

SÕNA ESINEMISSAGEDUS JA TÄHENDUSTE ERISTAMISE VAJADUS HÄÄLDUSE MÕJUTAJANA

Mari-Liis Kalvik ja Liisi Piits

Eesti Keele Instituut

Ülevaade. Artikkel käsitleb kaht eesti keeles varieeruvat nähtust: *h* hääldamist sõna alguses ning *i*-tüveliste (C)VVC-struktuuriga sõnade lõpukonsonantide palatalisatsiooni. Mõlema nähtuse puhul uuritakse, milline on sõna esinemissageduse ning tähenduste eristamise vajaduse mõju hääldusele. Uurimismaterjaliks on 42 *h*-algulist sõna ja 30 pika vokaaliga ühesilbilist *i*-tüvelist sõna, mis on keelejuhtide poolt ette loetud. 94 keelejuhi 3945 hääldusjuhu uurimine andis tulemuseks, et a) *h* välja-hääldamine oli valdav (92% juhtudest); b) suurema esinemissagedusega sõnades oli sõnaalguline *h* enamasti lühem ja see jäeti välja hääldamata sagedamini kui väiksema esinemissagedusega sõnades; c) esines tendents, et sõnades, kus *h*-l on tähendust eristav ülesanne, oli *h* veidi sagedamini ja veidi pikemalt välja hääldatud. 40 keelejuhi 1280 hääldusjuhu analüüs näitas, et a) väiksema esinemissagedusega sõnu ei palataliseeritud märgatavalt vähem kui suurema esinemissagedusega sõnu ning b) sõnu, kus palatalisatsiooni esinemine eristab tähendusi, palataliseeriti rohkem.

Märksõnad: sõnaalguline *h*, palatalisatsioon, esinemissagedus, tähenduste eristamine, lugemiseksperiment, keetus

DOI: <https://doi.org/10.12697/jeful.2019.10.1.04>

1. Sissejuhatus

Artiklis uuritakse sõna esinemissageduse ja tähenduste eristamise vajaduse mõju kahele eesti keeles varieeruvale nähtusele: sõnaalgulise *h* hääldusele ja palatalisatsiooni esinemisele (C)VVC-struktuuriga *i*-tüvelistes sõnades.

Sõnaalgulise *h* kirjutamine põhineb ajaloolistel kokkulepetel, kuid selle väljahääldamine on tänapäeval kujunenud vabaks. „Eesti õigekeelsuskäsiraamatus“ selgitatakse, et „*h* kirjutamine sõna algul on traditsioonipärane, häälduses ei ole see nõutav“ (Mäearu 2014–2018). Varasemad uurimused on toonud esile, et *h* hääldamine on seotud ametliku situatsiooniga, mispuhul hääldatakse sõnaalguline *h* märgatavalt

rohkem välja (Cui 1999, Mesipuu 2007). Neid tulemusi kinnitab ka eesti keele spontaanse kõne foneetilise korpuse¹ materjal. Korpuse põhjal tehtud uurimuste tulemustest on selgunud, et kui keskmiselt on *h* sõna alguses välja hääldatud 46% kõigist juhtudest, siis ametlikumates monoloogides on seda tehtud 76% hääldusjuhtudest (Asu et al. 2016: 115). Väikses etteloeatud kõnel põhinevas uurimuses on varem leitud, et loetud kõnes on *h* väljahääldamise protsent veelgi kõrgem, keskmiselt 95% (Kalvik ja Piits 2015: 65).

Sõna esinemissagedust on nähtud eelkõige häälduskestuse mõjutajana (Bell et al. 2009) ja häälikukao põhjustajana (Coetzee ja Kawahara 2013, Bybee 2002). Varasemates eesti keele kohta tehtud sõnaalgulise *h* uuringutes on tähele pandud, et suurema esinemissagedusega sõnades jääb häälik sagedamini välja hääldamata kui väiksema esinemissagedusega sõnades (Cui 1999: 42). Siinses uurimuses tahtsimegi esmalt välja selgitada, kas sagedates sõnades on *h* kestus lühem ja kas seda jäetakse rohkem välja hääldamata kui harvades sõnades.

Eesti keeles üldjuhul ei ole rangelt fikseeritud hääldusstandardit (Asu et al. 2016: 11), siiski märgib kirjakeele normi alus „Eesti õigekeelsussõnaraamat ÕS 2018“ rõhu ja välte kõrval ka palatalisatsiooni. Hiljutine uurimus (Piits ja Kalvik 2019) tegi kindlaks, et ÕS-is palatalisatsioonimärgi saanud (C)VVC-struktuuriga *i*-tüveliste sõnade lõpu-konsonantide (*l, n, s, t/d*) palataliseerimine on igapäevases keelekasutuses ebajärjekindel. Kokkuvõtlikult öeldes on selles tüübis valdav palatalisatsioonita hääldus. Käesolevas uurimuses soovisime teada saada, kas palatalisatsiooni esinemine on seni käsitletud tegurite kõrval mõjutatud ka sõna esinemissagedusest.

