

KÕNETEMPO EELISTUSED JA AUDIOSÜSTEEM NÄGEMISPUUDEGA INIMESTELE

MEELIS MIHKLA, INDREK HEIN,
INDREK KIISSEL, MARGIT ORUSAAR, ARTUR RÄPP

Eriliste võimetega inimestest kuuleme-näeme massimeedia vahendusel peaaegu iga päev. Üldjuhul ei ole loodud kuvandid alati tõepärased või pole üleskiidetud võimed teaduslikult põhjendatud. 2010. aasta novembris USA-s San Diegos neuroloogide ülemaailmse kokkutuleku aruandes nenditi, et on leitud huvitav rühm üliinimlike vaimsete võimetega üksikisikuid – pimedad inimesed. Veelgi enam, tänu aju funktsionaalsele pildidiagnostikale (magnetresonantsuuringule) on nüüdseks ka selgunud, kuidas need erilised võimed on saavutatud. Üldlevinud arvamus on see, et pimedad inimesed on oma ülejäänud meeled arendanud võimalikult teravaks, et kaotatud nägemist kompenseerida. Pimedate muusikute Stevie Wonderi, Andrea Bocelli, Ray Charlesi jt näitel võib tunduda, et nad saavad muusikas silma paista tänu kõrgelt arenenud kuulmismeelele. Saksamaa Tübingeni Ülikooli teadlased on aju kliiniliste uuringutega leidnud sellele väitele teaduslikku tuge (Moos jt 2008; Hertrich jt 2011). Nimelt võib osa nägemispuudega inimestest hästi aru saada kõnest, mis on tavatempost kaks ja enam korda kiirem ning mis nägijatele tundub arusaamatu mulinana.¹ Nägemispuudega inimeste suulise teabe kiiret omandamisvõimet, kuuldu mõistmist välkkiirelt ja kõne ajalise kokkusurumise piire on käsitletud ka mitmes varasemas uurimuses (nt Asakawa jt 2003; Moos, Trouvain 2007). Ingo Hertrichi ja tema kolleegide uurimuses skaneeriti kiirkõne tajumise ajal pimedate ja võrdlusena nägijate aju ning avastati, et ülikiire kõne tajumisel aktiveerus pimedatel ajus kolju tagaosas süstemaatiliselt teatav piirkond, mida tavajuhul kasutatakse visuaalse info töötlemiseks (Hertrich jt 2011). Selline aju ressurside ümberjagamine selgitab pimedate inimeste erilist võimet kiirkõnest aru saada. Tavaliselt tegeleb suur osa aju hallist massist visuaalse info töötlemisega, oluliselt väiksem osa on teiste meelte teenistuses. Nii märkimisväärse aju piirkonna jõude seismine oleks raiskamine ning pimedate inimeste aju suudabki mingil moel kuulmispiirkonnad ühendada nägemise ajukoore piirkonnaga.²

Kirjeldatud fenomeniga puutusime teataval määral kokku pimedatele mõeldud elektrooniliste teabetekstide ettelugemise ja heliraamatute kuulamisüsteemi loomisel. Eesti Pimedate Raamatukogu teabetekstide audiosüsteemi <http://elte.eki.ee/> abil saavad pimedad ja nägemispuudega inimesed lasta endale ette lugeda uudiseid, ajalehti, ajakirju ja raamatuid ning kuulata heliraamatuid. Süsteemi esimese versiooni kasutamise käigus selgus, et paljud

¹ Tavatempoks peetakse harilikult kõnet, mille kiirus jääb vahemikku 120–140 sõna minutis. Slaididega illustreeritud loengu või konverentsi ettekande keskmiseks kiiruseks on umbes 100 sõna minutis, heliraamatute soovitatavaks kõnetempoks on 150–160 sõna minutis.

