

VARIEERUVA VÄLTEGA SÕNADE HÄÄLDUSUURINGUD KÕNESÜNTEESI TEENISTUSES

Liisi Piits, Mari-Liis Kalvik

Ülevaade. Artiklis tutvustame lugemiseksperimenti, mille põhjal uurime nn varieeruva vältega sõnade hääldust. Varieeruva vältega sõnade määratlemisel lähtume õigekeelsussõnaraamatu (ÕS 2013) normingutest – uurime sõnu, mida lubatakse hääldada nii teises kui kolmandas vältes. Analüüsime sõnade pearõhulise ja järgsilbi kestussuhteid ja võrdleme saadud andmeid kuuldelise hinnangu tulemustega. Uurime, kas sarnase silbistruktuuri ja sama sõnaliigilise kuuluvusega sõnade häälduses on sarnaseid jooni. Rakenduslikust aspektist kannustab uurimust vajadus leida lahendus probleemidele, mida varieeruvus tekitab tekst-kõne sünteesi protsessis. Uuringu tulemusel selgusid peamised trendid varieeruva vältega sõnade häälduseelistustes. Nii silpide kestussuhete kui kuuldelise hinnangu alusel moodustusid kindlad sõnarühmad, kus domineeris üks või teine välde. Sõnatüüpide kaupa analüüs võimaldas määrata ka välte varieerumise trende tüübiti, nt kõik kolmesilbilised *lik*-liitelised adjektiivid ühe erandiga hääldusid kolmandas vältes.*

Võtmesõnad: fonoloogiline varieerumine, välte varieerumine, tekst-kõne süntees, lugemiseksperiment, eesti keel

1. Sissejuhatus

Masinaid kõnelema pannes oleme kokku puutunud uute probleemidega, mis inimeste omavahelise suhtluse korraldamisel on osutunud ebaoluliseks. Häälduse varieeruvus, mis inimsuhtluses on äärmiselt loomulik ja peegeldab tihti nii piirkondlikku päritolu, inimese vanust, vanemate muredausta kui muid faktoreid, tekitab tekst-kõne sünteesis mitmeid lahendust vajavaid probleemseid olukordi. Moodsate tekst-kõne sünteesi meetoditega on tegeletud Eestis juba ligi 20 aastat (vt ülevaadet Mihkla 2009), selle aja jooksul on ühelt poolt juurde tulnud uusi sünteesimeetodeid

* Artikli valmimist on toetanud Euroopa Liit Euroopa Regionaalarengu Fondi kaudu (Eesti-uuringute Tippkeskus), see on seotud Eesti Haridus- ja Teadusministeeriumi uurimisprojektiga IUT35-1 "Kõnestiilid, lauseprosoodia ja fonoloogiline varieerumine: kirjeldus, teooria ja modelleerimine".

ja teisalt on arenenud ka teksti lingvistiline analüüs. Viimane sisaldab morfoloogilise analüüsi ja liitsõnade tuvastamise kõrval ka täht-häälik teisendust, millega ortograafilisele tekstile lisatakse juurde info häälduse kohta, nt mis vältes tuleb sõna hääldada, kus asub rõhk või milliseid häälikuid palataliseerida. Välte, rõhu ja palatalisatsiooni määramisel tuginetakse “Eesti õigekeelsussõnaraamatule” (ÕS 2013), mis on eesti kirjakeele normi alus. Välte märkimisel kasutab õigekeelsussõnaraamat võimalust ligi 300 sõna puhul lubada nii teise- kui kolmandavälte hääldust. See on ka loomulik, kuna välde on murdeti ja isikuti üks kõige kõikuvam ja vahelduvam sõnastruktuuri komponent ja seda vaatamata asjaolule, et teise ja kolmanda välte kontrastil on eesti keele süsteemis täita oluline roll. Siiski võib see kontrast mõningates morfoloogilise süsteemi osades funktsionaalselt neutraliseeruda, mis ei avaldu teise ja kolmanda välte vahepealse hääldusena, vaid võib avalduda teise- ja kolmandavälte paralleelvormide esinemisena (Hint 1968: 78–79). Kui inimkõnelejale ei tekita kahesugused vältenormingud probleeme (või äärmisel juhul tekib segadus muutevormide moodustamisel), siis automaatne teksti analüüs ei tule kahesuguse väljundiga toime. Vajalik on kahe võimaliku variandi puhul anda teksti analüüsiprotsessis ühele variandile eelistus, mistõttu on oluline välja selgitada kõnelejate seas levinum ja eelistatum variant.

Peamiselt on meil kaks uurimisküsimust.

1. Kas kuuldeliselt või kestussuhete põhjal saab välja tuua sõnad, mille hääldamisel domineerib üks või teine välde?
2. Kas sama silbistruktuuriga ning samasse sõnaliiki kuuluvate sõnade puhul avalduvad ühtsed hääldustrendid, st kas sarnase struktuuriga sõnade puhul on tendents hääldada neid rohkem ühes või teises vältes?

Varasemalt on varieeruva vältega sõnade vastu huvi tuntud eelkõige vajadusest muuta normeerimine süsteemsemaks (Hint 1968) või rahvapärasemaks (Nurkse 1938), huvist otsida seoseid sõna silbistruktuuritüübi ja välte omistamise vahel (Päll 1986). Meie uurimus sarnaneb vast kõige enam Nele Salveste (2006) väikese uurimusega, kus ta kahe keelejuhiga tehtud lugemiskatse põhjal analüüsis varieeruva vältega sõnade foneetilisi parameetreid ja hindas tulemusi võrdluses vältenormingutega.

Käesolev artikkel on jätkuks hääldamiseksperimenti tutvustavale käsitlusele (Kalvik, Piits 2015), milles kirjeldasime fonoloogilist varieerumist kolme nähtuse näitel. Siinses kirjutises keskendume teise ja kolmanda välte hääldamisele nn varieeruva vältega sõnades. Piloottuurimusega võrreldes on keelejuhtide arv kahekordistunud ning oleme uurimusse vokaalvälte sõnade kõrval kaasanud ka muu struktuuriga vältevahelduslikke sõnu.

2. Materjal ja meetod

Välteuringuteks valiti sõnad, mille kogu paradigma häälduses või mõne muutevormi häälduses lubab ÕS 2013 homograafset sõnavormi hääldada nii teises kui kolmandas vältes. Selliseid sõnu on ÕS-is 295. Nende hulgas on palju samast tüvest moodustatud liitsõnu ja tuletisi, nt sõna *kangelane* kõrval on esindatud ka *kangelasema*, *kangelaskalm*, *kangelaslik*, *kangelaslikkus*, *kangelashugu*, *kangelasmeel*. Kuna kõigi 295 sõna uurimine ei ole võimalik ega otstarbekas, siis valisime

sagedussõnastiku (Kaalep, Muischnek 2002) abil nende seast välja 47 kõige suurema esinemissagedusega sõna, arvestades ühtlasi, et sama tüvi ei korduks (erandina võeti uurimusse mõlemad suure esinemissagedusega sõnad *andetu* ja *andekas*). Sageduspõhimõtet rakendame eelkõige seetõttu, et uurimistulemuste kasutusest tuleb praktiline kasu teksti automaatsel analüüsil on seda suurema kaaluga, mida sagedamini kriitilise vältega sõnad tekstianalüüsi satuvad (Kalvik, Piits 2015). Siinne uurimus põhineb 23 keelejuhiga läbi viidud lugemiseksperimentil, mille käigus luges iga keelejuht ette 52 lauset. Need laused sisaldavad 47 varieeruva vältega sõna ehk sihtsõna. Kõigi lugejate peale kokku kogunes 1080 hääldusjuhtu, mida siinses artiklis analüüsitakse. Uuritavad sihtsõnad on paigutatud võimalikult loomulikku lausekonteksti (näites 1 on varieeruva vältega sihtsõnad paksus kirjas).

(1) Ka **soodus** asi võib olla **maitsekas** või heldima panev.

Sõnade seas on nii käänd- kui muutumatuid sõnu. Sama eksperimendiga kogusime materjali ka sõnaalgulise *h* häälduse ning varieeruva palatalisatsiooniga sõnade kohta. Üksikasjalikult on kogu eksperimenti ning sellesse valitud materjali kirjeldatud artiklis “Lugemiseksperiment fonoloogilise varieerumise uurimiseks” (Kalvik, Piits 2015).

