

SUBTIITRITE HELINDAMINE – KAS, KUIDAS, KELLELE JA MILLEKS?

MEELIS MIHKLA, INDREK HEIN, INDREK KIISSEL,
ARTUR RÄPP, RISTO SIRTS, TANEL VALDNA

1. Sissejuhatus

Kõnetehnoloogiliste vahendite areng ja digiteleviiooni levik on teinud võimalikuks pakkuda keeleliste erivajadustega inimestele – vaegnägijatele, -kuuljatele ja -lugejatele – erinevaid massimeediaga seotud teenuseid. 2013. aasta juunist on vaegnägijatel ja -lugejatel võimalus Eesti Televisiooni telekanalite (ETV ja ETV2) võrkeelsete saadete ja filmide jälgimisel kuulata lisahelikanali kaudu helindatud subtiitreid. Subtiitrite helindamise süsteem on valminud Eesti Keele Instituudi, Eesti Rahvusringhäälingu ja Eesti Pimedate Liidu ühistööna.¹

1.1. Kellele on vaja subtiitreid ette lugeda?

Eestlane on kuulnud oma suure lugemuse ja lugemisoskuse poolest. Loeme palju raamatuid ja ajakirjandust.² Dupleeritud filmidele ja pealeloetud telesaadetele eelistame originaalkeeles filme ja saateid, millele on lisatud subtiitrid teleekraani alt. Eestis on aga umbes 2000 vaegnägijat, kelle nägemislangus ei võimalda igapäevaeluks vajaliku teabe hankimiseks kasutada traditsioonilisi info hankimise vahendeid. Subtiitrite helindamine on mõeldud põhiliselt neile vaegnägijatele, kes näevad televiisoriekraanil toimuvat, aga ei näe ekraani alaosas väikeses kirjas subtiitrite teksti lugeda. Teise suure keeleliste erivajadustega sihtrühma – vaeglugejate – arv Eestis pole täpselt teada, aga mitmesuguste uurimuste ja PISA-testide tulemuste põhjal on düslektikute määr ühiskonnas küllalt täpselt prognoositav. Euroopa Komisjoni tellimusel 2012. aastal valminud aruandest (European Commission 2012) selgub, et Eesti on kirjaoskuse tasemelt Euroopa Liidus Soome järel teisel kohal. Aruandes sisalduva PISA-testi tulemused meie lugemisoskusest kõditavad küll eestlaste kui väikese rahva edevusnärvi, kuid rõõmustamiseks pole siiski põhjust, sest ka meil on lugemisoskus puudulik 13,3 % inimestest (<http://www.delfi.ee/teemalehed/lugemisoskus>). Euroopa Liidus tervikuna puudub igal viiendal 15-aastaselt noorel ja peaaegu 75 miljonil täiskasvanul elementaarne lugemis- ja kirjutamisoskus (European Commission 2012). Aruande tulemuste põhjal võib subtiitrite helindamise teleteenuse vajadus olla Eestis küllaltki suur: sellesse keelelise erivajaduse sihtrühma kuulub iga kuues või seitsmes eestlane. Vaegnägijatele ja düslektikutele lisaks võib subtiit-

¹ Ühisprojekti rahastati riikliku programmi „Eesti keeletehnoloogia (2011–2017)” toel.

² Seda kinnitavad mitmed OECD algatatud PISA-testid ja Euroopa Komisjoni aruanded, <http://ec.europa.eu/education/literacy/what-eu/high-level-group/documents/literacy-report.pdf> (8. X 2013).

rite ettelugemine olla abiks telesaadete vaatamisel eakatele inimestele, kelle nägemisteravus on langenud, ja ka lastele, kes ei oska veel piisavalt hästi lugeda. Ka eesti keele õppijad võiksid olla sellest teenusest huvitatud. Meenutame, et suur osa Põhja-Eesti elanikkonnast omandas teatud tasemel soome keele just subtiitrite lugemisega Soome telekanalilt eelmise sajandi 70-ndatel ja 80-ndatel aastatel. Siis ei olnud muidugi veel subtiitrid helilised, mis nüüd on keeleõppijate jaoks lisaväärtus.

