

FONOLOOGILINE VARIEERUMINE EESTI KEELES KOLME NÄHTUSE NÄITEL¹

LIISI PIITS, MARI-LIIS KALVIK

Annotsioon. Artikkel on kokkuvõtte tulemustest, milleni oleme jõudnud viie aasta jooksul eesti keele hääldust uurides. Oleme käsitlenud kolme fonoloogiliselt varieeruvat nähtust: sõnaalgulise *h* esinemist, palatalisatsiooni ühesilbilistes *i*-tüvelistes pika vokaaliga sõnades ning väldet varieeruva vältega sõnades. Artiklis kirjeldame fonoloogilise varieerumise kõnekorpust, et anda edaspidiste hääldusuuringute tarbeks põhjalik ülevaade kogutud kõnematerjalist, ning teeme kokkuvõtte, kuidas käsitletavad nähtused varieeruvad ja milliseid hääldusvariante lugejad eelistavad. Tulemustest ilmneb, et sõnaalgulist *h*-d hääldatakse eesti keeles ette lugedes peaaegu alati: keskmiselt 92% juhtudest ja sagedates sõnades veelgi enam. Lõpukonsonandid ühesilbilistes *i*-tüvelistes pika vokaaliga sõnades on palataliseeritud ligikaudu kolmandikul juhtudest. Ligi 50 sõna puhul tegime kindlaks, kumba väldet keelejuhid eelistasid; nelja tüübi puhul selgus hääldustrend. Saadud tulemused võimaldavad arutleda võimalike põhjuste üle ja teha praktilisi järeldusi, kuidas sünteeskõnet loomulikumaks muuta.

Võtmesõnad: eesti keel, fonoloogia, kõnesüntees, sõnaalguline *h*, palatalisatsioon, välde

Artikli üks eesmärk on anda ülevaade olulisematest tulemustest, milleni oleme fonoloogilist varieerumist uurides jõudnud, ja kirjeldada põhjalikumalt fonoloogilise varieerumise korpust, mis sai nähtuste uurimiseks salvestatud ja mis võiks pakkuda huvi ka teistele kõneuurijatele. Uurimuse teine eesmärk on selgitada välja, millist hääldusvarianti kasutatakse eri nähtuste (sõnaalgulise *h*, palatalisatsiooni ja välte) korral sagedamini,

¹ Artikkel on valminud Eesti-uuringute Tippkeskuse (TK145-CEES), Haridus- ja Teadusministeeriumi uurimisprojekti IUT 35-1 „Kõnestiilid, lauseprosoodia ja fonoloogiline varieerumine: kirjeldus, teooria ja modelleerimine“ ning EKT projekti „Väljendusriikas ja mitmekesine kõnesüntees“ toetusel.

mis seda kasutust võib mõjutada ning milliseid järeldusi saab tulemustest kõnesünteesi tarbeks teha.

Praktiline vajadus fonoloogilist varieerumist uurida tekkis tekst-kõnesünteesi jaoks korpust salvestades, kui hääledoonori ehk korpuse sisse-lugeja hääldus ühesilbilistes *i*-tüvelistes pika vokaaliga sõnades, nagu nt *roos*, lahknes palatalisatsiooni osas arvestataval määral sellest, mis oli kirjas „Eesti õigekeelsussõnaraamatus ÕS 2013“. Kuna kõnesüntees lähtub palatalisatsiooni (ja ka välte) määramisel sõnaraamatust, siis oleks hääledoonori palataliseerimata häälikud määratud korpuses ÕSi põhjal palataliseerituks. Selle vastuolu lahendamiseks oli kaks võimalust: kas suunata lugejat, et ta hääldaks sõnu nii, nagu ÕSi märgendus ette näeb, või muuta ÕSi märgendust. Enne valiku tegemist oli vaja välja selgitada, kas sõnaraamatus esitatud konsonantide palatalisatsioon uuritavas sõnatüübis vastab tegelikule hääldustavale. Kui varem on uuritava sõnatüübi palatalisatsiooni normimisele ette heidetud ebajärjekindlust ja süsteemipäratust (vt Hint 1968), siis hilisemates õigekeelsussõnaraamatutes on püütud sarnases foneetilis-fonoloogilises ümbruses konsonantide palatalisatsiooni käsitleda järjekindlamalt ning aastate jooksul on iga uue õigekeelsussõnaraamatuga suurenenud palataliseerituks märgitud sõnade hulk (vt Piits, Kalvik 2019: 519–521). Meie eesmärk oli kindlaks teha, kas tegelik hääldustrend on liikunud samas suunas.

Lisaks vajas kindlaks tegemist, kummas vältes peaks süntesaator hääldama sõnu, mida sõnaraamat lubab hääldada mõlemas vältes. Kuigi vältekontrastil on eesti keeles täita oluline roll ja erinevalt palatalisatsioonist sõltub sõna vältest sageli ka sõna muuttüüp, võib välde morfoloogilises süsteemis mõnel juhul neutraliseeruda. See ei avaldu teise ja kolmanda välte vahepealse hääldusena, vaid võib avalduda teise- ja kolmandavälte-liste paralleelvormide esinemisena (Hint 1968: 78–79). „Eesti õigekeelsussõnaraamatus ÕS 2018“ on 395 märksõna, millel on sõnastikus olemas ka ainult välte poolest erinev paralleelvorm. Kui inimkõnelejale näitavad need vaid varieerimisvõimalusi, siis kõnesüntesaator vajab teksti analüüsimisel üht sisendit ja juhusliku valiku tegemise asemel oleks parem, kui valituks osutuks kasutatavam välde.

Peale eeltoodud nähtuste uurisime, kas ja kuidas hääldatakse sõnalgulist *h*-d ette lugedes, et teada saada, millal jätta *h* sünteeskõnes hääldamata. Eelmise sajandi alguses lepiti küll kokku, mis sõnade algusesse *h* kirjapildis kinnistada, ja õigekeelsussõnaraamatud on seda järjekindlalt

esitanud, aga kõnes on olnud palju varieerumist. Soovisime kindlaks teha varieerumise ulatuse, et rakendada seda sünteeskõne loomulikumaks muutmisel.

Eelkirjeldatud fonoloogiliselt varieeruvate nähtuste analüüsimiseks salvestasime korpuse, mis koosneb 191 eestlase etteloetud lausetest.

1. Kõnekorpus

Varieeruvate nähtuste uurimiseks otsustasime salvestada keelejuhtide ette loetud lausetest koosneva kõnekorpuse. Ettlugemist oleme nimetanud ka lugemiseksperimendiks ning lausete lugejaid eksperimendis osalejaks või keelejuhtideks.

Etteloetud teksti salvestamine osutus kõige otstarbekamaks materjali-kogumisviisiks: tähenduslikult ja süntaktiliselt terviklikud laused loovad häälduse uurimiseks kontrollitud raamistiku ja võimaldavad koguda kompaktsemalt rohkem uurimismaterjali. Erinevused (hääldus, tempo) tulevad esile isikuti. Ettlugemine on erinevalt spontaankõnest lähedane ka tekst-kõne-sünteesile, milles soovime uuringu tulemusi rakendada. Teisalt ei saa jätta arvestamata teksti ettlugemise mõju. Salvestatav ettlugemine on ametlik esinemissituatsioon, mis võib mõjutada keelejuhti hääldama teistsuguselt, kui tal loomulikumas situatsioonis tavaks on.

Kolme aasta (2015–2017) jooksul salvestasime kõnekorpuse jaoks kokku 191 keelejuhi kõnet, millest 2021. aastaks on märgendatud 101 keelejuhi salvestised. Korpus on kasutatav teadusuuringuteks (vt Kalvik, Piits 2021).

1.1. Lausete koostamine

Otsustasime, et mitme varieeruva hääldusnähtuse uurimiseks sobib kõige paremini ette loetav lausestik, kuhu on peidetud nähtusi esindavad sihtsõnad. Lausestik asub artikli lisan. Selles on käesoleva ülevaate jaoks varieeruva hääldusega sõnad paksu kirjaga esile tõstetud ning neile on lisatud nähtusele vastav indeks.

Sõnaalguline *h*. Lisaks sellele, kuivõrd *h*-d lugedes hääldatakse, soovisime teada saada, kas teda mõjutavad sõna esinemissagedus ning tähendust eristav roll. Moodustasime paarid, kus esimene sõna on suure ja teine väikese esinemissagedusega ehk kus olid vastavalt sage ja harv sõna,

nt *hakkas-hakke, hing-hingus, hästi-hädine, homme-homaar*. Andmed sõna esinemissageduse kohta on pärit ilu- ja ajakirjanduskorpusel põhinevast sagedussõnastikust (Kaalep, Muischnek 2002). Sage sõna esineb korpuses 88–2178 korral, harv 0–15 korral. Lisasime ka sõnapaare, kus *h* eristab tähendust, nt *hall-all, hea-ea, hõng-õng* ja *hind-ind*, et välja selgitada, kas vastanduse olemasolul hääldatakse *h*-d tihemini ja pikemalt.