Katsel osales 15 naist ja 8 meest, kes kõik kõnelevad eesti keelt emakeelena. Keelejuhtide keskmine vanus on 45 aastat ning nemad või nende vanemad on pärit Eesti erinevatest piirkondadest. Salvastused tehti Eesti Keele Instituudi helistuudios salvastusprogrammiga KJIK, mis kuvas keelejuhtidele laused ükshaaval arvuti-ekraanile ja lõi ka heli kohe lausepikkusteks lõikudeks. Salvastused segmenteeriti eestikeelse kõne autosegmenteerijaga.¹ Programmis Praat (Boersma, Weenink 2014) lõi me lisaks hääliku-, sõna- ja lausetasandile uuritavate nähtuste jaoks omaette tasandid ja nihutasime vajadusel käsitsi segmentide piire. Vältetasandile märkisime kestussuhete arvutamiseks pearõhulise silbi tuuma, kooda ja järgneva silbi tuuma.

Kõigepealt arvutasime automaatselt sihtsõnades silpide kestussuhted ning rühmitasime sõnad kestussuhete järgi vältegruppideks. Gruppide moodustamisel otsustasime, et sõna kuulus ühte või teise vältegruppi, kui selge enamus ehk vähemalt $\frac{2}{3}$ keelejuhtidest hääldas sõna selles vältes. Saadud tulemusi kõrvutasime muudes uurimustes leiduvate andmetega. Seejärel võrdlesime kestussuhteid kuuldelise hinnangu tulemustega. Kolmandaks analüüsisime varieeruva vältega sõnade jagunemist tüüpidesse, mille loomisel arvestati sõna silbistruktuuritüüpi ja muid tunnuseid nagu sõnaliiki või tuletusliidete olemasolu.

3. Teise ja kolmanda välte kestussuhted

Uurimise all olevad sõnad on kolme erineva struktuuriga. Välte avaldumiselaks sõnas on pearõhulise silbi tuum – kooda olemasolul ka kooda – koos sellele järgneva rõhutu silbi tuumaga ning vältet iseloomustav kestussuhe arvutatakse vastava jagatisena. Allpool toodud silbistruktuurides on kestussuhte arvutamisel osalevad osised alla joonitud. Andmestus on kolme tüüpi sõnu:

- 1) vokaalvälteelised (C)VVCV(C)-struktuuriga sõnad, nt *kiire* ja *politseinik*. Pearõhulises silbis kannab vältet selles struktuuris pikk vokaal või diftong;

¹ <http://www.phon.ioc.ee/dokuwiki/doku.php?id=projects:tuvastus:est-align.et> (15.6.2016).

- 2) konsonantvärtelised (C)VCCV(C)-struktuuriga sõnad, nt *ammu*, *ahne*, *hirmus*, kus pearõhulises silbis on välte kandjaks lühike vokaal koos konsonandiga. Sõnades, kus silbipiiril esines geminaat (*mullu*, *enne* jt) arvestasime esimesse silpi kuuluvaks pool geminaadist²;
- 3) vokaal- ja konsonantvärtelised (C)VVCCV(C)-struktuuriga sõnad, nt *saa-tanlik*, *kaitsetu*, *äärmuslikult*, kus pearõhulises silbis kannab väldet pikk vokaal või diftong koos konsonandiga.

Kestussuhete arvutamisel ja võrdlemisel on aluseks võetud senised tulemused, mis uurimusiti varieeruvad. Tabelis 1 on esitatud uuemates uuringutes erinevate sõnastruktuuride põhjal arvatud keskmised kestussuhted. Klassikaliselt on peetud teisele vältele omaseks keskmiseks kestussuhteks 1,5 ning kolmandat väldet iseloomustab 2-st suurem väärtus (Lehiste 1997). Uuem, ülevaatlik käsitlus juhib neile suhtarvudele tähelepanu kui suurtele üldistustele; sõnastruktuuriti/väldeti on kestussuhete varieeruvus suur, hälbides antud arvudest (Asu jt 2016: 134–135).

Tabel 1. Teise ja kolmanda välte keskmised kestussuhted erinevates uurimustes. Sõnastruktuur näitab välte avaldumisala (pearõhulist ja sellele järgnevat rõhutat silpi)

Uurimus	Vokaalvärteline (C)VVCV(C)		Konsonantvärteline (C)VCCV(C)		Vokaal- ja konsonantvärteline (C)VVCCV(C)	
	2. välde	3. välde	2. välde	3. välde	2. välde	3. välde
Sahkai jt 2015	1,8–2,03	2,71–3,49	1,41–1,71	2,38–3,12		
Lippus jt 2013	1,8	2,6	2,3	3,4		
Eek, Meister 2003	1,49	2,92	1,37	2,96	1,9	3,9
Siinne uurimus	1,8	2,8	1,7	2,8	1,8	3,5

Kõik tabelis 1 esitatud andmed põhinevad etteloetud kõnel, v.a Lippuse jt (2013) uurimus, mille aluseks on eesti keele spontaanse kõne foneetiline korpus (EKSKFK)³. Sahkai jt (2015) uurimuses on keskmistatud kestussuhted esitatud vahemikuna, kuna kestussuhteid arvutati selles käsitluses vastavalt sihtsõna fookustingimustele ning sõnade arvule lauses. Tabeli andmed näitavad, et teise välte keskmised kestussuhted on neis uurimustes enamasti suuremad kui 1,5 ning kolmanda välte omad vastavalt suuremad kui 2,5, ulatudes ka üle kolme. Oma töös toetume ka Georg Liivi käsitlusele, mis põhineb etteloetud kõne vokaalvärtelistel sõnadel. Liivi analüüsi järgi esinevad teisevärteliste sõnade kestussuhted vahemikus 1,0–2,0 (Liiv 1961: 480). Erinevatele uurimustele toetudes võtsime oma töös hääldusjuhtude jagamisel välterühmadeks teise välte ülempiiriks 2,0 ja kui kestussuhe oli 2,1, pidasime seda kolmandavärteliseks häälduseks.

Pärast hääldusjuhtude rühmitamist teise- ja kolmandavärtelisteks arvasime struktuuritüüpide kaupa ka teises ja kolmandas vältes sõnade keskmised kestussuhted (tabelis 1 viimasel real). Vokaalvärteliste ja konsonantvärteliste sõnade keskmised kestussuhted tulid meie uurimuses väga sarnased, aga (C)VVCCV(C)-struktuuriga sõnade puhul võib märgata suuremat kolmanda välte keskmist kestussuhet. Ka Eek ja Meister (2003) ning Salvete (2006) on täheldanud selle struktuuriga sõnades suuremat kolmanda välte kestussuhet.

² Sellistes struktuurides on sageli silbipiiri asukohaks määratud teisevärteliste sõnade puhul pool, kolmandavärteliste puhul ¾ geminaadist (Asu jt 2016: 137). Kuna meie uurimuses on tegu varieeruva vältega sõnadega ning sõna välde ei ole ette teada (ega kuuldeliselt ette määratud), siis määrasime silbipiiri kõikidel juhtudel geminaadi keskele.

³ Eesti keele spontaanse kõne foneetiline korpus <http://www.keel.ut.ee/et/foneetikakorpus> (29.09.2016)

Nagu näha tabelist 2, on meie uurimuses teisevärteliste sõnade kokkuvõtlik keskmine kestussuhe 1,8 ja kolmandavärteliste oma 2,9. Need tulemused on sarnased meie pilootuurimuse tulemustega, mis põhinevad 11 keelejuhi vokaalvärtelistel sõnadel ja kus keskmised kestussuhted värteti olid vastavalt 1,8 ja 2,7 (Kalvik, Piits 2015: 63).