Lisaks oletasime, et kummagi nähtuse puhul võib varieerumine sõltuda ka sellest, kas konkreetsetes sõnas on sellel häälikul tähendust eristav funktsioon või mitte. Sõnaalguline *h* ja palataliseeritud häälik võivad esineda nii sõnades, kus nad eristavad tähendusi (nt *haru* ‘tervikust hargnev iseseisev osa’ vs. *aru* ‘mõistus’ ja *noot*² ‘noodikirjamärk’ vs. *noot* ‘kalapüügivahend’) kui ka sõnades, kus nad tähendusi ei erista. Tähele on pandud, et suurema naabrustihedusega sõnu ehk sõnu, millest ühe foneemi asendamisega on võimalik moodustada rohkem tähenduslikke tähejärjendeid, hääldatakse välja püüdlikumalt (Vitevitch ja Luce

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

2 Kriipsuke konsonandi järel märgib palataliseeritud häälikut.

2016). Me ei ole uuritavate sõnade puhul välja arvutanud, milline oleks nende naabrustihedus ehk kui palju täistähenduslikke sõnu tekiks ühe foneemi asendamisel teisega. Selle asemel oleme vaadanud, kuidas mõjutab hääliku hääldamist see, kas ta täidab fonoloogilist funktsiooni, st kas sõnaalgulise *h* hääldamata jätmine või lõpukonsonandi mittepalataliseerimine tekitab uue tähendusega sõna. Seega oleme materjalist välja otsinud sõnad, kus sõnaalguline *h* või palataliseeritud konsonant eristab tähendusi. Meid huvitas, kas tähendusi eristavatel juhtudel hääldatakse häälik püüdlikumalt välja. *h* puhul jälgisime, kas sellistel juhtudel hääldatakse *h* rohkem välja ja kas *h* kestus on siis pikem kui tähendusi mitteeristavatel juhtudel. Palatalisatsiooniga seoses vaatasime vaid selle esinemist või mitteesinemist.

Püstitasime järgmised uurimisküsimused:

1. Kas suurema esinemissagedusega sõnades on sõnaalguline *h* lühem ja kas see jäetakse sagedamini välja hääldamata kui väiksema esinemissagedusega sõnades?
2. Kas väiksema esinemissagedusega, keelejuhile eeldatavasti tundmatumaid sõnu palataliseeritakse vähem kui suurema esinemissagedusega sõnu?
3. Kas sõnades, kus sõnaalgulise *h* või palataliseeritud hääliku esinemine eristab tähendusi, hääldatakse vastav häälik korrektsemalt välja?

2. Materjal ja meetod

Materjal on kogutud lugemiseksperimentidega, mille käigus paluti keelejuhtidel ette lugeda 56 lauset, mis sisaldasid 42 *h*-algulist sõna ja 30 (C)VVC-struktuuriga *i*-tüvelist sõna. Lugemiseksperimente on tehtud 2015. aastast alates ja praeguseks on kogutud materjali 194 keelejuhilt üle Eesti. Materjali salvestamise põhimõtteid ja käiku on üksikasjalikumalt kirjeldanud Kalvik ja Piits (2015). Siinse artikli tarvis tegime sellest materjalist valiku. Salvestised segmenteerisime ja märgendasime, kasutades programmi Praat (Boersma ja Weenink 2014). Lisaks sõna- ja häälikutasandile lõime programmis nii sõnaalgulise *h* kui ka palatalisatsiooni jaoks lisatasandid. *h* puhul piiritlesime lisatasandile vastava segmendi ning kui häälik kuuldelise mulje ning helilaine vaatluse järgi puudus, tähistasime selle märkega *h0*.

(C)VVC-struktuuriga *i*-tüvelistes sõnades piiritlesime lõpukonsonandi, mille märgendasime vastavalt kuuldelisele hinnangule palataliseerituks või mittepalataliseerituks (vrd *paat'* vs. *paat*). Pärast märgendamist kasutasime Praati skripte, et automaatselt leida, kui tihti *h* või palataliseeritud konsonant oli välja hääldatud ning kui pikk oli sõnaalgulise *h* kestus väljahääldamise korral.

2.1. Sõnaalguline *h*

Sõna alguses asuva *h* analüüsimiseks võtsime uurimusse kõik senini märgendatud salvestised ehk 94 keelejuhi materjali. Igalt keelejuhilt oli 42 *h*-algulist sõna, v.a kolm juhtu, kus keelejuht oli jätnud sõna lugemata või oli lugenud soovitud sõna asemel midagi muud. Sõnade hulgas olid enamjaolt ainsuse nimetavas käändes olevad nimi- ja omadussõnad, määrsõnad ja neli erinevat verbivormi. Valdavalt olid need omasõnad. Kokku võtsime vaatluse alla 3945 hääldusjuhtu. Uurimaks sõna esinemissagedust varieerumise mõjutajana moodustasime sõnapaarid, kus üks paariline oli suure ja teine väikese esinemissagedusega sõna. Andmed lemma esinemissageduse kohta (vt tabelis 1) võtsime sagedussõnaraamatust (Kaalep ja Muischnek 2002), mis põhineb miljonisõnalisel ilu- ja ajakirjanduskorpusel. Sõnavalikul lähtusime sellest, et nii sagedates (esinemiskordi korpuses 88–2178) kui ka harvades (esinemiskordi korpuses 0–15) sõnades järgneks *h*-le sama vokaal. Selliseid sageduse alusel vastanduvaid *h*-alguliste sõnade paare oli 14, esimesena on esitatud suure esinemissagedusega sõna, järgneb väikse esinemissagedusega sõna: *hakkas* (*hakkama* kindla kv. lihtm. ains. 3. p.) – *hakke* (*hakk* ‘vareslane’ mitm. os.), *hing* – *hingus*, *hästi* – *hädine*, *hiljem* – *hilpe* (*hilp* mitm. os.), *hindas* (*hindama* kindla kv. lihtm. ains. 3. p.) – *hingestatult*, *homme* – *homaar*, *hotelli* (*hotell* ains. om.) – *hotentoti* (*hotentott* ains. om.), *hunt* – *huupi*, *heida* (*heitma* käskiva kv. oleviku ains. 2. p.) – *heldima*, *haigus* – *halastus*, *hind* – *hindu*, *hammas* – *harras*, *hoolimata* – *hoomamata*, *hirm* – *hirss*.

h eristab tähendusi sõnades *hais* (vs. *ais*), *hall* (vs. *all*), *hirve*, *hirm* ains. om. (vs. *irve*), *haru* (vs. *aru*), *harutama* (vs. *arutama*), *hõng* (vs. *õng*), *halastus* (vs. *alastus*), *harva* (vs. *arva*, *arvama* käskiva kv. oleviku ains. 2. p.), *hala* (vs. *ala*), *hea* (vs. *ea*, *iga* ains. om.), *hind* (vs. *ind*). Selleks et lugejal tekiks selgem vastandusvajadus, paigutasime viis *h*-algulist sõna samadesse lausetesse koos vokaalialgulise sõnaga.

