

Käändevormide varieerumine ja seda mõjutavad tegurid kirjalikus keeles indefiniitpronoomeni *keegi* näitel

ANNIKA KÄNGSEPP

Indefiniitpronoomeniteks nimetatakse asesõnu, mille peamine funktsioon on väljendada umbmäärasust (Haspelmath 1997: 11), teisisõnu viidata „mingile täpselt määratlemata või tundmatule asjale, nähtusele või tunnusele” (EKK 2020: 167). Vormiliselt koosnevad need maailma keeltes tavaliselt tüvest (nt ingl *-body*, sm *kuka-*) ja indefiniitsuse markerist (nt ingl *some-* ja sm *-kaan*). Indefiniitsuse marker on afiks (nt ingl *some-*, sm *-kin*) või partikkel, mis paikneb pronoomeni tüve kõrval. Sufiksaalne indefiniitsuse marker on tavaliselt sõnas viimasel kohal, kuid võib kasutuses mõnikord käändelõpuga kohad vahetada. Tihti esinevad indefiniitpronoomenid keeltes sarjadena, mille iga liige vastab mingisugusele tähenduskategooriale, nagu isik, asi, tunnus, koht, aeg, viis või kogus. Inglise keeles on näiteks *some*-sari, kuhu kuuluvad *somebody* 'keegi' (isik), *something* 'miski' (asi), *somewhere*¹ 'kuski' (koht), *sometime* 'kunagi' (aeg) ja *somehow* 'kuidagi' (viis) ning soome keeles eitust väljendav *-kAAn*-sari, kuhu kuuluvad *kukaan* (isik), *mikään* (asi), *missään* (koht), *milloinkin* (aeg), *mitenkään* (viis) ja *kumpikaan* (isik/asi). (Haspelmath 1997: 21–22, 293)

Eesti keeles saab moodustada *gi-/ki*-liiteliste indefiniitpronoomenite sarja, kuhu kuuluvad käändsõnadest *keegi*, *miski*, *kumbki*, *ükski* ja *mingi*. Siinse artikli seisukohast on olulised neist neli esimest. Uuritavate indefiniitpronoomenite käändevormides esineb tegelikult keelekasutuses varieeruvust, mida on kirjakeeles ja murretes (vt kaarti 1) täheldatud üle saja aasta. Varieeruvus puudutab *-gi/-ki* asukohta käändevormides: kuigi kirjakeele normi (ÕS 2018) järgi peaks see paiknema käändelõpu järel (näide 1), võib see asetseda ka norminguvastaselt käändelõpu ees (näide 2), kahe käändelõpu vahel (näide 3) või nii käändelõpu ees kui ka järel (näide 4).

- (1) Kerge on **kellelegi** silti külge kleepida [---] (ÜK 2021, perioodika)
- (2) Diplom ei anna **kellegile** mõistust. (ÜK 2021, foorumid)
- (3) Mina ei anna mitte **kellegile** raha. (ÜK 2021, perioodika)
- (4) Ma loodan, et sellest postitusest on vähemalt **kellegilegi** kasu. (ÜK 2021, blogid)

Aado Rull (1917: 86–87) on juba 1917. aastal viidanud kaua kestnud segadusele *-gi/-ki* paiknemise üle indefiniitpronoomenite käändevormides. Seejuures peab tema ainuõigeks liite paiknemist käändelõpu järel (Rull 1917: 86). Temaga samal

¹ Martin Haspelmath (1997: 10) kasutab terminit *pronoomen* laiemas tähenduses, st ta käsitleb pronoomenitena erinevaid asendavaid sõnu, ka proadverbe (nt *somewhere* 'kuski', *sometime* 'kunagi'), proadjektive ja proverbe.



Kaart 1. Pronoomeni keegi allatiivivormide varieerumine murretes Andrus Saareste sõnakaardi põhjal (Saareste 1955: 16).

seisukohal on Johannes Voldemar Veski (1933: 384), kelle sõnul kasutatakse Eesti keelealal valdavalt vorme, kus *-gi/-ki* paikneb käändelõpu järel. Andrus Saareste (1923: 116, 1933: 450–451, 1936: 553–554) seevastu peab indefiniitpronoomenites nii kõne- kui ka kirjakeeles tavapärasemaks *-gi/-ki* asukohta käändelõpu ees, lisades, et vormid, kus *-gi/-ki* paikneb käändelõpu järel, on kõnekeeles haruldased ja sobilikud pigem ametlikku keelekasutusse (Saareste 1923). *-gi/-ki* paiknemist käändelõpu ees põhjendab ta liite kunagise kinnitava tähenduse 'ka, isegi' kadumisega, mistõttu on *-gi/-ki* saanud neis pronoomenites tüve lahutamatuks osaks (Saareste 1933: 450–451). Seda väidet toetavad Saareste sõnul vormid, kus kinnitava tähenduse lisamiseks kasutatakse veel üht liidet, näiteks *kellegilgi*. Selliste vormide taunimist kirjakeeles põhjendab ta ÕS-i peatoimetaja Veski eelistusega lähtuda õigekeelsusotsuste tegemisel oma kodumurdest ja otstarbekohasuse põhimõttest, st kui kõne- ja kirjakeeles kasutatakse vormi *linna-le-gi*, peaks *-gi/-ki* ka indefiniitpronoomenites käändelõpule liituma. (Saareste 1936: 554) Ka Paul Alvre järgi (1980: 542) võiksid paralleelkujud olla kirjakeeles lubatud, sest normiks saanud vormid, kus *-gi/-ki* paikneb käändelõpu järel, on rahvakeeles levinud vaid Tartumaal ja vähesel määral rannikumurdes.

-gi/-ki paiknemisel käändelõpu ees on Alvre sõnul keeleajalooline põhjendus: pronoomenite *kes* ja *mis* allatiiv on olnud kunagi *ke-lle* ja *mi-lle*, millele lisati *-gi/-ki* ning saadi *ke-lle-gi*, *mi-lle-gi*. Kui kujunesid kahe lõpuga allatiivivormid *ke-lle-le* ja *mi-lle-le*, tajuti *-lle-d* koos eelneva sõnaosaga uue tüvena, mistõttu jäi *-gi/-ki* uutes vormides käändelõpu ette, st *kellegile*, *millegile*. (Alvre 1980: 542–543) Nagu eri keeleuurijate seisukohtadest nähtub, on *-gi/-ki* asukoha loomulikku varieerumist sihipäraselt vähendanud keelekorraldus, kus kuni 1970. aastateni järgiti valdavalt rangelt normivat ja varieerumist vältivat suunda (EKK 2020: 36–37).

Tänapäeva keeles on *-gi/-ki* asukoha varieerumist vähe uuritud. Minu varasemast korpusuuringust (Pant 2018, 2020) selgus, et pronoomenite *keegi*, *miski*, *kumbki* ja *ükski* käändevormid olid enamasti normingupärased: nelja sõna peale kokku moodustasid need 87,9% juhtudest. Vastuvõetavuse hindamise ja vaba moodustuse katsed (Pant 2020: 98–106) näitasid, et kuigi sageli pidasid katsealused kõige loomulikumaks normingupäraseid käändevorme, moodustasid nad nt pronoomeni *miski* puhul rohkem norminguvastast abessiivivormi *millegita* kui normingupärast *milletagi* ning pronoomeni *kumbki* puhul rohkem norminguvastast illatiivivormi *kummagisse* kui normingupärast *kummassegi*. Kokkuvõttes näitas katsetulemuste analüüs, et iga kolmas loomulikuks peetud vorm oli norminguvastane (Pant 2020: 99).

Selle artikli eesmärk on anda 2021. aasta eesti keele ühendkorpuse (ÜK 2021) põhjal ülevaade *keegi*, *miski*, *kumbki* ja *ükski* käändevormide varieerumise ulatusest ning pronoomeni *keegi* käändevormide varieerumist mõjutavatest teguritest. Uurimus on kvantitatiivne ning rakendatud on varieerumise analüüsi meetodeid (vt nt Pilvik jt 2021). Laiem eesmärk on kirjeldada ja seletada eesti keele käändevormistikus esinevat varieerumist.

Analüüsisin tegelikust keelekasutusest pärit korpusmaterjali, tuginedes kasutuspõhise lingvistika teooriale. Kasutuspõhise lähenemise peamine uurimisobjekt on keel, mida inimesed kasutavad, loovad ja mõistavad. Selle lähenemise järgi on keele struktuur ja kasutusjuhud tihedalt seotud, mistõttu on olulisel kohal keeleüksuste

esinemissagedus. (Barlow, Kemmer 2000: viii, x, xv) Suurema sagedusega keeleüksused kinnistuvad keelesüsteemi tugevamalt (Barlow, Kemmer 2000: x; Bybee 2010: 214; Langacker 1987: 59; von Mengden, Coussé 2014: 8), ent võivad ka rohkem ja kiiremini muutuda (nt lüheneda või leksikaliseeruda) kui harvemad üksused (Bybee 2010: 20; von Mengden, Coussé 2014: 7). Kasutuspõhise lähenemise järgi sobivad keelekasutuse analüüsimiseks kõik meetodid, mis vaatlevad keelt kasutuses (Bybee 2010: 10). Üks olulisemaid neist on korpusanalüüs, mis aitab tuvastada keelekasutuse üldisi mustreid (Barlow, Kemmer 2000: xvi).

1. -gi/-ki olemusest eesti keeles

-gi/-ki on eesti keele levinuim kliitik, millel on kõik tüüpilised kliitiku omadused: see on rõhutu, vajab peremeesüksust ja võib liituda eri klasside sõnadele (nt *lauljalgi*, *rohelinegi*, *leidsimegi*, *rõõmsaltki*, *ettegi*, *poolegi*) (Viht, Habicht 2019: 39). Eesti keele morfotaktika reeglite järgi paiknevad kliitikud sõnas viimasel kohal: tüvi-tuletusliited-vormitunnused-kliitik (Viht, Habicht 2019: 47). Et -gi/-ki võib tegelikult keelekasutuses esineda ka käändetunnuse ees (nt *kellegile*), on see eesti keele morfotaktika reeglite järgi hakanud sarnanema tuletusliitega. Selline keelemuutus viitab leksikaliseerumisele, mis tähendab, et keelekõnelejad tajuvad konstruktsiooni või sõna teatud kontekstis uue tähendusliku üksusena (Brinton, Traugott 2005: 96).