Tabel 2. Varieeruva värttega sõnade kestussuhted kokku võetuna (3. ja 5. veerg). Teises veerus on toodud, mitu keelejuhti 23-st on hääldanud sõna teises värttes (väiksema kestussuhtega kui 2,10). Neljas veerg näitab, mitu keelejuhti on hääldanud sõna kolmandas värttes (kestussuhe 2,10 või suurem)

Sõna	2. värtte		3. värtte	
	hääldajate arv	keskmine kestussuhe	hääldajate arv	keskmine kestussuhe
<i>teisal</i>	21	1,7	2	2,4
<i>mullu</i>	21	1,6	2	2,4
<i>kirju</i>	21	1,6	2	2,2
<i>ilma</i>	19	1,5	4	2,5
<i>võimekas</i>	18	1,7	5	2,5
<i>hilja</i>	17	1,7	6	2,7
<i>toimekas</i>	16	1,8	7	2,9
<i>kaine</i>	16	1,7	7	2,8
<i>soodus</i>	15	1,9	8	2,4
<i>täiuslik</i>	15	1,9	8	2,5
<i>terve</i>	15	1,6	8	2,5
<i>rõõmus</i>	14	1,8	9	2,8
<i>ungarlane</i>	14	1,8	9	2,3
<i>augustis</i>	14	1,8	9	2,5
<i>kiire</i>	12	1,7	11	2,4
<i>tallinlane</i>	12	1,6	11	2,5
<i>nõiduslik</i>	11	1,9	12	2,9
<i>ilmetu</i>	11	1,6	11	3,0
<i>politseinik</i>	10	1,9	13	2,7
<i>hirmus</i>	10	1,8	13	2,8
<i>võimetu</i>	10	1,7	13	2,5
<i>meetod</i>	7	1,8	16	3,0
<i>keiserlik</i>	8	1,6	15	2,8
<i>enne</i>	6	1,8	17	2,8
<i>jaanuar</i>	6	1,7	17	3,0
<i>ammu</i>	5	1,9	18	3,0
<i>piknikuga</i>	5	1,9	18	3,2
<i>ihne</i>	5	1,7	18	3,0
<i>teaduslikes</i>	5	1,7	18	3,2
<i>kärme</i>	4	1,9	19	3,0
<i>ahne</i>	4	1,6	19	3,1
<i>kuulus</i>	3	1,9	20	2,8
<i>andekas</i>	3	1,8	20	2,7

<i>nooruslik</i>	3	1,8	20	3,2
<i>looduslik</i>	3	1,6	20	3,0
<i>maitsekas</i>	2	2,0	21	3,4
<i>peenelt</i>	2	1,9	21	2,7
<i>ümbrikus</i>	2	1,9	21	3,1
<i>rahvuslik</i>	2	1,9	21	2,9
<i>kangelane</i>	2	1,9	21	3,2
<i>aulasse</i>	2	1,8	21	3,1
<i>kaitsetu</i>	2	1,7	21	3,1
<i>äärmuslikult</i>	1	1,8	22	4,1
<i>pealik</i>	0		23	3,7
<i>haiguslik</i>	0		23	3,4
<i>saatanlik</i>	0		23	3,4
<i>saatuslik</i>	0		23	3,9
Keskmine kestussuhe		1,8		2,9

Tabelis 2 esitatud sõnad võib kestussuhete ning hääldajate arvu järgi jagada kolme tinglikku rühma. **Esimene rühm** on kõige väiksem, sisaldades 8 sõna: *teisal, mullu, kirju, ilma, võimekas, hilja, toimekas, kaine*. Selles rühmas on sõnad, mille puhul eelistatakse teist välde. Siin ei ole küll ühtegi sõna, mida kõik keelejuhid hääldaksid vaid teisevärtelisena, kuid esile võib tuua, et sõnade *teisal, mullu* ja *kirju* 23 hääldusjuhust on 21 ehk valdav enamik teisevärtelised. Ülejäänud sõnade puhul on teisevärtelisi hääldusi vähem, aga need moodustavad ikka vähemalt $\frac{2}{3}$ kõigist selle sõna hääldusjuhtudest.

Kõrvutame siinse tabeli tulemusi eesti keele spontaanse kõne foneetilises korpus (EKSKFK) esinevate andmetega. Võrdlemisel tuleb meeles pidada, et sealse andmestuse pole vältemäärang saadud kestussuhteid arvestades, vaid selle on kuulde- lise mulje põhjal määranud märgendaja. Korpusel leiame järgmised andmed: *hilja* (5) on kõigil juhtudel märgitud teisevärteliseks ja *ilma* (44) on samuti ühe erandiga teisevärteline. Varieeruva vältega sõnu foneetiliselt analüüsinud Salveste (2006) on arvutanud välteid iseloomustavaid kestussuhteid kahe keelejuhi kõnes, sh ka samades sõnades, mida meie käsitleme. Salveste mõõtmised näitavad, et *teisal* on teisevärteline, *kirju* esineb nii teises kui kolmandas vältes (Salveste 2006: 58–63).

Teise rühma kuuluvad sõnad *soodus, täiuslik, terve, rõõmus, ungarlane, augustis, kiire, tallinlane, nõiduslik, ilmetu, politseinik, hirmus, võimetus, meetod* ja *keiserlik* (15 sõna). Nende, tabelis hallile taustale paigutatud sõnade puhul on raske rääkida ühe või teise välte eelistamisest, silpide kestussuhete arvutused näitavad, et keelejuhid on hääldanud sõnu üsna võrdselt mõlemas vältes. Nii näiteks esineb sõna *nõiduslik* 11 keelejuhil teise- ning 12-l kolmandavärtelisena, sõnade *kiire* ja *tallinlane* puhul on need arvud vastupidised. *ilmetu* esineb võrdselt mõlemas vältes (andmeid on selle sõna puhul 22 keelejuhilt, sest üks keelejuht hääldas *ilmetu* asemel *inetu*).

EKSKFK-s on andmeid sõnade *meetod, hirmus, kiire, tallinlane, rõõmus, ungarlane, augustis* ning *terve* kohta. Olgu järgmisena nimetatud, kummas vältes need sõnad korpusel rohkem esinevad. Sõna *hirmus* (18) esineb 12 juhul kolmandas vältes, liigitudes pigem kolmandavärteliseks. Kumbki välde ei domineeri sõnade

meetod, terve ja ungarlane hääldusjuhtudes. Sõna *terve* (61) esineb 39 korral ja *meetod* (5) esineb kolmel korral kolmandas vältes. Sõna *ungarlane* (8) on viiel juhul teisevälteline.

Teine välde domineerib sõnade *augustis, rõõmus, tallinlane* ja *kiire* häälduses. Sõna *augustis* (4) esineb kolmel korral teises vältes. Ainult teisevältelised on nii *rõõmus* (3) kui üksiku esinemisjuhuga *tallinlane*. Kui *kiire* on meie uurimuses kestussuhete järgi pea võrdselt mõlemas vältes hääldatav, siis EKSKFK-s esineb ta vaid kahel korral 21-st kolmandas vältes, olles ülejäänud juhtudel teisevälteline. Seega ei ole *kiire* korpuse andmeil mitte vahelduva vältega, vaid teisevälteline sõna. Ka Salvete keelejuhtide kõnes on *kiire* teisevälteline, samuti *rõõmus* ning *soodus; terve* varieerub, esinedes mõlemal keelejuhil mõlemas vältes (Salveste 2006: 58–63).

Kolmanda rühma, suurima, moodustavad sõnad, mida vähemalt $\frac{2}{3}$ keelejuhtidest on hääldanud kolmandas vältes. Rühma võib seega pidada tingimisi kolmandaväلتeliste sõnade rühmaks. Siia kuulub 24 sõna: *enne, jaanuar, ammu, piknikuga, ihne, teaduslikes, kärke, ahne, kuulus, andekas, nooruslik, looduslik, maitsekas, peenelt, ümbrikus, rahvuslik, kangelane, aulasse, kaitsetu, äärmuslikult, pealik, haiguslik, saatanlik ja saatuslik*. Neli viimast neist on sellised, mida kõik keelejuhid on hääldanud vaid kolmandaväلتelisenä. Ülejäänud sõnade puhul esineb kestussuhete järgi otsustades ka teiseväلتelisi hääldusi.