Palatalisatsioon (CC)VC-struktuuriga ühesilbilistes *i*-tüvelistes sõnades. Et uurida, kui sageli palatalisatsiooni seda tüüpi sõnades esineb, valisime lausetesse erineva konsonandiga (*l, n, s, t/d*) lõppevaid ja erineva pika vokaaliga (*aa, oo, uu, ee, öö* või *ää*) sõnu. Moodustasime siingi paarid sagedast ja harvast sõnast, et välja selgitada, kas harvem esinevat ja ilmselt keelejuhile tundmatumat sõna pigem palataliseeritakse või mitte. Paarikutele lisasime sarnase struktuuriga, ent teise tüvevokaaliga sõna, et akustiliseks uurimiseks oleks võrdlusmaterjali. Nii tekkisid kolmikud, nt *saal, raal* ja *vaal*; *veen, jeen* ja *seen*; *roos, oos* ja *soos*; *pruut, nuut* ja *ruut*. Osa *i*-tüvelisi sõnu moodustab teiste sõnade või sõnavormidega minimaalpaare, nt *kool' 'õppeasutus' vs. kool' 'ila'*; *luus' vs. luus (luu* ains seesütl). Soovisime nende abil kontrollida, kas selline vastandusvõimalus võib avaldada palatalisatsiooni esinemisele mõju.

Varieeruva vältega sõnad. ÕS 2018 sisaldab 395 märksõna, mida saab hääldata nii teises kui ka kolmandas vältes. Sageli kordub varieeruv tüvi erinevates liitsõnades (nt *mõisa|härä, -proua, -orjus, -sundiija, -tegu, -valitseja*) või tuletistes (nt *andekas, andekus, andetu, andetus*). Kordumatuid tüvesid on neist ligi kaks kolmandikku ehk 245. Et maksimeerida uurimusest saadavat praktilist kasu, valisime neist lugemiskatsesse 51 sagedamat. Valitud sõnad on erineva silbistruktuuriga ja kuuluvad eri muuttüüpidesse, sõnaliigiti esinevad muutumatud (nt *ammu, mullu*), omadus- (nt *kärme, hirmus, kaitsetu, andekas, rahvuslik*) ja nimisõnad (nt *pealik, kangelane, aula*).

Salvestuste käigus selgusid mõned probleemid: raskesti hääldatavad sõnad (nt *homöopaat*), liiga palju uuritavaid sõnu ühes lauses, puuduolevad sõnad (nt *terve* nimisõnana). Seetõttu tegime esimeste lugemiskatsete jooksul lausetes väikseid parandusi, aga alates 38. salvestusest lugesid kõik ette täpselt samad 56 lauset (vt lisa).

1.2. Salvestamisprotsess

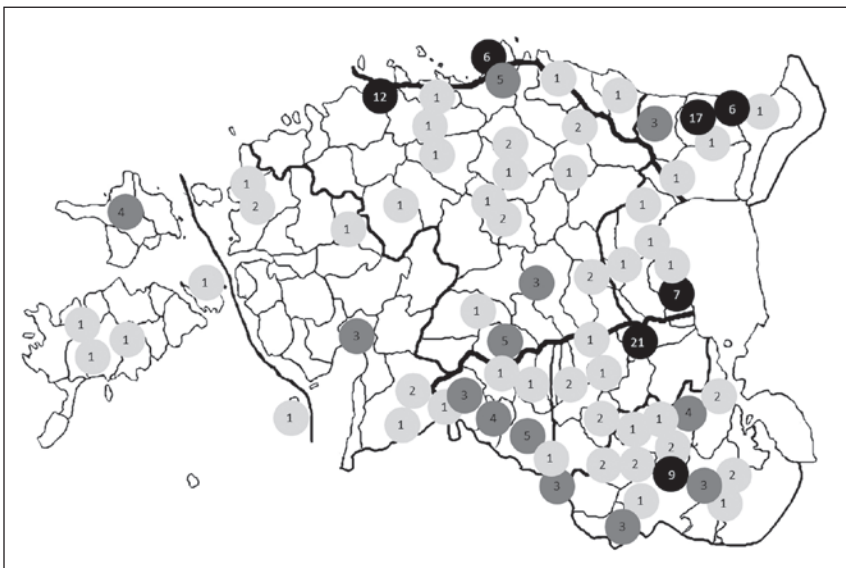
Lugemiseksperimentidega alustasime Eesti Keele Instituudi (EKI) helistuudios ning esimese 11 keelejuhi sõnade põhjal valmis ka esialgne uuring (vt Kalvik, Piits 2015). Selles selgus muu hulgas, et peamiselt Põhja-Eestist pärit lugejate kõnematerjali põhjal ei saa analüüsida nähtuste murdetasta. Seetõttu kavandasime edasised kogumisretked mujale Eestisse. Hilisemad salvestused toimusid Tartu Ülikooli foneetikalaboris ja Võru Instituudis, Alatskivil, Jõhvis, Karksi-Nuias, Kiviõlis, Palamusel, Toilas, Tõrvas, Meremäel ja Kolgas. Koduseid üksiksalvestusi tegime teiselgi.

Salvestused on tehtud kahel viisil: kas EKI helistuudios ja Tartu Ülikooli foneetikalabori stuudios (kokku 77) või väljaspool stuudioid (kokku 114). Stuudiosalvestused viisime läbi suletud ruumides, kasutades helipulti Mackie Onyx 1640 (Tallinnas), mikrofone Neumann TLM 102 (Tallinnas) ja Beyerdynamic MC 930 (Tartus), helikaarti Sound Devices USBPre (Tartus) ja programmi „Korpussalvestuste ja indekseerimise kiosk“ (KJIK). See on EKIs loodud programm, mille abil saab lauseid ükshaaval arvutiekraanile kuvada ja eraldi helifailidena salvestada. Väljaspool stuudiot kasutasime kas sülearvutit ja KJIK programmi või diktofoni Zoom H4 (nii mikrofoni kui helikaardina). Üldjuhul võttis salvestamine koos ankeeditäitmise ja lugemisjärgse mõttevahetusega aega keskmiselt 10 minutit.

1.3. Keelejuhid

Meie eesmärk oli salvestada võimalikult palju erinevaid keelejuhte ning kõik, kes soovisid lugemiseksperimentis osaleda, said salvestatud. Murdetasta võimaliku mõju uurimiseks tahtsime leida ka selliseid keelejuhte, kelle vanemad oleksid samuti pärit samast piirkonnast. Piirkonna all peame silmas ajaloolisi kohamurdealasid ehk saarte, lääne-, kesk-, ranna-, kirde- ja idamurdeala Põhja-Eestis ning Mulgi, Tartu ja Võru murdeala Lõuna-Eestis. Kõigi keelejuhtide kohta on ankeet, milles info soo, vanuse, hariduse ja elukoha kohta. Soovisime teada saada nii peamist elukohta esimese kümne eluaasta jooksul kui ka praegust elukohta, samuti vanemate peamist päritolupiirkonda. Keelejuhtide päritolu on näha alloleval joonisel 1.

Ankeedile lisasime keelejuhi jaoks loodud numbrikombinatsiooni, millel pole seost isikuandmetega. Ankeedis informeeriti lausete ettelugejat



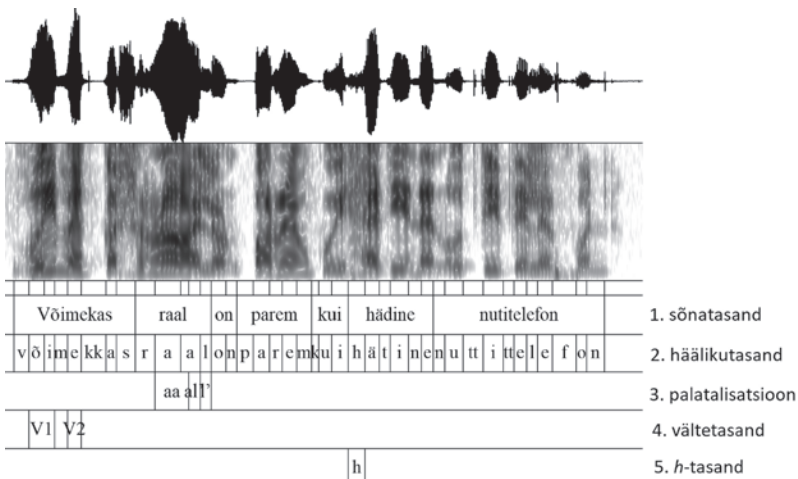
Joonis 1. Keelejuhtide päritolu. Number näitab, mitu keelejuhti on sellest piirkonnast pärit (mida tumedam ring, seda suurem arv keelejuhte)

ka sellest, et tema kõne salvestisi kasutatakse analüüsiks, kuulamistestides ja kõnekorpuste koosseisus.

191 keelejuhust 140 on naised ning 51 mehed. Naised on vanuses 18–82 (keskmine 50) ja mehed 20–87 aastat (keskmine 50). Enamik keelejuhte on kas kesk(eri)- või kõrgharidusega. Kokkuvõtlikult on korpuses enim esindatud kõrgharidusega naised. Haridusfaktorit me siiani tehtud uurimustes pole analüüsinud, aga info võib osutada kasulikuks edaspidi.