1.2. Õiguslikud alused ja digiteleviiooni levik

Meediateenuste seadus³ ja ÜRO puuetega inimeste õiguste konventsioon⁴ kohustavad tegema kättesaadavaks massimeedia teenused erivajadustega inimestele, kasutades selleks saate subtiitreid, eraldi audiokanaleid ja teisi vahendeid. Digiteleviiooni arenguga on puuetega inimeste võimalused televiisiooniprogrammidest osa saada laienenud: vaegkuuljatele emakeelsed subtiitrid, vaeglugejatele ja -nägijatele subtiitrite helindamine lisahelikanalis, pimedatele kirjeldustõlked jpm. Meediateenuste seaduse § 23-s loetletakse need kohustused üles, samas § 60 võtab hoogu tublisti maha, kehtestades, et teenuseid tuleb rakendada tasa ja targu ehk järk-järgult. Seega sõltub paljuski vaegnägijate, vaegkuuljate ja düslektikute ning neid esindavate organisatsioonide aktiivsusest, millal need teenused tarbijateni jõuavad. Eesti Rahvusringhääling on vaegkuuljatele mõeldes lisanud populaarsemate saadete („Pealtnägija”, „Osoon” jt) kordustele eestikeelsed subtiitrid, aga digiteleviiooni võimalusi arvestades selline teenus veel täies mahus ei toimi. Pimedatele mõeldud kirjeldustõlkega saateid Eesti telekanalites veel ei edastata. Alates 2013. aasta juunist on Eesti Rahvusringhäälingu ETV ja ETV2 lisahelikanali vahendusel võimalik kuulata võõrkeelsete saadete helindatud eestikeelseid subtiitreid. Selliste teleteenuste kättesaadavuse peamiseks eelduseks on digiteleviiooni levik. Esialgu oli meediateenuste seaduse rakendusotsusesse kirjutatud kaabelsideettevõtete kohustus 2012. aasta juuliks üle minna digivõrkudele. Paraku paljude kaabelfirmade lobitöö tulemusena lükati see tähtaeg ebamäärasesse tulevikku. Seega digiteleviiooni võimalused pole veel kõigis Eesti kaabelvõrkudes kättesaadavad.

1.3. Sarnased rakendused

Filmide ja saadete subtiitrite automaatne ettelugemine kõnesüntesaatori abil pole väga uus idee. Esimene regulaarne avalik subtiitrite helindamise teenus käivitus Hollandi rahvusringhäälingus 2002. aastal (Verboom jt 2002). 2005. aastal pakkus analoogilist teenust Rootsi telekanal STV (A-focus 2010). Meie töörühm hankis oma subtiitrite helindamise süsteemi luues infot eelkõige soomlaste YLE telekanalites töötava subtiitrite helindaja kohta (Kotisaari, Kilpi 2009). Subtiitreid loetakse ette ka Taani ja Tšehhi telekanalites. Kõigi riikide kohta andmeid ei ole, aga pigem on subtiitrite helindamine väiksemate rahvusriikide ja keelte eripära. Kõnelejate arvult suurte keeltega riikides, nagu USA-s, Suurbritannias, Prantsusmaal, Saksamaal, Venemaal jt, subtiitreid pole ja ka nende lugemisharjumus inimestel praktiliselt puudub, kuna

³ RT I, 06.01.2011, 1 (<https://www.riigiteataja.ee/akt/106012011001>).

⁴ RT II, 04.04.2012, 6. Eesti ratifitseeris selle riikidevahelise kokkuleppe 21. III 2012.

peaaegu kõik võõrkeelsed filmid ja saated on nende riikide telekanalites ja filmilevis kas dubleeritud või peale loetud kohalikus keeles. Küll aga edastavad osa neist (nt Suurbritannia, Saksamaa) lisahelikanalis pimedatele kirjeldustõlkeid (Looms 2010). Eestis ja teistes väiksemates riikides eelistatakse aga originaalkeeles filme ja saateid, millele lisatakse subtiitrid.