EKSKFK-s esinevad järgmised sõnad (sulgudes esinemisjuht): *äärmuslikult* (1), *rahvuslik* (1), *jaanuar* (1), *peenelt* (3), *kuulus* (12), *ammu* (31), mis kõik on märgitud kolmandaväلتelisteks. Sõna *enne* (170) on 108 korral märgitud kolmandaväلتeliseks. Lisaks esineb korpuses sõna *teaduslik*, mille kaheksast esinemisjuhust on viis kolmandas vältes. Sõna *ahne* leidub ühel korral teiseväلتelisenä. Salvete (2006: 58–63) uurimuses olid vaatluse all ka selle rühma *ammu, ahne* ning *aula* (sg g), mis esinesid mõlemal keelejuhil kolmandas vältes. Võrdlus näitab, et siinse uurimuse kolmanda rühma sõnu hääldatakse arvatud kestussuhete põhjal enamasti (v.a *ahne*) samamoodi kui teistes andmestutes.

Kokkuvõtlikult saab esitatud rühmade kohta öelda järgmist. Meie uurimuses kestussuhete ning häälduseelistuste järgi väلterühmadesse jagatud sõnade välde on üldjoontes sama vältega ka muudes andmestutes, kus küll meid huvitavate sõnade kohta andmeid piisavalt ei leidu. Leitud väلterühmade põhjal on võimalik teha otsustusi, millist väldet keelejuhid eelistavad ning rakendada neid otsustusi kõnesünteesis. Siiski on sellist meetodit kasutades sõnade hulk, mille puhul väلte-eelistust ei õnnestunud välja selgitada, liiga suur. Selline tulemus ärgitab otsima, milliseid tunnuseid või omadusi saaks veel väلte-eelistuste selgitamisel arvesse võtta.

4. Vältes kestussuhet ja kuuldelist hinnangut arvestades

Lisaks kestussuhete järgi väلte määramisele hindasime ka iga hääldatud sõna väldet kuuldeliselt. Kuuldeline hinnang võib sellisel kujul sisaldada mingil määral subjektiivsust, sest väلtemääramisel ei kasutanud me erinevaid hindajaid, vaid määrasime sõna väلte ise. See kannab endas riski eelkõige kõnetaju kategoriaalsuse tõttu (Schouten jt 2003, Salvete 2013) ehk väga palju on kuulaja keeletundest, kuidas ta vastavat akustilist signaali kategoriseerib. Sama hääldusjuhtu võis kuulata piiramatu arv kordi. Hindamise käigus selgus, et mõnel juhul olenes väلtemäärang

sellest, kas hindaja kuulab ainult määratavat sõna või sõna koos lausekontekstiga. Esines ka juhte, kus hindaja ei suutnudki vältet määrata, siis jäi see hääldusjuht kuuldelise hinnangu arvutustest välja. Probleemidele vaatamata annab kuuldeline hinnang kestuse automaatsele arvutamisele lisaks teatava kontrollivõimaluse ja võimaluse vältet hindamisel arvestada ka teisi parameetreid (nt põhitooni), mis temporaalse struktuuri kõrval on osutunud vältet tajus samuti oluliseks (Lippus 2011).

Enamike sõnade puhul langesid hinnangud arvatud kestussuhetega täiesti või enam-vähem kokku (vt tabel 3), nt sõna *mullu* häälduse kestussuhete arvutustest selgus, et 91% lugejatest hääldas seda sõna teisevälteliselt ning sarnane protsent ehk 95% hääldatud sõnadest määrati teisevälteliseks ka kuuldelise hinnanguga. Sarnaselt kokkulangevad on näiteks ka tabeli lõpus esinevatele sõnadele *pealik* ja *saatanlik* antud hinnangud: nii kestussuhte kui kuulmistaju järgi hinnatult hääldasid kõik lugejad neid sõnu kolmandavältelistena. Siiski võib välja tuua sõnad *maitsekas*, *andekas* ja *kaine*, mille puhul ainult silbikestuste järgi saadud hinnang osutus vastupidiseks kuuldelisele hinnangule.

Tabel 3. Väitemääranud kestussuhte ja kuuldelist hinnangut arvestades. Kolmandas veerus on esitatud 2. vältet hääldatud sõnade protsent, arvestades kuuldelist hinnangut ja neljandas veerus 2. vältet hääldatud sõnade protsent, arvestades kestussuhte. Hallil taustal on esitatud need väärtused, mille puhul ühes või teises vältet hääldusjuhud pole saavutanud $\frac{2}{3}$ suurus ülekaalu

	Sõna	Kuuldeliselt 2. vältet hääldatud (%)	Kestussuhte järgi 2. vältet hääldatud (%)
1.	<i>ilma</i>	100	83
2.	<i>teisal</i>	96	91
3.	<i>mullu</i>	95	91
4.	<i>võimekas</i>	91	78
5.	<i>toimekas</i>	91	70
6.	<i>soodus</i>	91	65
7.	<i>rõõmus</i>	91	61
8.	<i>tallinlane</i>	91	52
9.	<i>täiuslik</i>	91	65
10.	<i>hilja</i>	83	74
11.	<i>ungarlane</i>	83	61
12.	<i>maitsekas</i>	83	9
13.	<i>andekas</i>	81	13
14.	<i>kirju</i>	78	91
15.	<i>kiire</i>	78	52
16.	<i>augustis</i>	77	61
17.	<i>võimetu</i>	70	43
18.	<i>ilmetu</i>	64	50
19.	<i>hirmus</i>	41	43
20.	<i>politseinik</i>	38	43
21.	<i>kaitsetu</i>	36	9
22.	<i>ümbrikus</i>	35	9
23.	<i>teaduslikes</i>	32	22
24.	<i>keiserlik</i>	30	35

25.	<i>enne</i>	26	26
26.	<i>meetod</i>	24	35
27.	<i>kaine</i>	22	70
28.	<i>terve</i>	22	65
29.	<i>nõiduslik</i>	22	48
30.	<i>kärme</i>	17	17
31.	<i>aulasse</i>	17	9
32.	<i>piknikuga</i>	9	22
33.	<i>looduslik</i>	9	13
34.	<i>ammu</i>	9	22
35.	<i>jaanuar</i>	9	26
36.	<i>ihne</i>	9	22
37.	<i>ahne</i>	9	17
38.	<i>rahvuslik</i>	9	9
39.	<i>kangelane</i>	9	9
40.	<i>peenelt</i>	4	9
41.	<i>haiguslik</i>	4	0
42.	<i>saatuslik</i>	4	0
43.	<i>kuulus</i>	0	13
44.	<i>nooruslik</i>	0	13
45.	<i>äärmuslikult</i>	0	4
46.	<i>pealik</i>	0	0
47.	<i>saatanlik</i>	0	0

Sõnu *andekas* ja *maitsekas* võib kestussuhte järgi pidada kolmandavärtelisteks, samas kui kuuldelselt määrasime neid üle 80 protsendil juhtudest teisevärtelisteks. Iga hääldusjuhtu eraldi vaadates selgub, et sõna *maitsekas* 23-st hääldusjuhust läheb ainult neljal juhul kestussuhte järgi antud hinnang kokku kuuldelse hinnanguga. Nt sõna *maitsekas* kõik kuus suurima kestussuhtega (5,4; 4,4; 4,2; 3,9; 3,8; 3,7) hääldusjuhtu määrasime kuuldelselt ikka teisevärtelisteks. Sõna *kaine* võib pidada kestussuhte järgi teisevärtelisteks, samas kui kuuldelse hinnangu järgi on ta kolmandavärteline. Ligi pooltel hääldusjuhtudel ei lähe kestussuhte järgi määratud välde kokku kuuldelse vältetäringuga: nt klassikaliselt teisevärtelise kestussuhtega 1,5 hääldusjuhud määrati koguni neljal korral kuuldelselt kolmandavärtelisteks.

Sellised lahknevused, mis tulevad esile erineva struktuuriga sõnades, näitavad ilmekalt, et silbituumade kestussuhe on küll peamine, aga sugugi mitte ainus väldet määrav tunnus (Lehiste 1997, Lippus 2011 jt). Lisaks sihtsõna temporaalsele ja tonaalsele struktuurile võib kuuldelist hinnangut mõjutada aga ka kogu lause intonatsioon. Välte hindamisel panime tähele, et otsustus võib muutuda olenevalt sellest, kas kuulata ainult sihtsõna või kogu lauset.