1.4. Materjalitöötlus

Ette loetud laused on salvestatud wav-formaadis. Kui kasutasime KJIK programmi, salvestus iga lause eraldi helifailina. Diktofoniga salvestatu lõikasime ise lauseti helifailideks, kasutades programme Cool Edit Pro 2.1 või Praat (vt Boersma & Weenink 2015). Siis segmenteerisime laused sõna- ja häälikutasandil eestikeelse kõne autosegmenteerijaga (vt Alumäe jt 2018). Kõigi analüüsitavate sõnade märgendused kontrollisime üle ning



Joonis 2. Näide 47. lause märgenduse ja segmentatsiooni kohta

vajadusel korrigeerisime segmendipiire. Lause-, sõna- ja häälikutasandi alla lõime programmis Praat iga nähtuse segmentide jaoks lisatasandi (vt joonis 2). Sõnaalgulise *h* märgendasime 5. tasandile, hääliku puudumisel jätsime küll segmendi alles, aga märgendasime selle kui *h0*. (CC) \bar{V} C-struktuuriga sõnades märgendasime 3. tasandile lõpukonsonandi, mis sai vastavalt kuuldelsele hinnangule vajadusel palatalisatsioonimärgi ('). Piiritlesime ka lõpukonsonandile eelneva pika vokaali, milles määrasime stabiilse osa ja siirde ehk ülemineku konsonandile. Varieeruva vältega sõnades märgendasime 4. tasandile sõna pearõhulise silbi tuuma (V1), kooda olemasolul kooda (C1) ning järgneva silbi tuuma (V2).

Seejärel kasutasime Praadi skripte, arvutamaks järgmisi andmeid.

1. Sõnaalgulise *h*-ga sõnades arvutasime, kui tihti *h* esineb (*h* hääldamise %) ning kui pikk ta esinemise korral on.
2. (CC) \bar{V} C-sõnades arvutasime, kui suure osa (%) moodustab siire lõpukonsonandile eelnevast vokaalist. Võrdlesime, kas kuuldeline palatalisatsioonihinnang langeb kokku siirde kestuse järgi arvutatud määranguga.
3. Vahelduva vältega sõnades arvutasime pearõhulise silbi ja sellele järgneva silbi kestussuhte põhjal sõna välte. Lisaks võrdlesime, mil määral kattus välte kohta antud kuuldeline hinnang kestussuhte põhjal arvutatud määranguga.

2. Tulemused

Esitame nähtuste kaupa kokkuvõtlikud tulemused, milleni oleme jõudnud viie aasta jooksul. Analüüsitava nähtuste eelnevat uurimist oleme kirjeldanud põhjalikult varasemas artiklis (Kalvik, Piits 2015: 50–56).

2.1. Sõnaalguline *h*

Sõna alguses asuva *h* küsimus on mitmetahuline. Juba ajalooliselt on sõnaalgulise *h* hääldamisel olnud olulisi murdelisi erinevusi (Asu jt 2016: 13) ja ka tänapäevases eesti keeles ei ole selle hääldamine nõutav (EKK 2020: 60). Argisuhtluses *h*-d sõna alguses eesti keele spontaanse kõne foneetilise korpuse² andmeil pigem ei hääldatagi (Asu jt 2016: 115).

Tegurid, mis kindlasti mõjutavad häälikut hoolikamalt välja ütleva, on ametlik olukord (Asu jt 2016: 115; Teras 2019: 212–213) ning ettelugemine (Cui 1999: 30–31; Kalvik, Piits 2015: 65). Varasemad uuringud (nt Cui 1999; Mesipuu 2007) on lisaks tähelepanu juhtinud *h* keelelisele ümbrusele ja sõna rõhulisuse mõjule ning ka tõestanud, et väikese esinemissagedusega sõnades hääldatakse *h*-d rohkem. Uusim uurimus (Teras 2019) tõstab esile sõna lauserõhulist positsiooni, mis mõjutab kõnelejat *h*-d rohkem välja hääldama.

Meid huvitasid sõnaalgulise *h* puhul järgmised uurimisküsimused.

1. Kui palju ettelugemisel *h*-d hääldatakse?
2. Kas suure esinemissagedusega sõnades jääb *h* rohkem hääldamata / on lühem kui väiksema esinemissagedusega sõnades?
3. Kas sõnades, kus *h*-l on tähendust eristav roll, hääldatakse teda rohkem?

Analüüsime oma uuringus (Kalvik, Piits 2019) 94 keelejuhi ette loetud sõnu, kokku 3945 hääldusjuhtu. Arvutasime kõigepealt üldise sõnaalgulise *h* esinemisprotsendi. Selgus, et 92% juhtudest hääldati *h* sõna alguses välja (Kalvik, Piits 2019: 76). Hilisem analüüs näitas, et kõige vähem hääldasid *h*-d sõna alguses keelejuht N_141 (12% hääldatud juhtudest) ja N_020 (45% juhtudest). Ülejäänud keelejuhid hääldasid sõnaalgulist *h*-d üle 69% juhtudest ja neist ligi veerand ehk 22 keelejuhti hääldas kõik *h*-d välja.

² <https://www.keel.ut.ee/et/foneetikakorpus>.

Lisaks arvutasime *h* hääldamisprotsendi ning keskmise kestuse neljas sõnarühmas:

- 1) sagedates sõnades: *hakkas, hing, hästi, hiljem, hindas, homme, hotelli, hunt, heida, haigus, hind, hammas, hoolimata, hirm*;
- 2) harvades sõnades: *hakke, hingus, hädine, hilpe, hingestatult, homaar, hotentoti, huupi, heldima, halastus, hindu, harras, hoomamata, hirss*;
- 3) sõnades, kus *h* eristab tähendusi: *hais-ais, hall-all, hirve* (ains om) - *irve, haru-aruu, harutama-arutama, hõng-õng, halastus-alastus, harva-arva* (käskiva kv ains 2. p), *hala-ala, hea-ea, hind-ind*. Sõnapaarid *hall-all, haru-aruu, harutama-arutama, hõng-õng* ja *hind-ind* paigutasime samadesse lausetesse, et tähendusvastandus oleks ilmsem (vt lisa, allajoonitud sõnad);
- 4) sõnades, kus *h* ei erista tähendusi. Siia kuuluvad kõik 1. ja 2. rühma sõnad, v.a *hind*.

Kokkuvõtlikult selgus sõnarühmade 1 ja 2 võrdlusest, et harvades sõnades jäi *h* hääldamata 7% ja sagedates 10% juhtudest ning see oli statistiliselt oluline erinevus ($t = 2,759$, $df = 2630$; $p = 0,0059$).³ Sagedamas sõnas oli *h* kestus lühem ja harvas pikem (keskmised vastavalt 68 ms ja 79 ms). Statistiliselt olulisele erinevusele viitab ka *t*-test: $t = 9,642$, $df = 2414$; $p < 0,0001$. Üksiksõnuti leidis aga lahknevusi: harvas sõnas *hilpe* oli *h* lühem kui tema sagedas paarikus *hiljem*, samuti paaris *hingestatult-hindas*. (Kalvik, Piits 2019: 78) Erandlikud tulemused näitavad, et on mitmeidki uurimata lisamõjureid, nt sõna pikkus, asend lauses, rõhk.

Sõnarühmade 3 ja 4 võrdlusest selgus, et sõnades, kus *h* eristab tähendusi, jäeti see hääldamata 5% hääldusjuhtudest. Sõnades, kus *h* ei erista tähendusi, jäeti *h* hääldamata 9% hääldusjuhtudest. *t*-test osutas, et tegu on olulise erinevusega: $t = 4,082$, $df = 3943$; $p < 0,0001$. Sõnaalgulise *h* kestuste võrdlusel selgus, et sõnades, kus *h* eristab tähendusi, oli *h* keskmine kestus 77 ms, ja sõnades, kus tähenduseristus puudus, oli *h* keskmine kestus 73 ms. Kuigi statistiline erinevus oli *h* kestustes olemas ($t = 2,432$, $df = 3631$; $p = 0,0076$), ei olnud millisekundites mõõdetav erinevus nii suur, et saaks väita, nagu tähenduste eristamise vajadus tingiks *h* kestuse olulise pikenemise. (Kalvik, Piits 2019: 82) Need tulemused ei kinnita,

³ Rühmadevaheliste statistiliste erinevuste testimiseks rakendasime sõltumatute valimite *t*-testi, kasutades selleks Exceli ja StatPaci3 vahendeid.

et *h*-d hääldatakse pikemalt, kui on vaja tähendust eristada. Küll aga on kindel, et sellisel juhul hääldatakse *h* tihemini välja.

2.2. Palatalisatsioon

Konsonantide *l*, *n*, *s*, *t/d* palatalisatsioon on ühesilbilistes *i*-tüvelistes pika vokaaliga sõnades, nagu näiteks *kool*, *veen*, *kruus*, *vööt*, *ääs* ja *laad*, olnud varieeruv (Laugaste 1956: 77). Uuritavat tüüpi sõnu on uuemal ajal käsitlenud Teras ja Pajusalu (2014), kes kinnitasid varieeruvust ka akustilise uuringuga.

Meid huvitasid selliste sõnade puhul järgmised uurimisküsimused.

1. Kui sageli lõpukonsonanti palataliseeritakse?
2. Kas palatalisatsiooni esinemine on seotud kõneleja päritolupiirkonnaga?
3. Kas sageli esinevates ja lugejale eeldatavasti tuttavamates sõnades palataliseeritakse rohkem kui harvades?
4. Kas sõnu, kus palatalisatsioon eristab tähendusi, palataliseeritakse tihemini?

