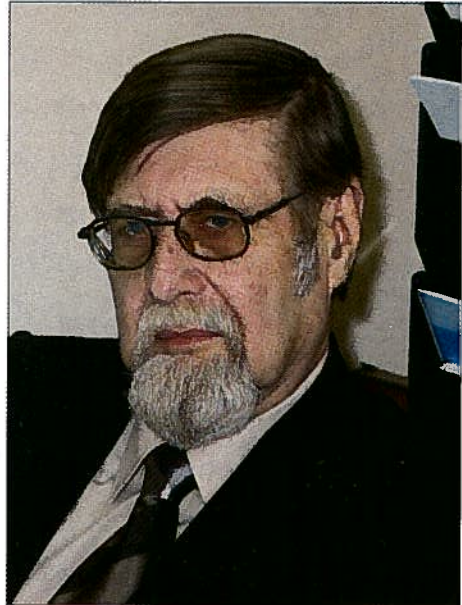


Teaduspreemia tehnikateaduste alal töö
"EESTI KEELE TEKST-KÕNE SÜNTEES" eest



Meelis Mihkla

Sündinud 6.07.1955 Tallinnas
1973 Tallinna Reaalkool
1978 Tallinna Tehnikaülikool,
automatiseeritud
juhtimissüsteemid
2001 MA, eesti keel, Tartu Ülikool
1977–1980 Eesti TA Keele ja Kirjan-
duse Instituudi vaneminsener.
1981–1992 Küberneetika Instituudi
Arvutustehnika Erikonstrueeri-
misbüroo insener-programmeerija.
Alates 1993 Eesti Keele Instituut:
haldusdirektor, infotehnoloogia
töörühmajuht, vanemteadur.
Rahvusvahelise Foneetikateaduste
Ühingu liige.
Avaldanud üle 20 teaduspublikat-
siooni.



Arvo Eek

Sündinud 13.10.1937 Pärnus
1952 Eidapere 7-klassiline Kool
1956 Viljandi Pedagoogiline Kool,
algkooliõpetaja
1963 Tartu Ülikool, eesti filoloogia
1971 filoloogiakandidaat, Eesti TA
1994 filoloogiadoktor, Tartu Ülikool
1966–1992 Eesti TA Keele ja Kirjan-
duse Instituut: nooremteadur, labora-
tooriumi juhataja, vanemteadur.
1992–1994 Stockholmi Ülikooli
külalisuurija foneetika alal.
Alates 1994 Tallinna Tehnikaülikooli
Küberneetika Instituudi vanemteadur.
Rahvusvahelise Foneetikateaduste
Ühingu liige
Avaldanud üle 100 teaduspublikat-
siooni.



Einar Meister

Sündinud 10.10. 1957 Võrumaal
Viitinas

1973 Viitina 8-klassiline Kool
1977 Tallinna Polütehnikum, elekt-
roonika-automaatika
1982 Tallinna Tehnikaülikool,
elektroonika
1998 tehnikateaduste magister, Tal-
linna Tehnikaülikool

1982–1984 Eesti TA Küberneetika
Instituudi vaneminsener.
1985–1987 Ajutise teaduslik-tehnili-
se kollektiivi START teadur.
1988–1989 Eesti TA (nüüd TTÜ)
Küberneetika Instituudi teadur, alates
1990 laboratooriumi juhataja.

Rahvusvahelise Foneetikateaduste
Ühingu liige.

Avaldanud üle 50 teaduspublikat-
siooni.



Heiki-Jaan Kaaley

Sündinud 19.05.1962 Tallinnas

1980 Tallinna 44. Keskkool
1985 Tartu Ülikool,
majandusküberneetika
1992 MSc, informaatika, Tartu
Ülikool
1999 PhD üldkeeleteadus, Tartu
Ülikool

1985– Tartu Ülikool: teadur, labori
juhataja, vanemteadur.

Avaldanud üle 20 teaduspublikat-
siooni.