² Pimedate inimeste aju eri piirkondade ümbersuunamisvõime pakub erilist huvi seetõttu, et seda võiks rakendada insuldihaigete raviks, suunates kahjustunud aju piirkondade infotöötluse kahjustamata aju piirkondadesse.

pimedatest soovivad uudiseid ja leheartikleid kuulata tavakiirusest tunduvalt kiirema kõnetempoga. Et süsteem on serveripõhine, siis suvalist kõnetempo muutmise võimalust kasutajatele välja pakkuda ei olnud võimalik: see oleks muutnud süsteemi liiga kohmakaks ja aeglaseks. Seega tekkis vajadus optimaalsete kõnetempode tuvastamiseks, pakkumaks paari tavatempost kiiremat – kiiret ja väga kiiret – kõnetempot. Ülikiires tempos kõnet eelistavad ja on võimelised sellest veatult aru saama nn treenitud inimesed: isikud, kes kasutavad ekraanilugejaga arvutit ja seega kuulavad kõnesüntesaatorit peaaegu iga päev (Moos jt 2008).

Oma uurimuses püüdsime selgitada, kas selline nähtus on üldine ja kui suurel hulgal pimedatel on selline võime välja arenenud. Kas on olemas nn optimaalne kõnetempo nii pimedatele kui ka nägijatele ja millisel treenituse astmel tekib nägemispuudega inimestel eriline võime kiirkõnest aru saada? Kõigile neile küsimustele püüdsime vastust saada pimedatele ja nägijatele korraldatud erinevate kõnetempode hindamise tajukatses.

1. Elektrooniliste tekstide ja heliraamatute audiosüsteem

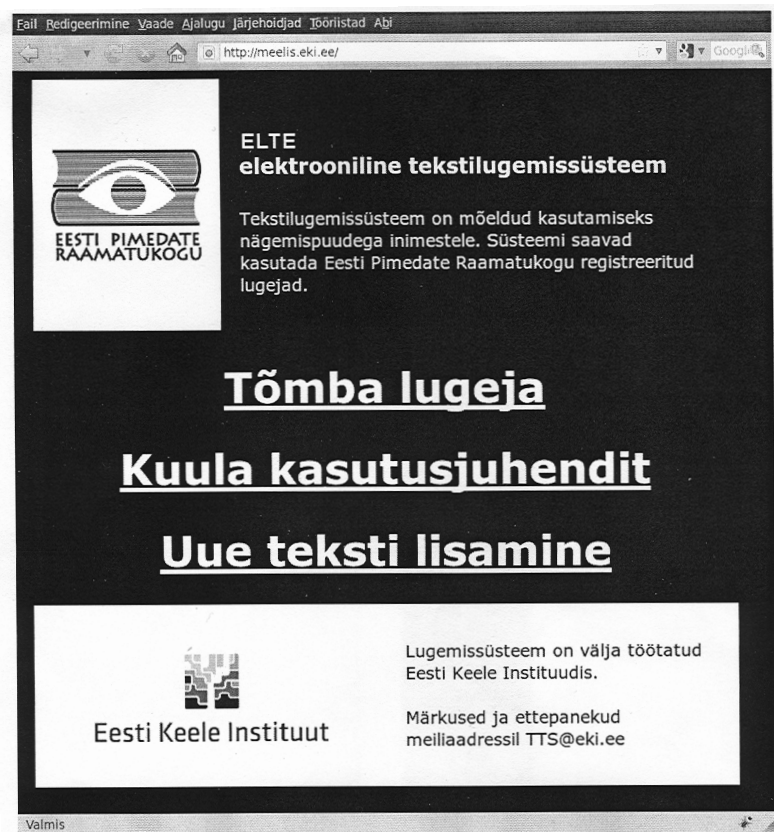
1.1. Taust

Eestis on umbes 2000 nägemispuudega inimest, kelle nägemislangus ei võimalda neil igapäevaeluks vajaliku teabe hankimiseks kasutada traditsioonilisi info hankimise vahendeid. Neist umbes 500 kasutab aktiivselt arvutit olemasolevate abivahendite (suurendus- ja mitmesuguste kõneprogrammide) toel. Eesti Pimedate Raamatukogus on ligikaudu 1500 heliraamatut. Arvutite abil informatsiooni hankimiseks ja vahetamiseks kasutavad nägemispuudega inimesed enamasti mitmesuguseid ekraanilugejaid (*Jaws* vm) ja kõnesüntesaatoreid. Ekraanilugejad on küllalt kallid ja need peab igaüks endale personaalselt soetama. Meie projekti eesmärgiks oli luua interneti vahendusel ühtne eestikeelsete elektrooniliste tekstide ettelugemise ja heliraamatute kuulamissüsteem, mis ei eeldaks kalli tarkvara soetamist. Audiosüsteemi peamiseks eesmärgiks oli luua elektroonilisest meediast osasaamise võimalus ka neile pimedatele ja vaegnägijatele, kes igapäevaselt arvutit ei kasuta.