Palatalisatsiooniuringusse kaasasime vaid 40 keelejuhi kõnematerjali, sest tahtsime, et keelejuhtide päritolupiirkondadena oleks võrdselt esindatud kõik kunagised murdealad. Kuna kõigist piirkondadest polnud võimalik leida üle nelja keelejuhi, kes ise ja kelle vanemad oleks pärit samalt murdealalt, siis valisime iga kunagist murdeala esindama neli keelejuhti. Uuritud sihtsõnu oli kokku 36 (vt tabel 1), hääldusjuhte kogunes kokku 1434. Sarnase struktuuriga, ent teise tüvevokaaliga kontrollsõnad lisasime materjali selleks, et oleks võimalik akustilisi tunnuseid paremini võrrelda. Palatalisatsiooni määramiseks analüüsisime palatalisatsiooni akustilisi tunnuseid, mõõtes siht- ja kontrollsõnade vokaalis esineva siirde kestust, kuid kasutasime ka kuuldelist määrangut (Piits, Kalvik 2019). Leidsime, et kuigi kuuldeline määrang oli üsna subjektiivne, ei osutunud ka siirde kestus eraldiseisva tunnusena usaldatavaks palatalisatsiooni tuvastajaks. Eri viisil määratud tulemuste kokkulangevus oli vaid 64% juhtudest (Piits, Kalvik 2019: 521). Käesolevas artiklis anname ülevaate tulemustest, mis põhinevad kuuldeliselt määratud andmetel.

Tabel 1. Sagedad ja harvad *i*-tüvelised sihtsõnad ning sama struktuuriga muutüvelised kontrollsõnad⁴

Vokaal	Konsonant	<i>i</i> -tüveline sihtsõna		Muutüveline kontrollsõna
		sage	harv	
<i>aa</i>	l	<i>saal</i>	<i>raal</i>	<i>vaal</i> (om <i>vaala</i>)
<i>aa</i>	n	<i>plaan</i>	<i>paan</i>	<i>saan</i> (kindla kv ol ains l. p)
<i>aa</i>	s	<i>vaas</i>	<i>paas</i>	<i>raas</i> (om <i>raasu</i>)
<i>aa</i>	d	<i>laad</i>	<i>raad</i>	<i>maad</i> (<i>maa</i> ains os)
<i>aa</i>	t	<i>paat</i>	<i>raat</i>	<i>laat</i> (om <i>laada</i>)
<i>ee</i>	l	<i>geel</i>		<i>keel</i> (om <i>keele</i>)
<i>ee</i>	n	<i>veen</i>	<i>jeen</i>	<i>seen</i> (om <i>seene</i>)
<i>ee</i>	s	<i>tees</i>		<i>mees</i> (om <i>mehe</i>)
<i>ee</i>	t	<i>peet</i>	<i>neet</i>	<i>keet</i> (om <i>keedu</i>)
<i>oo</i>	l	<i>kool</i>	<i>mool</i>	<i>sool</i> (om <i>soola</i>)
<i>oo</i>	n	<i>toon</i>	<i>foon</i>	<i>joon</i> (om <i>joone</i>)
<i>oo</i>	s	<i>roos</i>	<i>oos</i>	<i>soos</i> (<i>soo</i> ains seesütl)
<i>oo</i>	d	<i>kood</i>		<i>rood</i> (om <i>roodu</i>)
<i>oo</i>	t	<i>noot</i>	<i>toot</i>	<i>noot</i> (om <i>nooda</i>)
<i>uu</i>	l	<i>kuul</i>	<i>pruul</i>	<i>tuul</i> (om <i>tuule</i>)
<i>uu</i>	n	<i>pruun</i>	<i>uun</i>	<i>ruun</i> (om <i>ruuna</i>)
<i>uu</i>	s	<i>kruus</i>	<i>luus</i>	<i>puus</i> (om <i>puusa</i>)
<i>uu</i>	t	<i>pruut</i>	<i>nuut</i>	<i>ruut</i> (om <i>ruudu</i>)
<i>ää</i>	n	<i>lään</i>		<i>vään</i> (om <i>väänu</i>)
<i>ää</i>	s	<i>ääs</i>		<i>lääs</i> (om <i>lääne</i>)
<i>öö</i>	t	<i>vööt</i>		<i>sööt</i> (om <i>sööda</i>)

Palatalisatsiooni esines meie analüüsis kuuldelse hinnangu järgi 29% (Piits, Kalvik 2019: 522). Tulemus on kooskõlas eesti keele spontaanse kõne foneetilise korpuse alusel saadud näiduga, kus VVC sõnu palataliseeriti 30% (Teras, Pajusalu 2014: 263). Järelikult palataliseeritakse tänapäeval selles sõnatüübis lõpukonsonanti vähe. Täpsemalt oli meie uurimuses vaid viis sõna, mida enam kui pool keelejuhtidest hääldas

⁴ Mõnel juhul (*geel*, *tees* jt) sageduse alusel vastanduvad paarid puudusid. Paksus kirjas on sõnad, kus palatalisatsioon võib eristada tähendusi.

palataliseeritult: *i*-tüvelised *neet*, *tees* ja *paat* ning muutüvelised *keel* ja *keet*. Sõna *keel* võis palatalisatsiooni saada seetõttu, et nähtust on keskmurdes sellistes *e*-tüvelistes sõnades ajuti esinenud (Must, Univere 2002: 98); sõna *keet* (om *keedu*) puhul oletame seost sõnaga *keedis*, mis võis suunata palataliseerima. Välistatud pole seegi, et sõnas *tees* on palatalisatsiooni tinginud hoopis lausekontekst: lausete koostamisel jäi tähelepanuta, et sõnale järgneb lauses sidend *ja*, mis võis palatalisatsiooni esile kutsuda. Konsonantidest (*l*, *n*, *s*, *t/d*) osutus enim palataliseerituks fortisklusiil *t* (nt sõnas *neet* 70% juhtudest), vähim leenisklusiil *d* (15% juhtudest). Suur erinevus *t* ja *d* palatalisatsioonis võis tuleneda häälikute kestuserinevusest: *d* keskmine kestus oli 85 ms (standardhälve 25), *t*-l aga 182 ms (standardhälve 69). Kui hääliku kestus on pikem, siis on ka palatalisatsiooni avaldumiseks ilmselt rohkem aega. Vokaalidest palataliseeriti enim *ee* järel (49%), vähim *ää* järel (4%). (Piits, Kalvik 2019: 522–524)

Üks uurimuse hüpoteese oli, et palatalisatsiooni (CC) \bar{V} C-sõnades võib seostada keelejuhi päritolupiirkonnaga. Oletasime, et rohkem palataliseerivad Lõuna-Eesti päritolu inimesed, vähem põhjaeestlased. Ajalooliselt on Tartu ja Mulgi murdes vaadeldavas tüübis konsonant olnud tavaliselt palataliseeritud (Tanning 1961: 26; Keem 1970: 41), väga järjekindel on nähtus olnud Võru murdealal (Keem, Käsi 2002: 32). Valisime iga päritolupiirkonna kohta neli keelejuhti, kes ise ja kelle vanemad on pärit samalt ajalooliselt murdealalt. Palatalisatsiooni esinemisprotsendid murdealuti olid järgmised: Võru (v.a Setu) 42%, Setu murrakurühm 42%, rannamurre 42%, Mulgi (39%), Tartu (32%), kirdemurre (30%), läänemurre 25%, idamurre 24%, keskmurre 11%, saarte murre 3%. Keelejuhiti leidsime, et enamik palataliseeris sõnu vähem kui pooltel juhtudel ja ainult kaheksa keelejuhti palataliseeris rohkem (53–69%). Üldse ei esinenud palatalisatsiooni Muhust pärit keelejuhi sõnades, kõige rohkem (69%) palataliseerisid Meremäe ja Kolga keelejuht. (Piits, Kalvik 2019: 525–529)

Võib öelda, et tulemused on kunagise murde pärased saarlaste puhul. Võrukestel ja setudel esineb palatalisatsiooni küll uuringu lõikes enim, kuid märksa vähem, kui ajaloolise murdetausta järgi võiks oletada. Nt kui tänapäeval on Võru murdealalt pärit inimeste keeles VVC sõnades palatalisatsiooni 42% juhtudest, siis varem oli murdekeeles seda 86% (Pöld 2016: 29). Sama palju (42%) ehk eeldatust rohkem on palatalisatsiooni ka meie rannamurdealaga seotud keelejuhtidel; varem on seal nähtust esinenud 17% (Pöld 2016: 29). Kokkuvõtlikult võib väita, et kunagisest

murdepärasusest on häälduspildis jälgi ning päritolupiirkonda ei saa mõjutajana kõrvale jätta, kuigi üldiselt näib, et palatalisatsiooni piirkondlik esinemus selles sõnatüübis on ühtlustumas.

