

DISKURSUSEMARKER (MA) ARVAN (ET)¹

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

Annotatsioon. Artiklis analüüsitakse diskursusemarkeri (*ma*) *arvan* (*et*) variantide kasutust 15 tekstirühmas, mis koonduvad kolme registrisse: trükitekstid, netitekstid ja suulised tekstid. Materjal pärineb Keeleveebi, suulise keele ja netivestluste korpuselt. Uuritakse, mille poolest eristub diskursusemarker samakujulisest komplementlause pealausest ning kas markeri kasutussagedus, positsioon lauses ja ehitus erinevad eri tekstirühmades ja/või registrites. Analüüs näitab, et kõik kolm näitajat varieeruvad tekstirühmiti ja/või registriti. Partiklite kasutuse poolest eristuvad selgelt suuline keel ja mitteilukirjanduslikud trükitekstid, seevastu ilukirjandus ja netitekstid moodustavad heterogeense vaheala. Suulist keelt iseloomustavad markeri suurem kasutussagedus ja sagedasem tagapoolne asend, sagedasem subjekti *ma* säilimine ja sidendi *et* puudumine ning prosoodilise piiri puudumine markeri ja ülejäänud lause vahel. Mitteilukirjanduslike trükitekstide keelt iseloomustavad markerite eesasend, suur subjektita markerite hulk ja *et*-markerid. Kokkuvõttes seostuvad markerite kasutuse erinevused keelekasutuse spontaansuse või redigeeritusega, dialoogilisusega, tekstiehituse protsessuaalsusega, teksti personaalsuse või impersonaalsusega, kirjakeele reeglite mõjuga.

Võtmesõnad: eesti keel, pragmaatika, subjektiivsus, registriline varieerumine, grammatiseerumine, pragmatiseerumine

¹ Artikli valmimist on toetanud Euroopa Regionaalarengu Fond (Eesti-uuringute Tippkeskus) ja Eesti Teadusagentuur (projekt PRG341).

1. Sissejuhatus

Artikli teemaks on diskursusemarkeri² (*ma arvan* (*et*) kasutus eesti keele registrites. Marker võib esineda neljal kujul (*ma arvan*, *ma arvan et*, *arvan*, *arvan et*) ja kanda kahte keerukalt eristatavat põhitähendust: tugevalt isikliku arvamuse avaldamine ehk 'minu arvates' (näide 1) ja üldisem ebakindluse väljendamine propositsiooni tõeväärtuse suhtes ehk 'arvatavasti' (näide 2).³

(1) **Arvan**, sa oled lihtsalt rumal inimene. (KOM)

(2) **Arvan, et** igäüks on kokku puutunud oma elus inimestega, kellele väga meeldib igasugu õpetusi jagada, kes justkui teavad „kuidas asjad tegelikult käima peaksid“. (DOK)

Ma arvan kuulub lausekujuliste markerite sarja, mida esineb eri keeltes, nt inglise *I think, I guess* 'ma arvan' (Brinton 2017), prantsuse *je pense* 'ma arvan', *je crois* 'ma usun, ma arvan', rootsi *jag tycker, jag tror* 'ma arvan' (Mullan, Karlsson 2012; Auer, Lindström 2016), saksa *ich denke* 'ma arvan', *ich glaube* 'ma usun' (Breitkopf-Siepmann 2012; Auer, Lindström 2016). Ka eesti keeles on verbidest kujunenud mitmeid subjektiivsuse markereid, nt *ma arvan*, *ma usun*, *ma mõtlen*, *ma tean*, *ma kardan*, *ma ütlen* (Hennoste 2004, 2006; Keevallik 2003: 74–100, 2010). Eriti markerit *I think* on palju uuritud. Suurem osa uuringuid puudutab üksuse grammatiseerumist (nt Thompson, Mulac 1991; van Bogaert 2010; Degand, Simon-Vandenberg 2011; Beijering 2012; Heine 2013, 2018; Degand, Evers-Vermeul 2015; Brinton 2017). Meie artikli teema seisukohast on oluline, et grammatiseerunud/pragmatiseerunud (markeriks muutunud) üksuste kõrval on samal ajal kasutusel ka sama vormiga pealaused ning üksused, mis paiknevad reanalüüsiprotsessi keskel lausete ja markerite vahealas. Eri autorid näevad grammatiseerumist osalt erinevana (vrd Brinton 2017 ja Heine jt 2019), samal ajal tuuakse aga markeri määratlemisel välja suuresti samu tunnuseid (nt Fischer 2007; Fraser 2009; Heine 2013; Heine jt 2019).

² Neid üksusi on nimetatud eri viisil: diskursusepartiklid, pragmaatilised partiklid, siduvad/ühendavad määruused (Heine jt 2019: 2), episteemilised adverbid (Keevallik 2010), episteemilised kiilud (*epistemic parentheticals*) (Thompson, Mulac 1991, Kalteneböck 2007; Schneider 2007), hoiakumarkerid (*stance markers*) (Kärkkäinen 2003) jm.

³ Lühendite seletusi vt artikli lõpust.

Markeri kasutust eri tekstirühmades või registrites pole kuigivõrd uuritud ja tulemused paistavad vasturääkivad. Meile teadaolevad vähesed uurimused puudutavad üksnes suulist dialoogi. Näiteks on Aijmer (1997) osutanud, et *I think* on tüüpiline mitteformaalses (argises) vestluses, samas Simon-Vanderbergen (2000) on leidnud, et *I think* on poliitilises vestluses peaaegu kaks korda tavalisem kui mitteformaalses vestluses.

Artikli esimeses osas anname ülevaate sellest, kuidas me markerit määratlesime ehk mis eristab diskursusemarkerit (*ma arvan* (*et*) samavormilisest komplementause pealausest. Teises osas otsime vastuseid kolmele uurimisküsimusele.

- Kas markeri kasutussagedus erineb tekstirühmiti ja/või registriti?
- Kas markeri esinemissagedus eri süntaktilistes positsioonides (lause alguses, keskel, lõpus või järel) erineb tekstirühmades ja/või registrites?
- Kas markeri variandid (*ma arvan*, *ma arvan et*, *arvan*, *arvan et*) on seotud eri tekstirühmade ja/või registritega?

Materjali leidmiseks kasutasime portaali Keeleveeb⁴ korpusi, suulise keele korpus⁵ ja netivestluste korpusi, mis sisaldab aastatel 2010–2020 kogutud Skype'i, Messengeri jms vestlusi (95 000 sõnet). Kokku valisime uurimiseks 15 korpus⁶. Valitud korpused koonduvad kolme registrisse, milles omakorda saab eristada tekstirühmi.

I. Trükitekstid

- a. ilukirjanduslik proosa
- b. ajakirjandus: ajakiri Kroonika, nädalaleht Eesti Ekspress, päevalehed Öhtuleht, Eesti Päevaleht ja Postimees
- c. (populaar)teadus: teadustekstid, doktoritööd

II. Netitekstid

- a. reaaliajase argidialoogi: netivestlused, jututoad
- b. avalik mittereaaliajase suhtlus: foorumid, uudisgrupid, kommentaarid

⁴ keeleveeb.ee.

⁵ <https://keeleressursid.ee/et/220-suulise-est-kee-keele-korpus>.

⁶ Keeleveebis kättesaadavatest korpusetest jätsime välja väiksemad ajakirjandusväljaanded (Lääne Elu, Maaleht, Valgamaalane; Luup), teaduslikud ja populaarteaduslikud ajakirjad, kus *arvan*-verbi esines alla 50 korra (Horisont, Eesti Arst, Arvutustehnika ja Andmetöötlus), Eesti ja Euroopa seaduste korpuse, kus esines vaid üks *arvan*, samuti riigikogu stenogrammide kui üheselt määratletamatu tekstiliigi.

III. Suulised tekstid

- a. ametlik-avalik suhtlus
- b. argisuhtlus

Verbivormi *arvan* sisaldavaid lauseid oli valitud korpustes kokku 34 000. Põhjalikumaks uurimiseks võtsime igast korpusest juhuvalikuga 200 sõna *arvan* kasutuse näidet. Neljas korpuses oli näiteid vähem kui 200: netivestlustes 87, teadustekstides 55, doktoritöodes 71 ja suulises ametlikus suhtluses 159. Kokku analüüsisime 2572 näidet.

2. Diskursusemarkeri määratlemine

Analüüsi lähtekohaks on diskursusemarkeri (*ma*) *arvan* (*et*) tunnuste määratlemine. Lähtusime rahvusvahelises kirjanduses välja toodud tunnustest ning lisasime omalt poolt eesti keelega seotud erijooned. Varasemates uuringutes esitatud markeri tunnused koonduvad kuude rühma:

- marker on muutumatu ja ainult markeri ainest sisaldav üksus;
- marker ei moodusta suhtluses või tekstis pealiini;
- markeris puudub üldjuhul sidend/partikkel (nt *that*) või on see hääldatud verbiga kokku;
- markerina võib üksus paikneda nii lause alguses kui ka tagapool;
- marker on suulises keeles rõhutu subjektiga;
- marker moodustab suulises keeles prosoodiliselt eraldi üksuse.

