001>
- Biber, Douglas 1988.** Variation across Speech and Writing. New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511621024>
- Biber, Douglas; Conrad, Susan 2009.** Register, Genre, and Style. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511814358>
- Breitkopf-Siepmann, Anna 2012.** Hedging in German and Russian conference presentations: A cross-cultural view. – Subjectivity in Language and Discourse. (Studies in Pragmatics 10.) Toim Nicole Baumgarten, Inke Du Bois, Juliane House. Leiden: Brill, lk 295–318. https://doi.org/10.1163/9789004261921_014
- Brinton, Laurel 2017.** The Development of Pragmatic Markers in English: Pathways of Change. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781316416013>
- Chafe, Wallace L. 1982.** Integration and involvement in speaking, writing, and oral literature. – Spoken and Written Language: Exploring Orality and Literacy. Toim Deborah Tannen. Norwood: Ablex, lk 35–53.
- Heine, Bernd; Kaltenböck, Gunther; Kuteva, Tania; Long, Haiping 2021.** On the rise of discourse markers. – Studies at the Grammar-Discourse Interface: Discourse Markers and Discourse-related Grammatical Phenomena. (Studies in Language Companion Series 219.) Toim Alexander Haselow, Sylvie Hancil. John Benjamins, lk 24–55. <https://doi.org/10.1075/slcs.219.01hei>
- Hennoste, Tiit 2000.** Allkeeled. – Eesti keele allkeeled. (Tartu Ülikooli eesti keele õppetooli toimetised 16.) Toim T. Hennoste. Tartu: Tartu Ülikooli Kirjastus, lk 9–56.
- Hennoste, Tiit 2004.** *Et*-komplementlause peaverbide funktsioonid eestikeelses vestluses. – Keel ja Kirjandus, nr 7, lk 504–523; nr 8, lk 590–609.
- Hennoste, Tiit 2006.** *Et*-komplementlause peaverbide funktsioonid eestikeelses vestluses 2: *mõtlemä*. – Lähivertailuja 17. (Jyväskylä Studies in Humanities 53.) Toim Annekatrin Kaivapalu, Külvi Pruuli. Jyväskylä: Jyväskylän yliopisto, lk 119–134. <https://doi.org/10.5128/LV17.08>
- Hennoste, Tiit; Pajusalu, Karl 2013.** Eesti keele allkeeled. Tallinn: Eesti Keele Sihtasutus.
- Hennoste, Tiit; Habicht, Külli; Metslang, Helle; Prillop, Külli; Laanesoo, Kirsi; Ogren, David; Pärismaa, Liina; Pärt, Elen; Rumm, Andra; Rääbis, Andriela; Simmul, Carl Eric 2020.** Diskursusemarker (*ma arvan* (*et*)). – Emakeele Seltsi aastaraamat 65 (2019). Peatoim Mati Ereht. Tallinn: Teaduste Akadeemia Kirjastus, lk 63–90. <http://dx.doi.org/10.3176/esa65.03>