Uurisime ka esinemissagedust ning tähenduse eristamise vajadust kui palataliseerimise mõjutajaid. Jagasime sihtsõnad esmalt sageli (*saal, plaan, vaas, laad, paat, veen, peet, kool, toon, roos, noot, kuul, pruun, kruus, pruut*) ja harva esinevaiks (*raal, paan, paas, raad, raat, jeen, neet, mool, foon, oos, toot, pruul, uun, luus, nuut*). Sõnarühmade kõrvutamisel selgus, et sagedaid sõnu palataliseeriti veidi rohkem kui harvu. Kõigist palataliseeritud sõnadest oli 45% harva esinevaid ja 55% sageli esinevaid sõnu. Sõnapaaride (nt *saal-raal, kool-mool*) lähem vaatlus näitas, et sõna esinemissagedus mõjutab hääldust vähe: kui üht paarilist hääldati palataliseeritult, oli see vaid veidi suurema tõenäosusega sage sõna. (Kalvik, Piits 2019: 79–80)

Tähenduse eristamise vajadus võib meie materjalis olla nimetavas käändes sõnapaarides *noot* 'noodikirjamärk' (vs. *noot* 'kalapüügivahend'), *kruus* 'tass' (vs. *kruus* 'killustik') ja *kool* 'õppeasutus' (vs. *kool* 'ila'). Samamoodi võib selline vajadus olla erinevate sõnavormide korral sõnapaarides *kuul* 'kerakujuline ese' (vs. *kuul, kuu* ains alalütl), *kood* 'kokkuleppeline märk' (vs. *kood, kuduma* kindla kv ol ains 2. p), *tees* (vs. *tees, tee* ains seesütl), *toon* 'värvitoon' (vs. *toon, tooma* kindla kv ol ains 1. p) ja *luus* (vs. *luus, huu* ains seesütl). Vastandus on mõtteline, sest ühe erandiga esines lausetes neist minimaalpaaridest vaid esimene paarik. Seesütleval käändel põhinevaid vastandusi on veelgi, nt *roos* (vs. *roos* 'pilliroos'), *ääs* (vs. *ää-s, ä*-häälikus'), samuti osastaval põhinev *laad* (vs. *la*'d), kuid neid me uuringusse ei kaasanud. Sõnarühmade võrdlustest selgus, et sõnu, kus palatalisatsioon eristas tähendusi, palataliseeriti veidi enam (keskmiselt 37% juhtudest) kui sõnu, kus tähenduseristus puudus (keskmiselt 27% juhtudest); tegu oli olulise erinevusega: $t = 3,469$, $df = 1438$; $p = 0,0005$ (Kalvik, Piits 2019: 83–84). Seega esines suundumus suurema esinemissagedusega sõnu palataliseerida ning palatalisatsiooni esines rohkem juhtudel, kus tal oli tähendust eristav roll.

2.3. Välte

Varieeruva vältega sõnu oleme põhjalikumalt käsitlenud 2017. aastal kahes artiklis. Esimeses artiklis (Piits, Kalvik 2017) vaatlesime 23 keelejuhi andmete põhjal, millised on nende sõnade hääldustrendid, hinnates kuuldeliselt sõnade väldeid ja mõttes silbipikkuste suhet. Teises artiklis (Kalvik, Piits 2017) uurisime välte määramise probleeme ja leidsime, et kõigist kuuldelistest vältetüüpidest muutus teistkordsel kuulamisel 12%.

Siinses artiklis jätame kõrvale määramisküsimused ja teeme kokkuvõtte varieeruva vältega sõnade hääldustrendidest 101 keelejuhi materjali põhjal, tuginedes kuuldelisele hinnangule. Kokku on analüüsitud 49 sõna 5045 hääldusjuhtu. Kuna praktiliseks eesmärgiks on välja selgitada, mis vältes neid sõnu enamasti kasutatakse, püüdsime sõnu rühmitada, et uuringusse kaasatud sõnade põhjal saaks võimalusel järeldusi teha ka teiste sama tüüpi sõnade kohta.

Meid huvitasid varieeruva vältega sõnade puhul järgmised uurimisküsimused.

1. Mis vältes mingit sõna valdavalt kasutatakse?
2. Millised on hääldustrendid rühmade kaupa?

Peeter Päll (1986) on põhjalikult kirjeldanud välte sõltumist käändsõnade silbistruktuurist. Varieeruva vältega sõnade rühmitamisel olemegi lähtunud Pälli üld- ja alltüüpidest, mille loomisel on arvestatud sõna nimetava vormi silpide arvu, silbiliiki (pikk kinnine, pikk lahtine vms), sõnaliiki ja sõnamoodustust. Sel viisil jagunevad uuritavad sõnad kaheksasse tüüpi (vt allolev tabel 2).

Et näha, kas ja kuidas hääldustrendid ajas muutuvad, jagasime keelejuhid kahte võrdse suurusega vanusegruppi, võttes piiriks välteuuringusse kaasatud keelejuhtide mediaanvanuse (44 a). Seega on nooremas grupis 50 keelejuhti vanuses 18–43 aastat (keskmine vanus 35) ja vanemas grupis 50 keelejuhti vanuses 45–83 aastat (keskmine vanus 58).

Tabel 2. Varieeruva vältega sõnade valdav välde tüüpide kaupa

	Tüüp	Hääl- juhtude arv	Nooremate grupi 2. vältes sõnade %	Vanemate grupi 2. vältes sõnade %	Kokku 2. vältes sõnade %	Val- dav välde
toimekas	1	101	100	92	96	2
andekas	1	101	98	84	91	2
võimekas	1	100	92	82	87	2
maitsekas	1	101	94	72	83	2
võimetu	1	101	88	76	82	2
ilmetu	1	100	74	63	69	2
kaitsetu	1	100	67	42	55	2
ilma	2	99	96	94	95	2
mullu	2	101	96	92	94	2
hilja	2	101	92	66	78	2
eile	2	101	62	58	59	2
enne	2	101	40	42	41	3
ammu	2	101	26	32	29	3
kiire (subs)	3	101	74	78	75	2
kiire (adj)	3	65	41	60	51	2
terve (subs)	3	65	28	43	35	3
terve (adj)	3	101	16	32	24	3
kaine	3	101	16	42	29	3
ahne	3	100	14	31	22	3
ihne	3	100	18	27	22	3
kärme	3	101	16	28	22	3
aula	3	101	10	30	20	3
täiuslik	4	101	88	76	82	2
keiserlik	4	99	56	48	53	2
teaduslikes	4	101	28	50	39	3
nõiduslik	4	100	18	39	28	3
nooruslik	4	99	22	22	22	3
haiguslik	4	101	10	22	16	3
looduslik	4	101	8	20	14	3

äärmuslikult	4	101	8	10	9	3
rahvuslik	4	100	2	12	7	3
saatuslik	4	101	4	4	4	3
saatanlik	4	101	2	4	3	3
rõõmus	5	101	92	86	89	2
soodus	5	101	90	76	83	2
hirmus	5	101	34	40	37	3
kuulus	5	101	8	10	9	3
politseinik	6	86	89	74	83	2
augustis	6	100	78	73	76	2
ümbrikus	6	101	72	52	61	2
piknikuga	6	100	30	31	30	3
pealik	6	101	18	24	21	3
peenelt	6	101	6	10	8	3
tallinlane	7	101	96	80	88	2
ungarlane	7	101	74	62	68	2
kangelase	7	101	6	10	8	3
teisal	8	100	94	88	90	2
meetod	8	101	50	32	41	3
alles		101	98	86	92	2
kirju		101	86	88	87	2
jaanuar		98	8	15	11	3

Nelja esimese tüübi puhul saab välja tuua tüübile iseloomuliku välte-eelistuse, nelja ülejäänud puhul jääb see selgusetuks. Kahes esimeses tüübis eelistati ainult või pigem teist vältet. Esimesse tüüpi kuuluvad kolmesilbilised *kas-* ja *tu-*liitelised tuletised *mõte*-tüüpi sõnadest, kus tuletusliide liitub omastavas käändes kolmandavältelelise tüvele: *toimekas*, *andekas*, *võimekas*, *maitsekas*, *võimetu*, *ilmetu*, *kaitsetu*. Need sõnad on moodustanud varem kolmandavälteleised alltüübid, aga juba varasemad uuringud on näidanud, et üldkeeles on tugev tendents hääldada neid kolmanda välte asemel teisevälteleisena (Päll 1986: 33). Põhjuseks on pakutud nii sagedamate teisevälteleiste *kas*-liiteliste sõnade eeskujul kui ka kinnist teist silpi (Vare 1984: 20, 21). Kõigi selliste meie uuringusse kaasatud

sõnade puhul oli valdav teisevälteline hääldus. Sõnal *kaitsetu* ei olnud teisevälteliste hääldusjuhtude osakaal (55% kõigist hääldusjuhtudest) nii suur kui teistel sõnadel, aga märgatav oli tendents, et nooremas vanusegrupis oli teisevältelise häälduse protsent suurem kui vanemaealistel (67% vs. 42%). Seega võib selle rühma sõnade puhul eeldada, et ka kõik teised seda tüüpi varieeruva vältega sõnad (nt *ulmekas, ilmekas, mainekas*) on samamoodi pigem teisevältelised.