Joonisel 1 on toodud süsteemi veebileht, kust on võimalik alla laadida elektrooniline audiosüsteem, kuulata kasutusjuhendit ja samas on ka kõigil kasutajatel võimalik ise süsteemi üles laadida mitmesuguseid elektroonilisi tekste (käsiraamatuid, uudislugusid, teateid jms). Selline süsteem laiendab pimedate ja vaegnägijate ligipääsuvõimalusi vajalikule informatsioonile, sest seda on võimalik kasutada transpordis, traadita interneti levialas, raamatukogudes ja avalikes internetipunktides jm.

1.2. Süsteemi funktsionaalsus

Audiosüsteemi käsitlemine on püütud teha pimedatele võimalikult lihtsaks. Kõigi uudislugude, artiklite või raamatute valik, ajalehe sirvimine ja heliraamatukogus navigeerimine on teostatav üksnes numbriklahvide ja mõne funktsiooniklahvi abil. Vaid raamatute otsingul ja sisselogimisel on vaja täheklahve. Vajutatud klahvidest teavitatakse häälega. Süsteemi sisenemisel loetakse ekraanil olev peamenüü ette ning võimalik on kuulata ka süsteemi kasutusjuhendit (vt joonist 2).

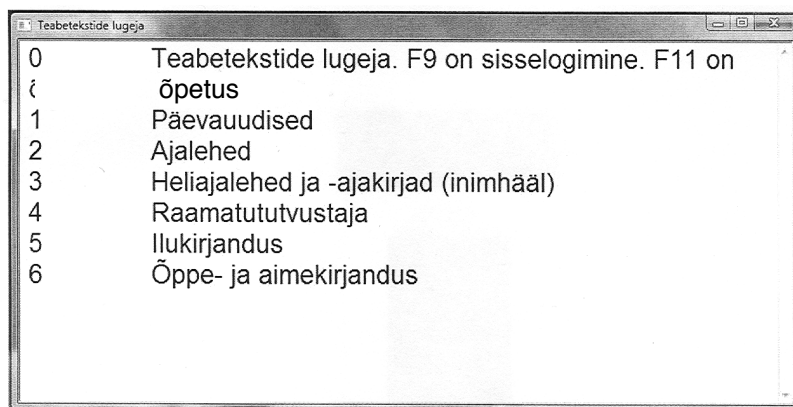


Joonis 1. Elektrooniliste tekstide ja heliraamatute audiosüsteemi veebilehekül (http://elte.eki.ee).

Uudiste ja mõningate ajalehtede lugemine on kõigile vaba. Heliraamatute ja -ajakirjade kuulamiseks peab sisse logima, vastavaid paroole jagab Eesti Pimedate Raamatukogu oma registreeritud lugejatele. Alammenüüdesse sisenemiseks tuleb vajutada vastavat numbriklahvi, seega võib menüüdes olla kuni üheksa jaotust. Kui näiteks raamatus on kaheksateist peatükki, siis esimese taseme menüüga saab välja valida peatükid ühest kuni üheksani (numbriklahv 1) või peatükid kümme kuni kaheksateist (2) ning järgneva taseme menüüs saab uuesti numbriklahvile vajutamiselega välja valida konkreetse peatüki (näiteks 13. peatüki valikuks tuleb esmalt vajutada klahvile „2” ja seejärel klahvile „4”). Tagasi kõrgema taseme menüüsse saab liikuda klahvile BACKSPACE või ESC vajutamisega.

Süsteemis on raamatuid kahel kujul: elektroonilise tekstina ja heliraamatuna. Elektroonilise tekstina olevaid raamatuid loeb kõnesüntesaator ette, genereerides vajalikud helifailid, heliraamatuid loevad ette professionaalsed diktorigid ning vastavad helifailid on serveris eelnevalt olemas. On võimalik kuulata ka eelmisel ja jooksva aastal raamatukokku saabunud ja süsteemi lisatud uute raamatute tutvustusi.