2.1. Muutumatu marker

Üldise arusaama järgi on diskursusemarker terviklik üksus, mis koosneb ainult markeri elementidest ja esineb ühel kindlal kujul (nt verbi sama aeg, isik, kõneviis, kõneliik). Oma uuringus oleme valinud analüüsimiseks üksnes markeri verbivormiga *arvan*, mitte nt *arvad*, *arvame*, *arvaks*, *arvaku*. See aga ei tähenda, et vaid *arvan* on markerina kasutusel. Markeriks võivad kujuneda sama verbi eri vormid, kuid need ei pea olema ühe markeri variandid. Näiteks eesti keeles on markerid nii *ma ütlen* kui ka *ma ütleks*, kuid nende tähendus ja roll on erinevad. *Ma ütlen* tugevdab lause referaatosa sõnumi tõsikindlust, samal ajal kui *ütleks* on sõnumi pehmendaja (Hennoste 2004: 514–518). Et tingiv kõneviis on ka üldiselt teise grammatilise tähendusega kui kindel kõneviis, jäävad konditsionaalvormid siinse artikli huvisfäärist välja.

Keerukam on olukord teiste kindla kõneviisi vormidega. Nagu iga-sugused kinnisüksused võivad ka markerid mingil määral varieeruda. Nii on võimalik, et mitmuse esimese isiku vorm *arvame* ning teise isiku verbivormid *arvad* ja *arvate* on *arvan*-markeri variandid (vt Helasvuo 2014), kuid selle väljaselgitamiseks on vaja eraldi kvalitatiivset analüüsi. Selles artiklis oleme piiranud analüüsitava üksuse *arvan*-vormiga. Pealegi esineb muid nimetatud vorme tekstides väga harva.

Markeriga ei ole tegu, kui lauses on lisaks muid lauseliikmeid (sh kõrvallauseid) või mittelauseliikmelisi elemente (*tegelikult*, *aga*, *no*, *ei no*, *ka*, *küll*, *siiski*, *kui*, *kuigi* jms, näited 3 ja 4). Samuti ei ole markeriga tegu juhul, kui *arvan* on rinnastatud teise lause või verbiga (näide 5).⁷

- (3) kui siin on sellise nimega nagu ruthe polvas siis vajutagu 999999 voi privagu <- **mingil mystilisel põhjusel ma arvan**, et ei ole siin kedagi sellise nimega vot (JT)
- (4) **ma=arvan`ka=et** seda saab `kiiremini teha. (SUUL AR)
- (5) Me peame kindlasti ametlikult aru pärima ja **ma arvan**, et kohtumine Läti peaministriga ongi üks selline võimalus. (P)

Eesti keele seisukohast tekitab markeri määratlemisel probleemi sõnajärg ja subjekti puudumine. *Ma arvan* esindab neutraalset sõnajärge, kuid teatud tingimustel toimub liitlause pealause aluse ja öeldise inversioon (Erelt 2017: 553). Selline muutus näitab, et üksus ei ole sõltumatu, vaid allub üldistele süntaksi reeglitele. Seega käitub inversiooniline *arvan ma* pigem pealausena, kuigi üksuses puudub alistusele viitav *et*.⁸ Samas on tegu piirinähtusega, mille puhul on võimalikud eri tõlgendused (vt Imo 2011; Posio 2014). Näiteks on Peter Auer ja Jan Lindström (2016: 77) rootsi ja saksa analoogiliste sõnajärjemuutustega üksusi uurides leidnud, et tegu ei ole n-ö puhta markeriga, nagu on inglise *I think*. Samas paigutavad nad pöördjärjega markerid siiski põhijärge esindavate üksustega kokku, tuues põhjenduseks nende sama semantika (samas).

Oma uurimuses oleme jätnud pöördjärjega üksused markerite hulgast välja (näide 6). Põhjuseks on V2-põhimõtte erinev rangus. Nimelt on rangema sõnajärjega saksa ja rootsi keeles V2-põhimõte palju tugevam kui eesti keeles. Aueri ja Lindströmi näidete seas on vaid üksikuid

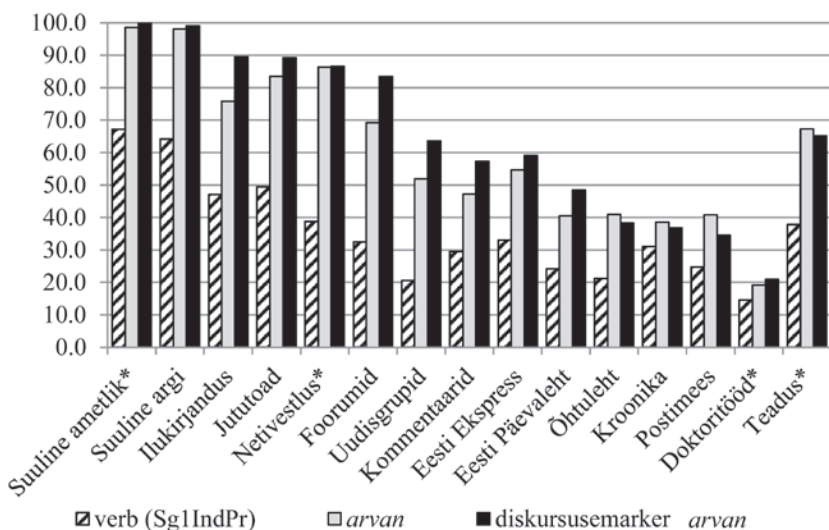
⁷ Suulise keele näidete transkriptsioonimärke vt artikli lõpust.

⁸ Liitlauseis võib sidend puududa, nt *Iga lollgi mõistab, [et] riigi majandus on oma-dega põhjas* (Erelt 2017: 687).

kõrvalekaldeid sellest põhimõttest (nt saksa verbide *denken* ja *glauben* puhul 3%). Eesti keeles on inversiooni kasutamine vabam ja sel juhul viitab pöördjärg tugevamalt kasutaja otsustusele. Üldist statistilist pilti see piirang ei muuda, sest pöördjärjega üksusi oli siinse materjali hulgas väga vähe.

(6) ”Staar on kuidagi seotud kassaeduga, **arvan** ma.” (P)

Eesti keeles on subjekt vähemalt osal juhtudel fakultatiivne lauseelement (Erelt jt 2017: 252–257) (näide 7). Meie hüpotees oli, et subjekti puudumine on pigem markeri tunnus. See eeldab, et markeriteks arvatute hulgas võiks olla palju subjektita üksusi ja lausete hulgas oleks neid minimaalselt. Katseks võrdlesime teiste tunnuste alusel markeriteks ja lauseteks määratud üksusi. Võrdlus näitas, et hüpotees ei kehti. *Ma* jäetakse *arvan*-vormi eest ära keskmiselt 39%-l ja markerina toimiva *arvan*-vormi eest keskmiselt 35%-l juhtudest. Seejuures kasutatakse subjekti kõigil *arvan*-juhtudel sagedamini kui verbivormidega keskmiselt (joonis 1).



Joonis 1. Subjekti esinemus (%) verbi 1. isiku vormi ees, *arvan*-vormi ees (kogu korpuse põhjal) ja markerina kasutatud *arvan*-vormi ees (200-lauseliste valimite põhjal). Tärniga on tähistatud korpused, milles *arvan* esines vähem kui 200 korral

Seda võib seletada mitme keele puhul täheldatud tendentsiga säilitada subjekti just kognitiivse tähendusega verbide korral ja nende grammatiseerunud konstruktsioonides (Posio 2014; Väänänen 2016: 190–191). Seetõttu jätsime subjekti esinemise või puudumise markeri tunnuste hulgast välja.

(7) **Arvan, et** peagi leitakse asjale lahendus. (SLÕ)

2.2. Suhtluse ja teksti kõrvalliin

Suhtlus või tekst liigub tüüpjuhul edasi mööda pealauseid, ent markerina toimiva formaalse pealause korral liigub see mööda grammatilist kõrvallauset. See tunnus ilmneb eriti selgelt suulises dialoogis. Näites (8) lõpetab V voo *millal*-küsimus, mille vastuseks on H voo *siis kui=mul seal `korter `üle antakse* (read 02–05), mis on formaalselt kõrvallause. *Ma=arvan et* on marker, mille võib tekstiliini lõhkumata ära jätta, kuigi see muudab lause mõtet (8).

- (8) 01 V: siis me `kuuendal ok`toobril sinna `sisse läheme tead=`noh?
02 (0.8) aga noh seal=et `siis e (0.3) et ma=i `tea ütleme siis=et
03 siis et et `millal sa siis võtmed saad=ju.
04 (0.6)
05 H: **ma=arvan et siis kui=mul seal `korter `üle antakse ju.**
(SUULAR)

Sama kasutus on leitav ka kirjalikus dialoogis ja monoloogis. Näites (9) vastab B küsimusele *Kas näiteks RAM-i suurendamine annab mingit efekti?* formaalse kõrvallausega *lisamälust oleks abi küll*. Näites (10) jätkub tekst pärast *ma arvan, et* osa mööda kõrvallauset: *Kirikuid on rohkem kui koole ja haiglaid > see viga seal parandatakse*.