Lisaks neile kolmele sõnale, millele antud hinnang oli kestussuhte arvatades ja kuuldelselt hinnates vastupidine, esineb ka sõnu, mis kuuldelselt on kindlalt teisevärtelised (nt *soodus*, *rõõmus*, *tallinlane*, *täiuslik*, *ungarlane*) või kindlalt kolmandavärtelised (nt *terve*, *nõiduslik*, *meetod*), aga kestussuhte järgi jäävad nn hallile alale ehk kummaski vältes olevad hääldusjuhud ei ole saavutanud $\frac{2}{3}$ suurust ülekaalu. Seega kui teises vältes hääldatud sõnade protsent on üle 66,6%,

on sõna määratud teisevärteliseks, ja kui teises välteliseks. Kui võrrelda erinevate sõnade liigitumist ühes või teises välteliseks. Kui võrrelda erinevate sõnade liigitumist ühes või teises välteliseks, siis selgub, et kuuldelise hinnangu põhjal jagunevad sõnad selgemalt teise- või kolmandavärteliseks kui kestussuhte järgi arvutades. Kui kestussuhte arvutades jäävad nn hallile alale 15 sõna (vt tabel 2), siis kuuldelise hinnangu põhjal ei liigitu ühe või teise välteliseks alla ainult 5 sõna: *ilmetu, hirmus, politseinik, kaitsetu, ümbrikus*. Neist kaks viimast liigituvad kestussuhte järgi kindlalt kolmandavärteliseks, aga kuuldelise hinnangu järgi jäävad kolmanda välteliseks mõned protsendid puudu, et 66% suurus ülekaalu saavutada.

Kuuldelise hinnangu järgi domineeris **teisevärteline** hääldus sõnades: *ilma, teisel, mullu, võimekas, toimekas, soodus, rõõmus, tallinlane, täiuslik, hilja, ungarlane, maitsekas, andekas, kirju, kiire, augustis, võimetu* ja **kolmandavärteline** hääldus sõnades *teaduslikes, keiserlik, enne, meetod, kaine, terve, nõiduslik, karme, aulasse, piknikuga, looduslik, ammu, jaanuar, ihne, ahne, rahvuslik, kangelane, peenelt, haiguslik, saatuslik, kuulus, nooruslik, äärmuslikult, pealik, saatanlik*.

Kindlasti peab kuuldelise hinnangu puhul arvestama, et vältelise määratlus võib sõltuda hindaja keeletajust ja päritolupiirkonnast. Näiteks on selgunud, et Lõuna-Eestist pärit inimestel on vältelise eristamisel oluline vaid silpide kestussuhe ehk temporaalne struktuur, Kesk- ja Lääne-Eestist pärit inimestele on aga vältelise tajumisel oluliseks kriteeriumiks ka toonikontuur (Lippus, Pajusalu 2009). Seega võiks kuuldelist vältelise määrangut usaldada enam, kui iga hääldusjuhtu hindaks mitu hindajat.

5. Varieeruva vältelise sõnad silbistruktuuritüüpide kaupa analüüsilt

Järgnevalt vaatleme, kuidas vältelise omistamine sõltub sõna silbistruktuuritüübist ja muudest tunnustest nagu tuletusliidetest või sõnaliigist. Otsime vastust küsimusele, kas ühte tüüpi kuuluvad varieeruva vältelise sõnad käituvad sarnaselt, nt kas saab välja tuua sarnaseid varieerumise tendentse ühe tüüpi sõnades. Ühtse hääldussuundumuse korral saaks uuritud sõnade puhul märgatud trendi laiendada tervele tüübile ning kõnesünteesis rakendada.

Peeter Pälli 30 aastat tagasi ilmunud raamatus “Eesti noomeni silbistruktuur ja aktsent” (Päll 1986) kirjeldatakse väga põhjalikult vältelise sõltumist käändsõnade silbistruktuurist. Siinses artiklis lähtumegi Pälli loodud üld- ja alltüüpidest, mille loomisel on arvestatud sõna nimetava vormi silpide arvu, silpide tüüpi (nt pikk kinnine, lühike lahtine), sõnaliiki, sõna sisehäälikuid ja sõnamoodustust.

Meie artiklis kirjeldatud sõnad saab jagada kaheksasse gruppi.

1. **Kahesilbilised pika esisilbiga lühivokaaliga lõppevad käändsõnad** saavad *lahja*-üldtüübis teise vältelise. Pälli käändsõnadele loodud tüüpide alusel võib liigitada siia sama silbistruktuuriga muutumatud sõnad: *ammu, enne, hilja, ilma, mullu*. Nende määrsõnade vältelise muudatused ajas (nt *enne* ja *hilja* on normitud varasemalt kolmandavärteliseks) näitavad, et tüübis on varieeruvusi. Meie uuringuga selgub, et sõnu *hilja, ilma* ja *mullu* võib pidada teisevärteliseks ja sõnu *ammu* ning *enne* kolmandavärteliseks (vt tabel 3).

2. **Kahesilbilised e-lõpulised adjektiivid** *ahne, ihne, kaine, kiire, kärme* ja *terve* kuuluvad koos teiste e-lõpuliste adjektiividega kolmandavärtelistena *lahja*-üldtüübi *raske*-alltüüpi (Päll 1986: 22). Normimise traditsioonis on neid varasemates ÕS-ides normitud teisevärtelisteks (ÕS 1960; ÕS 1976). Alates ÕS 1999 on lubatud mõlemad välted. Rein Nurkse (1938a; 1938b) on juba 1938. aastal selle sõnagrupi vältel küsitavale teisevärtelisusele tähelepanu pööranud. Ta toob välja põhjuse, miks see väikesearvuline e-lõpuliste adjektiivide grupp on suurema ja produktiivsema kolmandavärtelise e-lõpuliste adjektiivide grupi kõrval veel säilitanud osaliselt teise vältel. Erinevuse on tinginud häälikuajaloolised tegurid: kui kolmandavärtelised e-lõpulised adjektiivid (nt *haige, raske, tahke*) on tekkinud pikematest *eda*-lõpulistest vormidest ja saanud oma kolmandavärtelisuse järgsilbivokaalide kontraktsiooni tõttu, siis meid huvitav väike grupp adjektiive põlvneb algselt konsonandiga lõppenud sõnavormidest, mis nimetavas käändes tingis teise vältel (Nurkse 1938a: 46). Meie uurimusest selgub, et vaatamata häälikuajaloolistele põhjustele kuuluvad kuuldelise hinnangu järgi kõik sõnad v.a *kiire* kolmandavärteliselt häälduvate sõnade hulka. Silbituumade kestussuhte järgi antud hinnangud lähevad küll kahe sõna puhul kuuldelise hinnanguga vastuollu: sõnade *terve* ja *kaine* esisilbi ja järgsilbi kestussuhete arvutuste põhjal on suurem osa katses osalenuid hääldanud neid sõnu teisevärtelistena. Sõna *kiire* esineb aga mõlema määratluse kohaselt eelkõige teisevärtelisena. Võimalik põhjus peitub selles, et kui ülejäänud sõnad esinevad lugemiskatses adjektiividena, siis *kiire* on lauses substantiivina (vt näidet 2, kus *kärme* esineb adjektiivina, aga *kiire* substantiivina).

(2) Muidu nii **kärme** mees astus aegamisi, hoomamata, et teistel on **kiire** ning külm.

Nurkse (1938b: 202) kirjutabki, et häälikuajaloolistel põhjustel teisevärteliste adjektiivide väikese grupi, sõnade *ahne, ihne, kaine, kiire, kärme* ja *terve* siirdumist kolmandavärteliste hulka mõjutab vajadus vahet teha adjektiividel ja substantiividel. Tema sõnul on e-lõpulised substantiivid eelkõige teisevärtelise nominatiiviga (nt *hoone, raibe, saade, võte*), samal ajal kui adjektiivid on enamasti kolmandavärtelised (nt *kange, lääge, sirge, uhke, värsk*). Seega peaksid adjektiivid *ahne, ihne, kaine* ja *kärme* liikuma produktiivsesse kolmandavärteliste sõnade rühma ning sõnade *kiire* ja *terve* puhul peaks välde sõltuma sellest, kas sõna kasutatakse adjektiivi või substantiivina. Niisiis võib e-lõpuliste adjektiivide puhul ennustada liikumist kolmandavärteliste sõnade gruppi ning kui lugemiseksperimenti oleks olnud kaasatud sõna *kiire* adjektiivina, siis võiks see samuti olla kolmandas vältel. Kõrvutades saadud tulemusi Salvete väikese lugemiskatsega, kus ta leidis, et *kiire* on teisevärteline (vt alapeatükk 3), tuleb siiski rõhutada, et kuigi ta oma sõnul uurib *kiire* kui adjektiivi vältel, on ta lugemisteksti paigutanud selle sõna substantiivina (Salveste 2006: 56).