Audiosüsteemis on osal funktsiooniklahvidest erilised tähendused: F7 – kõnetempo muutmine, F8 – süntesaatori hääle valik, F9 – sisselogimine,



Joonis 2. Audiosüsteemi peamenüü.

F10 – programmi info, F11 – kasutusjuhendi ettelugemine, F12 – kellaaja teade. Lisaks on kuulamisel võimalik kasutada järgmisi klahve: tühik peatab lugemise, uuesti tühikule vajutamisel lugemine jätkub, vasaku nooleklahviga saab helifailis 20 sekundit tagasi liikuda ja parema nooleklahviga 20 sekundit edasi, BACKSPACE ja ESCAPE lõpetavad antud helifaili esituse ja pöörduetakse tagasi eelmisesse menüüsse. Klahvi F8 abil saab valida erinevaid sünteeshääli. Süsteemis on üks difoonsünteeshääli Mbrola-baasil (Mihkla, Meister 2002) ja 2–3 erinevat kõneüksuste valikul (Mihkla jt 2008) põhinevat häält. Võimalik on ka muuta kõnetempot: 1 – normaalne, 2 – kiire ja 3 – väga kiire.

2. Kõnetempo tajumine

Süsteemi esimese versiooni kasutamise käigus selgus, et paljud pimedatest soovivad uudiseid ja leheartikleid kuulata tavakiirustest tunduvalt kiirema kõnetempoga. Serveripõhise süsteemiga pole kõnetempo kõiki erinevaid väärtusi võimalik välja pakkuda. Seega oli tarvis kõnetempo tajuteste. Ehkki internetipõhine audiosüsteem on mõeldud eelkõige nägemispuudega inimestele, oli kõnetempo tajutesti kaasatud võrdluseks ka nägijaid. Selline võrdlusbaas aitab leida vastust järgmistele küsimustele: „Milliseid kõnetemposid eelistavad pimedad ja milliseid nägijad?“, „Kas muut nägemispuudega inimeste ülikiire kõnetempo eelistamisest peab paika?“, „Kas on olemas optimaalne kõnetempo?“.

2.1. Testis osalejad ja materjal

Testis osales 58 pimedat või raske nägemispuudega inimest (29 naist ja 29 meest, vanuses 14–79 aastat) ning 56 nägijat (41 naist ja 15 meest, vanuses 18–58 aastat). Kõigi testis osalenute emakeeleks oli eesti keel.

Kõnetempo tajutesti stiimulid genereeriti kahe heliraamatu (meeshääle loetud T. Dreiseri „Ameerika tragöödia“ ja naishääle loetud E. Stein-Fisheri „Esimest korda ja enamgi“) ning kõnesüntesaatori loetud uudisloolõikude

põhjal.³ Kõnesüntesaatoriks oli difoonidel põhinev eesti keele tekst–kõne-süntesaator (Mihkla, Meister 2002). Sünteeshääle genereerimisel kasutati kaht erinevat prosoodiamudelit: reeglipõhist (SYNT1) ja statistilist (SYNT2). Seda tehti nende nägemispuudega inimeste arvamuse kontrollimiseks, kelle meelest statistilise prosoodiamooduliga süntesaatori väljundkõne kvaliteet halveneb kiirema tempo puhul.

Testi kaasatud naishääle loomulikuks lugemiskiiruseks oli keskmiselt 135 sõna minutis, meeshäälel 122 sõna minutis. Süntesaatorid häälestati naishääle kõnetempole vastavaks. Iga hääle kohta genereeriti kaheksa erineva kõnetempoga 35–55-sekundilist kõnelõiku (vt tabelit 1).

Tabel 1.