- (9) A: Kas näiteks RAM-i suurendamine annab mingit efekti? Või on veel mingeid võimalusi?
B: **Ma arvan et lisamälust oleks abi küll** ja Windowsil tuleks tõesti kõik üleliigne taustal jooksev soft selleks ajaks keelata. (UG)
- (10) Kirikuid on rohkem kui koole ja haiglaid. **ma arvan, et** see viga seal parandatakse üpris kiiresti Putini poolt. (KOM)

Näited, mille korral vestluse või teksti pealiin liigub mööda grammatilist pealauset, jätsime markerite alt välja. Näites (11) kasutab E lauset *ma arvan=et*, partner reageerib sõnale *arvan* ja teeb asenduse *arvan > kindlasti* (read 03–04). Näites (12) esitatakse küsimus otseselt arvamise kohta ning vastus pakub sama verbi öeldisena (*arvate > arvan*). Nende kõrval on näiteid, milles (*ma*) *arvan* (*et*) ei seostu üheselt suhtluse pea- või kõrvalliiniga. Sellisel juhul pole seda tunnust võimalik üksuse määramisel kasutada.

(11)01 I: `see `sort on `ilus.

02 (2.8)

03 E: no ma `arvan=**et** ta `teab selle sordi nime[tust]

04 T: [ei] `kindlasti teab.

(SUUL AR)

(12) Mida **arvate** tervetest inimestest kes liiguvad alla 10 tuhande sammu päevas?

Arvan, et neil pole aega iga päev 2 tundi kõndimise alla panna ja nad hoiavad end vormis näiteks tennis käimisega (ajakulu sellele on märksa väiksem). (FOO)

2.3. Partikkel *et* ja markeri asukoht

Mitmes mudelis on esmaseks grammatiseerumise märgiks peetud sidendi kadumist (näide 13). Sidendita üksused arvatakse markerite hulka, kuid samal ajal ei väideta, et ainult need on markerid. Näiteks osa inglise keele uurijaid leiab, et sidendi *that* olemasolu või puudumine ei anna alust määratleda üksus kas markeriks või lauseks (nt Aijmer 1997; Kärkkäinen 2003; vt ka Dehé, Wichmann 2010: 38–39).

(13) **Ma arvan**, kui ta seda kirjutas, ei teadnud ta veel, kui tõsi see on. (ILU)

Sidendita marker kaotab oma kindla lausealgulise positsiooni ja võib paikneda peale (osa)lause alguse ka keskel või lõpus, samuti eraldi üksusena (inkremendina) lause järel (Thompson 2002; Brinton 2017: 168; Heine 2018: 37) (näide 14). Sidendit *et* sisaldavad markerid esinevad peaaegu ainult lausete või osalausete alguses, tagapool leidub vaid mõni juhunäide (näide 15).

- (14) Alguses: **Arvan**, sa oled lihtsalt rumal inimene. (KOM)
 Keskel: Koeramaks, **arvan**, küsimust ei lahenda. (EE)
 Lõpus: Ühe söögikorra võid sa ju sanatooriumis vahele ka jätta, **ma arvan!** (ILU)
 Järel: vajuta teise nupuga peale ja disable. **ma arvan.** (FOO)
- (15) .hh mm pidulikule `üritusele **ma arvan et** ta peaks olema väemalt `kuue`kilone. (SUUL AM)
 Ja tegelikult seisis mul risti ees mingi, **ma arvan et** kristliku sisuga keeld .. (ILU)

Lause alguses olevad ja sidendit *et* sisaldavad üksused võivad olla nii markerid kui ka pealused. Suulises keeles eristab neid hääldus. Grammatiseerunud verbivorme iseloomustab verbi ja partikli kokkuhääldus, nii *et* verb ja *et* moodustavad prosoodilise terviku (näide 16). Selline *et* ei ole enam kõrvallauset alustav sidend, vaid verbiga kokku kuuluv partikkel, millele järgnev lause on tõlgendatav iseseisvana (*`äkki se ei `ole Pissi `ekspert*) (Hennoste 2004: 510; Seppänen, Laury 2007: 556).

- (16) **ma=arvan=et** see kaks tuat `kümme > tuleb meil mõlemal `selline nagu < `rahulik (.) selline `ülesehitamise aasta. (SUUL AR)
 e **ma `arvan=et** `äkki se ei `ole Pissi `ekspert. (SUUL AR)

Piirirühma jäävad juhtumid, milles *et* ei ole hääldatud verbiga kokku, kuid pole sellest ka prosoodilise piiriga eraldatud (näide 17). Sellised näited arvasime siinses analüüsis samuti markerite hulka.

- (17) mhmm .hh **ma arvan et** ma võibolla=sis `meilin et=sis=on `mõtlemisaega `ka. (SUUL AM)

Kokkuvõttes arvasime markerite hulka partiklita *et* üksused ja sellised suulise keele üksused, milles partikkel *et* ei ole verbist prosoodilise piiriga eraldatud. Samal ajal ei ole partikli *et* esinemine või eraldi hääldus põhjus, miks üksus markeri alt välja arvata. Sellisel juhul toetusime analüüsis teistele markeri tunnustele.

2.4. Marker ja selle osade rõhud

Inglise keeles on leitud, et konstruktsiooni *I think* funktsioonide eristamisel suulises suhtluses on olulised komponentide rõhud: kogu üksuse mitterõhulisus viitab diskursusemarkerile, rõhuline *I* aga osalausele (Dehé,

Wichmann 2010). Oma analüüsis järgisime sama määratlust. Sealjuures olid analüüsitud näidetes kõik lause algusest tagapool paiknevad üksused tervikuna rõhutatud (näide 18).

(18) no: sealt=tules tuleb ` suur=raha tuleb sealt **ma arvan**. (SUUL AR)

Eesti suulises materjalis rõhulisi *ma*-vorme ei leidunud. Rõhulist varianti esindab asesõna pikk vorm *mina* (näide 19), mis esineb ka kirjalikus tekstis (näide 20). Sellised üksused jätsime markerite alt välja.

(19)01 H: .hhhhhh ted ` **mina arvan**=et me [võiksim]=ned `parimad
02 M: [* jah *]
03 H: `joogid ära proovida praägu,
04 A: jah (SUUL AR)

(20) „**Mina arvan, et** odaviskele ei aita doping üldse kaasa,“ turtsatas Värnik.
(SLÖ)

Samuti on leitud, et rõhuline *think* osutab episteemilisele kommentaarosalausele (*epistemic comment clause*; Dehé, Wichmann 2010). Selliselt nimetatud üksus on funktsioonilt võrreldav eesti markeriga tähenduses 'minu arvates', mis väljendab kasutaja ebakindlust öeldu/kirjutatu tõeväärtuse suhtes.

Rõhulise verbiga üksused esinesid meie näidetes üksnes lausete alguses. Nende seas eristub kaks varianti. Ühe rühma moodustavad üksused, milles *arvan* ja *et* on kokku hääldatud ning mis moodustavad sellest seisukohast ühtse markeri. Näites (21) on tegu tervikuks hääldatud üksusega ` *arvan=et* (rida 04). Sellised üksused oleme arvanud markerite hulka. Teise rühma kuuluvad juhtumid, milles esineb rõhuline *arvan* ja *et* ei ole hääldatud verbiga kokku. Sellised üksused jäävad markerite piirile. Siin oleme lähtunud analüüsis sellest, kas teksti pealiin liigub mööda formaalset pea- või kõrvallauset. Näites (22) on H vastus V küsimusele *ma` räägin sis kui ma` tullen* (rida 03). Sellele eelnev *ma` arvan et* toob esile kasutaja ebakindluse järgneva vastuse tõeväärtuse suhtes, millele viitavad pikk paus enne vastamist (rida 02), pikalt venitatud üneem *mmmmm* ja sellele järgnev paus (rida 03).

(21)01 H: tõõ-q @ tead=se=oli nagu [nimodi] @ se=li ästi `lahja `ka:.=
02 M: [hmh]
03 M: =`oli=jah
04 H: ma ` **arvan=et** `ästi lahja võibla `seitse kaheksa `kraadi. (SUUL AR)

- (22) 01 V: .hh kas=sa=tad `Retiga ka `rääkida.
 02 (1.7)
 03 H: mmmmm (1.6) **ma `arvan et** ma `räägin sis kui ma `tulen.
 04 (0.3)
 05 V: oki. (SUUL AR)

2.5. Marker ja lauseprosoodia

Eraldi probleemiks on prosoodia. Rahvusvahelises kirjanduses tuuakse esile, et markerid moodustavad üldjuhul eraldi prosoodilise üksuse (nt O'Grady 2017). Samal ajal ei ole see kindel markeri tunnus. Nii on inglise keele, aga ka saksa, rootsi ja eesti keele puhul osutatud, et marker on ülejäänud lausega valdavalt prosoodiliselt integreeritud (Keevallik 2010; Auer, Lindström 2016). Ka meie analüüsitud suulises materjalis on peaaegu kõik lauses markeriteks määratud üksused prosoodiliselt integreeritud ümbritseva lausega (näide 23a), leidub vaid üksikuid eraldatud juhtumeid (näide 23b). Prosoodiliselt eraldatud on aga markerid, mis paiknevad inkrementina lause järel (näide 24).