-gi/-ki-liitelistest indefiniitpronoomenitest kliitikute ja tuletusliidete valguses on kirjutanud Arnold M. Zwicky (1977) ja Joel A. Nevis (1984). Zwicky sõnul on emfaatiline morfeem -gi/-ki tüüpiliselt järelkliitik (ingl *enclitic*), ent kui see lisatakse indefiniitsuse väljendamiseks küsisõnadele, võib liide esineda nii käändelõpu ees (nt *kellegile*) kui ka järel (nt *kellelegi*), mistõttu nimetab ta küsisõnadele liituvat liidet sisekliitikuks (ingl *endoclititic*). Selline esinamus viitab Zwicky sõnul käimasolevale morfoloogilisele protsessile, kus liidet tajutakse üha enam tüvele liituva sufiksina. (Zwicky 1977: 8)

Nevis (1984: 213) arvates on emfaatiline -gi/-ki ja küsisõnadele lisatav -gi/-ki kaks eri morfeemitüüpi. Esimene neist väljendab rõhku, saab liituda eri liiki sõnadele (nt *nainegi*, *räägibki*, *suurigi*) ega esine kunagi käändelõpu ees. Teine väljendab indefiniitsust ja liitub ainult asesõnade tüvedele, mis tähistavad inimest, aega, kohta või tüüpi (nt *miski*, *millalgi*). Seetõttu peab ta indefiniitsust väljendavat liidet pigem tuletusliiteks kui kliitikuks. (Nevis 1984: 213) Pronoomenites *keegi* ja *miski* on -gi/-ki leksikaliseerunud tüve osaks ja sõnad on tervikuna omandanud indefiniitse tähenduse (Nevis 1984: 214–215). Seega on Zwicky (1977: 8) viidatud käimasolev morfoloogiline protsess Nevis järgi peaaegu lõpuni jõudnud, sest indefiniitsust väljendav liide on tõlgendatav tuletusliitena, mis on enamikul juhtudest liidetud tüvele (Nevis 1984: 221).

2. Indefiniitpronoomenite käändevormide varieerumise ulatus

Varasemates uurimustes (Pant 2018, 2020) olen analüüsinud indefiniitpronoomenite käändevormide varieerumist veebitekstide korpuse etTenTen 2013 põhjal. Siinses uurimuses kasutasin palju ulatuslikumat 2021. aasta eesti keele ühendkorpust, mis koosneb 11 allkorpusest ja sisaldab 2,4 miljardit sõna (Koppel, Kallas 2022: 209). Päringu tegin korpuspäringusüsteemi Sketch Engine (Kilgarriff jt 2014) abil kogu korpuse ulatuses. Et *-gi/-ki* asukoht varieerub peamiselt semantilistes ja vähem grammatilistes käanetes (nt vorm *kellegile* on sagedasem kui *kedagit*), jätsin nominatiivi, genitiivi ja partitiivi sellest uurimusest välja. Seega otsisin korpusest pronoomenite *keegi*, *miski*, *kumbki* ja *ükski* konkordantse ehk käändevorme koos tekstiümbrusega 11 käändes alates illatiivist. Päringu tegemiseks kasutasin Sketch Engine'i korpuspäringukeelt (CQL) (Jakubíček jt 2010), mille abil otsisin järgemööda iga pronoomeni iga käände kõiki vorme. Selline päring võimaldas otsida korraga väike-seid ja suuri tähti sisaldavaid käändevorme (nt *kellelegi*, *Kellelegi*, *KELLELEGI*, *kel-leLEGI*, *Kellelegi*, *KelleLEGI*), *-gi* ja *-ki* ortograafiaveaga käändevorme (nt *kelleleki*, *kellekile*) ning topelt käändelõpu (nt *kellelegile*), topelt *-gi/-ki* (nt *kellegilegi*) või muu lisalõpuga (nt *kellelegid*, *kellelegigi*) käändevorme. Saadud tulemused jagasin nelja vormimalli: **TLgi** (tüvi, käändelõpp, *-gi/-ki*), **TgiL** (tüvi, *-gi/-ki*, käändelõpp), **TLgiL** (tüvi, käändelõpp, *-gi/-ki*, käändelõpp) ja **TgiLgi** (tüvi, *-gi/-ki*, käändelõpp, *-gi/-ki*).

Järgnevalt esitan vormimalle järgides iga pronoomeni kohta käändevormide esinemise sageduse ja osakaalu tabelina. Muu lisalõpuga vormide esinemust kommenteerin tabeli all.

Tabel 1. Pronoomeni *keegi* käändevormide esinemise sagedus ja osakaal mallide järgi 2021. aasta eesti keele ühendkorpuses.

	TLgi	TgiL	TLgiL	TgiLgi	KOKKU
Illatiiv	1501 (83,6%)	294 (16,4%)	0 (0%)	1 (0,1%)	1796 (0,4%)
Inessiiv	1522 (91,3%)	144 (8,6%)	1 (0,1%)	0 (0%)	1667 (0,4%)
Elatiiv	9636 (89,1%)	1166 (10,8%)	7 (0,1%)	5 (0%)	10 814 (2,5%)
Allatiiv	128 349 (84,6%)	23 131 (15,2%)	196 (0,1%)	16 (0%)	151 692 (34,9%)
Adessiiv	147 209 (71,4%)	55 870 (27,1%)	2918 (1,4%)	36 (0%)	206 033 (47,4%)
Ablatiiv	12 659 (85,7%)	2104 (14,2%)	4 (0%)	7 (0%)	14 774 (3,4%)
Translatiiv	2846 (95,1%)	142 (4,7%)	4 (0,1%)	1 (0%)	2993 (0,7%)
Terminatiiv	57 (50,4%)	56 (49,6%)	0 (0%)	0 (0%)	113 (0%)
Essiiv	185 (50,1%)	183 (49,6%)	0 (0%)	1 (0,3%)	369 (0,1%)
Abessiiv	15 (34,9%)	28 (65,1%)	0 (0%)	0 (0%)	43 (0%)
Komitatiiv	38 050 (85,1%)	6652 (14,9%)	2 (0%)	1 (0%)	44 705 (10,3%)
KOKKU	342 029 (78,6%)	89 770 (20,6%)	3132 (0,7%)	68 (0%)	434 999 (100%)

Tabelist 1 on näha, et pronoomeni *keegi* puhul esines korpuses ootuspäraselt enim kirjakeele normi kohaseid käändevorme (78,6%). Norminguvastaseid vorme esines kokku 21,3%, millest enamuse (20,6%) moodustasid malli TgiL järgivad käändevormid. Kõige rohkem normingupäraseid käändevorme oli translatiivis (*kel-*

lekski; 95,1%) ning kõige rohkem norminguvastaseid vorme abessiivis (*kellegita*; 65,1%). Abessiivi puhul tuleb tulemise tõlgendamisel siiski arvesse võtta vormide vähesust. Peale tabelis esitatud käändemallide järgi moodustatud vormide esinesid korpuses sõnakujud *kellestkid* (1 kord), *kellestkigi* (1 kord), *kellelegid* (13 korda), *kellegigi* (9 korda), *kellekskigi* (1 kord), *kelleginigigi* (3 korda), *kellegagid* (4 korda) ja *kellegagigi* (3 korda). Eriti tähelepanuväärsed on selles loetelus vormid *kellestkigi*, *kellelegigi*, *kellekskigi* ja *kellegagigi*, kus rõhutamiseks on kasutatud ühe -gi/-ki asemel kaht. Märkimist väärt on ka vorm *kelleginigigi*, kus liidet kasutatakse koguni kolmes kohas. Sellised näited annavad kinnitust Nevise (1984: 213) seisukohale, et indefiniitpronoomenites esinev -gi/-ki väljendab indefiniitsust, sest juba olemas-olevale liitele saab lisanduda rõhukliitik.

Tabelist 2 nähtub, et ka pronoomeni *miski* puhul esines korpuses kõige rohkem normingupäraseid käändevorme (92,5%). Võrreldes pronoomeniga *keegi* esines *miski* puhul normingupäraseid vorme 13,9% võrra rohkem. Norminguvastaseid käändevorme oli kokku 7,5%, millest enamuse (7,2%) moodustasid taas malli TgiL järgivad käändevormid. Sarnaselt pronoomeniga *keegi* esines kõige enam normingupäraseid käändevorme translatiivis (*millekski*; 97,2%) ning norminguvastaseid käändevorme abessiivis (*millegita/millegitag*; 63,9%). Muudest vormidest olid esindatud lisarõhutusega vormid *millessegigi* (1 korral), *millestkigi* (14 korral), *millestikikiki* (2 korral), *millekskigi* (3 korral), *millenigigi* (1 korral), *millegagigi* (14 korral) ning murdetaustaga vormid *millestkid* (2 korral) ja *millegagid* (1 korral).

Tabel 2. Pronoomeni *miski* käändevormide esinemise sagedus ja osakaal mallide järgi 2021. aasta eesti keele ühendkorpuses.