Kõnelõikude stiimulid erinevate kõnetempode puhul
(loomulik kõnetempo 100 % = 135 s/min)

81	108	135	162	189	216	243	270	sõna/minutis
60	80	100	120	140	160	180	200	%

Kõne ajalise kokkusurumise piirid lepiti eelnevalt kokku „treenitud isikutega”, kel on iga päev sünteeskõne kuulamispraktikat. Nende arvates on tavalisest kaks korda kiirem kõnetempo maksimaalne kiirus, millega ilukirjandust kuulata võiks. Testi ettevalmistamise käigus avaldasid mõned vanemad nägemispuudega inimesed arvamust, et nemad tahaksid mõnda lõiku ka tavatempost aeglasemalt kuulata. Seetõttu sai lisaks genereeritud kaks aeglasemat kõnetempo stiimulit (0,8 ja 0,6 korda tavatempost aeglasemad). Heli- raamatute kõnelõikude kõnetempo muutmiseks kasutati signaalitöötlusprogrammi Adobe Audition 3 kõrge täpsusega ajalist kokkusurumist, säilitades põhitooni.

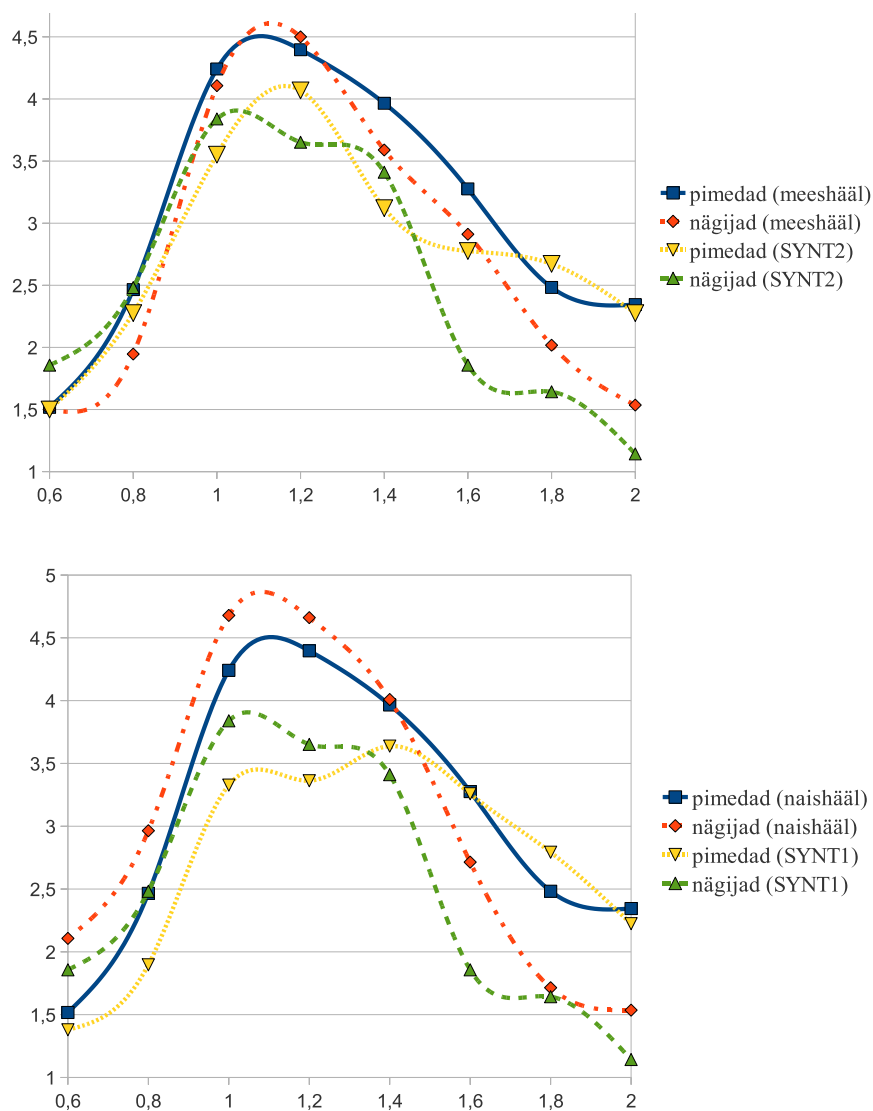
Kõnetempo stiimulid esitati testitavatele juhuslikus järjekorras hääleseeriade kaupa. Kõnetempo sobivust tuli hinnata viiepallisüsteemis (5 – sobivaim kõnetempo, 4 – hea kõnetempo, 3 – talutava kiirusega kõne, 2 – ebameeldiv kõnetempo, 1 – sobimatu kõnetempo, s.o arusaamatu või liiga aeglane/kiire kõne).