Põõsa all kasvas pruun punakas seen, mitte hall.

Ma saan väga hästi aru, kuhu selle tee haru viib.

Enne ma arutasime, kuidas võiks vanu hilpe kasutada ja kampsuneid üles harutada.

Meie nooruslik ind rauges, sest selle asja hind jäi äärmuslikult kõrgeks.

Õng käes kõndisin mere poole, kui mulle löi vastu soolane hõng ning nõiduslik hingus.

Sõnaalgulise *h* uurimiseks arvutasime välja kõigi sõnade *h* keskmise kestuse ja *h* väljahäldamise protsendi. Võrdlesime nimetatud näitajaid sageduse alusel vastanduvates sõnapaarides, et selgitada välja, kas suurema esinemissagedusega paarikus on *h* enamatel kordadel välja häälstatud ja pikema kestusega. Seejärel võrdlesime tulemusi suuremate rühmade kaupa. Sõnarühmi oli kokku neli: harva esinevad ja sageli esinevad sõnad ning sõnad, kus *h* eristab tähendusi, ja sõnad, kus *h* ei erista tähendusi. Arvutasime iga sõnarühma kohta *h* keskmise kestuse ning *h* väljahäldamise protsendi, et võrrelda omavahel sageduse alusel ning tähenduseristuse alusel vastanduvaid sõnarühmi. Rühmadevaheliste statistiliste erinevuste testimiseks rakendasime sõltumatute valimite *t*-testi, kasutades selleks Exceli ja StatPaci³ vahendeid.

2.2. Palatalisatsioon (C)VVC-struktuuriga *i*-tüvelistes sõnades

Palatalisatsiooni uurimismaterjaliks olid 40 keelejuhi ühesilbilised pika vokaaliga *i*-tüvelised sõnad. Nendes sõnades oli pikk vokaal *aa*, *oo*, *uu* või *ee* ning nad lõppesid kas *l*-i, *n*-i, *s*-i, *d* või *t*-ga, mis peakisid ÕS 2018 alusel olema häälstatud palataliseeritult. Kasutasime uurimismaterjali, mille põhjal käsitleti lõpukonsonantide palatalisatsiooni varieerumist keelejuhi päritolupiirkonnast lähtuvalt (vt Piits ja Kalvik 2019). Valimis esindas iga ajaloolist murdeala võrdne arv keelejuhte, kes olid pärit samast piirkonnast kui nende vanemad. Igalt keelejuhilt oli 15 sagedusel vastanduvat sama silbiriimiga sõnapaari, esikohal suure esinemissagedusega sõna, seejärel väikse esinemissagedusega sõna: *saal – raal*, *plaan – paan*, *vaas – paas*, *laad – raad*, *paat – raat*, *veen – jeen*, *peet – neet*, *kool – mool*, *toon – foon*, *roos – oos*, *noot – toot*,

3 StatPaci programmi vt <https://www.statpac.org/>.

kuul – *pruul*, *pruun* – *uun*, *kruus* – *luus*, *pruut* – *nuut*. Sõnade hulgas oli nii oma- kui ka võõrsõnu, kõik nad olid ainsuse nimetavas käändes.

Kuigi palatalisatsioon on eesti keeles olnud ajalooliselt koartikulaatorne nähtus, on see aja jooksul saanud ka tähendust eristava ülesande (Teras ja Pajusalu 2014: 259). Palataliseeritud häälik eristab kirjakeele normi järgi tähendusi meie uurimuse seitsmes (C)VVC-struktuuriga *i*-tüvelises sõnas: *noot* ‘noodikirjamärk’ (vs. *noot* ‘kalapüügivahend’), *kruus* ‘tass’ (vs. *kruus* ‘killustik’), *kool* ‘õppeasutus’ (vs. *kool* ‘ila’), *kuul* ‘kerakujuline ese’ (vs. *kuul*, *kuu* ains. alalü.), *kood* ‘kokkuleppeiline märk’ (vs. *kood*, *kuduma* kindla kv. oleviku ains. 2. p.), *tees* ‘juhtmõte’ (vs. *tees*, *tee* ains. seesü.), *toon* ‘värvitoon’ (vs. *toon*, *tooma* kindla kv. oleviku ains. 1. p.), *luus* (vs. *luus*, *luu* ains. seesü.). Lause-tes, mida keelejuhid ette lugesid, esines enamasti ainult *i*-tüveline sõna, seega ei olnud lauseid lugedes võimalik näha otsest vastandust esitatud sõnapaaride vahel, vaid vastandus oli mõtteline.

Kokku oli mõlemas vastanduses olevaid (C)VVC-struktuuriga *i*-tüvelisi sõnu 37, nende hääldusjuhte 1280. Siingi oli sõnarühmi neli: harvade ja sagedate sõnade rühmad ning rühmad, kus palatalisatsioonil oli tähendust eristav funktsioon ja kus tähendust eristav funktsioon puudus. Arvutasime, mil määral (%) sõnu rühmades palataliseeriti, et võrrelda omavahel sageduse alusel ning tähenduseristuse alusel vastanduvaid sõnarühmi. Sageduse alusel vastanduvaid sõnu võrdlesime ka paaride kaupa: iga hääldatud sõnapaari puhul hindasime, kas keelejuht palataliseeris mõlemat sõna, jättis mõlemad sõnad palataliseerimata või palataliseeris ainult sagedat / ainult harvaesinevat sõna.