- (23a) **ma=arvan=et** tööhõive ` ametis peaks ennast=ää `kirja panema. (SUUL AM)
 ` hinnad on niuke **ma=arvan** siuke `keskmine siuke ` hinnatase on ` sama
 mis Eesti ` mõni asi `kallim, ` mõni asi ` odavam. (SUUL AR)
- (23b) **ma `arvan, et** ma katsun mõlemad `ära mahutada. (SUUL AM)
- (24) 01 V: aa, ma võin aint `vägesi `juhatada.
 02 (1.1)
 03 H: no=aga `hää `toogi. (0.3) **ma arvan.** (SUUL AR)

Samal ajal on kirjalikes tekstides olukord vastupidine: peaaegu kõik markerid on eraldatud sõna *arvan* järel paikneva komaga (näide 25a). Selline kasutus vastab ühemõtteliselt kirjakeele komareeglitele. Leidub vaid üksikuid komata näiteid (näide 25b).

- (25a) **Ma arvan**, kui ta seda kirjutas, ei teadnud ta veel, kui tõsi see on. (ILU)
 (25b) No erilisi eeliseid **ma arvan** ei olegi pigem on mingi välimuse vahe
 vms ... või siis mingi käikude vahe, aga muid asju nagu ei tohiks olla.
 (FOO)

Eelneva põhjal leidsime, et kõiki keeleregistreid arvestades ei saa prosoodilist seotust/eraldatust käsitleda markeri tunnusena, kuid see on oluline suuliste ja kirjalike registrite eristaja.

Kokkuvõttes valisime analüüsimiseks ainult konstruktsiooni (*ma*) *arvan* (*et*), milles *ma* ja *et* võivad esineda või mitte. Verbi grammatilised variatsioonid (konditsionaal, mitmus, teine isik, pöördõnajärg) jätsime kõrvale. Markerite hulka arvasime kõik näited, milles teksti või dialoogi pealiin liikus mööda formaalset kõrvallauset. Juhul kui (*ma*) *arvan* (*et*) ei seostunud üheselt suhtluse pea- või kõrvalliiniga, kasutasime teisi tunnuseid. Markerite hulka arvasime a) üksused, milles puudus sidend *et*; b) suulise keele üksused, milles *et* ei olnud verbist prosoodilise piiriga eraldatud; c) üksused, mis esinesid (osa)lause algusest tagapool; d) üksused, milles subjekt on asesõna *ma*, mitte selle rõhuline variant *mina*.

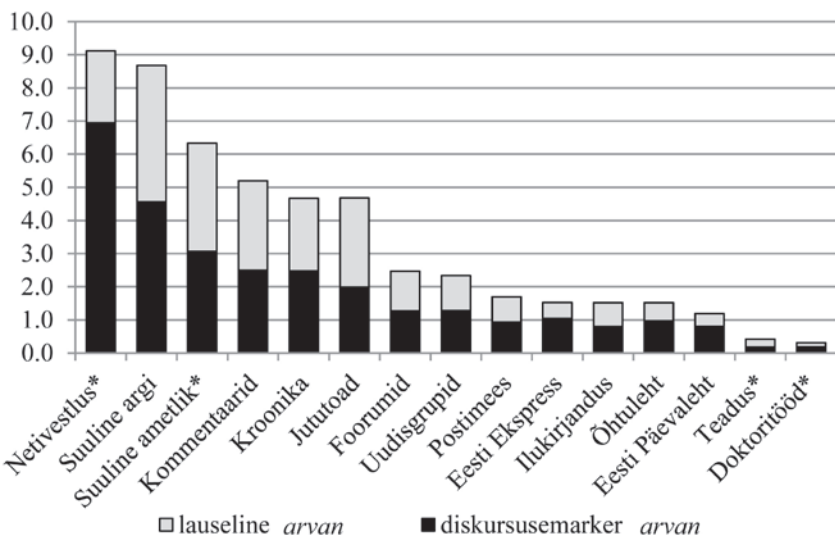
3. Diskursusemarkeri kasutus

Eelmises osas tõime välja diskursusemarkeri (*ma*) *arvan* (*et*) tunnused, mille alusel analüüsisime tekste. Selles osas esitame analüüsitulemused. Esmalt vaatame markerite üldist esinemissagedust. Seejärel uurime, kas markeri esinemissagedus erineb eri süntaktilistes positsioonides: (osa)lause alguses, keskel, lõpus või lause järel. Lõpuks vaatame markerivariantide esinemust eri tekstirühmades.

3.1. Markeri esinemissagedus

Esimene küsimus oli, kas markeri sagedus erineb tekstirühmiti ja/või registriti. Arvutasime markeri hinnangulise sageduse korpustes 10 000 tekstisõna kohta. Selleks võtsime igast korpusest juhuvalikuga 200 *arvan*-lauseid ja leidsime nende põhjal markeri osakaalu korpustes. *Arvan*-vormi üldsageduse ja diskursusemarkerite osakaalu järgi arvutasime markeri hinnangulise sageduse korpustes (joonis 2).

Analüüs näitas, et markerite kasutussagedus eri registrites ja tekstirühmades on erinev. Skaala ühes otsas eristuvad argised reaaliajased netivestlused ja suulised argivestlused, kus markeri sagedus on 10 000 sõne kohta suurem kui 4, ning teises otsas doktoritööd ja teadustekstid, kus markeri kasutussagedus on alla 0,2 korra.

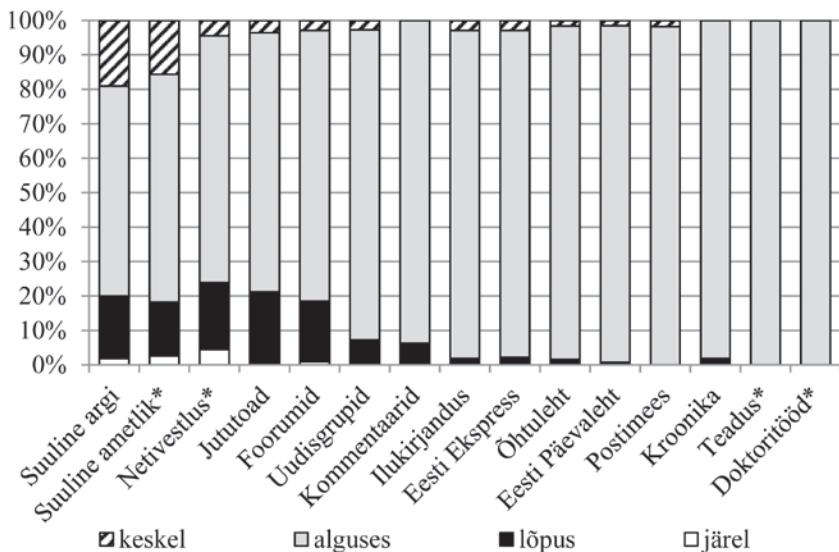


Joonis 2. Markeri hinnanguline sagedus 10 000 tekstisõna kohta

Kahe otsa vahele jäävad tekstirühmad jagunevad kaheks: sagedasema markerikasutusega rühma moodustavad ametlikud suulised vestlused, netikommentaariid, jututoad ja ajakiri Kroonika, harvema kasutusega rühma uudisgrupid ja foorumid ning ajalehed ja ilukirjandus. Eripärane paistab siin Kroonika positsioon, milles esineb vaatlusalust diskursusemarkerit oluliselt rohkem kui muudes ajakirjandusväljaannetes. Siin on usutavasti põhjuseks asjaolu, et Kroonika materjalis esines diskursusemarker (*ma arvan* (*et*) üksnes intervjuudes, samuti paistab silma nende intervjuude vähene toimetamine.

3.2. Markeri positsioon

Teine küsimus oli, kas markeri esinemissagedus eri süntaktilistes positsioonides erineb tekstirühmades ja/või registrites. Analüüs näitas, et üldiselt valitseb eesti keeles alguspositsioon (sama on toodud esile ka nt saksa ja rootsi suulise keele kohta, vt Auer, Lindström 2016: 77). Siiski eristus markeri positsioonide alusel selgelt kaks tekstide rühma (joonis 3).