2.2. Tulemused

Joonisel 3 on toodud pimedate ja nägijate keskmised hinded erinevate hääle kõnetempode kohta: ülemisel graafikul on hinnangud meeshäälele ja sünteeskõnele SYNT2, alumisel naishäälele ja sünteeskõnele SYNT1. Tumedama joonega graafikute põhjal võib öelda, et pimedad eelistavad enim kiirusi vahemikus 1,1–1,4, mis on 10–40 % tavatempost kiiremad. Nägijate eelistuste maksimum jääb piiridesse 1,0–1,2, s.o kas normaalne kõnetempo (135 sõna/minutis) või kuni 20 % tavatempost kiirem.

Meeshääle puhul hindasid nii nägijad kui ka pimedad parimaks kõnetempoks 1,2. Selline hinnang tuleneb ilmselt meeshääle tavatempost, mis on umbes 10 % teistest häälest aeglasem (122 s/min vs. 135 s/min).

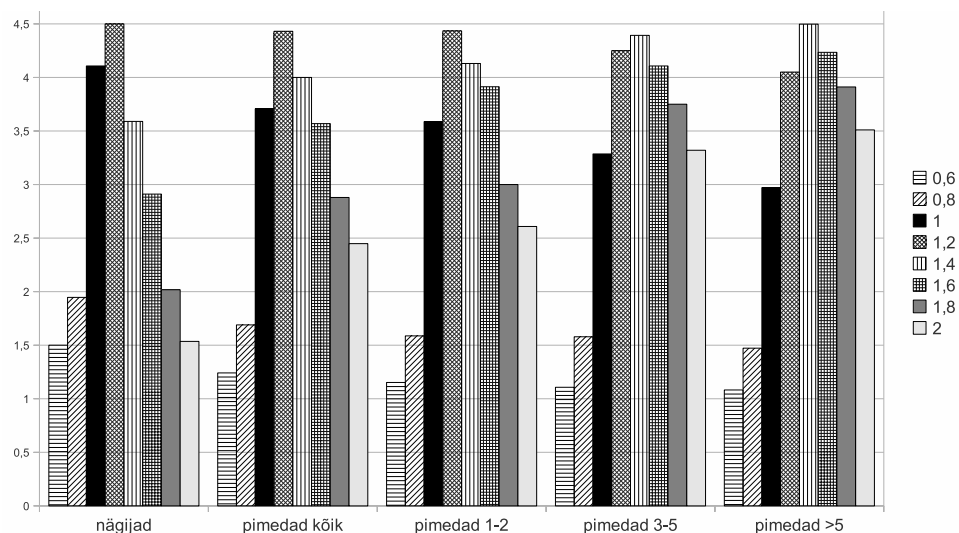
³ Kõnetempo suurtes piirides muutmise võimalus on üks väheseid kõnesüntesaatori omadusi, milles ta inimehäält edestab.



Joonis 3. Pimedate ja nägijate antud erinevate kõnetempode keskmised hinnangud.

Ehkki testis osalejail paluti hinnata ainult kõnetempo sobivust, mitte aga hääle meeldivust, on sünteeskõne kokkuvõttes ikkagi saanud inimkõnest ligi palli võrra madalamaid hindeid. Aga kiirema kõnetempo puhul pole süntesaaatori prosoodiamoodulist tingitud hinnete erinevusi SYNT1 ja SYNT2 juures oluliselt märgata. Tulemused ei toeta nägemispuudega inimeste varasemat arvamust, et kiirema tempo puhul SYNT2 hääle kvaliteet halveneb. Pigem ilmnes vastupidiselt, et SYNT2 puhul hinnati kõrgemalt just kiiremaid kõnetemposid (vrd joonisel 3 SYNT1 ja SYNT2 graafikuid).