	TLgi	TgiL	TLgiL	TgiLgi	KOKKU
Illatiiv	4047 (90%)	448 (10%)	0 (0%)	1 (0%)	4496 (2,7%)
Inessiiv	19 997 (95,5%)	774 (3,7%)	178 (0,8%)	0 (0%)	20 949 (12,7%)
Elatiiv	56 123 (95,1%)	2712 (4,6%)	172 (0,3%)	35 (0,1%)	59 042 (35,7%)
Allatiiv	15 515 (90,7%)	1554 (9,1%)	29 (0,2%)	2 (0%)	17 100 (10,3%)
Adessiiv	2303 (90,9%)	201 (7,9%)	29 (1,1%)	0 (0%)	2533 (1,5%)
Ablatiiv	241 (89,9%)	26 (9,7%)	0 (0%)	1 (0,4%)	268 (0,2%)
Translatiiv	19 965 (97,2%)	544 (2,6%)	12 (0,1%)	12 (0,1%)	20 533 (12,4%)
Terminatiiv	193 (65%)	104 (35%)	0 (0%)	0 (0%)	297 (0,2%)
Essiiv	214 (65,6%)	112 (34,4%)	0 (0%)	0 (0%)	326 (0,2%)
Abessiiv	161 (36%)	284 (63,5%)	0 (0%)	2 (0,4%)	447 (0,3%)
Komitatiiv	34 319 (87%)	5117 (13%)	2 (0%)	6 (0%)	39 444 (23,8%)
KOKKU	153 078 (92,5%)	11 876 (7,2%)	422 (0,3%)	59 (0%)	165 435 (100%)

Tabelist 3 on näha, et ka *kumbki* puhul esines korpuses enim normingupäraseid käändevorme (94,6%). Võrreldes pronoomenitega *keegi* ja *miski* esines neid vastavalt 16% ja 2,1% võrra rohkem. Norminguvastaseid vorme oli kokku 5,4%, millest enamiku (4,4%) moodustasid taas malli TgiL järgivad käändevormid. Malli TgiLgi järgi moodustatud vorme oli vaid üks (adessiivis), mistõttu ei ole seda tabelis kajastatud. Enim normingupäraseid vorme esines inessiivis (*kummaski*; 97,9%) ning enim norminguvastaseid vorme sarnaselt pronoomenitega *keegi* ja *miski* abessiivis (*kum-*

magita; 39,4%). Ka siin tuleb norminguvastaste vormide esinemuse tõlgendamisel arvesse võtta abesiivvormide vähesust. Muude lisalõppudega vorme korpusel ei esinenud.

Tabel 3. Pronoomeni *kumbki* käändevormide esinemise sagedus ja osakaal mallide järgi 2021. aasta eesti keele ühendkorpusel.

	TLgi	TgiL	TLgiL	KOKKU
Illatiiv	631 (86%)	103 (14%)	0 (0%)	734 (1,8%)
Inessiiv	8113 (97,9%)	149 (1,8%)	25 (0,3%)	8287 (20%)
Elatiiv	3284 (96,8%)	107 (3,2%)	2 (0,1%)	3393 (8,2%)
Allatiiv	7844 (93,2%)	553 (6,6%)	15 (0,2%)	8412 (20,3%)
Adessiiv	16 486 (93,9%)	702 (4%)	365 (2,1%)	17 553 (42,4%)
Ablatiiv	1838 (96,6%)	63 (3,3%)	1 (0,1%)	1902 (4,6%)
Translatiiv	99 (97,1%)	3 (2,9%)	0 (0%)	102 (0,2%)
Terminatiiv	11 (91,7%)	1 (8,3%)	0 (0%)	12 (0%)
Essiiv	4 (66,7%)	2 (33,3%)	0 (0%)	6 (0%)
Abessiiv	20 (60,6%)	13 (39,4%)	0 (0%)	33 (0,1%)
Komitatiiv	824 (86,1%)	133 (13,9%)	0 (0%)	957 (2,3%)
KOKKU	39 154 (94,6%)	1829 (4,4%)	408 (1%)	41 391 (100%)

Tabelist 4 on näha, et ka *ükski* puhul esines korpusel enim normingupäraseid käändevorme (99,1%). Norminguvastaseid käändevorme oli kokku vaid 0,9%, neist enim (0,5%) malli TLgiL järgi moodustatud vorme (nt *ühelgil*), mitte malli TgiL järgivaid vorme nagu teiste uuritud pronoomenite puhul. TgiLgi järgi moodustatud vorme oli vaid kaks (üks elatiivis ja üks komitatiivis), mistõttu ei ole neid tabelis kajastatud. See tähendab, et *ükski* puhul esines kirjakeele normi järgi moodustatud käändevorme sagedamini kui *keegi*, *miski* ja *kumbki* puhul – vastavalt 20,5%, 6,6% ja 4,5% võrra rohkem. Erinevus tuleb eriti välja pronoomenite *keegi* (78,6%) ja *ükski* (99,1%) normingupärase vormide osakaalu võrdluses. Ainult normingupäraseid vormid esinesid *ükski* puhul translatiivis (*ühekski*), terminatiivis (*ühenigi*)

Tabel 4. Pronoomeni *ükski* käändevormide esinemise sagedus ja osakaal mallide järgi 2021. aasta eesti keele ühendkorpusel.

	TLgi	TgiL	TLgiL	KOKKU
Illatiiv	1187 (98%)	24 (2%)	0 (0%)	1211 (1,3%)
Inessiiv	26 705 (99,3%)	54 (0,2%)	122 (0,5%)	26 881 (28,7%)
Elatiiv	11 304 (99,6%)	20 (0,2%)	28 (0,2%)	11 352 (12,1%)
Allatiiv	17 336 (99,1%)	95 (0,5%)	58 (0,3%)	17 489 (18,7%)
Adessiiv	32 381 (98,9%)	102 (0,3%)	271 (0,8%)	32 754 (34,9%)
Ablatiiv	2268 (99,5%)	11 (0,5%)	0 (0%)	2279 (2,4%)
Translatiiv	500 (100%)	0 (0%)	0 (0%)	500 (0,5%)
Terminatiiv	13 (100%)	0 (0%)	0 (0%)	13 (0%)
Essiiv	1 (100%)	0 (0%)	0 (0%)	1 (0%)
Abessiiv	4 (57,1%)	3 (42,9%)	0 (0%)	7 (0%)
Komitatiiv	1168 (93,5%)	81 (6,5%)	0 (0%)	1249 (1,3%)
KOKKU	92 867 (99,1%)	390 (0,4%)	479 (0,5%)	93 736 (100%)

ja essiivis (*ühenagi*). Enim norminguvastaseid käändevorme esines ka pronoomeni *ükski* puhul abessiivis (*ühegita*; 42,9%). Nii terminatiivi, essiivi kui ka abessiivi puhul tuleb siiski vormivariantide esinemuse tõlgendamisel arvesse võtta vormide vähesust. Muude lisalõppudega vorme korpuses ei esinenud.

Tabelist 5 nähtub, et nelja pronoomeni peale kokku esines ootuspäraselt kõige rohkem normingupäraseid käändevorme (85,3%). Norminguvastaste käändevormide osakaal oli 14,7%. Neist suurema osa (14,1%) moodustasid malli TgiL järgivad käändevormid. Malli TLgiL järgisid 0,6% ja malli TgiLgi ümardatult 0% vormidest. Võrreldes tulemusi varem tehtud korpusuuringutega (Pant 2018, 2020), on näha, et 2021. aasta eesti keele ühendkorpuses esineb norminguvastaseid käändevorme 2,6% võrra rohkem.

Tabel 5. Pronoomenite *keegi*, *miski*, *kumbki* ja *ükski* käändevormide jagunemine malliti.

	Arv	Protsent
TLgi	627 128	85,3%
TgiL	103 865	14,1%
TLgiL	4441	0,6%
TgiLgi	130	0,0%
KOKKU	735 564	100%

Tulemused näitavad, et vaatamata *gi/-ki*-liiteliste indefiniitpronoomenite käänamise järjepidevale normimisele (Pant 2018: 26–29, 2020: 14–16), on kirjalis keeles peaaegu iga seitsmes vorm ootuspärasest erinev. Käändevormide pikaaegne varieerumine ja lisarõhutusega vormide olemasolu näitavad, et normingud ei ole keelde kinnistunud ega loomulikud, ning kinnitavad ühtlasi Zwicky (1977) ja Nevis (1984) väiteid, mille järgi indefiniitpronoomenitele liituv *-gi/-ki* on pigem tuletusliide kui kliitik ja väljendab neis vormides indefiniitsust, mitte emfaatilisust.

3. Pronoomeni *keegi* käändevormide varieerumist mõjutavad tegurid: materjal ja meetod

Nagu on selgunud varasematest tähelepanekutest ja uurimustest ning eelmises peatükis kirjeldatud korpusuuringust, varieeruvad indefiniitpronoomenite *keegi*, *miski*, *kumbki* ja *ükski* käändevormid tegelikult keelekasutuses, sest *-gi/-ki* võib paikneda nii käändelõpu järel kui ka ees ning harvem kahe käändelõpu vahel või käändelõpu ees ja järel. Et minu varasemad uurimused (vt Pant 2018, 2020) on keskendunud *-gi/-ki* asukohale käändevormides, analüüsin siinses peatükis tegureid, mis võivad normingupärase ja -vastaste vormide varieerumist mõjutada.

Võtsin vaatluse alla pronoomeni *keegi* käändevormid, sest korpusuuring näitas, et selle pronoomeni vormid varieeruvad kõige enam (vt ptk 2). Käänded valisin vormide esinemissageduse järgi korpuses nii, et esindatud oleks vähemalt üks sisekoha ja üks väliskohakääne ning kaks muud kõige sagedamini esinevat semantilist käänet. Kuna väliskohakäänetest olid nii allatiiv kui ka adessiiv kõrge sagedusega, valisin

uurimiseks mõlemad. Täpsemalt uurisin *keegi* käändevorme elatiivis, allatiivis, adessiivis, translatiivis ja komitatiivis. Nii nagu käändevormide sagedusandmed, kogusin uuritava materjali 2021. aasta eesti keele ühendkorpusest.

Päringuga otsisin pronoomeni *keegi* normingupäraste ja -vastaste käändevormide konkordantse kolmest tekstižanrist (perioodika, blogid, foorumid), selleks et esindatud oleks nii toimetatud kui ka toimetamata tekstid. Päringu tegemiseks kasutasin CQL-i, mille abil otsisin normingupäraseid ja -vastaseid käändevorme eraldi ja ühe käände kaupa. Norminguvastaste käändevormide alla kuulusid nii mallide TgiL, TLgiL kui ka TgiLgi järgi moodustatud vormid.² Samuti jätsin valimisse eespool mainitud väike- ja suurtähti sisaldavad käändevormid ning -gi/-ki ortograafiaveaga vormid. Iga käände kohta otsisin korpusest 200 juhuslikku normingupärast ja 200 juhuslikku norminguvastast käändevormi sisaldavat lauset (v.a translatiivis, kus vormide vähesuse tõttu oli lausete arv vastavalt 100 ja 100). Seega on valim tasakaalustatud: 1800 käändevormist 900 on normingupärased ja 900 norminguvastased.