Joonis 3. Markeri positsioon (osa)lauses

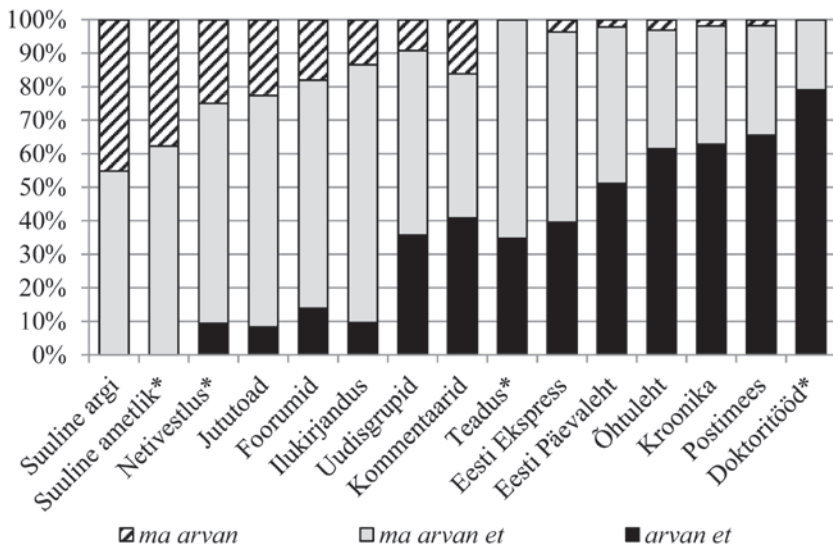
Ühe rühma moodustavad suulised tekstid, netivestlused, jututoad ja foorumid, milles lause keskel, lõpus ja järel paiknevaid markereid on kokku vähemalt 20% ja seega oluliselt rohkem kui mujal. Ülejäänud tekstirühmades jääb väljaspool lause algust paiknevate markerite sagedus alla 10%. Sinna kuuluvad kõik trükitekstid ja netitekstidest need, mille eesmärk on pigem info või arvamuse edastamine kui vestlus (uudisgrupid ja kommentaarid). Lisaks selgus, et lause keskel paiknevad markerid esinevad oluliselt enam suulises keeles: seal on neid üle 15%, muudes tekstirühmades alla 5%.

3.3. Markeri ehitus

Kolmandaks küsisime, kas markeri vormivariandid on tekstirühmade ja/või registritega seotud. Analüüsis ilmnas, et eesti keeles on kasutusel neli markerivarianti: *ma arvan et* esines 716, *arvan et* 494, *ma arvan* 183 ja *arvan* 18 korral. Üldpilt näitab *et*-markerite valitsemist kõigis tekstirühmades (joonis 4), neid on kokku 86% kõigist markeritest. Suulises keeles, netikeeles ja ilukirjanduses on ilma *et*-ita markerite osakaal

marginaalsest suurem (vähemalt 10%). Enim leidub neid suulises keeles, kus nende kasutus jääb 38–46% vahele.

Analüüs tõi esile olulise erinevuse eesti ja inglise keele vahel, kus sidendiga/partikliga markeri esinemine on minimaalne. Samas on eesti keelel sarnasusi saksa, rootsi ja taani keelega, kus samuti leidub palju sidendi/partikliga kasutusjuhte (Beijering 2012: 183–185; Auer, Lindström 2016: 79–80).



Joonis 4. Markeri ehitus (jooniselt puudub *arvan*)

Analüüsi tulemusel leidsime, et seesmiselt moodustub markeri ehituse järgi kolm tekstide rühma. Esimese rühma moodustab suuline keel. Seda iseloomustab pronoomenit *ma* sisaldavate markerite domineerimine ja pronoomenita markerite puudumine. *Et* võib sealjuures markeri koosseisus olla või mitte. Ilma *et*-ita variandi *ma arvan* suurem sagedus suulistes tekstides on seotud markeri asukohaga lauses: lause lõpus ega inkrementina pole (*ma*) *arvan et* võimalik ja lause keskel esineb see vaid üksikutel juhtudel. Teine, sellele vastanduv rühm on ajakirjanduse ja (populaar)teaduse keel, kus valitsevad *et*-markerid ja sidendita *et* markerite kasutus on minimaalne (alla 6%). *Ma* võib sealjuures esineda või mitte. Heterogeense

vaherühma moodustavad netitekstid ja ilukirjandus, mille võib omakorda jagada kaheks ennekõike *ma* esinemise sageduse alusel.

4. Tulemused

Analüüsisime diskursusemarkeri (*ma*) *arvan* (*et*) kasutust eri registrites ja tekstirühmades. Analüüs näitas, et kasutusel on neli markeri varianti: *ma arvan*, *ma arvan et*, *arvan* ja *arvan et* ning nii markerite sagedus, positsioonid lauses kui ka vorm varieeruvad tekstirühmiti ja/või registriti. Tulemused on kokku võetud tabelis.

Tabel. Diskursusemarkerite kasutus tekstirühmades.
(Tähistus: + jah, (+) pigem jah, (-) pigem ei, - ei)

	Suulised tekstid		Netitekstid					Trükitekstid		
	AR	AM	NV	JT	FOO	KOM	UG	ILU	AJA	TEA
Lauses tagapool	+	+	(+)	(+)	(+)	(-)	(-)	(-)	-	-
Sagedane	+	(+)	+	(+)	(-)	(+)	(-)	(-)	(-)	-
<i>ma</i> peab olema	+	+	(+)	(+)	(+)	(-)	(-)	(+)	(-)	-
<i>et</i> võib puududa	+	+	(+)	(+)	(+)	(+)	(+)	(+)	(-)	-

Markeri hinnanguline sagedus 10 000 tekstisõna kohta erineb tekstirühmiti ja registriti. Laiema üldistusena saab välja tuua skaala ühes otsas suulised tekstid, keskel netitekstid ja teises otsas trükitekstid. Trükitekstide seas eristuvad omakorda teadustekstid, milles markerite kasutus on eriti vähene.

Markeri positsioonide analüüs näitas olulisi erinevusi lause algusest tagapool esinevate markerite kasutuses. Siin eristusid suulised tekstid, netivestlused, jututoad ja foorumid, milles lause algusest tagapool paiknevaid markereid oli kokku vähemalt 20%. Ülejäänud tekstirühmades jäi kaugemal paiknevate markerite sagedus alla 10%.

Kõige üllatavam oli markeri ehituse ja tekstirühmade seos. Üldpilt näitas, et kõigis tekstirühmades valitsevad statistiliselt *et*-markerid. Samas selgus, et eri tekstirühmades ja registrites eelistatakse markeri eri vormivariante.

Sõna *et* esinemise/puudumise seisukohalt moodustavad skaala ühe otsa mitteilukirjanduslikud trükitekstid, milles ilma *et*-ita markereid kasutatakse minimaalselt või üldse mitte. Skaala teise otsa jäävad suulised tekstid, mida iseloomustab sage ilma *et*-ita markerite kasutus. Netitekstid ja ilukirjandus moodustavad vaheala.

Ma esinemise/puudumise seisukohalt on pilt osalt erinev. Siin moodustavad ühe poole tekstid, milles ilma *ma*-ta markerite kasutus on väga sage: mitteilukirjanduslikud trükitekstid ning ka uudisgrupid ja kommentaarid. Teisele poole jäävad suulised tekstid, milles ilma *ma*-ta markerite kasutus analüüsitud materjalis puudus, ning netivestlused, jututoad, foorumid ja ka ilukirjandus, kus see oli vähene.

Et-i ja *ma* kasutust kokku võttes saame markerite kasutuse põhjal eristada kolme tekstide rühma:

- 1) mitteilukirjanduslikud trükitekstid, mida iseloomustavad *et*-markerite valitsemine ja palju ilma *ma*-ta markereid;
- 2) suulised tekstid, milles esineb palju ilma *et*-ita markereid ja ilma *ma*-ta markerid puuduvad;
- 3) ülejäänud tekstid, milles esineb suhteliselt palju ilma *et*-ita markereid ja mis jagunevad kaheks:
 - 3a) netivestlused, jututoad, foorumid ja ilukirjandus, milles on vähe ilma *ma*-ta markereid;
 - 3b) uudisgrupid ja kommentaarid, milles esineb palju ilma *ma*-ta markereid.

Võttes kokku sageduse, positsiooni ja vormi, saame diskursusemarkeri (*ma*) *arvan* (*et*) kasutuses neli rühma: suulised dialoogid, mitteilukirjanduslikud trükitekstid, ilukirjandus ja netitekstid.

Kujutletava skaala ühe otsarühma moodustab suuline register oma kahe variandiga. Siin esinevad markerid muudest rühmadest oluliselt sagedamini. Ainuvalitsev on *ma*-kasutusega marker, lisaks esineb muudest rühmadest rohkem markereid, milles *et* puudub. Tavaline on verbi ja *et*-i kokkuhääldus, mis teeb markerist prosoodilise terviku. Markerid paiknevad sageli lause algusest tagapool ja on peaaegu alati prosoodiliselt lausega integreeritud. Registri sees on aga erinevused ametliku ja argisuhtluse vahel: argisuhtluse korral on markerite esinemissagedus, tagapool esinevate markerite sagedus ja ilma *et*-ita markerite kasutus suuremad kui ametlikus suhtluses.

Skaala teise otsarühma moodustavad mitteilukirjanduslikud trükitekstid ehk ajakirjandus ja (populaar)teadus. Seda iseloomustavad markerite eesand, suur subjektita markerite hulk ja *et*-markerid, millel on säilinud komaga märgitud ortograafiline piir sõna *et* ees. Markerid sageduse järgi jaguneb see ajakirjanduse ja (populaar)teaduse rühmaks, milles on markereid eriti vähe. Omaette vaheühmaks jääb ilukirjandus. Samamoodi nagu teistes trükitekstides, on siin markeri kasutus harv ja markerid paiknevad valdavalt lause alguses. Muudest trükitekstidest eristab ilukirjandust markeri ehitus: siin leidub sarnaselt suulise keelega rohkem ilma *et*-ita ja *ma*-kasutusega markereid.