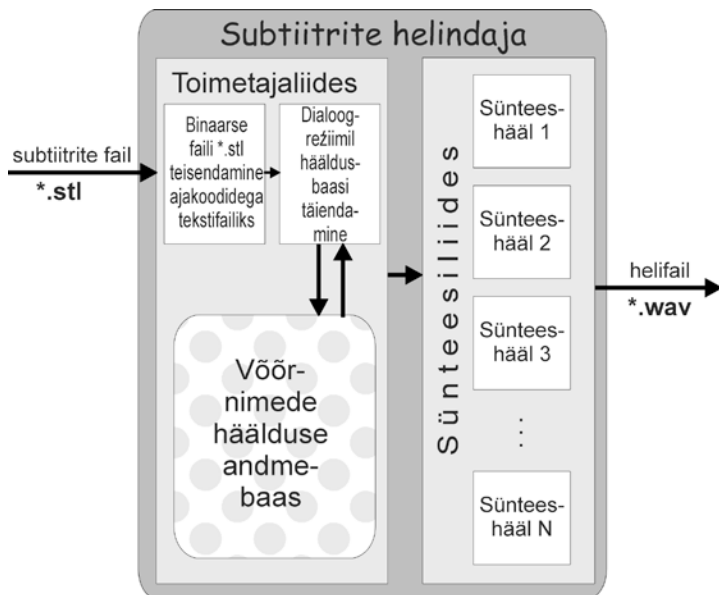
Lisaks lisahelikanalis pakutavale subtiitrite helindamise teleteenusele on püütud kasutada märkide optilist tuvastamist (OCR, *optical character recognition*), et tuvastada subtiitreid televiisori- või arvutiekraanil. Need OCR-il põhinevad helindamissüsteemid rakendavad märkide tuvastamist ja subtiitrite ekstraheerimist otse videosignaalist ning nende ettelugemist tekst-kõne süntesaatoriga (Nielson, Bothe 2007; Jønsson, Bothe 2007). Sellised kasutajale orienteeritud helindajad pole seotud konkreetse telekanali pakutava teenusega, vaid ainult omakeelsete subtiitritega (nt juba praegu varustavad digivõrgus ligi 20 telekanalit saateid ja filme eestikeelsete subtiitritega, aga subtiitrite helindamise teenus on kättesaadav vaid kahel kanalil). Paraku on OCR-il põhinevad süsteemid jäänud seni prototüübi tasemele ja pole avalikku kasutusse jõudnud. Subtiitrite helindajate praktilisel kasutamisel esineb küllalt sageli kattuvusi, st subtiitrite tekst vaheldub kiiremini, kui kõnesüntesaator neid ette lugeda jõuab – kuni 44 % kõigist juhtudest (Hanzlíček jt 2008). Teksti ja kõne paremaks sünkroniseerimiseks on tšehhid välja pakkunud kõne ajaliskokkusurumist ja teksti lihtsustamist (Matoušek jt 2010).⁵ Rootsis on välja töötatud vabavara internetist allalaaditavate subtiitrifailide helindamiseks, samuti on neil kinodes düslektikutele välja pakutud võimalus subtiitreid kuulata Wi-Fi võrgus mobiiltelefoniga (Ljunglöf jt 2012).

2. Subtiitrite helindaja

Meie töörühma subtiitrite helindamise ning tele-ee-trisse edastamise tarkvaralahenduse eesmärk on televisioonis kasutatavate subtiitrifailide alusel genereerida helifaile kõnesüntesaatoriga ning edastada need eraldi helikanalis digiteleviisiooni eetrisse. Subtiitrite helindamise tarkvaralahendus koosneb kahest loogilisest komponendist: 1) subtiitrite helindamise ja toimetamise tarkvara, mille sisendiks on STL-vormingus subtiitrifail ning väljundiks ajakooditapne helifail (WAV või MP3); 2) subtiitri- ja helifailide haldamise, eetrisse planeerimise ning eetrisse edastamise tarkvara. Eesti Keele Instituut on välja töötanud subtiitrite helindaja koos toimetajale mõeldud liidesega, mis teisendab subtiitrifailist täpse ajakoodiga helifaili (vt joonist 1).