- Hennoste, Tiit; Metslang, Helle; Habicht, Külli; Prillop, Külli 2021.** Kuue (inter)subjektiivsuspartikli kasutus eesti keele registrites. – Emakeele Seltsi aastaraamat 66 (2020). Peatoim Mati Erelt. Tallinn: Teaduste Akadeemia Kirjastus, lk 91–123. <https://doi.org/10.3176/esa66.04>
- Jonsson, Ewa 2015.** Conversational Writing: A Multidimensional Study of Synchronous and Supersynchronous Computer-Mediated Communication. (English Corpus Linguistics 16.) Frankfurt am Main jt: Peter Lang Edition. <https://doi.org/10.3726/978-3-653-06512-1>
- Kass, Gordon V. 1980.** An exploratory technique for investigating large quantities of categorical data. – Journal of the Royal Statistical Society. Series C (Applied Statistics), kd 29, nr 2, lk 119–127. <https://doi.org/10.2307/2986296>
- Keevallik, Leelo 2010.** Clauses emerging as epistemic adverbs in Estonian conversation. – Linguistica Uralica, kd 46, nr 2, lk 81–101. <https://doi.org/10.3176/lu.2010.2.01>
- Kärkkäinen, Elise 2003.** Epistemic Stance in English Conversation: A Description of its Interactional Functions, with a Focus on 'I think'. (Pragmatics & Beyond New Series 113.) Amsterdam–Philadelphia: John Benjamins. <https://doi.org/10.1075/pbns.115>
- Laanesoo, Kirsi 2022.** Diskursusemarker (*ma*) *arvan* (*et*) direktiivides. – Keel ja Kirjandus, nr 1–2, lk 151–167. <https://doi.org/10.54013/kk770a9>
- Leech, Geoffrey; Short, Mick 2007.** Style in Fiction: A Linguistic Introduction to English Fictional Prose. 2. tr. London: Longman.
- Mortensen, Janus 2012.** Subjectivity and intersubjectivity as aspects of epistemic stance marking. – Subjectivity in Language and Discourse. (Studies in Pragmatics 10.) Toim Nicole Baumgarten, Inke Du Bois, Juliane House. Bingley: Emerald, lk 229–246. https://doi.org/10.1163/9789004261921_011
- Mullan Kerry 2012.** Saying what you think: An analysis of French and Australian English non-native speaker expression of subjectivity. – Subjectivity in Language and Discourse. (Studies in Pragmatics 10.) Toim Nicole Baumgarten, Inke Du Bois, Juliane House. Bingley: Emerald, lk 193–227. https://doi.org/10.1163/9789004261921_010
- Mullan, Kerry; Karlsson, Susanna 2012.** Subjectivity in contrast: A cross-linguistic comparison of 'I think' in Australian English, French and Swedish. – Subjectivity in Language and Discourse. (Studies in Pragmatics 10.) Toim Nicole Baumgarten, Inke Du Bois, Juliane House. Bingley: Emerald, lk 271–294. https://doi.org/10.1163/9789004261921_013
- O'Grady, Gerard 2017.** "I think" in political discourse. – International Review of Pragmatics, nr 9, lk 269–303. <https://doi.org/10.1163/18773109-00901006>
- Prillop, Külli; Hennoste, Tiit; Habicht, Külli; Metslang, Helle 2021.** Ei saa me läbi „Pragmaatika” korpuseta. Korpuspragmaatika ja pragmaatikakorpus. – Mäetagused, nr 81, lk 161–176.
- Rühlemann, Christoph 2007.** Conversation in Context: A Corpus-Driven Approach. London: Continuum.
- Thompson, Sandra 2002.** "Object complements" and conversation: Towards a realistic account. – Studies in Language, kd 26, nr 1, lk 125–164. <https://doi.org/10.1075/sl.26.1.05tho>

Tiit Hennoste (snd 1953), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi kaasprofessor (Jakobi 2, 51005 Tartu), tiit.hennoste@ut.ee

Külli Habicht (snd 1964), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi kaasprofessor (Jakobi 2, 51005 Tartu), kulli.habicht@ut.ee

Helle Metslang (snd 1950), PhD, Tartu Ülikooli emeriitprofessor (Jakobi 2, 51005 Tartu), helle.metslang@ut.ee

Küllil Prillop (snd 1974), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi teadur (Jakobi 2, 51005 Tartu), kulli.prillop@ut.ee

Kirsi Laanesoo (snd 1984), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi teadur (Jakobi 2, 51005 Tartu), kirsi.laanesoo@gmail.com

Liina Pärismaa (snd 1988), MA, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi nooremteadur (Jakobi 2, 51005 Tartu), liinaparismaa@gmail.com

Elen Pärt (snd 1988), MA, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi doktorant (Jakobi 2, 51005 Tartu), elen.p2rt@gmail.com

Andra Rumm (snd 1991), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi lektor ja teadur (Jakobi 2, 51005 Tartu), andra.rumm@ut.ee

Andriela Rääbis (snd 1966), PhD, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi teadur (Jakobi 2, 51005 Tartu), andriela.raabis@ut.ee

Carl Eric Simmul (snd 1991), MA, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi nooremteadur (Jakobi 2, 51005 Tartu), simmulman@gmail.com

Complement-taking predicate markers in different registers in Estonian

Keywords: Estonian language, corpus pragmatics, discourse markers, register variation