Sama *raske*-alltüübi alla kuulub *aula*, mis *la*-liiteliste tuletiste grupis peaks saama kolmanda vältel nagu *haigla, parkla* ja *käimla* (Päll 1986: 23). Meie uuringust selgub, et sõna hääldatigi 91% juhtudest kolmandavärtelisena.

3. **Kahesilbilised lühivokaali ja -konsonandiga lõppevad käändsõnad** saavad üldjuhul kolmanda välte (*noomen*-tüüp), aga ÕS-is on leenisklusiiliga lõppevad sõnad (nt *keerub*, *hertsog* ja *talmud*) saanud ka teise välte. Siia gruppi kuuluv *meetod* on veel 1976. aasta ÕS-is olnud määratud ainult teisevärteliseks. Meie uuringus on sõna nii kuuldelise kui kestussuhete määrangu järgi kolmandas vältes.

Sõna silbistruktuuri arvestades võiks *noomen*-üldtüüpi kuuluda adverb *teisal*, mis peaks siis saama kolmanda välte (kuigi sõnaliigist sõltuvalt võib tekkida ka eraldi alltüüp, mis käitub teistmoodi). Veel 1976. aasta ÕS-is ongi lubatud vaid kolmandavärteline hääldus. Meie katse näitab, et sõna on tänapäeval teisevärteline.

4. **Kahesilbilised us-lõpulised adjektiivid** *hirmus*, *kuulus*, *soodus* ja *rõõmus*. Kuigi 15-st *us*>*sa* lõpuvaheldusega adjektiivist on suurem osa Pälli sõnul kolmandavärtelised (v.a *ihnus*, *ainus* ja *hõlbus*) ja siin välja toodud sõnad on sõnaraamatus ÕS 2006 olnud veel ühtviisi kolmandavärteliseks märgitud, on Päll (1986: 25, 27) juba 30 aastat tagasi rõhutanud, et *us*-lõpuliste adjektiivide välde on kõikuv. Need nominatiivis *us*-lõpulised adjektiivid kuuluvad omakorda suuremasse omastavas käändes *sa*-lõpuliste adjektiivide hulka, millest nominatiivis *es*- ja *as*-lõpulised on selles käändes teisevärtelised (nt *tõrges*, *õõnes*, *hoolas*, *armas*, *valvas*) ja *us*- ning *is*-lõpuliste rühmas esineb mõlemas vältes nominatiive. Nurkse võrdles ligi 80 aastat tagasi norminguid murdekeele kogudega ja leidis, et *us*-lõpuliste adjektiivide teisevärteline nominatiiv on kirjakeele aluseks olevas Põhja- ja Kesk-Eesti murretes domineeriv, aga “rahvakeeles päris ühtlane *sa*-adjektiivide sari on kirjakeeles kaheks kärjstatud” (Nurkse 1938a: 45). Ta soovitas kõnelejate sarjatunnet ja rahvakeeles käibivaid tendentse arvestada ja kogu *sa*-adjektiivide sarjas pöörduda teisevärtelise nominatiivi juurde. Kas aga saab öelda, et kõik *sa*-adjektiivid, sealhulgas *us*-lõpulised, käituvad välte koha pealt ühise sarjana? Meie eksperimendi tulemusena selgub, et mingit kindlat, kogu *us*-lõpulisi haaravat trendi tänapäeval ei esine. Sõnad *rõõmus* ja *soodus* on kuuldelise hinnangu põhjal teisevärtelised ja kestussuhete järgi pigem teisevärtelised. Sõna *hirmus* varieeruvus on lugejati kõige suurem, eelistust ei selgunud. Sõna *kuulus* on kindlasti kolmandavärteline.

5. **Kahesilbilised pika konsonandi või konsonantühendiga lõppevad sõnad** *pealik*, *piknik*, *ümbrik*, *politseinik*, *august* ja *peenelt*. Ainult silbistruktuuri arvestades peaksid need sõnad kuuluma kõik *õpik*-üldtüüpi ja saama teise välte. Siiski kuuluvad *lik*-, *nik*-, *rik*-liitega tuletised eraldi alltüüpi, sest selliste tuletusliidetega sõnades võib esineda kaks kõrvutist rõhusilpi (Päll 1986: 27–28). Meie uuring näitab, et selles grupis on valdav kolmas välde. Sõnad *pealik*, *piknik*, *ümbrik* on kolmandavärtelised ja *politseinik* pigem kolmandas vältes hääldatud.

Lisaks tuletistele kuulub siia üldtüüpi *august*, millel kaksikrõhku ei esine. Veel 2006. aasta ÕS-is on sõna esitatud ainult kolmandavärteliseks, aga meie uurimus näitab, et kuuldelise hinnangu järgi hääldatakse teda kindlalt teisevärteliseks ja kestussuhete järgi arvestatult hääldatakse pigem teisevärteliseks.

Sõna silbistruktuurilt samasse rühma peaks kuuluma veel muutumatu sõna *peenelt*, mis veel 1976. aasta ÕS-is oli märgitud teisevärteliseks, aga mis meie uurimustulemuste järgi kuulub kolmandavärteliste hulka.

6. **Kolmesilbilised *tu-* ja *kas-*liitelised tuletised *mõte-tüüpi sõnadest***, kus tuletusliide liitub omastavas käändes tüvele: *ilmetu, kaitsetu, võimetu, andekas, maitsekas, toimekas* ja *võimekas*. Kinnine teine silp peaks tingima automaatselt esisilbi teise välte, aga need tuletised moodustavad kolmandavärtelise alltüübi. Siiski on *tu-* ja *kas-*liiteliste sõnade kohta juba 1986. aastal öeldud, et nende välde kõnes nõrgeneb, eriti sagedatel sõnadel (Päll 1986: 33). Silvi Vare (1984: 20–21) peab just tugevneva astmevaheldusega *e-lõpulistest* sõnadest moodustunud tuletisi välteprobleemi allikaks, kuna tema sõnul on üldkeelses pruugis tugev tendents hääldada neid kolmanda välte asemel teisevärtelistena. Eelkõige toob ta põhjenduseks teisevärteliste tuletiste analoogia mõju, sest teise- ja kolmandavärteliste tuletiste suhe ÕS-is olevat 15:1. Meie uurimusest selgub, et kuuldelise hinnangu järgi on kõik *kas-*liitelised sõnad – *andekas, maitsekas, toimekas* ja *võimekas* – hääldatud kindlalt teisevärtelisena. Sõnade *andekas* ja *maitsekas* silpide kestussuhe viitab aga mõlema sõna puhul selgelt kolmandavärtelisele hääldusele. *tu-*karitiividest on sõna *kaitsetu* säilitanud kolmandavärtelisuse, kuuldelise hinnangu (aga mitte kestussuhete) järgi on teisevärteline *võimetu*. Sõna *ilmetu* hääldamisel ühtset trendi ei tekkinud.
7. **Kolmesilbilised *lik-*liitelised adjektiivid *äärmuslik, haiguslik, keiserlik, looduslik, nõiduslik, nooruslik, rahvuslik, saatanlik, saatuslik, täiuslik* ja *teaduslik*** moodustavad lugemiseksperimentis olevatest sõnadest kõige suurema grupi. Need sõnad kuuluvad silbistruktuuritüübilt *veerandik-tüübi lik-*liiteliste omadussõnade alltüüpi (Päll 1986: 38). Kui üldjuhul saavad pika lõpusilbiga kolmesilbilised käändsõnad teise välte, siis *lik-*liitelised sõnad moodustavad alltüübi, kuhu kuuluvad sõnad saavad kolmanda välte. Meie uuringust selgub, et kõigi selle tüübi sõnade häälduses ongi valdav kolmas välde, ainult sõna *täiuslik* on teises vältes.
8. **Neljasilbilised *lane-*liitelised substantiivid *kangelane, tallinlane* ja *ungarlane***. Sõnade silbistruktuuri arvestades on võimalik sõnu hääldada nii kolmandas vältes (nagu *moskvalane*) kui ka teises vältes (nagu *õpilane* või *tatarlane*). Pälli (1986: 46) andmetel kuulub üle kolmveerandi sellise struktuuriga sõnadest siiski teisevärtelistesse tüüpidesse. Meie uurimusest selgub, et kinnise teise silbiga sõnade *tallinlane* ja *ungarlane* hääldus hinnati kuuldeliselt teisevärteliseks, kuigi silbikestuste arvutuste järgi ei ole kumbki hääldusviis ülekaalus. Lahtise teise silbiga *kangelane* on kolmandavärteline.