Uuritava ehk sõltuva tunnuseks vaatasin normingupäraste või -vastaste käändevormide esinemust ning varieerumist seletavate ehk sõltumatute tunnustena žanri, kõneliiki, pronoomeni asukohta osalauses, rõhumäärsõna olemasolu osalauses, pronoomeni esinemist laiendina, eitussõna *mitte* esinemist osalauses, pronoomeni funktsiooni osalauses ja käänat nimetatud tunnuste kontekstis. Tunnuseid kirjeldan lähemalt peatükis 4. Seletavad tunnused valisin välja 354 lause põhjal tehtud prooviuuringu alusel ning materjali põhjal. Prooviuuringus olid seletavad tunnused kommunikatiivne lausetüüp, osalause tüüp, pronoomeni asukoht osalauses, pronoomeni esinemine laiendina, rõhumäärsõna olemasolu osalauses ja kõneliik.

Tunnuste mõju hindamiseks kasutasin hii-ruut-testi (Agresti 2007: 34–41). Kui andmestikus oli teoreetilisi sagedusi, mis olid väiksemad kui 5, kasutasin Fisheri täpset testi (Agresti 2007: 45–46). Nii hii-ruut-testi kui ka Fisheri täpse testi tulemust lugesi statistiliselt oluliseks, kui testi p-väärtus oli väiksem kui 0,05. Tunnustevahelise seose tugevuse hindamiseks kasutasin seosekordajat Craméri V (Cramér 1946). Ühtlasi tegin tingimusliku otsustuspuu mudeli (Hothorn jt 2006), et hinnata visuaalselt tunnuseid ja nende koosmõju normingupäraste ja -vastaste käändevormide varieerumisele. Mudeli headust testisin algandmestikust juhuslikult moodustatud sama suurusega valimite põhjal, milles 70% vaatlustest pärines algandmestikust, 30% lisati aga olemasolevate vaatluste juhusliku kordamise teel.

Materjali analüüsisin rakendustarkvaraga R (R Core Team 2022). Tulemuste visualiseerimiseks kasutasin pakette ggplot2 (vt Wickham 2016), partykit (vt Hothorn, Zeileis 2015) ja ggparty ning puumudeli headuse hindamiseks paketti pROC (vt Robin jt 2011).

² Valimisse sattus kokku 19 malli TLgiL järgi moodustatud käändevormi (kolm elatiivis, neli allatiivis, kümme adessiivis ja kaks translatiivis) ja kaks malli TgiLgi järgi moodustatud vormi (üks elatiivis ja üks translatiivis).

4. Pronoomeni *keegi* käändevormide varieerumise ühetunnuseline analüüs

Annan alapeatükkide kaupa detailse ülevaate vaid statistiliselt olulistest tunnustest: žanr, pronoomeni esinemine laiendina ja pronoomeni funktsioon osalauses. Tunnused, mis normingupäraste või -vastaste vormide kasutust statistiliselt oluliselt ei mõjutanud, olid kõneliik, pronoomeni asukoht osalauses, rõhumäärsõna olemasolu osalauses ja sõna *mitte* esinemine osalauses. Neist annan järevalt vaid põgusa ülevaate.

Tunnusel *kõneliik* oli kaks võimalikku väärtust: „eitav” ja „jaatav”. Kõneliiki märgendasin selle osalause alusel, kus pronoomen esines. Tulemustest selgus, et normingupäraseid ja -vastaseid käändevorme esines lausetes üsna võrdselt: eitava kõneliigiga vastavalt 52% ja 48% ning jaatava kõneliigiga vastavalt 48,8% ja 51,2%. Siiski nähtus käändeid eraldi analüüsides, et adessiivis sõltub normingupäraste või -vastaste vormide kasutus kõneliigist: $\chi^2(1, N = 400) = 10,3, p = 0,001$. Täpsemalt esines eitavates osalauetes rohkem normingupäraseid ja jaatavates rohkem norminguvastaseid adessiivivorme. Craméri V oli 0,17, mis tähendab nõrka seost.

Tunnusel *pronoomeni asukoht osalauses* oli kolm väärtust: „alguses”, „keskel” ja „lõpus”. Ka selle tunnuse analüüsitulemustest selgus, et normingupäraseid ja -vastaseid käändevorme esines üsna võrdselt nii osalause alguses (vastavalt 48,6% ja 51,4%), keskel (vastavalt 50,8% ja 49,2%) kui ka lõpus (vastavalt 48,1% ja 51,9%).

Nii tunnusel *rõhumäärsõna olemasolu osalauses* kui ka sõna „mitte” esinemine osalauses oli kaks väärtust: „ei” ja „jah”. Rõhumäärsõnadeks märgendasin muutumatud sõnad, mis lause põhistruktuuri kuulumata väljendasid modaalhinnanguid, sh tõenäosust ning sündmuse või selle osaliste tuntust ja olulisust kuulaja jaoks (nt *kahjuks, vist, hoopis*; vt Erelt 2017a: 61). Nagu eelnevategi tunnuste puhul, esines rõhumäärsõnu normingupäraste ja -vastaste käändevormidega samas osalauses üsna võrdsel määral, vastavalt 48,4% ja 51,6%. Sarnaselt kõneliigiga selgus rõhumäärsõna puhul käändeid eraldi analüüsides, et adessiivis sõltub normingupäraste või -vastaste vormide kasutus rõhumäärsõna olemasolust samas osalauses: $\chi^2(1, N = 400) = 7,36, p = 0,007$. Täpsemalt leidis norminguvastaseid adessiivivorme rohkem neis osalauetes, kus oli rõhumäärsõna. Craméri V oli 0,1, seega on seos siiski nõrk.

Tunnusega sõna „mitte” esinemine osalauses märgendasin laused, kus eitussõna *mitte* esines vahetult pronoomeni *keegi* käändevormi ees. Valimis oli kokku vaid 51 lauset, kus sõna *mitte* selles positsioonis esines. Kuigi *mitte* esines sagedamini koos norminguvastaste käändevormidega (31 korral ehk 60,8%) kui normingupäraste vormidega (20 korral ehk 39,2%), oli sageduste vahe väike (11 sõnavormi), mistõttu ei olnud seegi tulemus käändevormide varieerumise seletamisel statistiliselt oluline.

4.1. Žanr

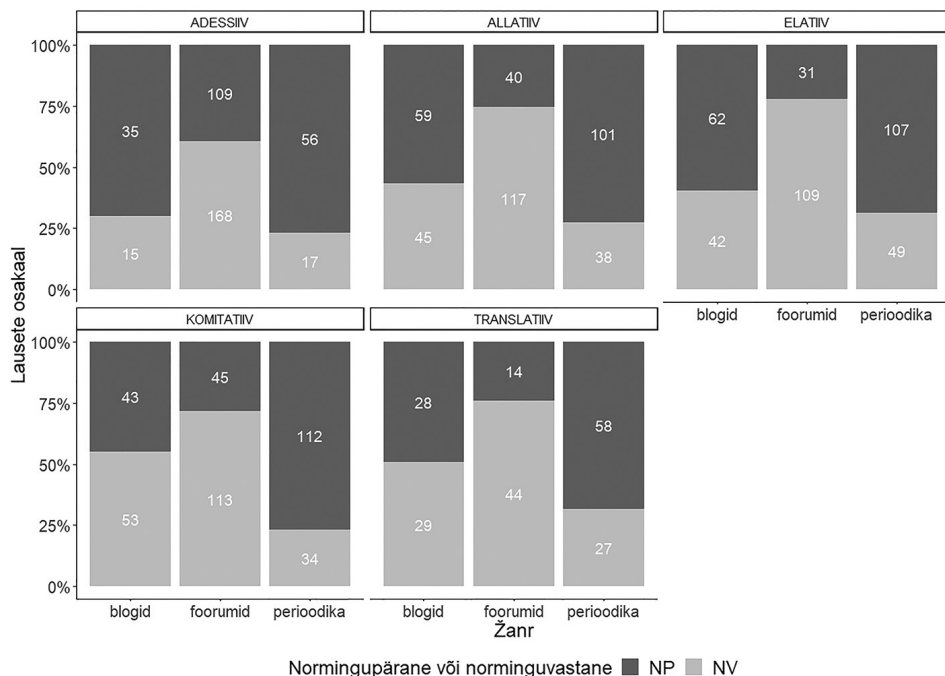
Analüüsisin korpusandmeid kolmest žanrist: perioodika, blogid, foorumid. Perioodika alla kuuluvad näiteks Õhtulehe, Sirbi ning Anne ja Stiili tekstid, blogide alla Malluka, Marimelli ja Päevakera tekstid ning foorumite alla Lapsemure, Soccerneti

ja Rahafoorumi tekstid. Ühtlasi on need kolm žanri ühendkorpuses kõige mahukamad. (Koppel, Kallas 2022: 214) Žanrid valisin selle järgi, et esindatud oleksid nii toimetatud kui ka toimetamata tekstid. Ühendkorpuses on teksti žanr juba määratud (vt Koppel, Kallas 2022: 213–217), mistõttu ei märgendanud ma seda selle uurimuse jaoks käsitsi. Tabelis 6 olen esitanud sõna *keegi* normingupäraste (edaspidi tabelites ja joonistel NP) ja norminguvastaste (edaspidi tabelites ja joonistel NV) käändevormide esinemise sagedus ja osakaal žanri järgi kõigis käänetes kokku.

Tabel 6. *Keegi* käändevormide esinemise sagedus ja osakaal žanri järgi.

	Blogid	Foorumid	Perioodika	Kokku
NP	227 (55,2%)	239 (30,3%)	434 (72,5%)	900
NV	184 (44,8%)	551 (69,7%)	165 (27,5%)	900
Kokku	411 (100%)	790 (100%)	599 (100%)	1800

Tabelist 6 on näha, et perioodikas on normingupäraste käändevormide osakaal 72,5%, samal ajal kui blogides ja foorumites jääb selliste vormide esinemus vastavalt 55,2% ja 30,3% juurde. See tähendab, et normingupärasteid vorme esineb ootuspäraselt enim toimetatud tekstides (perioodika), vähem osaliselt toimetatud tekstides (blogid) ja kõige vähem toimetamata tekstides (foorumid). Ka hii-ruut-testi tulemuse järgi sõltub normingupäraste või -vastaste vormide kasutus žanrist. Hii-ruut-testi tulemus on statistiliselt oluline: $\chi^2(2, N = 1800) = 248,52, p < 0,001$. Craméri V on 0,37, mis näitab tugevat seost. Et näha täpsemalt, mis käänetes sõltub normingupärasus või -vastasus žanrist, tegin hii-ruut-testi iga käände kohta eraldi.