Netitekstid moodustavad heterogeense koosluse, mille seesmised piirid on hajusad. Siin saame eristada omaette rühmana spontaanseid reaalarajas toimuvaid netivestlusi, milles marker esineb väga sageli nagu ka suulises keeles, kuid markeri kasutus on eripärane. Selle kõrvale paigutuvad jututoad. Üldistavalt on need rühmad lähedased suulisele keelele. Skaala teise otsa jäävad uudisgrupid, mis on kõige lähemal trükitekstidele. Foorumitekstid ja kommentaarid moodustavad heterogeense vaheala.

5. Arutus ja kokkuvõte

Artiklis analüüsisime diskursusemarkeri (*ma*) *arvan* (*et*) kasutust eesti keele registrites ja tekstirühmades. Analüüs näitas, et eesti keeles on kasutusel neli markeri varianti: *ma arvan*, *ma arvan et*, *arvan* ja *arvan et*. Markerid määratlemise lähtekohaks olid teistes keeltes leitud analoogiliste markerite tunnused, mille puhul valitseb uurijate seas suhteliselt suur üksmeel. Eesti keele seisukohast analüüsisime lisaks nelja tunnust: sõnajärge, subjekti *ma* rõhulisust ja vormi (*ma* ~ *mina*), subjekti esinemist või puudumist ja markeri prosoodilist integreeritust.

Lausegrammatika inversioonireegleid järgiva variandi *arvan ma* liigitasime lauseks, mitte markeriks (vrd teistsugust käsitlust Auer, Lindström 2016). Rõhulist *ma*-vormi uuritud materjalis ei esinenud (vrd inglise keele *I* kohta Dehé, Wichmann 2010). Samas on nii kõnes kui ka kirjas kasutusel konstruktsioon *mina arvan*, milles *mina* on rõhuline. Need üksused jätsime markerite hulgast välja. Eesti keeles on võimalikud nii subjektiga kui ka subjektita konstruktsioonid (*ma arvan* / *arvan*). Analüüs näitas, et mõlemad variandid võivad esineda samavõrra markeritena ning subjekti puudumine või esinemine ei ole käsitlev diskursusemarkeri tunnusena.

Eri seisukohti on esitatud selle kohta, kas marker on prosoodiliselt lausega integreeritud või mitte. Meie analüüs kinnitas väidet, et suulises eesti keeles on marker peaaegu alati prosoodiliselt lausega integreeritud (vt ka Keevallik 2010). Samas, kirjalikes tekstides on marker peaaegu alati eraldatud piiri markeeriva komaga. Sellisena ei saa ka seda joont kasutada markeri tunnusena.

Leitud tunnustele toetudes analüüsisime statistiliselt markerite kasutust. Varasemas teaduskirjanduses on meie teada uuritud üksnes markerite sagedust suulises institutsionaalses ja argisuhtluses ning tulemused on vasturääkivad (vrd Aijmer 1997 ja Simon-Vandenberg 2000). Meie võrdlesime markeri kasutust 15 tekstirühmas, mis koonduvad kolme registrisse: trükitekstid, netitekstid ja suulised tekstid. Otsisime vastuseid kolmele uurimisküsimusele.

- Kas markeri sagedus erineb tekstirühmades ja/või registrites?
- Kas markeri esinemissagedus eri süntaktilistes positsioonides erineb tekstirühmades ja/või registrites?
- Kas markeri vormivariandid on seotud eri tekstirühmade ja/või registritega?

Analüüs näitas, et nii markerite kasutussagedus, positsioonid lauses kui ka vorm varieeruvad tekstirühmiti ja registriti. Sellest tuleneb küsimus, kuidas neid erinevusi seletada.

Markeri sageduse poolest moodustavad kujuteldava skaala ühe otsa teadustekstid, kus markereid kasutatakse väga vähe. Siin on ilmselt piiravaks teguriks markeri semantika. *Ma arvan* väljendab tähendusi 'arvatavasti' ja 'minu arvates'. 'Arvatavasti' ei sobi teaduse tõekeskusega ja 'ma arvan' teaduse kaldega impersonaalsusele. Skaala teise otsa moodustavad tekstirühmad, kus markerit kasutatakse väga sageli: suulised ja netis peetavad spontaansed argivestlused. See kasutus on kooskõlas nende allkeelte tüüpiliste joontega. Spontaanses dialoogis kasutatakse palju erinevaid pehmendajaid, mis väljendavad ebakindlust öeldu tõeväärtuse suhtes (nt Biber, Conrad 2009: 236–241). Selle taustaks on vajadus kiiresti reageerida partneri jutule, mis ei võimalda alati leida täpset väljendust. Spontaansuse vähenemine mõjutab markerite hulka. See paistab silma juba suulises ametlikus suhtluses, mis on argivestlusest vähem spontaanne.

Markeri vormi seisukohast tuleb käsitleda eraldi sõnu *ma* ja *et*. Subjekti *ma* sagedus on eri tekstirühmades erinev. Skaala ühte otsa kuuluvad

siin suulised tekstid, netivestlused, foorumid, jututoad ja ilukirjandus, kus *ma* kasutus on valitsev. Soome keele näitel on ühe seletusena pakutud, et suuline keel kasutabki palju rohkem esimese isiku subjekti kui kirjakeel. Helasvuo (2014: 64) väidab, et soome kõnekeeles on esimeses isikus pronoomeni kasutus norm, samal ajal kui normikeel eelistab subjektita kasutust. Eesti keele kohta pole selliseid arvutusi tehtud. Meie analüüsitud materjal näitab ka veidi teistsugust pilti. Nimelt on jututoad, foorumid ja vestlused lähemal suulisele keelele, kasutades subjektita markereid alla 10%. Samasse rühma paigutub ka ilukirjandus. Samal ajal kuuluvad kommentaarid ja uudisgrupid kokku trükitekstide keelega. Teiseks seletuseks võib ka pakkuda, et suulises suhtluses väljendab kõneleja kergemini oma personaalsust, enesekeskset subjektiivset positsiooni, mida *ma* kasutus esile tõstab. Seda toetavad andmed ilukirjandustekstidest, kus on rohkem *ma*-kesksust. See võib olla seotud tegelaste dialoogiga, mis markeerib suulist dialoogi. Samas ei ole ametlik suuline keel alati enesekeskne ega subjektiivne. Samuti toetab seda *ma* harv esinemine impersonaalsust eelistavates doktoritöodes, kuid vastu räägib selle oluliselt suurem sagedus muudes analüüsitud teadustekstides.

Sõna *et* kasutus on seotud markeri positsiooniga, kuna *et*-i puudumine võimaldab markeril liikuda lause algusest tahapoole (vastupidiseid näiteid on vaid üksikuid). Samas ei ole *et*-i esinemine lause alguses paikneva markeri korral kohustuslik. Positsiooni ja sõna *et* seisukohast jaguneb markerite kasutus kolmeks.

Ühe rühma moodustavad mitteilukirjanduslikud trükitekstid, milles valitseb lause alguses paiknev *et*-marker. Selle oluline mõjur on suure tõenäosusega kirjakeelsuse ja grammatika primaat. Kuigi tegu on markeriga, on selle struktuuriomadustes grammatika oma alguspositsiooni eelistusega olulisem kui pragmaatika. Trükitekstide puhul on ka aega ja võimalusi teksti redigeerida ning ei saa eirata võimalust, et toimetajad on autoreid korrigerinud.

Teise rühma moodustavad suuline keel, netikeel ja ilukirjandus, milles ilma *et*-ita markerid on sagedasemad. Markerid positsiooni seisukohast eristub siin kõigepealt ilukirjandus, kus markerid paiknevad sarnaselt teiste trükitekstidega peaaegu alati lause alguses. Teiseks eristuvad suulised tekstid, milles tagapool paiknevaid markereid on kirjanduse ja netikeelega võrreldes oluliselt rohkem. Sealjuures on selle erinevuse põhipõhjjustajaks lause keskel paiknevate markerite oluliselt suurem kasutus.

See tulemus tõstatab küsimuse, miks esinevad suulises dialoogis markerid teistest registritest sagedamini tagapool. Üks võimalik seletus toetub suulise suhtluse kiirele tempole ja protsessuaalsusele. Kuna kõnelejad peavad dialoogis kiiresti reageerima, on võimalik hilisemaid markereid tõlgendada kui alguspositsioonis olevate markerite viivitusega variante. Protsessuaalsuse tõttu pole suulises keeles võimalik ka markeri asukoha hilisem korrigeerimine, kuna see nõuaks kogu lause uuesti ütlemist. Samas ei kehti need tegurid samavõrra kirjalikus netikeeles, kuna seal on alati võimalik enne sõnumi ärasaatmist marker lause algusesse tõsta. Teisalt on lauseperifeeriade uuringute põhjal täheldatud, et lause alguses (vasakus perifeerias) kujunenud markerid on järk-järgult nihkunud ka lause lõppu (paremasse perifeeriasse), kus nad võivad omandada uusi funktsioone, nagu intersubjektiivsus ja seose loomine järgneva tekstiga (vt nt Traugott 2012; Metslang jt 2014; suulise dialoogi kohta Haselow 2012; Helasvuoto 2014; Auer, Lindström 2016).