Testi tulemused üldisemalt näitasid, et pimedatel ja nägijatel erinevad kõnetempode keskmised hinnangud palju vähem, kui testi kavandades eeldada võis (vrd joonise 4 kaht vasakpoolset diagrammi). Pimedate hulgas on aga



Joonis 4. Keskmised hinnangud kõnetempode kohta nägijatelt ja pimedatelt, kolmel vasakpoolsel diagrammil on nägemispuudega inimeste eelistused sõltuvalt arvutikasutamiskogemusest koos ekraanilugejaga (aastates).

ka küllalt palju neid, kes kasutavad arvutit harva või ei kasutagi. Seega puudub neil erineva kõnetempoga sünteeskõne kuulamise kogemus. Joonise 4 kolmel parempoolsel diagrammil on toodud keskmised hinnangud kõnetempodele sõltuvalt nägemispuudega inimeste eelnevast arvuti- ja ekraanilugeja kasutamiskogemusest aastates. Tulemustes on selgelt märgatav tendents, et koos nn treenituse kasvu ehk pikema sünteeskõne kuulamise praktikaga kasvab nägemispuudega inimesel võime kiirkõnest aru saada. Aga see võime ei avaldu mitte kõigil. Tübingeni Ülikoolis tehtud uurimuses tuvastati, et vanus, mil inimene kaotab nägemise, onaju kuulmispiirkondade visuaalse info töötluspiirkondadesse ümbersuunamisel tõenäoliselt kriitiline. Nooremad inimesed, kes kaotasid nägemise 2–15 aasta vanuses, on võimelised paremini ajajuotusi ümber suunama kui nägemisest vanemas eas ilma jäänud või pimedana sündinud inimesed (Hertrich jt 2011). Meie pimedate testirühma hulgas oli 11 inimest (19 % testitud nägemispuudega inimestest), kellel oli selgelt väljaarenenud võime kiirkõnest aru saada. Nende eelistatuid kõnetempod olid 2,0- (tavatempost kaks korda kiirem) või 1,8-kordne kõnetempo. Kõigi nende „treenitud pimedate” arvuti ja ekraanilugeja kasutamise kogemus ületas viit aastat. Kas neil kiirkõnest arusaamiseks aktiveerub ka kindel ajupiirkond, nõuab omaette magnetresonantsuuringut. Et meie testid toimusid varem, kui ilmus saksa neuroloogide vastav uurimus, siis polnud neis kahjuks küsimust katseisikute nägemiskaotuse eakohast.

Nn optimaalset kõnetempot polnud testitulemusi analüüsides võimalik leida ei pimedatele ega nägijatele, sest kõnetempo eelistused olid väga individuaalsed ja sõltusid väga paljudest asjaoludest. Küll aga otsustati audiosüsteemi kõnetempo valikusse lisada kiirem kõnetempo 1,3 (176 sõna minutis) ja väga kiire kõnetempo 1,6 (216 sõna minutis). Rakendatud väga kiire kõnetempo ei rahulda võib-olla alati „kiirkõne supermänni”, aga selliseid isikuid

on pimedate hulgas vaid viiendik. Et tavatempost aeglasemale kõnetempole (0,6 ja 0,8) andsid üpris madalaid hindamispalle nii pimedad kui ka nägijad, siis otsustati need kõnelemiskiirused audiosüsteemist välja jätta.

Kokkuvõtteks

Audiosüsteem avar dab nägemispuudega inimeste juurdepääsu vajalikule informatsioonile. Kasutamise lihtsus võimaldab interneti vahendusel mitmesugustest elektroonilistest tekstidest ja heliraamatukogust osa saada ka neil pimedatel, kellel ei ole igapäevast arvutikasutamise kogemust. Kõnetempo tajutestil päris kõigile püstitatud küsimustele ühest vastust ei olnud võimalik saada. Pimedate ja nägijate kõnetempo eelistustes on teatavaid erinevusi, aga erinevuse suurus ei sõltu mitte niivõrd nägemispuudest, kuivõrd pigem arvuti ja ekraanilugeja kasutamise kogemusest. Võime kiirkõnest aru saada võib tekkida viieaastase igapäevase arvutikasutamisega ja sünteeskõne kuulamisega. Meie nägemispuudega inimeste testirühma hulgas ilmsid erilised võimed igal viiendal pimedal inimesel. Ehkki kõigile sobivat optimaalset kõnetempot tajutestil ei õnnestunud tuvastada, rakendati audiosüsteemis lisaks tavatempole kaks kiiremat kõnetempot: kiire, 30 % normaaltempost kiirem ja väga kiire (60 % kiirem).