3. Tulemused

3.1. Sõna esinemissagedus sõnaalgulise *h* häälduse mõjutajana

Kõigist 3945 hääldusjuhust hääldati 92% juhtudel sõnaalguline *h* välja.

Tabel 1. *h*-algulised sõnad paaridena, nende sagedus, *h* keskmine kestus (ms), standardhälve ja väljahäldamata *h*-de protsent.

Sõna	Sage (S) / Harv (H)	Esinemis- sagedus	<i>h</i> keskmine kestus (ms)	Standard- hälve	<i>h</i> -ta häälstatud juhud (%)
haigus	S	94	71	23	3
halastus	H	12	85	25	2
hakkas	S	2178	62	22	18
hakke	H	6	81	30	1
hammas	S	148	74	30	5
harras	H	6	101	34	0
heida	S	258	58	20	3
heldima	H	7	74	27	1
hiljem	S	449	76	23	9
hilpe	H	6	58	20	6
hind	S	398	76	26	2
hindu	H	0	78	33	2
hindas	S	182	68	21	4
hingestatult	H	0	54	29	35
hing	S	302	69	28	4
hingus	H	6	89	28	1
hirm	S	184	74	26	6
hirss	H	0	86	23	7
homme	S	169	69	20	5
homaar	H	0	97	41	7
hoolimata	S	88	80	35	7
hoomamata	H	15	89	36	9
hotelli	S	146	61	20	13
hotentoti	H	0	67	24	13
hunt	S	115	64	22	17
huupi	H	10	76	26	3
hästi	S	488	56	13	36
hädine	H	6	74	22	5

Märkus. Sõnad on esitatud paaridena esinemissageduse alusel: S – sagedad sõnad, H – harvad sõnad. Viimane veerg näitab, kuivõrd sõnu ilma *h*-ta häälstatuti. Joonega ümbritsetud väärtused märgivad juhte, kus sagedamas sõnas on *h* pikema kestusega või enamatel kordadel välja häälstatud kui harvemas sõnas.

Esinemissagedusel vastanduvate sõnapaaride võrdlus näitas, et 14 paarist oli enamik ehk 12 sellised, kus suurema esinemissagedusega sõna algul olev *h* oli lühemalt hääldatud kui väiksema esinemissagedusega sõnas. Neist 12 paarist olid kaks sellised, kus erinevus ei olnud suur: sage *hind* (*h* keskmine kestus 76 ms, standardhälve 26) vs. harv *hindu* (78 ms, standardhälve 33) ning sage *hotelli* (61 ms, standardhälve 20) vs. harv *hotentoti* (67 ms, standardhälve 24). Ka *h*-ta hääldatud juhtude protsent oli neis paarides ühesugune. Kuid esitatuile lisaks leidis ka kaks sõnapaari, kus olukord oli kestuse poolest eeldatule vastupidine. Nii oli sagedas sõnas *hiljem* sõnaalgulise *h* kestus 76 ms ning harvas sõnas *hilpe* 58 ms (standardhälbed vastavalt 23 ja 20). Tähelepanuväärne oli paar *hindas* – *hingestatult* (*h* keskmised kestused vastavalt 68 ja 54 ms, standardhälbed 21 ja 29), kus oli märgatav erinevus ka *h* hääldamata jätmisel: 4% juhtudest (sage *hindas*) vs. 35% juhtudest (harv *hingestatult*). Selliseid paare, kus *h*-ta hääldusjuhtude protsent oli vastupidiselt eeldusele harva sõna puhul suurem, oli teisigi, näiteks *hirm* – *hirss*, *homme* – *homaar* ja *hoolimata* – *hoomamata*. Neis oli erinevus siiski väga väike, paari protsendi vahemikus, ning keskmiste kestuste andmed olid ootuspärased.

Üksiksõnade kaupa *h* keskmisi kestusi vaadates on näha, et suure esinemissagedusega sõnades oli *h* enamasti lühem kui väikse esinemissagedusega sõnades. Kui *h* keskmine kestus oli 74 ms (standardhälve 29), siis keskmisest lühema *h* kestusega olid enamasti sagedad sõnad, kuigi ka kolmes harvaesinevas sõnas – *hingestatult* (54 ms), *hilpe* (58 ms) ja *hotentoti* (67 ms) – oli *h* kestus keskmisest lühem. Keskmisest pikema *h* kestusega olid enamasti harvad sõnad, kuid ootustele vastupidiselt oli *h* pikem ka sõnades *hind* (76 ms), *hiljem* (76 ms) ning *hoolimata* (80 ms).

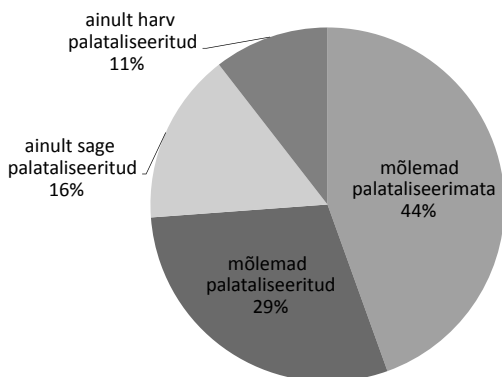
Kokkuvõtlikult näitas esinemissagedusel vastanduvate sõnade uurimine, et sagedamas sõnas oli *h* kestus lühem (keskmiselt 68 ms) ning harvas pikem (keskmiselt 79 ms). Gruppide statistiliselt olulisele erinevusele viitab ka *t*-test: $t = 9,642$, $df = 2414$; $p < 0,0001$. Väikse esinemissagedusega sõnades jäeti *h* välja hääldamata 7% hääldusjuhtudest ja suurema esinemissagedusega sõnades 10% hääldusjuhtudest, mis osutus statistiliselt oluliseks erinevuseks: $t = 2,759$, $df = 2630$; $p = 0,0059$.