2002. a on Eesti keeletehnoloogia arengus märkimisväärne – eduka uurimis- ja arendustöö tulemusena on välja töötatud rakendusteks sobiv eestikeelse kõnesünteesi tarkvara. Saavutuse aluseks on eestikeelse kõne foneetilis-fonoloogiliste uuringute tulemuste ja teksti lingvistilise töötamise meetodite oskuslik integreerimine tänapäeva-infotehnoloogiliste võimalustega. Eestikeelse kõnesünteesi tarkvara olemasolu avab tee uute kommunikatsioonitehnoloogiate arendamiseks ja kasutuselevõtuks ning on asendamatuks abivahendiks puuetega inimestele suhtlemisel arvutiga.

MIKS ON SEE TULEMUS OLULINE?

Info- ja kommunikatsioonitehnoloogia kiire areng viimasel aastakümnel on olulisel määral muutnud meie elu- ning tööstiili – personaalarvuti, Internet ja mobiiltelefon on saanud meie igapäevasteks kommunikatsioonivahenditeks. Tehnoloogiaarendus liigub vaieldamatult arvutiga loomulikus keeles suhtlemise suunas ja *Star Treki* visiooni realiseerumine on suurte keelte (eelkõige inglise, saksa, prantsuse) puhul lähiaastate küsimus. Selline tehnoloogiline progress toob endaga kaasa keelelise diskrimineerimise – keeled, millel puudub vastav tehnoloogiline tugi, so. kõnesüntees ja -tuvastus, pole samaväärselt kasutatavad tuleviku infotehnoloogilises keskkonnas. Paratamatult tingib see keele järk-järgulise asendumise erinevates ühiskonnaelu valdkondades mingi teise, tugevama keelega. Sellist ohtu väiksematele keeltele ennustas 1992. a Euroopa Komisjoni XIII peadirektoraadi tellimisel valminud nn Danzini raport¹; UNESCO andmetel hävib maailmas iga kuu keskmiselt kaks keelt², hinnanguliselt on 100 aasta pärast maailmas kasutusel vaid 50–10%³ täna eksisteerivast enam kui 6000 keelest.

Kui sada aastat tagasi oli keele säilimise tingimuseks kirjakeele tekkimine, siis nüüd on selleks saamas keele arvutitoe, st keeletehnoloogia olemasolu⁴.

Ka eesti keel kuuluks nende ohustatud keelte hulka, kui keeletehnoloogilist uurimis- ja arendustööd Eestis ei tehtaks.

Loodud kõnesünteesi tarkvara on eesti keele arvutitoe üheks oluliseks komponendiks.

EESTIKEELSE KÕNESÜNTEESI ARENG

Eestikeelse kõnesünteesi uuringuid on tehtud TTÜ Küberneetika Instituudis ja Eesti Keele Instituudis vahelduva aktiivsusega juba paarkümmend aastat. Teadmiste kogunedes ja tehnoloogia arenedes on liigutud samm-sammult edasi parema sünteeskõne kvaliteedi suunas.

1997. a alustati uuringuid difoonidel baseeruva kõnesüntesaatori väljatöötamiseks. DIFOONID on inimkõne segmendid, mis koosnevad kahest järjesti-

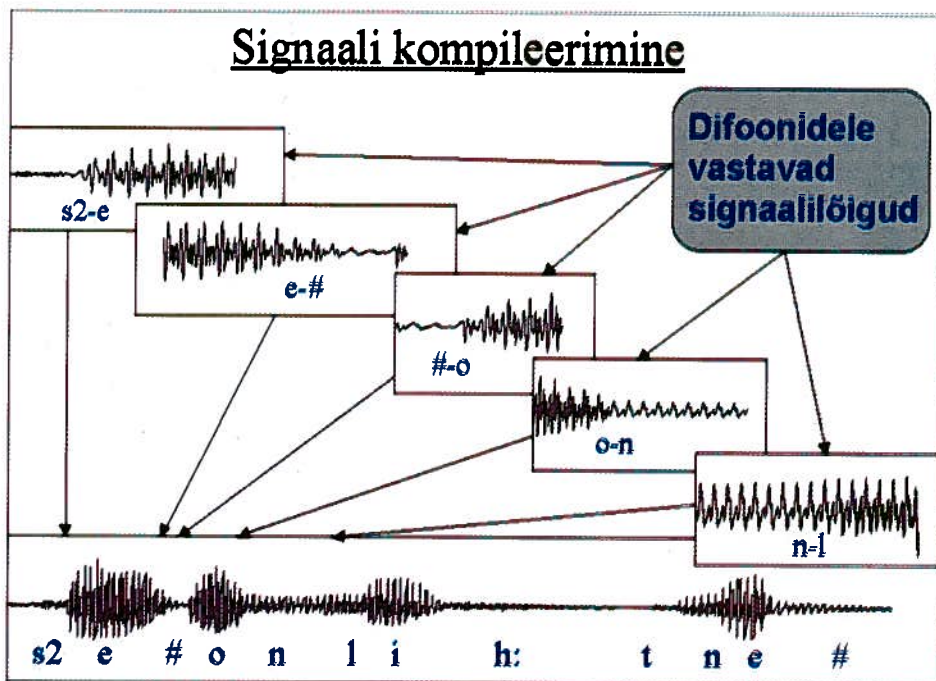
¹ Towards a European Language Infrastructure. Report by A. Danzin and the Strategic Planning Study Group for the Commission of European Communities (DG XIII). 31 March 1992.