Joonis 1. *Keegi* käändevormide esinemise sagedus ja osakaal žanri järgi käänete kaupa.

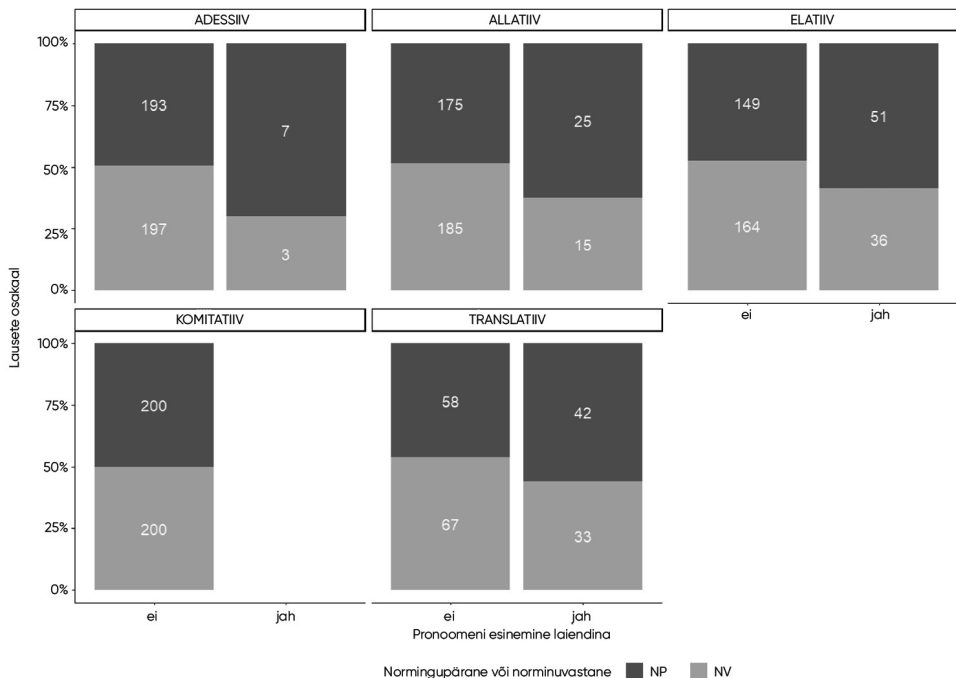
Normingupäraste ja -vastaste käändevormide sagedused ja osakaalud on esitatud joonisel 1. Hii-ruut-testide järgi sõltub normingupäraste või -vastaste vormide kasutus žanrist kõigis käänetes: adessiivis $\chi^2(2, N = 400) = 41,4$, $p < 0,001$, allatiivis $\chi^2(2, N = 400) = 68,2$, $p < 0,001$, elatiivis $\chi^2(2, N = 400) = 68,87$, $p < 0,001$, komitatiivis $\chi^2(2, N = 400) = 71,98$, $p < 0,001$ ja translatiivis $\chi^2(2, N = 200) = 26,84$, $p < 0,001$. Craméri V on adessiivi puhul 0,32, mis näitab keskmist seost. Kõigis teistes käänetes jäävad väärtused 0,37 ja 0,42 vahele ehk seos žanriga on tugev.

4.2. Pronoomeni esinemine laiendina

Tunnusega *pronoomeni esinemine laiendina* märgendasin laused, kus pronoomen ühildus laiendina käändsõna põhjaga käändes ja arvus (näited 5 ja 6). Tunnusel on kaks väärtust: „ei” ja „jah”. Tabelis 7 olen esitanud normingupäraste ja -vastaste vormide sageduse ja osakaalu kõigi käänete peale kokku vastavalt sellele, kas pronoomen esines laiendina („jah”) või mitte („ei”).

(5) [---] või muul juhul peab lubama **kellelgi teisel** vaadata [---] (ÜK 2021, perioodika)

(6) Ära hakka ennast muutma **kellegiks teiseks** [---] (ÜK 2021, foorumid)



Joonis 2. Keegi käändevormide sagedus ja osakaal pronoomeni laiendina esinemise järgi käänete kaupa.

Tabelist 7 on näha, et kui pronoomen *keegi* esineb laiendina, on käändevormid sagedamini normingupärased (59%) ja harvem norminguvastased (41%). Hii-ruut test kinnitab, et seos pronoomeni laiendina esinemise ja vormide normingupärasuse või -vastasuse vahel on statistiliselt oluline: $\chi^2(1, N = 1800) = 7,32, p = 0,007$, kuigi Craméri V on 0,07, mis näitab, et seos on väga nõrk.

Tabel 7. *Keegi* käändevormide sagedus ja osakaal pronoomeni laiendina esinemise järgi.

	Ei	Jah	Kokku
NP	775 (48,8%)	125 (59%)	900
NV	813 (51,2%)	87 (41%)	900
Kokku	1588 (100%)	212 (100%)	1800

Et hinnata, millistes käänetes on statistiliselt oluline seos normingupärasuse või -vastasuse ja selle vahel, kas pronoomen on laiend või mitte, tegin iga käände kohta hii-ruut-testi (v.a komitatiivi puhul, sest selles käändes vormid laiendina ei esine nud). Hii-ruut-testide tulemuste põhjal võib öelda, et üheski käändes ei sõltu normingupärane või -vastane kasutus sellest, kas pronoomen on laiend või mitte (vt ka joonist 2).

4.3. Pronoomeni funktsioon osalauses

Haspelmathi järgi on maailma keeltes indefiniitpronoomenitel üheksa põhifunktsiooni või konteksti, milles need esinevad. Esimeses kolmes funktsioonis viitab indefiniitpronoomen kellelegi/millelegi kindlale ja kõnelejale teadaolevale, kellelegi/millelegi kindlale ja kõnelejale mitteteadaolevale ning kellelegi/millelegi mittekindlale ja irrealsele. Ülejäänud kuues funktsioonis esineb indefiniitpronoomen kontekstides, kus väljendatakse üldküsimust, *kui*-konditsionaalsust, võrdlust koos võrdlusalusega, otsest eitust, kaudset eitust (ingl *indirect negation*) ning vaba valikut. (Haspelmath 1997: 2, 31–48)

Indefiniitpronoomenite funktsioonide määratlemisel huvitasid Haspelmathi (1997: 2) indefiniitpronoomenite sarjade funktsionaalsed erinevused. Et eesti keeles saab üldistavalt moodustada ainult *gi-/ki*-liiteliste indefiniitpronoomenite sarja, kus üks pronoomen täidab peaaegu kõiki funktsioone, nagu näitan järgnevas analüüsis, mugandasin teooriat oma analüüsi tarbeks, eesmärgiga vaadata, kas pronoomeni *keegi* käändevormide normingupärasus või -vastasus sõltub pronoomeni funktsioonist. Vaba valikut ma siinses peatükis ei käsitle, sest seda ei ole võimalik pronoomeni *keegi* abil väljendada. Eesti keeles võiks vaba valikut väljendada fraasiga *ükskõik kes*, inglise keeles väljendatakse seda näiteks indefiniitpronoomeni *anybody* abil. Seega on tunnusel *pronoomeni funktsioon osalauses* siinses analüüsis kaheksa võimalikku väärtust.

Kindlale ja kõnelejale teadaolevale viitavaks (KT) märgendasin indefiniitpronoomeni lausetes, 1) kus viidatav on olemas, teda on võimalik ainuliselt identifitseerida ja ta on lugejale teada (näide 7); 2) kus pronoomenit kasutatakse kellelegi tagasiviitavalt (näide 8); 3) kus kõneleja soovib kuulaja eest viidatava isikut varjata (näide 9).

- (7) [---] ja sel hooajal on selleks **kellegiks** näiteks **Mikel Merino** [---] (ÜK 2021, foorumid)
- (8) [---] **ta** pole isegi kena ja ma saan aru, et keegi kes on nagu mina ei peaks isegi mõtlema **kellestki temasugusest** aga ikkagi... (ÜK 2021, foorumid)
- (9) Ma pean **kellegile** väga tänulik olema mõne asja eest [---] (ÜK 2021, blogid)

Kindlale ja kõnelejaile mitteteadaolevale viitavaks (KET) märgendasin indefiniitpronoomeni lauses, kus viidatav on olemas ja teda on võimalik ainulisena identifitseerida, kuid ta ei ole lugejale teada (näited 10 ja 11). Näidet 11 saaks tõlgendada ka nii, et kõneleja teab, kellest ta räägib, kuid tal on viidatava kohta negatiivne aramus. Samuti märgendasin indefiniitpronoomeni KET-iks lauses, kus kasutati sõna *kindel* (näide 12) või fraasi *üks ainus* (näide 13).

- (10) Mõne aja eest edastasid agentuurid teate **kellestki 18-aastasest koolipoisist**, kes maakerale üksipäini tiiru ümber tegi [---] (ÜK 2021, perioodika)
- (11) [---] mis algas **kellestki Goldbergist** [---] (ÜK 2021, perioodika)
- (12) Ilmselgelt rääkis ta **kellegist kindlast** [---] (ÜK 2021, foorumid)
- (13) [---] mis paneb meid mõtlema **kellegist ühest ainsast**... (ÜK 2021, blogid)

Mittekindlale ja irrealsele viitavaks (IMT) märgendasin indefiniitpronoomeni lauses, kus viidatavat ei ole võimalik kindlaks määrata, st ta pole kõnelejaile teada (näited 14 ja 15).

- (14) Veel mõni foto on tuvastatult tehtud hoopis **kellestki teisest**. (ÜK 2021, perioodika)
- (15) Mis teeb **kellegist** hea rulluisutaja? (ÜK 2021, perioodika)

Üldküsimust väljendavas kontekstis esinevaks (KÜ) märgendasin indefiniitpronoomeni lauses, kus väljendatakse ainult üldküsimust. Arvestasin nii küsisõnaga markeeritud (näide 16) kui ka markeerimata (näide 17) küsimusi.