Subtiitrite helindaja toimetajaliides teisendab binaarse STL-subtiitrifaili ajakoodidega varustatud tekstifailiks. Seejärel analüüsitakse subtiitrite teksti ja koostatakse arvatavatest võõrnimedest, mida ei leidu võõrnimede häälde andmebaasis, ja tundmatutest lühenditest ning märgijadadest nimekiri, millele subtiitrite toimetaja saab hääldebaasist lisada ja ka subtiitrite hääldebaasist kontrollida. Selline toimetajaliides koos võõrnimede hääldebaasiga eristab meie subtiitrite helindajat Soome vastavast süsteemist (Kotisaari, Kilpi 2009). Soome süsteemis toimub kogu subtiitrite töötlus *online*'is, kuna Soo-

⁵ Meie subtiitrite helindajal oleme täheldanud kattuvusi alla 10 % juhtudest, seetõttu subtiitrite teksti lihtsustamiseks vajadus puudub. Suuremate teksti ja kõne mittesünkroonsuste korral oleme pisut tõstnud kõnetempot.



Joonis 1. Toimetaja- ja sünteesiliidesega subtiitrite helindaja.

mes on tavaks hääldada ka võõrnimesid soome keele hääldusreeglite kohaselt. Eesti keele- ja kultuurikeskkonda selline lähenemisviis ei sobi, kuna meil on kombeks kirjutada ja hääldada võõrnimesid lähtekeelele võimalikult omaselt. Võõrnimede hääldusbaas täieneb pidevalt, sisaldades juba praegu üle 10 000 võõrnime, lühendi ja märgijada hääldust. Seriaalides nimed küll korduvad, aga siiski toob peaaegu iga võõrkeelne saade või film hääldusbaasi uusi kirjeid.

Joonisel 2 on väljavõte võõrnimede häälduse andmebaasist, mis sisaldab võõrnimede ja tekstis esinevate lühendite interpretatsioone ning sõnadena lahtikirjutamisi (nt ASAP → võimalikult kiiresti). Kuna eestikeelsed kõnesüntesaatorid eeldavad sisendis eesti keele tähestikus kirjutatud tekste, siis ka hääldusbaasis pole kasutatud spetsiaalseid foneetilisi transkriptsioonimärke,

Sonastik				
ID	Sõne	keel	Hääldus ↑	
749	Whitby		Vitbi	🗑️ 🗑️ 🗑️
34	Vladimir		vladimir	🗑️ 🗑️ 🗑️
13074	Vladikavkaz		Vladikavk'az	🗑️ 🗑️ 🗑️
275	ASAP		võimalikult kiiresti	🗑️ 🗑️ 🗑️
19306	Võssotski		Võss'otaki	🗑️ 🗑️ 🗑️
10986	Vyyryläinen		Vüüriläinen	🗑️ 🗑️ 🗑️
5461	Jean-Paul		Zaan Pool	🗑️ 🗑️ 🗑️
1793	Jacques		Žak	🗑️ 🗑️ 🗑️
24	Zarudnoi		Zar'udnoi	🗑️ 🗑️ 🗑️
1344	Geraard		Žeraar	🗑️ 🗑️ 🗑️
568	Zealand		Ziland	🗑️ 🗑️ 🗑️
13147	Zvezda		Zvesd'na	🗑️ 🗑️ 🗑️

Joonis 2. Väljavõte võõrnimede häälduse andmebaasist.