² Atlas of the World Languages in Danger of Disappearing by Stephen A. Wurm. 2001, ISBN 92-3-103798-6.

³ M. Krauss 1992. The world's languages in crisis. – Language 68, 4-10.

⁴ H. Öim 2001. Keeletehnoloogiast ja eesti keelest. – Keel ja Kirjandus 2001, nr 7, lk 499-501.

kusest häälikust – nende algus fikseeritakse esimese hääliku stabiilse osa keskel ja lõpp järgmise hääliku stabiilses osas. Seega sisaldavad difoonid loomulikke häälikutevahelisi üleminekuid, mis lihtsustab nende rittaühendamist ja kokkusobitamist (vt joonis 1). Difoonide andmebaasi loomisel on lähtutud eesti vokaalide [Eek, Meister, 1998] ja konsonantide [Eek, Meister, 1999a] akustilispertseptiivsetest omadustest ning eesti keele foneetilisest andmebaasist [Eek, Meister, 1999b] esitatud hääliku-, rõhu- ja prosoodilise süsteemi foneetilis-fonoloogilistest kirjeldustest. Eesti keele difoonide andmebaas hõlmab kõiki võimalikke häälikuühendeid: konsonant-vokaal, vokaal-konsonant, vokaal-vokaal, konsonant-konsonant, paus-vokaal, paus-konsonant, vokaal-paus, konsonant-paus – kokku umbes 2000 difooni [Mihkla jt, 1998].



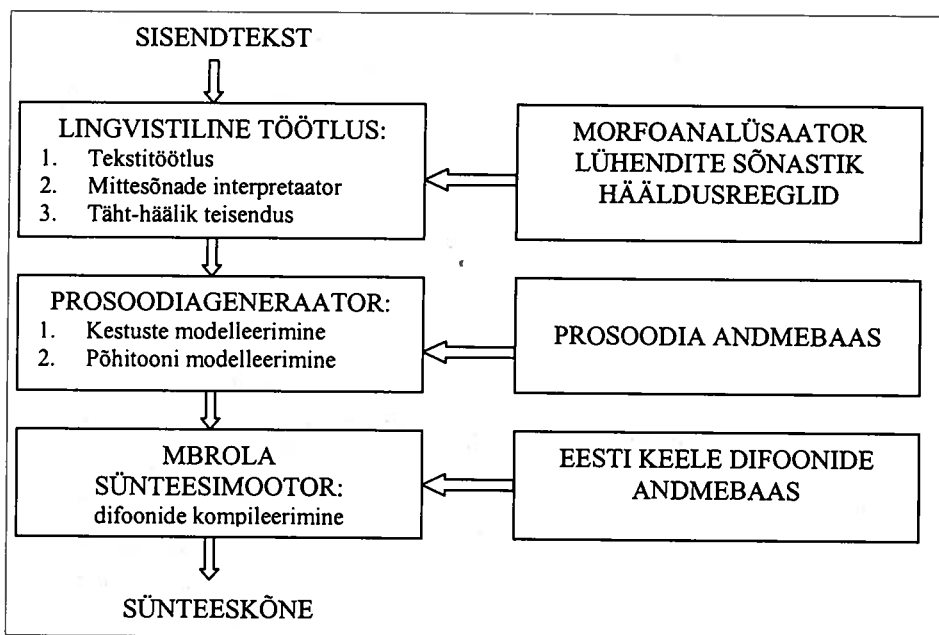
Joonis 1. Kõnesünteesil difoonide põhjal kõnelaine kompileerimine.