3.2. Sõna esinemissagedus palatalisatsiooni mõjutajana

Palatalisatsiooni ja sõna esinemissageduse seoseid analüüsid peaks vastanduse sage/harv asemel arvestama vastandusega tuttav/tundmatu. Kuigi keelejuhtidelt ei küsitud, kui hästi nad lausetes esinevaid sõnu tunnevad, eeldasime, et sellised harvem esinevad sõnad nagu *raat, oos, mool, uun* ja *toot* on lugejatele tundmatumad. Mõnel juhul mõõnsid ka lugemiskatses osalenud, et nad ei tea mõne sõna tähendust. Tahtsime välja selgitada, kas väiksema esinemissagedusega ja kuulajale eeldatavasti mitte nii tuttavate sõnade hääldamisel lõpukonsonanti pigem palataliseeritakse või jäetakse see palataliseerimata. Uuritavad suure esinemissagedusega sõnad olid *saal, plaan, vaas, laad, paat, veen, peet, kool, toon, roos, noot, kuul, pruun, kruus* ja *pruut*. Vastavad väikse esinemissagedusega sõnad olid *raal, paan, paas, raad, raat, jeen, neet, mool, foon, oos, toot, pruul, uun, luus* ja *nuut*.

Kõigi 1200 hääldusjuhu hulgas esines palataliseerimata sõnu 71%, suure esinemissagedusega palataliseeritud sõnu oli 16% ja väikse esinemissagedusega palataliseeritud sõnu oli 13%. Arvestades, et materjalis oli sagedaid ja harvaesinevaid sõnu võrdselt, oli näha, et suurema esinemissagedusega sõnu palataliseeriti kolme protsendi võrra rohkem.

Sageduse alusel vastanduvaid sõnu võrreldi ka paaride kaupa: iga hääldatud sõnapaari puhul hindasime, kas palataliseeriti mõlemat sõna, jäeti mõlemad sõnad palataliseerimata või palataliseeriti ainult sagedat / ainult harvaesinevat sõna. Võrdlesime palatalisatsiooni sageduse alusel vastanduvate sõnapaaride kaupa. Jooniselt 1 on näha, et kõigist 600 sõnapaarist oli kõige rohkem selliseid, kus mõlemad paarid jäeti palataliseerimata (44%), veidi vähem oli selliseid paare, kus palataliseeriti mõlemad paarilisi (29%). Kõigist sageduse alusel vastanduvatest paaridest moodustasid 16% need paarid, kus ainult sagedam paariline oli palataliseeritud, ja 11% moodustasid need, kus ainult harvem esinev paariline oli palataliseeritud. *t*-test näitas, et sõnapaaride arv, kus ainult sage paarik on palataliseeritud, erines oluliselt sõnapaaride arvust, kus ainult harv paarik oli palataliseeritud: $t = 2,368$, $df = 599$; $p = 0,0182$.



Joonis 1. Palatalisatsioon sageduse alusel vastanduvates sõnapaarides. Esitatud on sõnapaaride osakaal (%), kus a) mõlemad paarilised on palataliseeritud, b) mõlemad paarilised on palataliseerimata, c) ainult sage paariline on palataliseeritud ning d) ainult harv paariline on palataliseeritud.

Siiski näitavad tulemused, et sõna esinemissagedus või sõna tuntus mõjutas keelejuhti vähe. Kui sõnapaarist palataliseeriti ainult üht paari- list, siis oli see pisut suurema tõenäosusega sage sõna, aga palju suurem oli võimalus, et palataliseeriti mõlemaid paarilisi või jäeti mõlemad paarilised palataliseerimata.

3.3. Tähduse eristamise vajadus sõnaalgulise *h* häälduse mõjutajana

Selleks et analüüsida tähenduse eristamise vajaduse mõju sõna- algulise *h* hääldamisele, võrdlesime nii *h* väljahääldamist kui ka *h* kestust tähendust eristavatel ja mitteeristavatel juhtudel. Tabelis 2 on *h*-algulised sõnad järjestatud alates sõnast *harras*, mille hääldamisel *h*-ta hääldatud juhud puudusid, kuni sõnani *hästi*, mille hääldamisel jäeti 36% juhtudest *h* hääldamata. Tabelist on näha, et sõnad, kus *h* eristab tähendusi (hallil taustal), esinevad rohkem tabeli algusosas ehk nendes sõnades jäeti *h* harva hääldamata. Teistest eristub tabeli lõpuosas paiknev sõna *hirve*, kus küll tähenduse eristamise vajadus võiks tingida *h* esinemise, aga milles *h* on jäetud välja hääldamata 19% juhtudest.

Tabel 2. Sõnaalgulise *h*-ga sõnad järjestatuna *h*-ta hääldatud hääldusjuhtude protsendi järgi.

Sõna	<i>h</i> tähendust eristavas asendis	Hääldus- juhtude arv	<i>h</i> -ta hääldatud juhte (%)	<i>h</i> keskmine kestus (ms)
harras	–	94	0	101
hall	+	94	1	78
heldima	–	94	1	74
hakke	–	94	1	81
hingus	–	94	1	89
hind	+	94	2	76
hõng	+	94	2	79
hea	+	94	2	82
hai	+	94	2	97
hindu	–	94	2	78
holpsas	–	94	2	84
harva	+	93	2	69
halastus	–	93	2	85
heida	–	94	3	58
haigus	–	94	3	71
huupi	–	94	3	75
hais	+	94	4	72
hala	+	94	4	77
himur/a	–	94	4	79
hindas	–	94	4	68
hing	–	94	4	69
homme	–	94	5	69
hädine	–	94	5	73
hammas	–	94	5	74
haru	+	94	6	77
hilpe	–	94	6	58
hirm	–	94	6	74
harutada	+	94	7	74
hoolimata	–	94	7	79
hirss	–	94	7	86
homaar	–	94	7	97

Sõna	<i>h</i> tähendust eristavas asendis	Hääldusjuhtude arv	<i>h</i> -ta hääldatud juhte (%)	<i>h</i> keskmine kestus (ms)
hiljem	–	94	9	76
hoomamata	–	93	9	88
hotelli	–	94	13	61
hotentoti	–	94	13	67
hunt	–	94	17	63
hakkas	–	94	18	61
hirve	+	94	19	62
hoia	–	94	22	73
hommikul	–	94	23	60
hingestatult	–	94	35	52
hästi	–	94	36	54

Märkus. Hallil taustal on sõnad, kus *h* eristab tähendusi (tumehallil taustal sõnad, kus tähendusvastandus oli esindatud ka ettelõetud lauses).