The use of Estonian CTP-markers (*ma arvan (et)* ‘I think that’, *(ma) usun (et)* ‘I believe that’, *(ma) loodan (et)* ‘I hope that’, *(mulle) näib (et)* ‘it seems to me that’, *(mulle) tundub (et)* ‘it feels to me that’, *(mulle) paistab (et)* ‘it looks to me that’ was analysed in everyday conversation, oral institutional communication, prose fiction, the press, scientific research, online chats and online comments. It was revealed that:

- 1) Compared to others, *(ma) arvan (et)* occurs very often.
- 2) The marker *(ma) arvan (et)* dominates in oral registers, exceeds the frequency of others in online registers, and appears relatively equally in print registers.
- 3) As for co-occurring pronouns, *arvan* usually takes one, while *loodan* and *usun* are split in half, and *tundub*, *näib* and *paistab* prefer dropping the pronoun altogether.
- 4) CTP-markers consisting of a verb only are extremely rare.
- 5) In oral interaction, *arvan* and *tundub* can often be used further down the sentence or clause, whereas the rest are usually found in sentence beginning, whatever the register.

Across registers, the use of CTP-markers falls into three types:

- 1) Oral communication uses plenty of markers, which are dominated by *(ma) arvan (et)* and many of which have been shifted further down the sentence.

2) Online texts make profuse and varied use of markers, but they split in two as far as marker shift is concerned, notably, online chats, which have relatively many shifted markers, vs online comments, where marker shift is relatively rare.

3) Print texts use a small number but diverse markers, being also divided in half shiftwise: prose fiction with more marker shift vs journalistic and science texts.

The main consituative conditions affecting the frequency, morphological choices and position of the markers are text production in dialogue and in real time, which enhance the number of markers, the tendency of keeping to a single marker, and marker shift further down the sentence (or clause). In addition, there are certain factors depending on text structure such as *conversation and imitation of dialogue* in fiction, which also contribute to the number and onward shift of markers.

Morphologically speaking, the marker with the highest pragmatic status is *arvan*, while *usun* and *loodan* are less high and the lowest on the scale are *tundub*, *näib* and *paistab*. In terms of positional variation it is *arvan* and *tundub* that seem to enjoy the highest pragmatis status. The sequence (*ma*) *arvan* (*et*) is a general marker, which has many functions and can be used in place of other CTP-markers.

Tiit Hennoste (b. 1953), PhD, University of Tartu, Institute of Estonian and General Linguistics, Associate Professor (Jakobi 2, 51005 Tartu), tiit.hennoste@ut.ee

Küllli Habicht (b. 1964), PhD, University of Tartu, Institute of Estonian and General Linguistics, Associate Professor (Jakobi 2, 51005 Tartu), kulli.habicht@ut.ee

Helle Metslang (b. 1950), PhD, University of Tartu, Professor Emerita (Jakobi 2, 51005 Tartu), helle.metslang@ut.ee

Küllli Prillop (b. 1974), PhD, University of Tartu, Institute of Estonian and General Linguistics, Researcher (Jakobi 2, 51005 Tartu), kulli.prillop@ut.ee

Kirsi Laanesoo (b. 1984), PhD, University of Tartu, Institute of Estonian and General Linguistics, Researcher (Jakobi 2, 51005 Tartu), kirsi.laanesoo@gmail.com

Liina Pärismaa (b. 1988), MA, University of Tartu, Institute of Estonian and General Linguistics, Junior Researcher (Jakobi 2, 51005 Tartu), liinaparismaa@gmail.com

Elen Pärt (b. 1988), MA, University of Tartu, Institute of Estonian and General Linguistics, Doctoral Student (Jakobi 2, 51005 Tartu), elen.p2rt@gmail.com

Andra Rumm (b. 1991), PhD, University of Tartu, Institute of Estonian and General Linguistics, Lecturer and Researcher (Jakobi 2, 51005 Tartu), andra.rumm@ut.ee

Andriela Rääbis (b. 1966), PhD, University of Tartu, Institute of Estonian and General Linguistics, Researcher (Jakobi 2, 51005 Tartu), andriela.raabis@ut.ee

Carl Eric Simmul (b. 1991), MA, University of Tartu, Institute of Estonian and General Linguistics, Junior Researcher (Jakobi 2, 51005 Tartu), simmulman@gmail.com