Nii nagu viimase analüüsitud tüübi puhul, võib ka ülejäänud tüüpide kohta öelda, et kuigi mõne tüübi puhul avaldusid teatud suundumused, ei käitunud üheski tüübis kõik sõnad ühtemoodi.

6. Kokkuvõte

Kokku oli selles uurimuses vaatluse all 47 varieeruva vältega sõna 1080 hääldusjuhtu. Arvutasime sõnade kohta välteid iseloomustavad kestussuhted ning hindasime sõnade väldeid ka kuuldeliselt. Saadud andmete põhjal selgusid erinevad trendid nii üksiksõnade häälduseelistuste kohta kui natuke suuremat üldistust võimaldav vältegruppidesse jagunemine sõnastruktuuri tüübiti.

Analüüsisime oma uurimuses nii vokaal- kui konsonantvälteisi sõnu. Teisevälteiste sõnade kokkuvõtlik keskmine kestussuhe oli 1,8 ja kolmandavälteiste oma 2,9. Need tulemused sarnanevad nii meie pilootuurimuse tulemustega kui ka andmetega teistest samalaadsetest uurimustest. Vokaal- ja konsonantvälteiste sõnade kõrval pakkusid mõnevõrra erinevaid tulemusi (C)VVCCV(C)-struktuuriga sõnad, kus pearõhulises silbis kannavad väldeid nii pikk vokaal või diftong kui ka osa konsonandist. Selliste sõnade puhul on märgata suuremat kolmanda välte keskmist kestussuhet. Samuti näitas antud struktuuriga *maitsekas* vastukäivaid tulemusi, kui võrdlesime häälduseelistusi kestussuhete ning kuuldelise määranagu põhjal.

Sihtsõnade kestussuhete ning sõnu ühes või teises vältes hääldajate arvu järgi jagasime sõnad 3 välterühma: teine välde domineeris 8 sõna häälduses, kolmas välde domineeris 24 sõna häälduses ning 15 sõna hääldati selge välte-eelistusega. Kuna viimane rühm osutub liialt suureks, tuleb otsida lisatunnuseid, mille alusel oleks võimalik sõnu selgemalt kahte vältesse jagada. Usutavasti on edaspidi võimalik täpsemaid andmeid saada, kui uurida sõnade tonaalseid tunnuseid.

Kuuldelise hinnangu põhjal saime välja tuua 17 sõna, mille häälduses domineeris teine välde ja 25 sõna, mille häälduses domineeris kolmas välde. Vaid 5 sõna puhul ei moodustanud ühes vältes hääldatud juhud $\frac{2}{3}$ kõigist hääldusjuhtudest ja seetõttu ei saanud nende sõnade puhul hääldustrendi esile tuua. Kuuldelise hinnangu järgi jagunesid sõnad selgemalt välterühmadesse. Siiski tuleks objektiivsema hinnangu saamiseks edaspidi kaasata hindajatena rohkem inimesi, kes soovivatult oleks erinevat piirkondlikku päritolu.

Edasi otsisime vastust küsimusele, kas sama silbistruktuuriga ning samasse sõnaliiki kuuluvate sõnade puhul avalduvad ühtsed hääldustrendid: kas ühte tüüpi kuuluvad varieeruva vältega sõnad käituvad sarnaselt. Enamasti samasse gruppi kuuluvad sõnad ühtviisi ei käitunud, tihti jaguneti kaheks. Näiteks silbistruktuurilt ja sõnaliigilt ühte gruppi kuulunud adverbide *hilja*, *ilma* ja *mullu* häälduses oli valdav teisevältelisus, aga samasse gruppi kuuluvaid sõnu *ammu* ja *enne* hääldati kolmandavälteeliselt. Samuti ei käitu ühtse grupina *us*-lõpulised adjektiivid. Kui sõna *kuulus* on kolmandavälteeline, siis *rõõmus* ja *soodus* on teisevälteelised ning sõna *hirmus* puhul eelistust ei selgunud.

Siiski võib tuua välja mõned grupid, kus avaldusid veidi selgemad trendid. Nii kahe- kui kolmesilbiliste *lik*-liiteliste adjektiivide (v.a *politseinik* ja *teaduslik*) häälduses oli ülekaalus kolmas välde. 15-st seda tüüpi tuletisest jäi ainult kahe sõna häälduses valdavaks teine välde. Ka *e*-lõpulisi adjektiive võib mööndusega pidada kolmandavälteelisteks. Kuigi sõna *kiire* hääldus on adjektiivina täpselt uurimata, selgus, et ajalooliselt kunagised teisevälteelised sõnad *ahne*, *ihne*, *kaine*, *kärme* ja *terve* moodustavad küllalt ühtse grupi, milles valitseb kolmandavälteeline hääldus. Kuuldelise hinnangu järgi võib ühtseks grupiks pidada ka *kas*-liitelisi sõnu, kuigi

vajab edasist uurimist, miks kuuldeline hinnang ja silpide kestussuhete arvutused sõnades *maitsekas* ja *andekas* olid vastandlikud. Paistab, et nendes omadussõnalistes tuletistes avaldub üks kindel suundumus: algselt kolmandaväetelise tuletusalusega sõnad kipuvad analoogia mõjul teiste samast sõnaliigist tuletiste eeskujul muutuma teiseväetelisteks.

Sõnastruktuuri tüübiti trendide leidmine võimaldab teha järeldusi ka teiste sama tüüpi sõnade kohta, mis meie uurimuses ei osalenud. Nii saab kõnesünteesi teksti analüüsi etapis suurema hulga varieeruva vältega sõnade puhul, mis ÕS-is omavad kahetist hääldust, teha otsustuse levinuma variandi kasuks. Loomulikult on võimalus ka üksikute sõnade kaupa soovitusi anda, eriti sagedamate sõnade puhul võib sellest kõnesünteesi kvaliteet märgatavalt paraneda.

Selgusid esialgsed suundumused varieeruva vältega sõnade hääldamisel, mida soovime kinnitada suuremal ning erinäolisemal kõnematerjalil, samuti rakendada mitmekülgsemat välte akustilist analüüsi ning hindamist. Seejärel saaks aga juba tulemusi kasutada tekst-kõne sünteesi häälduse parandamisel.