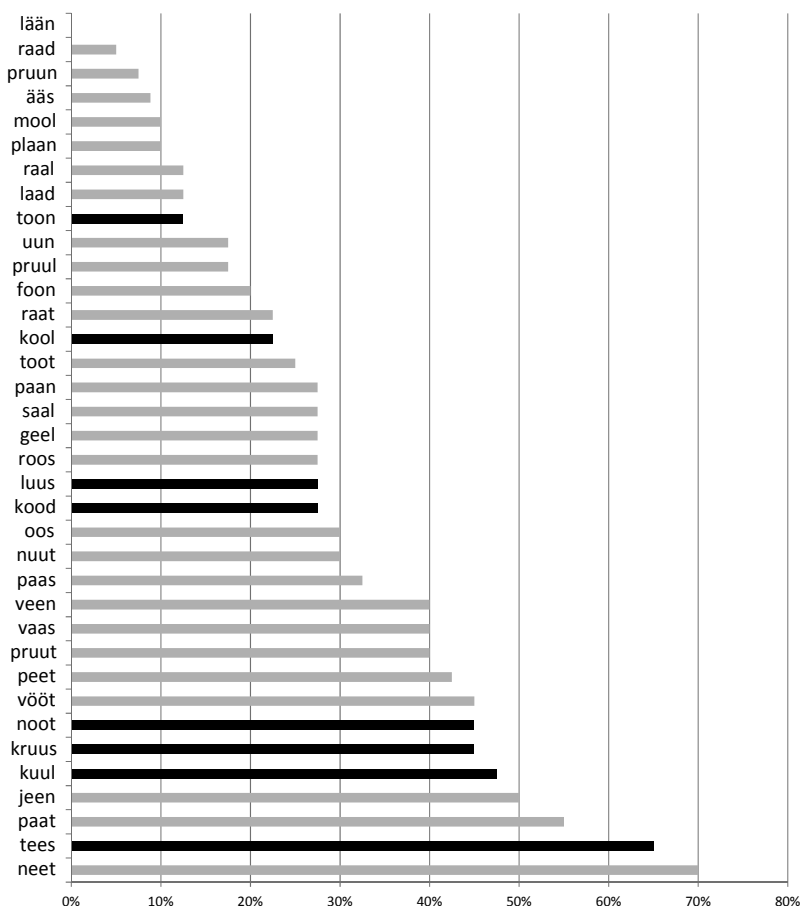
Sõnaalgulise *h* väljahääldamata juhtude võrdlusest selgus, et sõna-des, kus *h* eristab tähendusi (11 sõna, kokku 1033 hääldusjuhtu), jäeti *h* välja hääldamata 5% hääldusjuhtudest. Sõna-des, kus *h* ei erista tähendusi (31 sõna, kokku 2912 hääldusjuhtu), jäeti *h* välja hääldamata 9% hääldusjuhtudest. *t*-test osutas, et tegu on olulise erinevusega: $t = 4,082$, $df = 3943$; $p < 0,0001$.

Vaatasime ka eraldi sõnu (*hall*, *hing*, *hõng*, *haru*, *harutada*), kus *h* eristab tähendusi ja kus tähendusvastandus esineb ühes ja samas lauses (nt *Ma saan väga hästi aru, kuhu selle tee haru viib*). Nende sõnade hääldamisel jäeti *h* välja hääldamata 4% esinemisjuhtudest. Seega tähendust eristaval juhul hääldati *h* sõna alguses rohkem välja.

Sõnaalgulise *h* keskmisi kestusi võrreldes selgus, et sõna-des, kus *h* eristab tähendusi (11 sõna, kokku 1033 hääldusjuhtu, neist *h* välja hääldatud 982 juhul), oli *h* keskmine kestus 77 ms (standardhälve 30) ja sõna-des, kus *h* tähendusi ei erista (31 sõna, kokku 2912 hääldusjuhtu, neist *h* välja hääldatud 2650 juhul), oli *h* keskmine kestus 73 ms (standardhälve 29). 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.

3.4. Tähenduse eristamise vajadus sõnade palatalisatsiooni mõjutajana

Siinses uurimuses on sõnu, kus minimaalpaari moodustavad nime-
tavas käändes sõnad: *noot* vs. *noot'*, *kruus* vs. *kruus'* ja *kool* vs. *kool'*.
Lisaks on uurimuses sõnu, kus minimaalpaari moodustavad erinevad
sõnavormid: *kuul* vs. *kuul'*, *kood* vs. *kood'*, *tees* vs. *tees'*, *toon* vs. *toon'*
ja *luus* vs. *luus'*. Võrdlesime, kas sõnu, kus palatalisatsioon eristab
tähendusi, palataliseeritakse rohkem kui neid sõnu, kus nähtusel tähen-
dust eristav funktsioon puudub.



Joonis 2. Palataliseeritult hääldatud juhtude osakaal (%) kõigist sõna hääldejuhtudest. Mustad tulbad näitavad sõnu, kus palatalisatsioonil on tähendust eristav funktsioon.