Teise tüüpi kuuluvad kahesilbilised pika esisilbiga ja lühivokaaliga lõppevad sõnad *eile, hilja, ilma, mullu, ammu* ja *enne*, mis peaksid Pälli käändsõnadele loodud *lahja*-üldtüübi eeskujul saama teise välte. Meie tulemuste järgi oligi esimese nelja sõna puhul valdav teine välde, kuid sõnu *ammu* ja *enne* hääldati pigem kolmandas vältes.

Kahes tüübis (kolmas ja neljas) eelistati pigem kolmandat väldet. Kolmandas tüübis on kahesilbilised *e*-lõpulised adjektiivid *kaine, terve, ahne, ihne, kärke* ja *kiire*, mis Pälli järgi kuuluvad *lahja*-üldtüübi *raske*-alattüüpi ja peaksid saama kolmanda välte. *e*-lõpulised adjektiivid on häälikuajaloolistel põhjustel olnud kunagi pigem teisevältelised (Piits, Kalvik 2017: 133), aga aja jooksul on sõnad muutunud suurema ja produktiivsema kolmandaväلتeliste rühma (nt *raske, julge, uhke*) mõjul kolmandaväلتeliseks. Hääldusuuring näitas, et kõik sõnad (v.a *kiire*) kuuluvadki kolmandas vältes häälduvate sõnade hulka, ja võrdlus vanusegrupiti osutab, et trend on eriti tugev noorematel. Sõna *kiire* jääb aga nii adjektiivu kui substantiivina teiseväلتeliseks. Kuigi nooremas vanusegrupis oli *kiire* adjektiivne kasutus pigem kolmandaväلتeline (teises vältes vaid 41%), siis *kiire* substantiivne kasutus jäi kindalt teiseväلتeliseks. Varem ongi arvatud, et kui *e*-lõpulised adjektiivid liiguvad kolmandaväلتeliste sõnade hulka, võivad substantiivid jääda teiseväلتeliseks (vt rohkem Piits, Kalvik 2017: 133). *Lahja*-üldtüübi substantiivide hulka kuuluv *aula* osutus kolmandaväلتeliseks nagu *la*-liiteliste tuletiste alltüübi sõnadele kohane.

Neljandasse tüüpi kuuluvad kolmesilbilised *lik*-liitelised adjektiivid *haiguslik, looduslik, nooruslik, nõiduslik, rahvuslik, saatanlik, saatuslik, teaduslik, äärmuslik, keiserlik* ja *täiuslik*. Pälli (1986: 38) sõnul moodustavad *lik*-liitelised sõnad *veerandik*-tüübi kolmandaväلتelise alltüübi. Meie uuring näitas, et üldjuhul ongi neis kolmandaväلتelistest sõnadest tuletatud sõnades valdav kolmas välde, ainult sõnad *keiserlik* ja *täiuslik* on pigem teiseväلتelised. Mõlema sõna puhul oli teiseväلتelisuus nooremas

vanusegrupis suurem kui vanemas. Vanemas vanusegrupis oli *keiserlik* veel valdavalt kolmandavärtliline.

Ülejäänud neljas tüübis või sõnades, mis rühmi ei moodustanud, ei esinenud ühtset tendentsi ning valdavalt teises ja valdavalt kolmandas vältes sõnu leidus võrdset. Viiendasse tüüpi kuuluvad kahesilbilised *us*-lõpulised adjektiivid *hirmus*, *kuulus*, *soodus* ja *rõõmus*. Kuigi *us* > *sa* lõpuvaheldusega adjektiividest on Pälli (1986: 27) sõnul suurem osa kolmandavärtlised, on siiski ka varasemaid tähelepanekuid, et *us*-lõpuliste adjektiivide välde on kõikuv (Piits, Kalvik 2017: 134). Meie uurimistulemus kinnitab, et mingit kindlat trendi selles grupis ei esine. Sõnu *hirmus* ja *kuulus* hääldati valdavalt kolmandas vältes ning sõnu *soodus* ja *rõõmus* valdavalt teises vältes.

Kuuendasse tüüpi kuuluvad kahesilbilised pika konsonandi või konsonantühendiga lõppevad sõnad *pealik*, *piknik* ja *ümbrik*, *august*, *peenelt* ja *politseinik*⁵. Silbistruktuuri arvestades peaksid kõik need sõnad kuuluma *õpik*-üldtüüpi ja saama teise välte, aga *lik-*, *nik-* ja *rik*-liitelised sõnad moodustavad eraldi alltüübid. Neis sõnades võib esineda kaks kõrvutist rõhusilpi, mis muudab tüübi ebastabiilseks (Päll 1986: 29). Kindlat välte-trendi selles tüübis ei esine: sõnu *pealik*, *peenelt*, *piknik* hääldati rohkem kolmandas vältes ja sõnu *august*, *ümbrik*, *politseinik* rohkem teises vältes.

Seitsmendasse tüüpi kuuluvad neljasilbilised *lane*-liitelised substantiivid *kangelane*, *tallinlane* ja *ungarlane*. Pälli (1986: 46) andmetel võiks sõna silbistruktuuri arvestades hääldada neid nii teises kui kolmandas vältes, aga siiski kuulub kolmveerand sellise struktuuriga sõnadest teisevärtlisesse tüüpidesse. Sama vahekord esineb ka meie uuritud sõnades, kus sõna *kangelane* hääldatakse peamiselt kolmandavärtlisena, aga sõnu *tallinlane* ja *ungarlane* teisevärtlisena.

Kaheksanda tüübi kahesilbilised lühivokaali ja -konsonandiga lõppevad käändsõnad (nt *meetod*) kuuluvad *noomen*-tüüpi ja peaksid olema üldjuhul kolmandas vältes, kuigi nõrga klusiiliga lõppevad sõnad (nt *talmud*, *hertsog* ja *keerub*) saavad ÕSis ka teise välte. Sõna *meetod* esines meie uuringus peamiselt kolmandas vältes. Samasse tüüpi võib mööndustega lisada ka adverbi *teisal*, mis veel 1976. aasta ÕSis ongi kolmandas vältes. Meie tulemus aga näitab, et tänapäeval hääldatakse sõna valdavalt teisevärtlisena.

⁵ Sõna *politseinik* kuulub *õpik*-tüüpi, kui pearõhk langeb kolmandale silbile.

Seega kui esimese nelja tüübi puhul võis märgata ühtsemat hääldustrendi, siis nelja viimase tüübi puhul esines peaaegu võrdselt teises ja kolmandas vältes hääldust ning ühtset trendi ei saanud välja tuua. Ühtse trendi puudumisel saab üksiksõnades leitud välte-eelistusi rakendada siiski sünteeskõnes. Samamoodi tegime kindlaks ka kolme eraldi sõna *alles*, *kirju* ja *jaanuar* hääldustava. Kaks esimest esinesid peamiselt teises vältes ja viimane kolmandas vältes.

3. Tulemuste kasutamine kõnesünteesis

Sõnaalguline *h* hääldatakse eesti keeles ette lugedes peaaegu alati välja: keskmiselt 92% juhtudest. Sageli esinevate sõnade puhul on *h* küll lühem ja seda hääldatakse veidi vähem, aga harvem esinevates sõnades ja sõnades, kus *h* eristab tähendusi, hääldatakse seda veelgi sagedamini. Seetõttu on kaheldav, et tekst-kõne-süntees muutuks sellest loomulikumaks, kui sünteeskõnes jätta 8% juhtudel sõna alguses *h* välja hääldamata. Samas ei saa unustada, et meie materjal on kogutud ettelugemisega, mis on tavalisest suhtlussituatsioonist mittedespontaanssem ja ametlikum olukord. Seetõttu võib oletada, et samasuguses ametlikus situatsioonis ning sama-laadse teksti korral (nagu nt rahvusringhäälingu uudistes) soovitakse *h*-d ka sünteeskõnes kuulda, kuid oleks ennatlik sarnast järeldust teha kõigi sünteeskõne kasutusolukordade kohta. Sõnaalgulise *h* hääldamata jätmine võiks sobida näiteks mõne vestlusroboti kõnesse, et rõhutada spontaansust ja mitteametlikku olukorda. Sel juhul võiks just sagedamates sõnades *h* olla lühem või hoopis välja hääldamata ning arusaadavuse huvides võiks *h* jätta alles nende sõnade algusesse, kus tal on selge tähendust eristav funktsioon.