Viidatud kirjandus

- Asu, Eva Liina; Lippus, Pärtel; Pajusalu, Karl; Teras, Pire 2016. Eesti keele hääldus. Eesti keele varamu II. Tartu: Tartu Ülikooli Kirjastus.
- Boersma, Paul; Weenink, David 2014. Praat: doing phonetics by computer. [Arvutiprogramm.] <http://www.praat.org/> (6.10.2016).
- Eek, Arvo; Meister, Einar 2003. Foneetilisi katseid ja arutlusi kvantiteedi alalt (I). Häälikukestusi muutvad kontekstid ja välde. – Keel ja Kirjandus, 11, 815–837; 12, 904–918. Eestikeelse kõne autosegmenteerija. [Arvutiprogramm.] <http://www.phon.ioc.ee/dokuwiki/doku.php?id=projects:tuvastus:est-align.et> (15.6.2016).
- Eesti keele spontaanse kõne foneetiline korpus. <http://www.keel.ut.ee/et/foneetikakorpus> (29.9.2016).
- Hint, Mati 1968. Ortoeopia normeerimise probleeme. – Keel ja struktuur 2. Töid struktuurse ja matemaatilise lingvistika alalt. Tartu: Tartu Riiklik Ülikool, 2–123.
- Kaalep, Heiki-Jaan; Muischnek, Kadri 2002. Eesti kirjakeele sagedussõnastik. Tartu: Tartu Ülikooli Kirjastus.
- Kalvik, Mari-Liis; Piits, Liisi 2015. Lugemiseksperiment fonoloogilise varieerumise uurimiseks. [Reading experiment for discovering phonological variation.] – Aspects of Speech Studies. ESUKA / JEFUL, 6 (3), 49–77. <http://dx.doi.org/10.12697/jeful.2015.6.3.02>
- Lehiste, Ilse 1997. Search for phonetic correlates in Estonian prosody. – Ilse Lehiste, Jaan Ross (Eds.), Estonian Prosody: Papers from a Symposium. Proceedings of the International Symposium on Estonian Prosody. Tallinn: Institute of the Estonian Language, 11–35.
- Liiv, Georg 1961. Eesti keele kolme vältusastme kestus ja meloodiatüübid. – Keel ja Kirjandus, 7, 412–424; 8, 480–490.
- Lippus, Pärtel; Pajusalu, Karl 2009. Regional variation in the perception of Estonian quantity. – Martti Vainio, Reijo Aulanko, Olli Aaltonen (Eds.), Nordic Prosody. Proceedings of the Xth Conference, Helsinki 2008. Frankfurt: Peter Lang, 151–157.
- Lippus, Pärtel 2011. The Acoustic Features and Perception of the Estonian Quantity System. Dissertationes philologiae estonicae Universitatis Tartuensis 29. Tartu: Tartu University Press.
- Lippus, Pärtel; Asu, Eva Liina; Teras, Pire; Tuisk, Tuuli 2013. Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian. – Journal of Phonetics, 41 (1), 17–28. <https://doi.org/10.1016/j.wocn.2012.09.005>

- Mihkla, Meelis 2009. Kõnesüntees kui keele iseäralik võlu. – Jaan Ross, Arvo Krikmann (Toim.), Humanitaarteadused. Teadusmõte Eestis V. Tallinn: Eesti Teaduste Akadeemia, 55–62.
- Nurkse, Rein 1938a. Õige keelega kimbus. – Looming 1. Tartu, 39–46.
- Nurkse, Rein 1938b. Õige keelega kimbus. – Looming 2. Tartu, 202–214.
- Päll, Peeter 1986. Eesti noomeni silbistruktuur ja aktsent. Tallinn: Eesti NSV Teaduste Akadeemia ühiskonnateaduste osakond.
- Sahkai, Heete; Mihkla, Meelis; Kalvik, Mari-Liis 2015. Emfaas ja fookus eesti keeles. [Emphasis and focus in Estonian.] – Aspects of Speech Studies. ESUKA / JEFUL, 6 (3), 97–122. <http://dx.doi.org/10.12697/jeful.2015.6.3.04>
- Salveste, Nele 2006. Eesti keele normitud välited ja nende tegelik hääldus. Bakalaureusetöö käsikiri Tartu Ülikooli eesti ja üldkeeleteaduse instituudis.
- Salveste, Nele 2013. Kõnetaju kategoriaalsus ehk hüpotees sellest, kuidas me keelelisi üksusi tajume. [Categorical perception or the hypothesis of how we perceive linguistic units.] – Theoretical linguistics in Estonia III. ESUKA / JEFUL, 4 (1), 127–143. <http://dx.doi.org/10.12697/jeful.2013.4.1.08>
- Schouten, Bert; Gerrits, Ellen; Hossen, van Arjan 2003. The end of categorical perception as we know it. – Speech Communication, 41 (1), 71–80. [http://dx.doi.org/10.1016/S0167-6393\(02\)00094-8](http://dx.doi.org/10.1016/S0167-6393(02)00094-8)
- Vare, Silvi 1984. Omadussõnaliited tänapäeva eesti kirjakeeles. Tallinn: Valgus.
- ÕS 1960 = Õigekeelsuse sõnaraamat. Tallinn: Eesti Riiklik Kirjastus, 1960.
- ÕS 1976 = Õigekeelsussõnaraamat. Tallinn: Valgus, 1976.
- ÕS 1999 = Eesti keele sõnaraamat. Tallinn: Eesti Keele Sihtasutus, 1999.
- ÕS 2006 = Eesti õigekeelsussõnaraamat ÕS 2006. Tallinn: Eesti Keele Sihtasutus, 2006.
- ÕS 2013 = Eesti õigekeelsussõnaraamat ÕS 2013. Tallinn: Eesti Keele Sihtasutus, 2013.

Liisi Piits on Eesti Keele Instituudi teadur kõnetehnoloogia alal. Ta osaleb kõnetehnoloogia ja kõneuurimise projektides ning uurib fonoloogilist varieerumist eesti keeles. Varem on ta tegelenud ka korpuslingvistiliste sõnavarauringutega.
Roosikrantsi 6, 10119 Tallinn, Estonia
Liisi.Piits@eki.ee

Mari-Liis Kalvik on Eesti Keele Instituudi vanemleksikograaf ja eksperimentaalfoneetika teadur. Ta on uurinud välited nii murde- kui ühiskeeles ning osaleb kõneuurimise projektis. On "Eesti murrete sõnaraamatu" üks toimetajaid.
Roosikrantsi 6, 10119 Tallinn, Estonia
Mari-Liis.Kalvik@eki.ee

WORDS OF VARIABLE QUANTITY DEGREES AS A PROBLEM FOR SPEECH SYNTHESIS

Liisi Piits, Mari-Liis Kalvik

Institute of the Estonian Language

Estonian text-to-speech synthesis relies in its determination of pronunciation on the Dictionary of Standard Estonian (ÕS 2013), which is the basis of standard Estonian. However, for roughly 300 words, this dictionary allows pronunciation with both the second and third quantity degree. This causes problems in the text-to-speech synthesis system, since the automatic text analysis cannot handle multiple outputs. 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 in actual speech.

For the studies of quantity degrees, words were chosen which ÕS 2013 lists as being pronounced with both the second and third quantity degree. This study is based on a reading experiment conducted with 23 informants (15 women and 8 men), in which each informant read 52 sentences aloud. These sentences contained 47 target words, i.e. words of variable quantity degrees; in total, the study yielded 1080 pronunciation instances to examine. The group of target words includes those with varying vowel quantity degrees as well as those with varying consonant quantity degrees; the duration ratios characteristic of each quantity degree were calculated on the basis of the primary-stress syllable and the unstressed syllable following it. The average duration ratio for the second quantity degree is 1.8, and for the third quantity degree 2.9. These results are similar to those obtained in previous studies.

On the basis of the informants' pronunciation, the words were grouped into three categories: second quantity degree, variable quantity degree (where neither the second nor the third quantity degree accounted for more than 2/3 of all pronunciations) and third quantity degree. Based on the duration ratio, 8 words fell into the second quantity degree group; however, based on auditory assessment, this group increased to 17 words. The variable quantity degree group contained 15 words based on duration ratios, but only 5 words based on auditory assessment. The third quantity degree group contained 24 words by duration ratio and 25 words by auditory assessment.

Finding trends in word structure among words in the same quantity degree groups would make it possible to draw inferences about other words of the same type as well, which would increase the applied value of the study. Generally, though words of the same syllable structure and part of speech did not exhibit the same pronunciation patterns. However, it can at least be stated that the third quantity degree dominated among both two- and three-syllable adjectives formed with the suffix *-lik*. Of the 15 such words analysed in the study, only two were pronounced predominantly with the second quantity degree.

In order to apply these results to speech synthesis, the preliminary discovered trends should be checked on the basis of larger and more varied speech material,

and more diverse approaches should be used to determine which quantity degree is used in any given pronunciation instance.

Keywords: phonological variation, variation of quantity degrees, text to speech synthesis, reading experiment, Estonian