vaid kõik hääldused on esitatud eesti kirja pildis. Lisamärgiks on vaid rõhumärk järgsilbi sõnarõhu tähistamiseks. Toimetaja saab hääldusbaasis sõna hääldust kontrollida ja ka jooksvalt parandada. Mõne sõna hääldus sõltub mingil määral konkreetsest sünteeshäälest (teatud hääldusnüansse tuleb eri sünteeshäälele kirja panna natuke erinevalt), samuti võib üksiksõna hääldus erineda sama sõna hääldusest sidusas kõnes. Keelt tähistav veerg hääldusbaasis on vajalik seetõttu, et eristada sama kirja pildiga sõnade hääldusi (nt vaimne liikumine *New Age* hääldub kui [njuu eidž], aga eesti isikunime *Age* juures ingliskeelset hääldust ei tohi rakendada). Sellised hääldusprobleemid tekivad siis, kui võörnimi või selle osa kattub eesti keele mingi sõnavormiga (ingl *road* [rõud] 'tee, sõidutee, maantee', ee *roog* : pl nom *road*; ingl *sea* [sii] 'meri', ee *sig* : sg gen *sea*; ingl *lake* [leik] 'järv', ee sg nom *lake*). Näiteks kahe inglise keelest tõlgitud saate pealkirja WATERLOO ROADI KOOL ja JAMIE OLIVERI 30 MINUTI ROAD (saadete ja filmide nimetused on tiitrites tavaliselt kirjutatud suurtähtedega) ettelugemisel peab esimeses pealkirjas hääldama tähejärjendit ROAD → [rõud], aga teisel juhul on tegu eestikeelse sõnavormiga *road*. Vahel on häälduse eristajaks nime, nimetuse või pealkirja esisuurtäht, kuid isikunimedele või lühenditele hääldamisel pole ka sellest abi (nt Roger [rodžer] v [roger]; CV [tseevee] v [siivii]; Jane [džein] v [jane]; Paul [pool] v [paul]).

Subtiitrite helindamise oluliseks eelduseks, lisaks digiteleviioonile, on emakeelne kõnesüntesaator, mille abil sünteesi liides teisendab subtiitrite teksti automaatselt ortoepiliseks kõneks. Muuhulgas võimaldab sünteesi liides subtiitrite helirežissööril valida erinevaid sünteeshääli ja muuta konkreetse hääle parameetreid. Eestikeelsete kõnesünteesi moodulite ja nende põhjal genereeritud sünteeshääle omaduste kirjeldus on tabelis 1.

Eri omadustega kõnesüntesaatoritel on oma spetsiifilised kasutusala. Kõne segmentaalne kvaliteet on parim üksuste valikul põhineva sünteese meetodi (Üksuste_valik-et) korral. Kõne ladusus ja prosodia on parim Markovi mudelitel põhineval sünteetil (HTS-et). Tekstitöötlemise võimalused on kõige piiratumad eSpeak_et süsteemis, samas on see sünteesimoodul kõige kompaktsim. Sünteesi protsessi juhtimise ja jälgimise tase on madal üksuste valikul, samas on kõne variatiivsus sellel meetodil kõige rikkalikum. Kõigi nende moodulite (peale MBROLA-et mooduli, mille arendus kahjuks lõppes mõni aasta tagasi) edasiarendus toimub pidevalt, genereerime uusi ja viimistleme vanaid hääli.

Tabel 1.

Eestikeelse kõnesünteesi moodulid ja nende omadused

Omadused	Kõnesünteesi moodulid			
	Mbrola-et	Üksuste_valik-et	eSpeak-et	HTS-et
Kompaktsus	mõõdukas	madal	kõrge	mõõdukas
Segmentaalne kvaliteet	mõõdukas	kõrge	madal	mõõdukas
Kõne ladusus	mõõdukas	mõõdukas	madal	kõrge
Tekstitöötlemise tase	kõrge	kõrge	madal	kõrge
Sünteesi protsessi juhtimine	mõõdukas	madal	kõrge	kõrge
Kõne variatiivsus	mõõdukas	kõrge	madal	mõõdukas

Subtiitrite helindamisel kasutatakse praegu kõige sagedamini Markovi peidetud mudelitel põhinevaid HTS-süntheeshääli. Subtiitrite helindaja kasutuspraktika on näidanud, et sagedane sünteeshääle vaheldumine pole õigustatud, kuna sünteeshääle kuulamine vajab harjumist. Seega sünteeshäält vahetatakse vaid juhul, kui mingi kõnesüntesaatori väljundkõne pole paljudele teenuse tarbijatele vastuvõetav.