Lõpukonsonandid ühesilbilistes *i*-tüvelistes pika vokaaliga sõnades on palataliseeritud vaid ligikaudu kolmandikul juhtudel. Tulemus, et (CC)VC-sõnade lõpukonsonanti eesti keeles enamasti ei palataliseerita, lahkneb normingutest, mida on esitanud õigekeelsussõnaraamatud ning mis praegugi on kirjas uusimas ÕSis (2018). Tundub, et sõnaraamatutes on juhitud sellest, et sarnase struktuuriga ja samasse muuttüüpi kuuluvates *i*-tüvelistes sõnades peaks lõpukonsonant olema palatalisatsioonimärgiga. Meie uuring aga näitab, et keskmiselt esineb selles sõnatüübis palatalisatsiooni harva: vaid kaheksa keelejuhti 40st palataliseeris sõnu rohkem kui 50%. Sellele vaatamata esineb ka üksikuid sõnu, mis üldisest

trendist erinevad ja mida siiski üle poole keelejuhtidest palataliseeris (nt *neet*, *tees* ja *paat*). Vaja oleks laiemat arutelu, millel peaks põhinema palatalisatsiooni märkimine kõnealuses sõnatüübis. Näib, et kui palataliseeritakse, siis kas teatud konsonante (nt *t*) või teatud vokaali järel (nt *e*), võib-olla ka lihtsalt teatud sõnu (nt *neet*, *tees* ja *paat*). Sõna sagedus ja vajadus tähendusi eristada tunduvad palatalisatsiooni esinemist mõningal määral mõjutavat, samuti sõltub palatalisatsiooni olemasolu sellest, kust kõneleja on pärit. Unustada ei saa asjaolu, et kui kõnelejatele, kes uuritud sõnatüübis ei palataliseeri, oleks palataliseeritud variant lihtsalt kõlalisel võõras, siis kõnelejatele, kes palataliseerivad, on palatalisatsioon sageli seotud ka tähendust eristava rolliga. Seega on esimestele palatalisatsioon selles sõnatüübis pelgalt kõlaküsimus, aga teistele seotud ka semantikaga, st see on neile juba palju olulisem nähtus kui ainult kõlaline.

Olenemata sellest, milliseid soovitusi peaks ÕS andma keelekasutajatele, saab kõnesünteesi tarbeks teha järelduse, et selles sõnatüübis lõpukonsonanti ühiskeelne sünteeshääli palataliseerima ei peaks. Samas oleks ühiskeelse kõnesünteesi kõrval muidugi võimalik luua piirkondlike aktsentidega sünteeshääli, mille palatalisatsioon nii kirjeldatud sõnatüübis kui ka muudel juhtudel oleks iseloomulik just selle piirkonna ajaloolisele keelekasutusele.

Varieeruva vältega sõnade uurimisel selgus, mis väldet keelejuhid valdavalt eelistasid. Mõne sõna puhul ei olnud vahe väga suur, nt *kiire* esines vaid 51% ja *keiserlik* vaid 53% keelejuhtidel teises vältes, aga ülejäänud sõnade puhul oli eelistus selgem. Seega õnnestus ligi 50 sõna puhul teha kindlaks, kumba väldet keelejuhid eelistasid. Keeruline on nende sõnade põhjal teha aga järeldusi kõigi u 400 sõna kohta, mida ÕS 2018 lubab hääldada nii teises kui ka kolmandas vältes, aga mitme tüübi puhul saab siiski arvestada hääldustrendi ja rakendada seda kõnesünteesis ka teistel samasse tüüpi kuuluvatel sõnadel. Näiteks peaksid kolmesilbilised *kas-* ja *tu-*liitelised tuletised *mõte*-tüüpi sõnadest ning kahesilbilised pika esisilbi ja lühivokaaliga lõppevad sõnad saama teise välte ning kahesilbilised *e-*lõpulised ja kolmesilbilised *lik-*liitelised adjektiivid kolmanda välte. Kuigi kolmes viimati nimetatud tüübis esines ka mõni sõna, mida enamasti hääldati tüübi trendile vastupidises vältes, on tüübis esineva trendi arvestamine sünteeskõnes siiski parem, kui praegune juhuslik teise või kolmanda välte rakendamine varieeruva vältega sõnades.

Kirjandus

- Alumäe, Tanel, Ottokar Tilk, Asadullah 2018.** Advanced rich transcription system for Estonian speech. – Human Language Technologies – The Baltic Perspective. Proceedings of the Eighth International Conference, Baltic HLT 2018. Ed. by Kadri Muischnek, Kaili Müürisep (= Frontiers in Artificial Intelligence and Applications 37.) Amsterdam: IOS Press. <http://doi.org/10.3233/978-1-61499-912-6-1>
- Asu, Eva Liina, Pärtel Lippus, Karl Pajusalu, Pire Teras 2016.** Eesti keele hääldus. (= Eesti keele varamu II.) Tartu: Tartu Ülikooli Kirjastus.
- Boersma, Paul, David Weenink 2021.** Praat: Doing Phonetics by Computer. <https://www.fon.hum.uva.nl/praat/> (02.05.2021).
- Cui, Kaili 1999.** Sõnaalguline *h* eesti keeles. Bakalaureusetöö käsikiri Tartu Ülikooli eesti ja üldkeeleteaduse instituudis.
- Eesti õigekeelsussõnaraamat ÕS 2006.** Koost. Tiiu Erelt, Tiina Leemets, Sirje Mäearu, Maire Raadik. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus.
- Eesti õigekeelsussõnaraamat ÕS 2013.** Koost. Tiiu Erelt, Tiina Leemets, Sirje Mäearu, Maire Raadik. Eesti Keele Instituut. Tallinn: Eesti Keele Sihtasutus.
- Eesti õigekeelsussõnaraamat ÕS 2018.** Koost. Tiiu Erelt, Tiina Leemets, Sirje Mäearu, Maire Raadik. Eesti Keele Instituut. Tallinn: EKSA.
- EKK 2020** = Mati Erelt, Tiiu Erelt, Kristiina Ross 2020. Eesti keele käsiraamat. Uuendatud väljaanne. Eesti Keele Instituut. Tallinn: EKSA.
- Hint, Mati 1968.** Ortoepia normeerimise probleeme. – Keel ja struktuur 2. Töid struktuurse ja matemaatilise lingvistika alalt. Tartu: Tartu Riiklik Ülikool, 2–123.
- Kaalep, Heiki-Jaan, Kadri Muischnek 2002.** Eesti kirjakeele sagedussõnastik. Tartu: Tartu Ülikooli Kirjastus.
- Kalvik, Mari-Liis, Liisi Piits 2015.** Lugemiseksperiment fonoloogilise varieerumise uurimiseks. – Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics 6 (3), 49–77. <https://doi.org/10.12697/jeful.2015.6.3.02>
- Kalvik, Mari-Liis, Liisi Piits 2017.** Varieeruva vältega sõnad: häälduseelistused ja määramisraskused. – Mäetagused 68, 83–100. https://doi.org/10.7592/MT2017.68.kalvik_piits
- Kalvik, Mari-Liis, Liisi Piits 2019.** Sõna esinemissagedus ja tähenduste eristamise vajadus häälduse mõjutajana. – Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics 10 (1), 71–88. <https://doi.org/10.12697/jeful.2019.10.1.04>
- Kalvik, Mari-Liis, Liisi Piits 2021.** Fonoloogilise varieerumise kõnekorpus. <https://doi.org/10.15155/3-00-0000-0000-0000-08971L>

- Keem, Hella 1970.** Tartu murde tekstid. Eesti murded III. Toim. Salme Nigol. Eesti NSV TA Keele ja Kirjanduse Instituut. Tallinn: Valgus.
- Keem, Hella, Inge Käsi 2002.** Võru murde tekstid. Eesti murded VI. Toim. Aldi Sepp. Eesti Keele Instituut. Tallinn: [Eesti Keele Sihtasutus].
- Laugaste, Gerda 1956.** Konsonantide palatalisatsioon eesti keeles. – TRÜ toimetised nr 43. Ajaloo-keeleteaduskonna töid. Tallinn: Eesti Riiklik Kirjastus, 74–85.
- Mesipuu, Margit 2007.** Sõnaalguline *h*. Magistritöö käsikiri Tartu Ülikooli eesti ja üldkeeleteaduse instituudis.
- Must, Mari, Aili Univere 2002.** Põhjaeesti keskmurre. Häälikulisi ja morfoloogilisi peajooni. Toim. Eevi Ross. (= Eesti Keele Instituudi toimetised 10.) Tallinn: Eesti Keele Sihtasutus.
- Piits, Liisi, Mari-Liis Kalvik 2017.** Varieeruva vältega sõnade hääldusuuringud kõnesünteesi teenistuses. – Eesti Rakenduslingvistika Ühingu aasta-raamat 13. Toim. Margit Langemets, Maria-Maren Linkgreim, Helle Metslang. Tallinn: Eesti Rakenduslingvistika Ühing, 123–140. <http://dx.doi.org/10.5128/ERYa13.08>
- Piits, Liisi, Mari-Liis Kalvik 2019.** Palatalisatsioonist ühesilbilistes *i*-tüvelistes pika vokaaliga sõnades. *Roos närtsis, sest vaas oli tühi*. – Keel ja Kirjandus 7, 513–533.
- Pöld, Enel 2016.** Pika vokaali järgse konsonandi hääldus *i*-tüvelistes sõnades eesti murretes. Bakalaureusetöö. Tartu: Tartu Ülikool. <https://dspace.ut.ee/handle/10062/51894> (02.05.2021).
- Päll, Peeter 1986.** Eesti noomeni silbistruktuur ja aktsent. Tallinn: Eesti NSV Teaduste Akadeemia ühiskonnateaduste osakond.
- Tanning, Salme 1961.** Mulgi murdetekstid. Eesti murded I. Toim. Aili Univere. Tallinn: Eesti Riiklik Kirjastus.
- Teras, Pire 2019.** Sõnaalgulise /h/ hääldus samadel kõnelejalatel avalikus ja argisuhtluses. – Eesti ja soome-ugri keeleteaduse ajakiri. *Journal of Estonian and Finno-Ugric Linguistics* 10 (1), 211–231. <https://doi.org/10.12697/jeful.2019.10.1.11>
- Teras, Pire, Karl Pajusalu 2014.** Palatalisatsioonist ja prepalatalisatsioonist spontaanses eesti keeles. – Keel ja Kirjandus 4, lk 257–269.
- Vare, Silvi 1984.** Omadussõnaliited tänapäeva eesti kirjakeeles. Tallinn: Valgus.
- Õigekeelsussõnaraamat 1976.** Toim. Rein Kull, Erich Raiet. Tallinn: Valgus.