- (16) **Kas** sa unistad **kellegist**, keda tead? (ÜK 2021, foorumid)
- (17) Siin pole pärast van Dijki **kellestki** räägitud? (ÜK 2021, foorumid)

Kui-konditsionaalsust väljendavas kontekstis esinevaks (KO) märgendasin indefiniitpronoomeni *kui*-tingimuslauses (näide 18) ja lauses, kus *kui* väljendas ajatähendust (näide 19), sest eesti keeles on neid raske üksteisest eristada.

- (18) **Kui kellelgi** on mõtteid, **siis** need võib siia kirja panna. (ÜK 2021, blogid)
- (19) [---] **koguaeg kui** ma **kellegiga** räägin **siis** ta hakkab ennast sinna vestlusesse segama [---] (ÜK 2021, foorumid)

Võrdlust koos võrdlusalusega väljendavas kontekstis esinevaks (V) märgendasin indefiniitpronoomeni lauses, kus see esines võrdlusalusena (näited 20 ja 21).

(20) **Ma ei ole kellestki üle.** (ÜK 2021, blogid)

(21) [---] et ikka on **keegi kellegist** tugevam [---] (ÜK 2021, foorumid)

Otsest eitust väljendavas kontekstis esinevaks (OE) märgendasin indefiniitpronoomeni lausetes, kus mingis käändes pronoomen *keegi* on eitavas vormis verbi komplement (näited 22 ja 23).

(22) **Ma ei sõltu kellestki** ega millestki [---] (ÜK 2021, perioodika)

(23) Sportlik sisu aga **ei paku kellegile** mingit huvi. (ÜK 2021, blogid)

Kaudset eitust väljendavas kontekstis esinevaks (KE) märgendasin indefiniitpronoomeni lausetes, kus eituse mõjualas oli üksnes pronoomeni *keegi* käändevormi sisaldav komplementlause ning tegu oli eituse tõstega (ingl *negative raising*) alistatud moodustajast põhilause, mille predikaat väljendab uskumist, soovimist või näimist (näide 24) (vt Erelt 2017b: 190–191; Haspelmath 1997: 32–33). Lisaks märgendasin sama funktsiooniga laused, kus kasutatakse pronoomeni *keegi* käändevormi sisaldava komplementlause ees sõna *ilma* (näide 25), samuti laused, kus infinitiivi- ja konverbitarindite eitust väljendatakse partikliga *mitte* (näide 26) või kus sõna *mitte* mõjualas on ainult pronoomeni *keegi* käändevorm (näide 27) (vt Erelt 2017b: 191).

(24) Siis ma ei usu, et **nad kellegile silma saaks paista** [---] (ÜK 2021, foorumid)

(25) Loodan, et langetate enda otsuse mõistlikult **ilma**, et **kellegist** teisest sõltuv või mõjutatud oleks. (ÜK 2021, foorumid)

(26) Kindlasti olen **kellestki mitte välja teinud** [---] (ÜK 2021, perioodika)

(27) [---] ja jääda **mitte kellegiks**. (ÜK 2021, perioodika)

Funktsioonide määratlemisel on Haspelmath (1997: 2–3) lähtunud nii infostruktuurist (kõnelejale teadaolev ja mitteteadaolev) kui ka lause grammatilisest struktuurist (lausetüüp, eitus), mistõttu on funktsioone lausetes kohati keeruline eristada. Nii võib indefiniitpronoomen esineda näiteks kontekstis, kus väljendatakse samal ajal üldküsimust ja võrdlust (näide 28), üldküsimust ja otsest eitust (näide 29), *kui*-konditsionaalsust ja otsest eitust (näide 30) või võrdlust ja otsest eitust (näide 31).

(28) [---] **kas keegi on kellestki** parem. (ÜK 2021, perioodika)

(29) **Kellegile** Tallinnas üks mehhaanik-fanaatik kaasaelaja **ei mahu** peale? (ÜK 2021, foorumid)

(30) [---] **kui kellestki** mööda **pole** vaja kupatada, pole vigagi. (ÜK 2021, perioodika)

(31) Nii sitasti **ei ole ma** veel kunagi **kellestki** arvanud kui **sinust**. (ÜK 2021, perioodika)

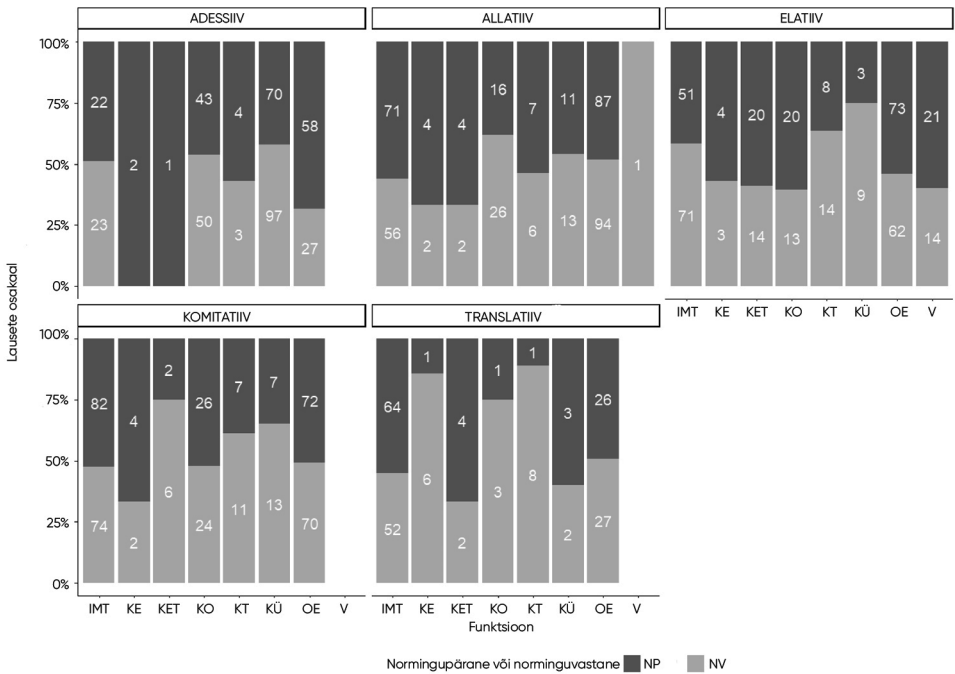
Juhul kui funktsioonid lausetes kattusid, võtsin arvesse väiksemat üksust, kus indefiniitpronoomeni käändevorm esines, mitte tervet osalauset. Näiteks märgendasin näidetes 28 ja 31 indefiniitpronoomeni võrdlust koos võrdlusalusega väljendavas

kontekstis (V) esinevaks, mitte üldküsimumst väljendavas kontekstis (KÜ) või otsest eitust väljendavas kontekstis (OE) esinevaks. Kui kattusid funktsioonid, mis mõlemad hõlmasid kogu osalauset, nagu näidetes 29 ja 30, eelistasin verbist lähtuval otsest eitust väljendavas kontekstis esinemisele muid kitsamaid funktsioone, nagu üldküsimumst väljendavas kontekstis või *kui*-konditsionaalsust väljendavas kontekstis esinemist.

Tabelis 8 olen esitanud normingupäraste ja -vastaste käändevormide esinemise sageduse ja osakaalu funktsioonide järgi kõigis käänetes kokku.

Tabel 8. Keegi käändevormide esinemise sagedus ja osakaal funktsioonide järgi.

	IMT	KE	KET	KO	KT	KÜ	OE	V	Kokku
NP	290 (51,2%)	15 (53,6%)	31 (56,4%)	106 (47,7%)	27 (39,1%)	94 (41,2%)	316 (53%)	21 (58,3%)	900
NV	276 (48,8%)	13 (46,4%)	24 (43,6%)	116 (52,3%)	42 (60,9%)	134 (58,8%)	280 (47%)	15 (41,7%)	900
Kokku	566 (100%)	28 (100%)	55 (100%)	222 (100%)	69 (100%)	228 (100%)	596 (100%)	36 (100%)	1800



Joonis 3. Keegi käändevormide esinemise sagedus ja osakaal funktsioonide järgi käänete kaupa.

Tabelist 8 on näha, et normingupäraste vormide esinemine on ülekaalus viies funktsioonis: indefiniitpronoomen viitab mittekindlale ja irrealsele (IMT) või kindlale ja kõnelejale mitteteadaolevale (KET), esineb otsest eitust väljendavas kontekstis (OE), kaudset eitust väljendavas kontekstis (KE) või kontekstis, kus on võrdlus koos võrdlusalusega (V). Norminguvastaste vormide esinemine on ülekaalus

kolmes funktsioonis: indefiniitpronoomen esineb *kui*-konditsionaalsust (KO) või üldküsimumst väljendavas kontekstis (KÜ) või viitab kindlale ja kõnelejale teadaolevale (KT). Enim normingupäraseid vorme esineb kontekstis, kus väljendatakse võrdlust koos võrdlusalusega (58,3%), ja enim norminguvastaseid vorme seal, kus viidatakse kindlale ja kõnelejale teadaolevale (60,9%). Hii-ruut-testi tulemus $\chi^2(7, N = 1800) = 15,28, p = 0,03$ on statistiliselt oluline ja näitab, et vormide normingupärasus või -vastasus sõltub pronoomeni funktsioonist osalauses. Craméri V on 0,1, mis tähendab nõrka seost. Et hinnata vormide normingupärasuse või -vastasuse sõltumist käändest, visualiseerisin tulpdiagrammiga normingupärase ja -vastaste käändevormide esinemise sageduse ja osakaalu kõigis uuritavates käänetes (vt joonist 3).