Konkreetsed sünteeshääle parameetreid muudetakse siis, kui ühte kaadrisse ilmuvad subtiitrid sisaldavad mitme inimese kõnet (mitu kõnevooru, vt joonist 3), keda on ühes plaanis näha. Kõnevoorud on eristatud subtiitrites kriipsuga. Erinevate inimeste kõnevoorude markeerimiseks muudetakse sünteeshääle parameetreid: tõstetakse või langetatakse häälekõrgust ja/või kõnetempot ning muudetakse pisut hääletämbrit. Antud näites on meestegelase küsimuse „Kas te tantsite?” sünteeshääli vaikeparameetritega: kõnetempo 155 sõna minutis, häälekõrgus 110 hertsi ja hääletämbrit iseloomustav koefitsient väärtusega 0,54. Naistegelase vastuse ajal „Ei saanud mahti.” tõuseb kõnetempo 170 sõnani minutis ja keskmine häälekõrgus 121 hertsini ning hääletämbri koefitsiendi väärtuseks on 0,56. Mehe vastuse ajaks „Kasutage juhust, kuni saab.” taastatakse sünteeshääle algseid parameetreid. Hääleparameetrite muutmine peab olema märgatav, samas ei tohi nendega liialdada. Eriti tundlik parameeter on häälekõrgus, kuna subtiitrite kõnevoorudes pole isiku sugu määratud ning mees- ja naistegelaste hääled võivad kergesti minna vastasfaasi, st võib juhtuda, et meeshäälel tõstame häälekõrgust ja naishäälel langetame. Ideaalne oleks naistegelaste puhul kasutada naise sünteeshääli ja meeste puhul seda hääli, mis on treenitud meeshääle kõnekorpuse põhjal. Paraku pole lähiajal oodata, et subtiitrite tekstifailis personifitseeritakse isikute kõnevoorud, st subtiitrifailis oleks märgitud kõneleja sugu ja vanus. Seega mees-, nais- ja lapshäälega kõnesüntesaatorite sünkronne kasutamine subtiitritega pole veel võimalik.

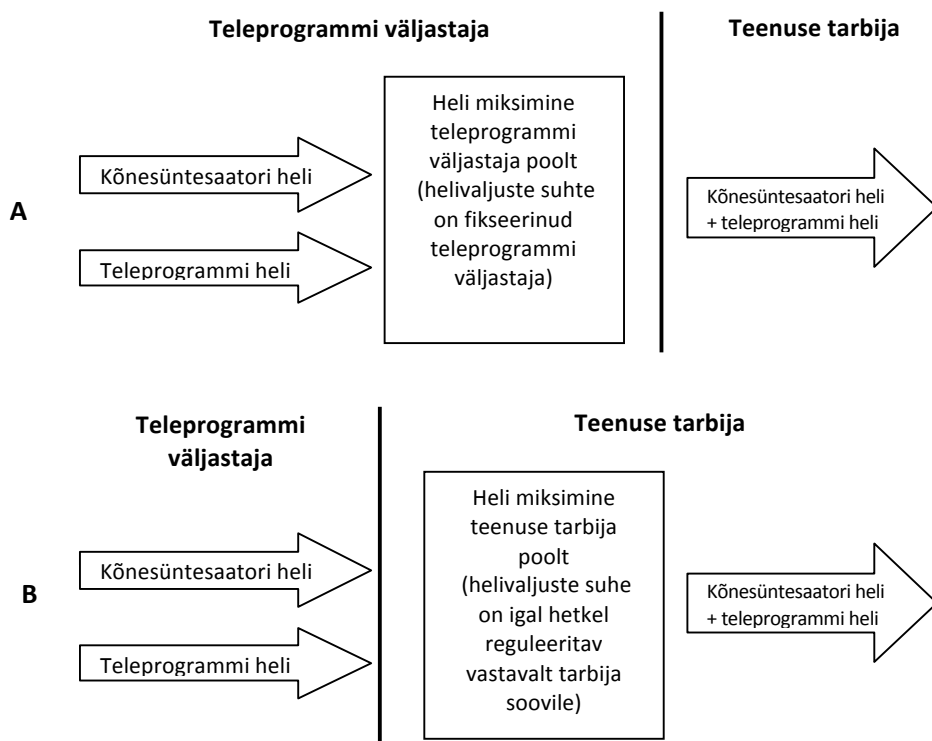


Joonis 3. Kõnevoorude vaheldumine ühes kaadris (seriaal „Maakonnahaigla”).