Jooniselt 2 on näha, et sõnade hulgas, kus palatalisatsioon eristab tähendusi, oli nii sõnu, mida palataliseeriti keskmisest enam (*noot, kruus, kuul* ja *tees*), kui ka sõnu, mida palataliseeriti keskmisest vähem (*toon, kood, luus* ja *kood*). Kõigi *i*-tüveliste sõnade hääldusjuhtude võrdlusest selgus, et sõnu, kus palatalisatsioon eristas tähendusi, palataliseeriti veidi enam (37% juhtudest) kui sõnu, kus palatalisatsioon ei eristanud tähendusi (27% juhtudest). *t*-test osutas, et tegu on olulise erinevusega: $t = 3,469$, $df = 1438$; $p = 0,0005$.

4. Kokkuvõte ja arutelu

Tulemused näitasid, et sõna esinemissagedus ja tähenduste eristamise vajadus mõjutasid mõningal määral *h* hääldust ja *i*-tüveliste pika vokaaliga ühesilbiliste sõnade palataliseerimist. Sõna sagedus osutus oluliseks *h* kestuse puhul, mõjutades mõõdukalt ka *h* väljahääldamist, aga vähem sõna palataliseerimist. Tähenduste eristamise vajadus mõjutas rohkem sõnade palataliseerimist ja *h* väljahääldamist, aga ei mõjutanud oluliselt *h* kestust.

Selgus, et loetud kõnes hääldati *h* välja 92% juhtudest. Kõrge väljahääldamise protsendi tingisid ilmselt lugemiseksperimentist tulenevad põhjused: mittespontaanne situatsioon (vt ka Asu et al. 2016) ja kirjapildi mõju hääldusele. Varasemad uurimused (Kalvik ja Piits 2015) on loetud kõnes täheldanud veelgi kõrgemat sõnaalgulise *h* väljahääldamise protsenti (95%), aga see võis tuleneda keelejuhtide valikust: nimetatud uuring põhines ainult 11 kõrgharidusega keelejuhi materjalil.

Tulemusi uurimisküsimuste põhjal analüüsidest saab väita:

1. Esines tendents, et suurema esinemissagedusega sõnades on sõnaalguline *h* lühem ja see jäetakse sagedamini välja hääldamata kui väiksema esinemissagedusega sõnades. See tendents avaldus rohkem sageduse alusel vastanduvate paaride juures. 14 sõnapaari võrdlusest selgus, et 12 puhul oli *h* kestus suurema esinemissagedusega sõnas lühem. Erandina oli harvaesinevates sõnades *hilpe* ja *hingestatult* *h* keskmine kestus lühem kui nende sagedamini esinevates paarilistes.

Kuigi etteloadetud kõnes on olenemata sõna esinemissagedusest *h* väljahääldamise protsent väga kõrge, kinnitas meie uuring varasemat tähelepanekut, et suurema esinemissagedusega sõnades on *h* väljahääldamist siiski vähem kui väiksema esinemissagedusega sõnades (vt Cui

1999: 42). Esines ka erandeid, kus *h* hääldati väga tihti välja ka suure esinemissagedusega sõnades (*hind, heida, hing* ja *hiljem*), milles võiks eeldada selle hääldamata jätmist. Samal ajal jäi mõnes väikse esinemissagedusega sõna vormis, nt *hotentoti* ja *hingestatult*, *h* paljudel juhtudel välja hääldamata. Ilmselt võisid mitmed üldistele tendentsidele vastupidised juhud tuleneda muudest hääldust mõjutavatest faktoritest, näiteks häälikuümbruse mõjust, sõna pikkusest, paiknemisest lauses ja/või rõhulisusest. Need faktorid võisid võimendada, kuna samu sõnu on salvestatud samas kontekstis. Näiteks sõna *hilpe* esines küll 94 erineva keelejuhi esituses, aga kõik lugesid seda sõna samas lausekontekstis. Seega, kui sõna juhtus olema selles lauses rõhulises positsioonis või eelnes mõni häälikukombinatsioon, mis võis hääldust mõjutada, siis see mõju eeldatavasti võimendus.

2. Selgus, et väiksema esinemissagedusega, keelejuhile eeldatavasti tundmatumaid sõnu ei palataliseeritud oluliselt rohkem kui suurema esinemissagedusega sõnu. Sagedust on varasemates uurimustes seostatatud eelkõige häälduskestuse lühenemise ja häälikukaoga (Bell *et al.* 2009, Bybee 2002 ning Coetzee ja Kawahara 2013), mis on häälduse ökonoomsusprintsiiibist tingitud nähtused. Kuigi palatalisatsioon on ajalooliselt olnud koartikulatoorne nähtus, siis praeguseks pole (C)VVC-struktuuriga sõnades palatalisatsioon enam seotud keele ökonoomsusprintsiiibiga. See võib olla põhjus, miks sõna esinemissagedus ei mõjutanud sõnade palataliseerimist.

3. Ilmnes, et sõnades, kus *h* või palataliseeritud häälik eristab tähendusi, hääldatakse see korrektsemalt välja. Erinevused ei olnud aga suured. Sõnades, kus *h* eristas tähendusi, hääldati see veidi sagedamini välja (95% juhtudest) kui tähendusi mitteeristavatel juhtudel (91% hääldusjuhtudest). Aga *h* kestust tähenduste eristamise vajadus oluliselt ei mõjutanud: sõnades, kus *h* eristab tähendusi, oli *h* keskmine kestus neli millisekundit pikem kui sõnades, kus *h* tähendusi ei eristanud. Sõnu, kus palatalisatsioon eristas tähendusi, palataliseeriti rohkem (37%) kui sõnu, kus palatalisatsioon tähendusi ei eristanud (27%).