Selle kõrval tõuseb ka küsimus, miks esineb lause keskel paiknevaid markereid enam just suulises dialoogis. Ka neid võib tõlgendada kui alguspositsioonis olevate markerite viivitusega variante (vt ka Auer, Lindström 2016: 88). See tõlgendus vajab siiski suulise dialoogi ja netidialoogi võrdlevat kvalitatiivset analüüsi. On võimalik, et need markerid täidavad teistest positsioonidest erinevat rolli, mis on oluline just suulises dialoogis.

Kolmanda rühma moodustavad *et* ja markeri positsiooni seisukohast netikeele tekstid, mis jagunevad tagapool paiknevate markerite sagedustelt kaheks. Vestlused, jututoad ja foorumid sisaldavad oluliselt enam tagapoolseid markereid kui uudisgrupid ja kommentaarid. Reaalajas netivestlused ja jututoad hõlmavad kiiretempolist spontaanset suhtlust, samas uudisgrupid ja kommentaarid on aeglasemad ja võimaldavad enam redigeerimist. Eripäraseks jäävad foorumid, kus kasutatakse suhteliselt palju tagapoolseid markereid, kuid kus ei vajata kiiret tempot. Seega on spontaansus oluline mõjur, kuid mitte piisav seletus. Tähtis on seegi, et netisuhtluse keelt iseloomustab suur eklektilisus: osa kasutajaid järgib kirjakeele norminguid, teine osa aga ei järgi neid üldse; osa redigeerib oma tekste, teine osa mitte. See toob kaasa keelekasutuse varieerumise, mis peegeldub statistikas.

Markerivariantide kasutuse analüüs näitab, et suulises keeles esineb palju enam markereid, mida iseloomustavad maksimaalsed markeritunnused: sidendi *et* puudumine ja paiknemine (osa)lause algusest tagapool. Tulemus on kooskõlas arusaamaga, et grammatiseerumist juhib tavaliselt

suuline keel (Metslang 2006: 187; Kortmann, Schneider 2011: 263). Suulise keele sees omakorda on argisuhtluse markerikasutus rohkem grammatiseerunud kui ametliku-avaliku suhtluse oma. Kirjalikus, ennekõike trükikeeles on markerite vorm peaaegu eranditult sarnane samavormilise lausega. See näitab, et pragmaatiline markeristumine on toimunud enne markeri grammatiliste tunnuste muutumist. Teisalt on mitmes uurimuses väidetud, et esimene grammatiseerumisega seostuv struktuuritunnus on sidendi ('et') kadumine, mis võimaldab ühtlasi hakata markeril liikuma lauses alguspositsioonist tahapoole. Meie analüüs näitas, et eesti keeles sisaldas 86% markeriteks analüüsitud üksustest sõna *et*. Sellisena erineb eesti (*ma*) *arvan* (*et*) oluliselt nt inglise markerist *I think*, sarnanedes samal ajal saksa, rootsi ja taani keele markeritega (Beijering 2012; Auer, Lindström 2016).

Kokkuvõttes seostuvad diskursusemarkeri (*ma*) *arvan* (*et*) kasutuse erinevused eri tekstirühmades mitme teguriga, mille hulgas on olulised suulisus/kirjalikkus, protsessuaalsus, dialoogilisus, spontaansus/redigeeritus, teksti personaalsus/impersonaalsus ja kirjakeele reeglite mõju.

Kvantitatiivses analüüsis ilmnevad ühelt poolt selged markerikasutuse tendentsid, teiselt poolt aga jääb vastuseta hulk küsimusi, mis vajavad lahendamiseks põhjalikku kvalitatiivset analüüsi.

Suuliste näidete transkriptsioonimärgid

.	langev intonatsioon
?	tõusev intonatsioon
,	poollangev intonatsioon
(.)	mikropaus: 0.2 sekundit või lühem
(0.8)	mõõdetud paus sekundites
`	rõhutatud sõna või silp
:	hääliku venitamine
.hhh	häälekas sissehingamine
@...@	hääletooni või hääle kvaliteedi muutumine
=	otsarääkimine (kahe üksuse vahel ei ole vaikust)
[pealerääkimise algus
]	pealerääkimise lõpp
>...<	kiirendatud tempoga lõik
...	vaiksem jutt
\$.\$.	naerva häälega öeldud sõna või lõik
hehe, mhemhe	naer

Lühendid

AJA – ajakirjandustekstid; DOK – doktoritööd; EE – Eesti Ekspress; ILU – ilukirjandustekstid; JT – jututoad; FOO – foorumitekstid; KOM – kommentaarid; NV – netivestlused; P – Postimees; SLÕ – SL Õhtuleht; (SUUL) AM – suuline ametlik suhtlus; (SUUL) AR – suuline argisuhtlus; TEA – teadus; UG – uudisgrupid

Kirjandus

- Aijmer, Karin 1997.** *I think* – an English modal particle. – Modality in Germanic languages: Historical and comparative perspectives. Ed. by Toril Swan, Olaf Jansen-Westvik. Berlin, New York: Mouton de Gruyter, 1–47. <https://doi.org/10.1515/9783110889932.1>
- Auer, Peter, Jan Lindström 2016.** Left/right asymmetries and the grammar of pre- vs. post-positioning in German and Swedish talk-in-interaction. – *Language Sciences* 56, 68–92. <https://doi.org/10.1016/j.langsci.2016.03.001>
- Beijering, Karin 2012.** Expressions of Epistemic Modality in Mainland Scandinavian: A Study into the Lexicalization-Grammaticalization-Pragmaticalization Interface. Groningen: Rijksuniversiteit Groningen.
- Biber, Douglas, Susan Conrad 2009.** Register, Genre, and Style. Cambridge: Cambridge UP. <http://dx.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*. Ed. by Nicole Baumgarten, Inke Du Bois, Juliane House. Bingley: Emerald, 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.
- Degand, Liesbeth, Jacqueline Evers-Vermeul 2015.** Grammaticalization or pragmaticalization of discourse markers? More than a terminological issue. – *Journal of Historical Pragmatics* 16 (1), 59–85. <https://doi.org/10.1075/jhp.16.1.03deg>
- Degand, Liesbeth, Anne-Marie Simon-Vandenberg 2011.** Introduction: Grammaticalization and (inter)subjectification of discourse markers. – *Linguistics* 49 (2), 287–294. <https://doi.org/10.1515/ling.2011.008>
- Dehé Nicole, Anne Wichmann 2010.** Sentence-initial *I think (that)* and *I believe (that)*: Prosodic evidence for uses as main clause, comment clause and discourse marker. – *Studies in Language* 34 (1), 36–74. <https://doi.org/10.1075/sl.34.1.02deh>
- Erelt, Mati 2017.** Liitlause. – *Eesti keele süntaks*. (= Eesti keele varamu III.) Toim. Mati Erelt, Helle Metslang. Tartu: Tartu Ülikooli Kirjastus, 647–755.

- Erelt, Mati, Helle Metslang, Helen Plado 2017.** Alus. – Eesti keele süntaks. (= Eesti keele varamu III.) Toim. Mati Erelt, Helle Metslang. Tartu: Tartu Ülikooli Kirjastus, 240–257.
- Fischer, Olga 2007.** The development of English parentheticals: A case of grammaticalization? – Tracing English through Time: Explorations in Language Variation. A Festschrift for Herbert Schendl on the Occasion of his 65th Birthday. Ed. by Ute Smit *et al.* Vienna: Braumüller, 103–118.
- Fraser, Bruce 2009.** Topic orientation markers. – Journal of Pragmatics 41 (5), 892–898. <https://doi.org/10.1016/j.pragma.2008.08.006>
- Haselow, Alexander 2012.** Subjectivity, intersubjectivity and the negotiation of common ground in spoken discourse: Final particles in English. – Language & Communication 32 (3), 182–204. <https://doi.org/10.1016/j.langcom.2012.04.008>
- Heine, Bernd 2013.** On discourse markers: Grammaticalization, pragmaticalization, or something else? – Linguistics 51 (6), 1205–1247. <https://doi.org/10.1515/ling-2013-0048>
- Heine, Bernd 2018.** Are there two different ways of approaching grammaticalization? – New Trends in Grammaticalization and Language Change. Ed. by Sylvie Hancil, Tine Breban, José Vicente Lozano. (= Studies in Language Companion Series 202.) Amsterdam/Philadelphia: John Benjamins Publishing Company, 23–54. <https://doi.org.10.1075/slcs.202.02hei>
- Heine, Bernd, Gunther Kaltenböck, Tania Kuteva 2019.** On the rise of discourse markers. https://www.researchgate.net/publication/333783353_On_the_rise_of_discourse_marker (09.04.2020).
- Helasvuo, Marja-Liisa 2014.** Agreement or crystallization: Patterns of 1st and 2nd person subjects and verbs of cognition in Finnish conversational interaction. – Journal of Pragmatics 63, 63–78. <https://doi.org/10.1016/j.pragma.2013.11.011>
- Hennoste, Tiit 2004.** *Et*-komplementlause peaverbide funktsioonid eestikeelses vestluses. – Keel ja Kirjandus 7–8, 504–523; 590–609.
- Hennoste, Tiit 2006.** *Et*-komplementlause peaverbide funktsioonid eestikeelses vestluses 2: *mõtlemine*. – Lähivertailuja 17. (= Jyväskylä Studies in Humanities 53.) Toim. Annekatrin Kaivapalu, Külvil Pruuli. Jyväskylä: Jyväskylän yliopisto, 119–134.
- Imo, Wolfgang 2011.** Clines of subordination – Constructions with the German complement-taking predicate *glauben*. – Subordination in Conversation. Ed. by Ritva Laury, Ryoto Suzuki. Amsterdam: John Benjamins, 165–190. <https://doi.org/10.1075/slsi.24.08imo>
- Kaltenböck, Gunther 2007.** Spoken parenthetical clauses in English. – Parentheticals. (= Linguistics Today 106.) Ed. by Nicole Dehé, Yordanka