3. Subtiitrite edastamine tele-eeetrisse

Eesti Rahvusringhääling haldab ja juhib subtiitrite helindaja andmevooge ja tagab subtiitrite helifailide ja saadete originaalheli kokkumiksimise ning eetrisse edastamise (nn eetriliides).

Joonisel 4 on toodud erinevad variandid, kuidas eetriliidese poolt subtiitrite heli teleprogrammiheligaga kokku miksitakse. Praegu on Eesti Rahvusringhäälingu tele-eeetris kasutusel variant A, st teleprogrammi väljastaja võtab originaalsaate heli 12dB võrra vaiksemaks ja lisab sinna kõnesüntesaatori heli.



Joonis 4. Programmi- ja subtiitriheli miksimisvariandid: A – teleprogrammi väljastaja poolt eelnevalt kokkumiksitud heli, B – sobiva balansiga helikanalite kokkumiksimine televaataja poolt.

Põhimõtteliselt on uuemates televiisorites ja digiboksidest võimalus, et televaataja saab ise kahest helikanalist tuleva heli sobiva balansiga kokku liita. Kuna aga kasutusel olevate digibokside tase on väga erinev, siis praegu sellist praktilist võimalust välja pakutud ei ole.

Kui teil tekkis huvi või kui teil on tuttav või sugulane, kellel helindatud subtiitrite kuulamisest võiks televiisori vaatamisel abi olla, siis võõrkeelsete saadete või filmide korral tuleb digiboksi seadete audiovalikust puldiga aktiveerida vaegnägijatele mõeldud lisahelikanal.

4. Kokkuvõte ja edasised tööd

Subtiitrite helindamine annab suurele hulgal keeleliste erivajadustega inimestele (nagu vaegnägijad, düslektikud, eakad inimesed ja väikesed lapsed) võimaluse osa saada vöörkeelsetest filmidest ja telesaadetest. Ehkki praegu on lisahelikanalis kõnesüntesaatoriga helindatud subtiitrid kuulatavad vaid Eesti Rahvusringhäälingu telekanalites, on loodud tarkvaralahendus vabavaraline, mida alates 2014. aastast saavad tasuta kasutada kõik Eesti meediaettevõtted. Süsteemi arendamine ja täiustamine toimub ka edaspidi, nt funktsionaalsuse osas võiks kasutajatele anda võimaluse helikanaleid ise sobiva balansiga kokku miksida. Oluline on ka sünteeskõne kvaliteedi parandamine nii olemasolevate sünteeshäälte viimistlemisega kui ka uute loomulikuma kõlaga sünteeshäälte genereerimisega. Samuti võiks subtiitrite helindamisel katsetada erinevaid filmi pealelugemistehnikaid ühes võtteplaanis erinevate kõnevoorude eristamiseks. Kuna XXI sajandil on väga suur osa tavalistest meediakanalitest (raadio, televisioon, filmilevi) saadaval ka internetis, siis võiks analoogilise subtiitrite helindamise süsteemi välja pakkuda avaliku vabavaralise rakendusena internetist allalaaditavate eestikeelsete subtiitri-failide helindamiseks.

Kõnesünteesi ja subtiitrite helindamise alast arendus- ja uurimistööd on toetanud riiklik programm „Eesti keeletehnoloogia (2011–2017)” ja sihtfinantseerimisteema SF0050023s09.