Uurimis- ja arendustööd on rahaliselt toetanud Hasartmängumaksu Nõukogu, ETF-i grant ETF7998 ja sihtfinantseerimisteema SF0050023s09.

Kirjandus

- Asakawa, Chieko, Takagi, Hironobu, Ino, Shuichi, Ifikube, Tohru 2003. Maximum Listening Speeds for the Blind. – Proceedings of the 2003 International Conference on Auditory Display. Boston, MA, lk 276–279.
- Hertrich, Ingo, Dietrich, Susanne, Ackermann, Hermann 2011. Cross-Modal Interactions During Perception of Audiovisual Speech and Nonspeech Signals. An MRI Study. – Journal of Cognitive Neuroscience, nr 23 (1), lk 221–237.
- Mihkla, Meelis, Meister, Einar 2002. Eesti keele tekst–kõne–süntees. – Keel ja Kirjandus, nr 2, lk 88–97; nr 3, lk 173–182.
- Mihkla, Meelis, Piits, Liisi, Nurk, Tõnis, Kiissel, Indrek 2008. Development of a Unit Selection TTS System for Estonian. – Proceedings of the Third Baltic Conference on Human Language Technologies. Kaunas Lithuania, October 4–5, 2007. F. Čermak, R. Marcinkevičienė, E. Rimkutė, J. Zabarskaitė (toim). Vilnius, lk 181–187.
- Moos, Anja, Hertrich, Ingo, Dietrich, Susanne, Trouvain, Jürgen, Ackermann, Hermann 2008. Perception of Ultra-Fast Speech by a Blind Listener – Does He Use His Visual System? – Proceedings of the 8th Seminar on Speech Production, ISSP. Strasbourg, lk 297–300.
- Moos, Anja, Trouvain, Jürgen 2007. Comprehension of Ultra-Fast Speech – Blind vs. „Normally Hearing” Persons. – J. Trouvain, W. J. Barry (toim), Proceedings of the 16th International Congress of Phonetic Sciences. Saarbrücken, 6–10 August 2007. Saarbrücken, lk 677–680.

Preferences of Speech Rate and an Audio System for the Visually Impaired

Keywords: speech rate, perception test, audio system, visually impaired people

The article describes an audio system of electronic texts and audio books, designed for the visually impaired people to help them read the news, newspapers, magazines and books and listen to audio books over the Internet. The audio system enhances access of the visually impaired to necessary information. Simplicity of use enables them, through the Internet, to benefit from various electronic texts and the sound library, even if they do not have a daily experience of computer use. In order to select appropriate speech rates perception tests were carried out for the blind and the sighted. The speech rate perception tests did not provide unequivocal answers to all questions posed. Although the speech rate preferences of the blind and the sighted displayed certain differences, the size of the difference does not seem to depend so much on the visual impairment than on the subject's experience of using the computer and screen reader. The ability to understand high rate speech may develop after the fifth year of daily use of the computer and of listening to synthesized speech. Among our test group of visually impaired people such a capacity was manifested in every fifth blind person. Although we did not manage to establish the optimum speech rate suitable for all who took the perception test, we applied, besides the regular rate, two higher speech rates in the audio system: a high rate (30% higher than normal) and a very high rate that was 60% higher than the normal rate.

Meelis Mihkla (b. 1955), PhD, senior researcher, Department of Linguistics and Language Technology, Institute of the Estonian Language, meelis@eki.ee

Indrek Hein (b. 1963), software developer, Department of Linguistics and Language Technology, Institute of the Estonian Language, indrek.hein@eki.ee

Indrek Kiissel (b. 1968), software developer, Department of Linguistics and Language Technology, Institute of the Estonian Language, indrek@eki.ee

Margit Orusaar (b. 1967), information specialist, Estonian Library for the Blind, margit@epr.ee

Artur Räpp (b. 1968), board member, North-Estonian Association for the Blind, rtr@planet.ee