Kuna kõigis käänetes on mõne väärtuse teoreetiline sagedus väiksem kui 5, tegin sõltuvuse hindamiseks kõigi käänete puhul Fisheri täpse testi. Testide tulemus näitas, et vormide normingupärasus või -vastasus sõltub pronoomeni funktsioonist adessiivis ja translatiivis, kus testi tulemused on statistiliselt olulised: adessiivis $p = 0,001$ ja translatiivis $p = 0,05$. Et käändevormide funktsioonid lausetes kohati kattuvad ja neid on raske kindlate parameetrite alusel eristada, jätsin tunnuse *pronoomeni funktsioon osalauses* järgnevas peatükis esitatud mitmetunnuselisest analüüsist välja.

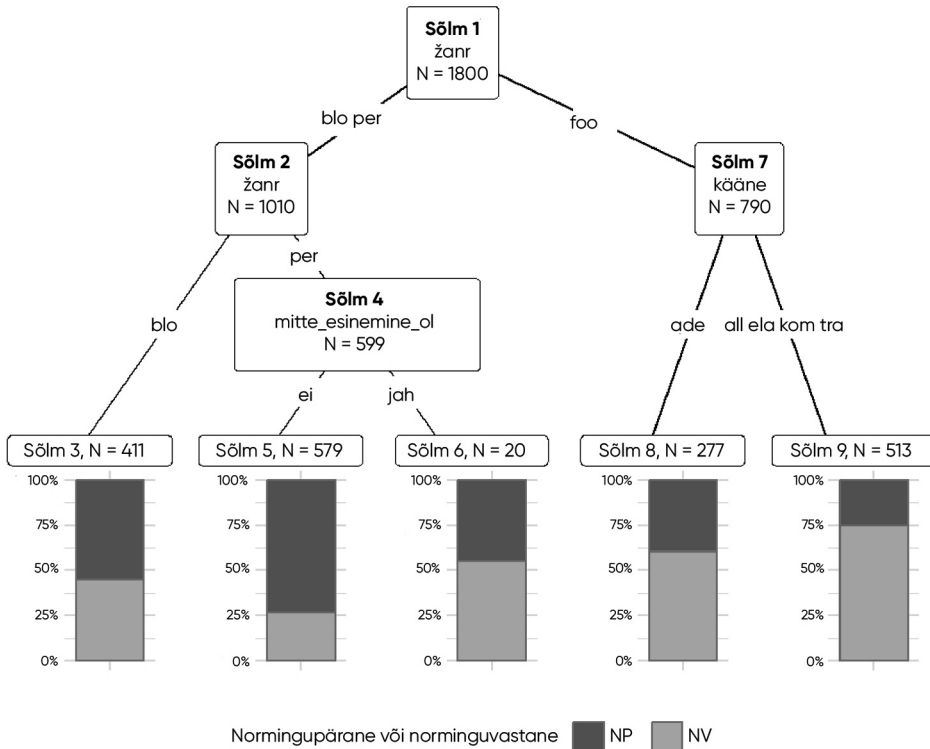
5. Pronoomeni *keegi* käändevormide varieerumise mitmetunnuseline analüüs

Tegin tingimusliku otsustuspuu mudeli, et selgitada välja, kuidas eri tunnused koos käändevormide kasutust mõjutavad ning millised neist tunnustest on varieerumise seisukohast statistiliselt olulised. Tingimuslike otsustuspuude eelis on see, et need ei eelda andmetelt jaotusi, ei ole tundlikud korreleeruvate tunnuste suhtes ja võimaldavad hästi visualiseerida tunnustevahelisi koosmõjusid.

Jooniselt 4 on näha, et normingupärase ja -vastaste käändevormide varieerumisega on statistiliselt kõige tugevamalt seotud žanr. Selle alusel jagunevad esimeses sõlmes blogid ja perioodika ühte harusse ning foorumid teise harusse. Teises sõlmes jagunevad blogid ja perioodika omakorda žanri alusel eri harudesse, mille järgi käändevormid, mis esinevad blogitekstides (sõlm 3), on pigem normingupärased (55,2%), ning perioodikatekstides (sõlm 4) sõltub vormide normingupärasus või -vastasus sellest, kas samas osalauses esineb eitussõna *mitte*: kui seda ei esine (sõlm 5), on käändevormid pigem normingupärased (73,4%), ning kui esineb (sõlm 6), on käändevormid pigem norminguvastased (55%). Foorumites sõltub vormide normingupärasus või -vastasus käändest (sõlm 7). Täpsemalt on foorumites kõigi uuritud käänete vormid pigem norminguvastased, kuid erineval määral: adessiivis (sõlm 8) on norminguvastaseid vorme 60,6%, allatiivis, elatiivis, komitatiivis ja translatiivis (sõlm 9) kokku aga 74,7%.

Kokkuvõttes mõjutab pronoomeni *keegi* käändevormide normingupärasust või -vastasust statistiliselt kõige tugevamalt teksti žanr ja selle taustal toimetatuse mõju. Kui tekstid on toimetatud enamjaolt (perioodika) või osaliselt (blogid), on käände-

vormid pigem normingupärased (v.a perioodikas, kui vahetult käändevormi ees esineb eitussõna *mitte*), ning kui tekstid on toimetamata (foorumid), on käändevormid pigem norminguvastased, seda käändest sõltumata. Tingimusliku otsustuspuu klassifitseerimistäpsus on 67% ja C-indeksi väärtus 0,72, mis viitab ühelt poolt mudeli rahuldavale võimele normingupärasteid ja -vastaseid vorme üksteisest eristada, ent teisalt sellele, et vormide kasutust võivad mõjutada paljud muudki tegurid, mida analüüsis ei arvestatud.



Joonis 4. Normingupäraste ja -vastaste käändevormide varieerumise seletamisel olulised tunnused.³

6. Kokkuvõte

*gi-/ki-*liiteliste indefiniitpronoomenite käändevormide kasutus on varieeruv: liide võib peale kirjakeele normi kohase positsiooni käändelõpu järel paikneda ka käändelõpu ees, kahe käändelõpu vahel või käändelõpu ees ja järel. Sellist varieerumist on täheldatud juba üle saja aasta nii kirjakeeles (vt Rull 1917) kui ka murretes (vt Saareste 1955: 16).

³ `puu <- ctree(np_nv ~ kääne + asukoht_ol + esineb_laiendina + kõneliik + mitte_esinemine_ol + modalsõna_ol + žanr, data = andmed)`

Sellest lähtuvalt andsin artiklis 2021. aasta eesti keele ühendkorpuse põhjal ülevaate indefiniitpronoomenite *keegi*, *miski*, *kumbki* ja *ükski* semantiliste käändevormide varieerumise ulatusest kirjalikus keeles. Uuringust selgus, et kõigi nelja pronoomeni puhul esines ootuspäraselt enim normingupäraseid käändevorme: *keegi* (78,6%), *miski* (92,5%), *kumbki* (94,6%) ja *ükski* (99,1%). Nelja pronoomeni peale kokku oli normingupäraseid käändevorme 85,3% ja norminguvastaseid 14,7%, millest enamiku (14,1%) moodustasid malli TgiL järgivad vormid, kus tüvele järgneb kõigepealt liide ja seejärel käändelõpp. See tähendab, et peaaegu iga seitsmendat vormi kasutatakse ootuspärasest erinevalt.

Hii-ruut-testi ja Fisheri täpse testi tulemused näitasid, et käändevormide normingupärane või -vastane kasutus sõltub statistiliselt oluliselt žanrist nii kõigis uuritud käänetes (elatiiv, allatiiv, adessiiv, translatiiv, komitatiiv) eraldi kui ka kõigi käänete peale kokku. Uuritud käänete peale kokku sõltus käändevorm ka pronoomeni esinemisest laiendina ja pronoomeni funktsioonist osalauses. Igat käänet eraldi vaatav analüüs näitas, et vormide kasutus sõltub adessiivis ja translatiivis ka pronoomeni funktsioonist osalauses. Tingimusliku otsustuspüü mudeli järgi on normingupäraste ja -vastaste käändevormide varieerumisega statistiliselt kõige tugevamalt seotud žanr ja sellest tulenevalt ka teksti toimetatus. Blogi- ja perioodika-tekstides, mis on enamjaolt või osaliselt toimetatud, on käändevormid tõenäoliselt normingupärased, ning foorumitekstides, mis on toimetamata, tõenäoliselt norminguvastased.

Varieerumise ulatus olukorras, kus peaaegu igas seitsmendas käändevormis esineb *-gi/-ki* käändelõpu ees, moodustatakse lisarõhutusega vorme ning varieerumist mõjutava tegurina esildub kõige tugevamalt žanr, näitab, et vaatamata järjepidevale keele normimisele ei ole normingud keelde kinnistunud ega tundu märgatavale osale keekekasutajatest loomulikud. Sellised uurimistulemused kinnitavad, et indefiniitpronoomenitele liituv *-gi/-ki* ei ole kliitik, vaid pigem tuletusliide, sest väljendab neis sõnades indefiniitsust, mitte emfaatilisust (vt Zwicky 1997; Nevis 1984). Tuletusliitena käsitamist toetavad ka eesti keele morfotaktika reeglid, mille järgi tuletusliited paiknevad tüve järel ja vormitunnuse ees ning kliitikud sõnas viimasel positsioonil (vt Viht, Habicht 2019: 47). Kirjakeele normi kohase vormi ebaloomulikkusele ajaloolises vaates viitab ka Saareste (1955: 16) sõnakaart (vt kaarti 1) *keegi* allatiivivormide varieerumise kohta, kus enamikus murretes paikneb *-gi/-ki* käändelõpu ees.

Siinne uurimus on näide sellest, kuidas statistiliste meetoditega süsteemselt uurida eesti keele käändevormistiku varieerumist. Käändevormide varieerumist mõjutava tegurina esildus statistiliselt kõige tugevamalt žanr, kuid seda võivad mõjutada ka teised tegurid. Kirjalik keel registrina võimaldab vaadata vaid teatud hulka varieerumist mõjutada võivaid tunnuseid, mistõttu tuleks uurida ka suulist keelt, et näiteks lausungi infostruktuuri ja prosoodiat vormikasutuse varieerumisel täpselt arvesse võtta.

Artikli valmimist on toetanud teadusprojekt EKKD-TA2 „Eesti keele morfosüntaktiline varieerumine”.