Mõlema nähtuse puhul esines ka erandeid. Sõnavormis *hirve* jäi *h* olenemata tähenduse eristamise vajadusest sagedamini välja hääldamata, kui oleks oodanud. (C)VVC-struktuuriga *i*-tüveliste sõnade hulgas oli näha, et sõnu *toon, kood, luus* ja *kood*, kus küll palatalisatsioon eristab tähendusi, ei palataliseeritud hoolikamalt. Põhjuseks võib olla,

et minimaalpaari vahel olev tähendusopositsioon on teoreetiline ja lause lugemisel ei pruukinud tähendusvastandus keelejuhile üldse meenuda. Eriti juhul, kui tähendusopositsioonis sõna ka vormiliselt lausekonteksti ei sobi. Sõnades *noot*, *kruus*, *kuul* ja *tees*, kus tähendusvastandus võis olla ilmsem, palataliseeriti lõpukonsonanti rohkem. Sõna naabrustiheiduse mõõtjad on samuti rõhutanud, et naabrustiheidus on vaid teoreetiline näitaja ja pole selge, millised sõnad tegelikult kõneleja või kuulaja mentaalses leksikonis aktiveeruvad (Vitevitch ja Luce 2016: 7.3–7.4). Seega on võimalik, et mitmel juhul ei tajunud teksti lugeja üldse vajadust tähendusi eristada.

Edaspidi soovime uurida, mil määral mõjutavad sõnaalgulise *h* hääldamist ning (C)VVC-struktuuriga *i*-tüveliste sõnade palatalisatsiooni esinemist sellised sotsiolingvistilised tegurid nagu keelejuhi vanus ja sugu. Keelelistest mõjutajatest vajab mõlema nähtuse puhul uurimist emfaas.

Tänuavaldus

Täname kõiki keelejuhte üle Eesti. 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 „Väljendusrikas ja mitmekesine kõnesüntees“ toetusel.

Aadressid

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

Liisi Piits
Eesti Keele Instituut
Roosikrantsi 6
10119 Tallinn, Eesti
E-mail: liisi.piits@eki.ee

Kirjandus

- Asu, Eva Liina, Pärtel Lippus, Karl Pajusalu ja Pire Teras (2016) *Eesti keele häälde*. (Eesti keele varamu, 2.) Tartu: Tartu Ülikooli Kirjastus.
- Bell, Alan, Jason Brenier, Michelle Gregory, Cynthia Girand ja Dan Jurafsky (2009) „Predictability effects on durations of content and function words in conversational English“. *Journal of Memory and Language* 60, 1, 92–111.
- Boersma, Paul ja David Weenink (2014) *Praat: doing phonetics by computer*. [Arvuti-programm]. Kättesaadav aadressil <<http://www.praat.org>>. Alla laetud 06.07.2014.
- Bybee, Joan (2002) „Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change“. *Language Variation and Change* 14, 3, 261–290.
- Coetzee, Andries W. ja Shigeto Kawahara (2013) „Frequency biases in phonological variation“. *Natural Language & Linguistic Theory* 31, 1, 47–89.
- Cui, Kaily (1999) *Sõnaalguline h eesti keeles*. Bakalaureusetöö käsikiri Tartu Ülikooli eesti ja üldkeeleteaduse instituudis.
- Kaalep, Heiki-Jaan ja Kadri Muischnek (2002) *Eesti kirjakeele sagedussõnastik*. Tartu: Tartu Ülikooli Kirjastus.
- Kalvik, Mari-Liis ja Liisi Piits (2015) „Lugemiseksperiment fonoloogilise varieerumise uurimiseks“. *Eesti ja soome-ugri keeleteaduse ajakiri* (erinumber *Kõneuurimise suundi*) 6, 3, 49–77. <https://doi.org/10.12697/jeful.2015.6.3.02>
- Mesipuu, Margit (2007) *Sõnaalguline h*. Magistritöö käsikiri Tartu Ülikooli eesti ja üldkeeleteaduse instituudis.
- Mäearu, Sirje (2014-2018) „Täheortograafia“. Veebiväljaanne *Eesti õigekeelsuskäsiraamat*. Tallinn: Eesti Keele Instituut. Kättesaadav aadressil <<https://keeleabi.eki.ee/vikieksport/Taheortograafia.html>>. Vaadatud 03.07.2019.
- Piits, Liisi ja Mari-Liis Kalvik (2019) „Palatalisatsioon ühesilbilistes *i*-tüvelistes pika vokaaliga sõnades. Roos närtsis, sest vaas oli tühi“. *Keel ja Kirjandus* 7, 513–533.
- Teras, Pire ja Karl Pajusalu (2014) „Palatalisatsioonist ja prepalatalisatsioonist spontaanses eesti keeles“. *Keel ja Kirjandus* 4, 257–269.
- Vitevitch, Michael ja Paul Luce (2016) „Phonological neighborhood effects in spoken word perception and production“. *Annual Review of Linguistics* 2, 7.1–7.20.
- ÕS 2018 = *Eesti õigekeelsussõnaraamat ÕS 2018*. Tallinn: Eesti Keele Sihtasutus.

Abstract. Mari-Liis Kalvik and Liisi Piits: Word frequency and a meaning-distinguishing function of a phoneme as a reason for variation.

The article investigates the variation of word-initial /h/ and palatalization in *i*-stemmed monosyllabic words with a (C)VVC structure. Two possible causes of the variation were examined: word frequency and meaning-distinguishing function of a phoneme. Material was collected by means of a reading task. The text consisted of words with both the word-initial /h/ and palatalized consonants. Altogether, 5225 pronunciations were analysed. The results show

that: a) in 92% of occasions the word-initial /h/ was pronounced; b) in high frequency words, the word initial /h/ was shorter or absent more often than in low frequency words; c) the word-initial /h/ in a meaning-distinguishing position tended to be more often pronounced and to be longer; d) consonant palatalization in *i*-stemmed words with a (C)VVC structure did not depend significantly on word frequency; and e) in the words where palatalization had a meaning-distinguishing function, the palatalized consonant was slightly longer than in the words where the function did not exist.

Keywords: Estonian, word-initial /h/, palatalization, reading experiment, word frequency, meaning-distinguishing function, duration