Kavalova. Amsterdam: John Benjamins, 25–52. <https://doi.org/10.1075/la.106.05kal>

- Keevallik, Leelo 2003.** From Interaction to Grammar: Estonian Finite Verb Forms in Conversations. (= Acta Universitatis Upsaliensis. Studia Uralica Upsaliensia 34.) Uppsala.
- Keevallik, Leelo 2010.** Clauses emerging as epistemic adverbs in Estonian conversation. – *Linguistica Uralica* 46 (2), 81–101. <http://dx.doi.org/10.3176/lu.2010.2.01>
- Kortmann, Bernd, Agnes Schneider 2011.** Grammaticalization in non-standard varieties of English. – *The Oxford Handbook of Grammaticalization*. Ed. by Heiko Narrog, Bernd Heine. Oxford: Oxford University Press, 263–278. 10.1093/oxfordhb/9780199586783.013.0021
- Kärkkäinen, Elise 2003.** Epistemic Stance in English Conversation: A Description of its Interactional Functions, with a Focus on ‘I think’. Amsterdam, Philadelphia: John Benjamins. <http://dx.doi.org/10.1075/pbns.115>
- Metslang, Helle 2006.** Grammatiseerimise teooriaid tänase pilguga. – *Teoreetiline keeleteadus Eestis II*. (= Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 7.) Toim. Ilona Trigel, Haldur Õim. Tartu: Tartu Ülikooli Kirjastus, 176–195.
- Metslang, Helle, Karl Pajusalu, Külli Habicht 2014.** Koordinaatiivsed partiklid lause perifeerias. – *Emakeele Seltsi aastaraamat* 59 (2013). Peatoim. Mati Ereht. Tallinn: Emakeele Selts, 139–161. <http://dx.doi.org/10.3176/esa59.07>
- Mullan, Kerry, Susanna Karlsson 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.) Ed. by Nicole Baumgarten, Inke Du Bois, Juliane House. Bingley: Emerald, 271–294. https://doi.org/10.1163/9789004261921_013
- O’Grady, Gerard 2017.** “I think” in political speech. – *International Review of Pragmatics* 9 (2), 269–303. <https://doi.org/10.1163/18773109-00901006>
- Posio, Pekka 2014.** Subject expression in grammaticalizing constructions: The case of *creo* and *acho* ‘I think’ in Spanish and Portuguese. – *Journal of Pragmatics* 63, 5–18. <https://doi.org/10.1016/j.pragma.2013.07.001>
- Schneider, Stefan 2007.** Reduced parenthetical clauses in Roman languages. – *Parentheticals*. (= *Linguistics Today* 106.) Ed. by Nicole Dehé, Yordanka Kavalova. Amsterdam: John Benjamins, 237–258. <https://doi.org/10.1075/la.106.13sch>
- Seppänen, Eeva-Leena, Ritva Laury 2007.** Complement clauses as turn continuations: The Finnish *et(tä)*-clause. – *Pragmatics* 17 (4), 553–572. <https://doi.org/10.1075/prag.17.4.06sep>

- Simon-Vanderbergen, Anne-Marie 2000.** The functions of *I think* in political discourse. – *International Journal of Applied Linguistics* 10 (1), 41–63. <https://doi.org/10.1111/j.1473-4192.2000.tb00139.x>
- Thompson, Sandra 2002.** “Object complements” and conversation: towards a realistic account. – *Studies in Language* 26 (1), 125–164. <https://doi.org/10.1075/sl.26.1.05tho>
- Thompson, Sandra, A. Anthony Mulac 1991.** A quantitative perspective on the grammaticization of epistemic parentheticals in English. – *Approaches to Grammaticalization*. Ed. by Elizabeth Closs Traugott, Bernd Heine. Amsterdam: John Benjamins, 313–329.
- Traugott, Elizabeth Closs 2012.** Intersubjectification and clause periphery. – *English Text Construction* 5 (1), 7–28. <https://doi.org/10.1075/etc.5.1.02trau>
- Van Bogaert, Julie 2010.** A constructional taxonomy of *I think* and related expressions: Accounting for the variability of complement-taking mental predicates. – *English Language and Linguistics* 14 (3), 399–427. doi:10.1017/S1360674310000134
- Väänänen, Milja 2016.** Subjektin ilmaiseminen yksikön ensimmäisessä persoonassa: tutkimus suomen vanhoista murteista. (= Turun yliopiston julkaisuja, sarja C, osa 430.) Turku.

THE DISCOURSE MARKER (*MA*) *ARVAN* (*ET*) 'I THINK'

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

The article analyzes the usage of variants of the discourse marker (*ma*) *arvan* (*et*) in different registers of Estonian. The use of this marker is compared in 15 text groups, which can be grouped into three registers: printed texts, online texts and spoken texts. The research material comes from the corpora available on Keeleveeb as well as corpora of spoken language and online conversations and contains a total of 2572 usage instances. We seek answers to three research questions:

- does the frequency of the marker differ by text group and/or register?
- how does the frequency of variants of the marker differ according to syntactic position in different text groups and/or registers?
- are different variants of the marker associated with different text groups and/or registers?

Our analysis shows that there are four variants of the marker in use: *ma arvan*, *ma arvan et*, *arvan*, and *arvan et*, and registers and text groups differ in terms of the frequency of the marker, its typical position in the sentence, and its typical form. Based on the usage of the marker (*ma*) *arvan* (*et*), texts can be divided into four groups.

On one end of the scale lies the spoken language register. The frequency of the marker is substantially higher in spoken language than in other groups, and almost exclusively variants including the pronoun *ma* are used. Moreover, variants without the conjunction *et* are more common in spoken language than elsewhere. It is common for the verb and *et* to be pronounced together, such that the marker is in itself a prosodic whole. The marker often appears non-sentence-initially and is almost always prosodically integrated with the rest of the sentence.

The other end of the scale is formed by non-fiction printed texts, i.e. journalism and (popular) science. These groups are characterized by the sentence-initial

position of the marker, the high frequency of pro-drop variants (i.e. without the subject pronoun *ma* 'I'), and variants in which the orthographic boundary before the word *et* (marked by a comma) has been preserved. Fiction texts represent a middle ground, and are distinguished from other printed texts by the structure of the marker; similarly to spoken language, fiction texts feature more use of *ma* and less use of *et*.

Online texts form a heterogeneous group with vague internal boundaries. One identifiable subgroup is spontaneous real-time online conversations, in which the marker appears very frequently (as in spoken language), but its usage is peculiar. Next to this group on the scale are chat rooms. Generally, these groups behave similarly to spoken language. On the other end of the scale are news groups, which are most similar to printed texts. Forum texts and comments form a heterogeneous intermediate range.

In summary, the differences in marker usage in different text groups are related to numerous factors, among which are spoken/written, processuality, dialogicity, spontaneity/editedness, the personal/impersonal nature of the interaction, and the influence of the rules of the standard language.

Keywords: Estonian language, pragmatics, subjectivity, language variation, grammaticization, pragmatization

Tiit Hennoste, tiit.hennoste@ut.ee
Küllli Habicht, kulli.habicht@ut.ee
Helle Metslang, helle.metslang@ut.ee
Küllli Prillop, kulli.prillop@ut.ee
Kirsi Laanesoo, kirsi.laanesoo@gmail.com
David Ogren, dpogren@yahoo.com
Liina Pärismaa, liina.parismaa@gmail.com
Elen Pärt, elen.p2rt@gmail.com
Andra Rumm, andra.rumm@ut.ee
Andriela Rääbis, andriela.raabis@ut.ee
Carl Eric Simmul, simmulman@gmail.com

Eesti ja üldkeeleteaduse instituut
Tartu Ülikool
Jakobi 2
51014 Tartu