VEEBIVARAD

- etTenTen 2013.** <https://www.keeleeveeb.ee/dict/corpus/ettenten/about.html>
- ggplot2.** Create Elegant Data Visualisations Using the Grammar of Graphics.
<https://CRAN.R-project.org/package=ggplot2>
- ggparty.** ggparty: „ggplot” Visualizations for the „partykit” Package. R package version 1.0.0.
<https://CRAN.R-project.org/package=ggparty>
- partykit.** A Modular Toolkit for Recursive Partytioning in R.
<https://CRAN.R-project.org/package=partykit>
<https://doi.org/10.32614/cran.package.partykit>
- pROC.** Display and Analyze ROC Curves. <https://CRAN.R-project.org/package=pROC>
- R Core Team 2022.** R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Sketch Engine.** Lexical Computing CZ s.r.o. <https://www.sketchengine.eu>
- ÜK 2021** = Eesti keele ühendkorpus 2021. Koost Kristina Koppel, Jelena Kallas.
<https://doi.org/10.15155/3-00-0000-0000-0000-08D17L>

KIRJANDUS

- Agresti, Alan 2007.** An Introduction to Categorical Data Analysis. (Wiley Series in Probability and Statistics.) 2. tr. Hoboken, New Jersey: John Wiley & Sons, Inc.
- Alvre, Paul 1980.** *gi-liitelisist pronoomeneist.* – Keel ja Kirjandus, nr 9, lk 539–543.
- Barlow, Michael; Kemmer, Suzanne 2000.** Introduction: a usage-based conception of language. – Usage-Based Models of Language. Toim M. Barlow, S. Kemmer. Stanford, California: CSLI Publications, lk vii–xxviii.
- Brinton, Laurel J.; Traugott, Elizabeth Closs 2005.** Lexicalization and Language Change. (Research Surveys in Linguistics.) Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511615962>
- Bybee, Joan 2010.** Language, Usage and Cognition. New York: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511750526>
- Cramér, Harald 1946.** Mathematical Methods of Statistics. (Princeton Mathematical Series, Princeton Landmarks in Mathematics and Physics.) Princeton: Princeton University Press. <https://doi.org/10.1515/9781400883868>
- EKK 2020** = Mati Erelt, Tiiu Erelt, Kristiina Ross, Eesti keele käsiraamat. Uuendatud väljaanne. Tallinn: Eesti Keele Sihtasutus, 2020.
- Erelt, Mati 2017a.** Sissejuhatus süntaksisse. – Eesti keele süntaks. (Eesti keele varamu 3.) Toim M. Erelt, Helle Metslang. Tartu: Tartu Ülikooli Kirjastus, lk 53–89.
- Erelt, Mati 2017b.** Öeldis. – Eesti keele süntaks. (Eesti keele varamu 3.) Toim M. Erelt, Helle Metslang. Tartu: Tartu Ülikooli Kirjastus, lk 93–239.
- Haspelmath, Martin 1997.** Indefinite Pronouns. (Oxford Studies in Typology and Linguistic Theory.) Oxford: Clarendon Press.
- Hothorn, Torsten; Hornik, Kurt; Zeileis, Achim 2006.** Unbiased recursive partitioning: A conditional inference framework. – Journal of Computational and Graphical Statistics, nr 15, kd 3, lk 651–674. <https://doi.org/10.1198/106186006X133933>

- Hothorn, Torsten; Zeileis, Achim 2015. *partykit*: a modular toolkit for recursive partytioning in R. – Journal of Machine Learning Research, nr 16, lk 3905–3909.
- Jakubiček, Miloš; Kilgarrieff, Adam; McCarthy, Diana; Rychlý, Pavel 2010. Fast syntactic searching in very large corpora for many languages. – Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation. Toim Ryo Otoguro, Kiyoshi Ishikawa, Hiroshi Umemoto, Kei Yoshimoto, Yasunari Harada. Institute for Digital Enhancement of Cognitive Development, Waseda University, lk 741–747.
- Kilgarrieff, Adam; Baisa, Vít; Bušta, Jan; Jakubiček, Miloš; Kovář, Vojtěch; Michelfeit, Jan; Rychlý, Pavel; Suchomel, Vít 2014. The Sketch Engine: ten years on. – Lexicography, nr 1, lk 7–36.
- Koppel, Kristina; Kallas, Jelena 2022. Eesti keele ühendkorpuste sari 2013–2021: mahukaim eesti keele kogu. – Eesti Rakenduslingvistika Ühingu aastaraamat, nr 18, lk 207–228. <https://doi.org/10.5128/ERYa18.12>
- Langacker, Ronald W. 1987. Foundations of Cognitive Grammar. I kd. Theoretical Prerequisites. Stanford: Stanford University Press.
- Mengden, Ferdinand von; Coussé, Evie 2014. Introduction: The role of change in usage-based conceptions of language. – Usage-Based Approaches to Language Change. (Studies in Functional and Structural Linguistics.) Toim E. Coussé, F. von Mengden. Amsterdam: John Benjamins Publishing Company, lk 1–19.
- Nevis, Joel A. 1984. A non-endoclititic in Estonian. – Lingua, nr 64, kd 2–3, lk 209–224. [https://doi.org/10.1016/0024-3841\(84\)90017-2](https://doi.org/10.1016/0024-3841(84)90017-2)
- Pant, Annika 2018. Asesõnade *keegi*, *miski*, *kumbki* käändevormide varieerumine eesti kirja-keeles. Bakalaureusetöö. Tartu Ülikool. <http://hdl.handle.net/10062/60560>
- Pant, Annika 2020. Pronoomenite *keegi*, *miski*, *kumbki*, *ükski* käändevormide kasutus tänapäeva eesti keeles. Magistritöö. Tartu Ülikool. <http://hdl.handle.net/10062/68143>
- Pilvik, Maarja-Liisa; Plado, Helen; Lindström, Liina 2021. Murded, varieerumine ja korpus-andmed. Eitussõna paiknemine võru ja seto eituslausetes. – Keel ja Kirjandus, nr 8–9, lk 771–795. <https://doi.org/10.54013/kk764a7>
- Robin, Xavier; Turck, Natacha; Hainard, Alexandre; Tiberti, Natalia; Lisacek, Frédérique; Sanchez, Jean-Charles; Müller, Markus 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. – BMC Bioinformatics, nr 12:77. <https://doi.org/10.1186/1471-2105-12-77>
- Rull, Aado 1917. Kaaslõpp „gi” ja ta kidurad kaimud. – Eesti Kirjandus, nr 2, lk 82–88.
- Saareste 1923 = A. S. [Andrus Saareste], *Kellegile, mingit, kuskil*. – Eesti Keel, nr 4, lk 116–117.
- Saareste 1933 = Albert [Andrus] Saareste, Meie keele korraldamise põhimõttest ja nende kohaldamisest. – Eesti Kirjandus, nr 9, lk 446–454.
- Saareste, Andrus 1936. Eesti õigekeelsuse päevaküsimustest. Johannes Aaviku „Eesti õigekeelsuse õpiku ja grammatika” ilmumise puhul. 3. – Eesti Kirjandus, nr 12, lk 548–556.
- Saareste, Andrus 1955. Petit atlas des parlers estoniens. Väike eesti murdeatlas. (Skrifter utgivna av Kungl. Gustav Adolfs Akademien 28.) Uppsala: Almqvist & Wiksells.
- Zwicky, Arnold M. 1977. On Clitics. Bloomington, Indiana: Indiana University Linguistics Club. https://web.stanford.edu/~zwicky/on_clitics.pdf
- Veski, Johannes Voldemar 1933. Keelelist nipet-näpet. – Eesti Kirjandus, nr 8, lk 382–386.

- Viht, Annika; Habicht, Külli 2019. Eesti keele sõnamuutmine. (Eesti keele varamu IV.) Tartu: Tartu Ülikooli Kirjastus.
- Wickham, Hadley 2016. ggplot2: Elegant Graphics for Data Analysis. (Use R!) 2. tr. Cham: Springer. <https://doi.org/10.1007/978-3-319-24277-4>
- ÕS 2018 = Eesti õigekeelsussõnaraamat ÕS 2018. Toim Maire Raadik. Koost Tiit Erelt, Tiina Leemets, Sirje Mäearu, M. Raadik. Tallinn: Eesti Keele Sihtasutus.

Annika Kängsepp (sünd 1995), MA, Tartu Ülikooli eesti ja üldkeeleteaduse instituudi keeleteaduse nooremteadur (Jakobi 2-430, 51005 Tartu), annika.kangsepp@ut.ee

The variation in the case forms of the indefinite pronoun *keegi* 'someone' in written Estonian

Keywords: language variation, usage-based linguistics, corpus linguistics, language planning, indefinite pronouns

There is variation in the case forms of the indefinite pronouns *keegi* 'someone', *miski* 'something', *kumbki* 'either' and *ükski* 'none' in Estonian. In these pronouns, *-gi/-ki* can appear after the case ending (e.g., *kellelegi*), as specified in the standard language norm; before the case ending (e.g., *kellegile*); between two case endings (e.g., *kellelegile*); or both before and after the case ending (e.g., *kellegilegi*).

In this article, I used data from the Estonian National Corpus 2021 to provide an overview of the extent of variation in the case forms of *keegi*, *miski*, *kumbki* and *ükski*, and to describe the factors influencing the variation in the case forms of the pronoun *keegi*. The results indicate that for all four pronouns, *-gi/-ki* is most commonly placed after the case ending, accounting for 85.3% of all occurrences. The second most common forms are those where *-gi/-ki* precedes the case ending, making up 14.1% of occurrences.

A univariate analysis conducted to identify the factors influencing this variation showed that genre, the occurrence of the pronoun as an attribute, and the function of the pronoun in a clause significantly influenced the variation of case forms. A multivariate analysis indicated that genre – and, consequently, the impact of editing – was most strongly associated with the variation. In genres such as blogs and periodicals, which are typically edited, *-gi/-ki* is predominantly placed after the case ending. Conversely, in forums, which are often unedited, it is more common for *-gi/-ki* to appear before the case ending, between two case endings, or both before and after the case ending.

Annika Kängsepp (b. 1995), MA, University of Tartu, Institute of Estonian and General Linguistics, Junior Research Fellow in Linguistics (Jakobi 2-430, 51005 Tartu), annika.kangsepp@ut.ee