Lisa

Sihtsõnad on siinses lisas paksus kirjas. Ülaindeks 1 märgib sõnaalgulise *h*-ga sihtsõnu ning võrdluseks lisatud *h*-ta kontrollsõnu (allajoonitud), 2 märgib varieeruva palatalisatsiooniga *i*-tüvelisi (CC)VC-sihtsõnu ja võrdluseks lisatud muutüvelisi kontrollsõnu, 3 varieeruva vältega sihtsõnu. Esimesed kolm lauset on sissejuhatavad, ettelugemisega kohanemiseks, neid ei analüüsitud.

1. Merel möllas vali tuul ja paat kiikus kõvasti.
2. Hommik jõudis kätte, tuli avada plaan ja hakata arutama.
3. See politseinik on kangelane, ta tegutsemine on kiire ja täiuslik.
4. **Alles**³ olin ma **terve**³ inimene, täna valutavad nii parem **puus**² kui **hammas**¹.
5. **Hunt**¹ jooksis **soos**² selles suunas, kus kuulis **hirve**¹ vaikset astumist.
6. Jäta **hala**¹, probleemidest leidub pääs ning elu on jälle **hea**¹.
7. Matkajad läksid **huupi**¹, kuulamata **maad**², kuspool imeline **oos**² asub.
8. **Keiserlik**³ käsk kõlas nõnda, et **rood**² sõdureid kaitsku **homme**¹ paleed.
9. Üks **raat**² kükitab puu ladvas, välimuselt **ilmetu**³, ent laul on tal **harras**¹.
10. Kas **mullu**³ oli meie rühmas **ungarlane**³ või **hindu**¹?
11. Linnas möllas **saatanlik**³ **haigus**³ ja **raad**² oli **võimetu**³ midagi ette võtma.
12. **Jaauar**³ on plaanitud selleks, et välja mõelda uurimuse **tees**² ja **meetod**³ ning valida materjal.
13. Vere parandamiseks oleks **täiuslik**³ süüa üks **peet**² päevas.
14. **Toimekas**³ sepp tagus **kangelase**³ mõõka, kuni **ääs**² hõõgus.
15. **Hirss**¹, vesi ja **sool**² – selline **keet**² viib keele alla.
16. Jäta **nuut**² rahule, vii parem loomale **sööt**² ette ning **heida**¹ sõnnik välja.
17. Eksitus oli **saatuslik**³, **hoolimata**¹ sellest, et **kood**² osutus õigeks.
18. Pudenes veel üks **neet**² ning rõiva allärest pääses ripakil olnud **toot**² lahti.
19. Viiulist kuuldis **harva**¹ mõni **kiire**³ **noot**² ja kogu mängu **laad**² oli ilmetu.
20. **Rõõmus**³ **pruut**² astus rahva sekka, käes suur **kruus**² õllega.
21. Muidu nii karme **mees**² astus aegamisi, hoomamata, et teistel on **kiire**³ ning külm.
22. Kui pea ei püsi **kaine**³, võib tagajärjeks olla pikk **luus**² ning pohmell.
23. Läbinisti **ahne**³ või **ihne**³ inimene ei tea tihti, mis on **halastus**¹ ja suuremeelsus.
24. Oli **hirm**¹, et **paat**² läheb kummuli ning **noot**² on jälle kaladest tühi.
25. **Kuulus**³ kunstnik oli **andekas**³, aga laisk.
26. Järjekordne **rahvuslik**³ **laat**² toimub mardipäeval.
27. Ka **soodus**³ asi võib olla **maitsekas**³ või **heldima**¹ panev.
28. Riide mustris põimusid suur must **ruut**², jäme kollane **võöt**² ja peenikesed triibud.
29. **Hing**¹ kisub Ameerikasse, metsik **lääs**² ootab seiklejaid.
30. **Tuul**² **hakkas**¹ eile **hommikul**¹ tugevasti puhuma.

31. Põõsa **all**¹ kasvas **pruun**² punakas **seen**², mitte **hall**¹.
32. Ma **saan**² väga **hästi**¹ **aru**¹, kuhu selle tee **haru**¹ viib.
33. Kurvilisel teel on iga käänu taga uus **vään**² ootamas.
34. Selle **kirju**³ tapeedi **paan**² mulle meeldis, sest **toon**² sobis **aula**³ seinale hästi.
35. **Enne**³ me **arutasime**¹, kuidas võiks vanu **hilpe**¹ kasutada ja kampsuneid üles **harutada**¹.
36. Meie **nooruslik**³ **ind**¹ rauges, sest selle asja **hind**¹ jäi **äärmuslikult**³ kõrgeks.
37. **Politseinik**³ **hindas**¹ olukorda **peenelt**³ ja **himur**¹ pätt sai õiglase karistuse.
38. **Õng**¹ käes kõndisin mere poole, kui mulle lõi vastu soolane **hõng**¹ ning **nõiduslik**³ **hingus**¹.
39. **Hotentoti**¹ **pealik**³ tantsis **hingestatult**¹ ja **holpsas**¹ ümber lõkke.
40. Meie toidulauale ei satu just tihti **vaal**², **homaar**¹ või **hai**¹.
41. **Mool**² on mõõtühik, mida **teaduslikes**³ tekstides on **ammu**³ kasutatud.
42. **Kool**² algas **augustis**³ ühise **piknikuga**³ pargis.
43. **Roos**² närtsis, sest **vaas**² oli tühi.
44. **Eile**³ oli **ümbrikus**³ veel jaapani **jeen**² koos mõne euroga.
45. Mul oli **plaan**² soome **keel**² **hiljem**¹ **ilma**³ sinu abita ära õppida.
46. **Paas**² on Eesti **looduslik**³ ressurss.
47. **Võimekas**³ **raal**² on parem kui **hädine**¹ nutitelefoni.
48. Mulle ei meeldi mitte üks **raas**² selle **hotelli**¹ suur **saal**² ega restoran.
49. Seal on mingi **hirmus**³ **hais**¹ ja ebameeldiv **foon**², mis segab.
50. **Tallinlane**³ on **kaitsetu**³, kui ta **hilja**³ õhtul pimedas ringi kõnnib.
51. **Teisal**³ peletas **ruun**² lendu **hakke**¹.
52. See **geel**² ravib sind, kui sul on **haiguslik**³ **veen**² või mõni muu häda.
53. Õlle tegemisel **hoia**¹ **uun**² sees, kui tahad, et **pruul**² õige tuleb.
54. Soomes on **lään**², meil maakond.
55. Tõmba sinna **joon**², kuhu esimene **kuul**² veeres.
56. Kaks poolt on kokku **terve**³ ja kaks veerandit on kokku pool.

Liisi Piits
 Eesti Keele Instituut
 Roosikrantsi 6
 10119 Tallinn
 liisi.piits@eki.ee

Mari-Liis Kalvik
 Eesti Keele Instituut
 Roosikrantsi 6
 10119 Tallinn
 mari-liis.kalvik@eki.ee

Phonological variation in Estonian on the basis of three phenomena

LIISI PIITS, MARI-LIIS KALVIK

This paper examines phonological variation in Estonian, focusing on three main areas of variation: word-initial /h/, palatalization in *i*-stemmed words with a (CC) $\bar{V}C$ structure, and quantity degrees. The pronunciation fixed in the Dictionary of Standard Estonian (ÕS 2018) is subject to variation, and the variants reflected in the dictionary are not always the ones that seem to be preferred in the usage. Variation causes problems in the text to speech synthesis system, where it is necessary to give preference to one of the pronunciation variants in the text analysis process, and therefore it is important to identify which variant is more common among language users. The study is based on a reading experiments conducted with 191 informants. The article consists of two parts: a description of the speech corpus collected by experiments and an overview of the research results.

The study revealed that according to our results, word-initial /h/ was pronounced 92% of the time, probably due to the formal situation of the reading experiment and written text. So it is doubtful that text-to-speech synthesis would become more natural if, in 8% of cases, the /h/ is not pronounced at the beginning of the word.

Palatalization occurs among more than half of the pronunciation instances in only 5 out of the 36 words. The results confirm that the palatalization of the consonants *l*, *n*, *s*, *t*, *d* at the end of *i*-stemmed Estonian words with a (CC) $\bar{V}C$ structure is generally untypical. Thus, there is no reason to palatalize these words in a synthetic speech.

In the case of almost 50 examined words, it was possible to determine which quantity degree (second or third) the informants preferred the most. Additionally, words were divided into eight groups according to the part of speech and word syllable structure, in order to make it possible to draw inferences about other words of the same group as well. In two groups the third quantity degree dominated, in the other two the second quantity dominated. The other four groups did not exhibit unified pronunciation patterns.

Keywords: Estonian, phonology, text to speech synthesis, word-initial /h/, palatalization, quantity degree