Kirjandus

- A-focus. Utredning avseende TV-tillgänglighet för personer med funktionsnedsättning. Myndigheten för radio och TV. Stockholm, 2010.
- European Commission 2012 = Final report of EU high level group of experts on literacy. Luxembourg: Publications Office of the European Union, 2012.
- H a n z l í č e k, Zdenek, M a t o u š e k, Jindřich, T i h e l k a, Daniel 2008. Towards automatic audio track generation for Czech TV broadcasting: Initial experiments with subtitles-to-speech synthesis. – ICSP'08, 9th International Conference on Signal Processing, Beijing, China.
- J ø n s s o n, Morten, B o t h e, Hans Heinrich 2007. OCR algorithm for detection of subtitles in television and cinema. – CVHI'07, 5th Conference and Workshop on Assistive Technology for People with Vision and Hearing Impairments, Granada, Spain.
- K o t i s a a r i, Matti, K i l p i, Risto 2009. TV-ohjelmien tekstitys, selostus ja tulkkaus. (Liikenne- ja viestintäministeriön julkaisuja 40.) Helsinki: Liikenne- ja viestintäministeriö.
- L j u n g l ö f, Peter, D e r b r i n g, Sandra, O l s s o n, Maria 2012. <https://www.gu.se/english/research/publication/?publicationId=168968> (26. X 2013). A free and open-source tool that reads movie subtitles aloud. – SLPAT'12, 3rd Workshop on Speech and Language Processing for Assistive Technologies, Montreal, Canada, lk 1–4.
- L o o m s, Peter Olaf 2010. The future of DTV access services. – EBU technical review, nr Q4, lk 10. http://www.ebu.ch/en/technical/trev/trev_home.html (11. X 2013).

- Matoušek, Jindřich, Hanzlíček, Zdenek, Tihelka, Daniel, Ménéer, Martin 2010. Automatic dubbing of TV programmes for the hearing impaired. – 10th IEEE International Conference on Signal Processing, Beijing, China.
- Nielson, Simon, Bothe, Hans Heinrich 2007. SubPal: A device for reading aloud subtitles from television and cinema. – CVHI'07, 5th Conference and Workshop on Assistive Technology for People with Vision and Hearing Impairments, Granada, Spain.
- Verboom, Maarten, Crombie, David, Dijk, Evelien, Theunisz, Mildred 2002. Spoken subtitles: Making subtitled TV programmes accessible. – ICCHP'02, 8th International Conference on Computers Helping People with Special Needs, Linz, Austria.

Spoken subtitles – if, how, for whom and why

Keywords: spoken subtitles, speech synthesis, digital television, visual impairment, reading impairment

The article describes a system for automatic reading and broadcasting subtitles, based on synthetic speech files generated from TV subtitle files and transmitted to the television ether through a separate audio channel. Owing to the recent progress in the development of speech technology (speech synthesis and speech recognition) and the growing popularity of digital television, people with special linguistic needs such as those suffering from visual or reading impairments, can access various services connected with mass media. The software solution of spoken subtitles consists of two logical components, namely, an interface of automatic reading and voice editing and another interface for the management and transmission of subtitle and speech files. The editor interface analyses the subtitles and makes a list of those foreign names whose pronunciation is not available in the foreign names database, enabling the editor to add pronunciations and check the correctness of the spoken subtitles. In addition, the editor can, in principle, select among different synthetic voices and change their parameters, if necessary. An Estonian speech synthesizer translates subtitles into speech and synchronizes it with the film or broadcast. The ether interface manages and controls the data streams for automatic reading and is responsible for mixing the spoken subtitle files and the original broadcast sound as well as for their transmission into the ether. The spoken subtitles are, first and foremost, meant to help people with reading impairment and those of the visually challenged who can see what is going on in the TV screen, but cannot read the small font of the subtitles. Besides the dyslectics and the visually handicapped, spoken subtitles could be of help to the elderly TV-viewers whose vision has dropped, and to the children who cannot read really well yet. Also, students of Estonian as a foreign language could be interested in the service. Although at present the subtitles read by synthetic voice and transmitted through a separate sound channel can be used only with the ETV and ETV2 channels of Estonian Public Broadcasting, the software solution is freeware, available free of charge to all Estonian media companies.

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

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

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

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

Risto Sirts (b. 1980), MSc, technical director, Estonian Public Broadcasting, risto.sirts@err.ee

Tanel Valdna (b. 1987), technical manager, Estonian Public Broadcasting, tanel.valdna@err.ee