Difoonide rittaühendamiseks kasutatakse spetsiaalset MBROLA algoritmi, mis on välja töötatud Monsi ülikoolis Belgias.

Kuid kõnesünteesiks ei piisa ainult difoonide andmebaasist ja MBROLA algoritmist. Vajalikud on veel kaks moodulit: sünteesitava teksti lingvistiline töötlus ja prosoodiageneraator.

Eestikeelse tekst-kõne süntesaatori [Mihkla jt, 1999a; 1999b] struktuur on esitatud joonisel 2.

LINGVISTILISE TÖÖTLUSE käigus teisendatakse ortograafiline tekst hääldus-tekstiks. Kuigi kirjutatud tekst on eestlasele kergesti hääldatav, valmistab see raskusi nii mitte-eestlasele kui ka arvutile, sest eesti kirjaviis ei ole täiesti fo-



Joonis 2. Eestikeelse tekst-kõne süntesaatori struktuurskeem.

foneetiline. Kirjapildis ei ole 2. ja 3. välde üldjuhul eristatavad, eristamata on palataliseeritud konsonandid palataliseerimata konsonantidest ja palju muud. Lisaks välte ja palatalisatsiooni märkimisele leitakse lingvistilise töötlu käigus ka liitsõnapiirid, sõnarõhud ja silbipiirid – need on vajalikud prosoodiageneraatori tööks (joonis 3). Kirjalikes tekstides esineb ka suur hulk tähejärjendeid ja märke, mis enne süntesaatorile edastamist tuleb kas sõna- või tähthaaval lahti kirjutada. Lingvistilise töötlu põhimooduliks on OÜ Filosoft poolt loodud eesti keele morfoloogiline analüsaator [Kaalep, Vaino, 2001].

Erinevalt noodikirjast pole tavalises tekstis peale kirjavahemärkide ühtegi kõnetempot, helikõrgust, intonatsiooni, pausi või rõhku tähistavat märki. Teksti võib iga inimene küllalt vabalt interpreteerida, kuid arvutile on vaja ette anda reeglid, millise tempo ja intonatsiooniga tuleks üht või teist lauset lugeda. Kõnesünteesi üheks raskemaks ülesandeks ongi KÕNE PROSOODILISE STRUKTUURI MODELLEERIMINE, mis hõlmab häälikute kestuse ja lausetüübile vastava meloodiakontuuri genereerimist (joonis 4) [Mihkla jt, 2000a; 2000b].

2000. a valmis eestikeelse kõnesüntesaatori prototüüp, mis oskas juba päris korralikult rääkida, kuid ei olnud veel siiski sobilik laialdasemaks kasutuseks. Aastatel 2001–2002 teostatud mitmeetapilise sünteeskõne arusaadavuse testimise [Mihkla jt, 2001a; 2001b] ja sünteesimoodulite täiustamise järel ongi saavutatud tänane, kvalitatiivselt uus tase [Mihkla, Meister, 2002]. Võimaldamaks sünteesitarkvara kasutamist MS Windowsi keskkonnas, loodi SAPI (Speech Application Programming Interface) standardile vastav liidesprogramm ja testiti süntesaatorit erinevate programmide ning tekstidega.

Lingvistiline töötlus

Sisendtekst:

See on lihtne, sisesta tekst ja käivita
süntees.

Lingvistilise töötluse väljund:

s2e_on lih:tne3_slisesta tek:st_j7a k{ivita
syntee:s

Joonis 3. Teksti lingvistilise töötluse näide.

Prosoodia modelleerimine

s2 80	t 60	
e 150 0 129 66 131 100 122	e 85 0 126 33 128 100 127	s 60
# 2	k: 160	y 70 0 110 100 109
o 75 0 122 66 127 100 126	s 65	n 65 0 109 100 108
n 100 0 126 100 119	t 220	t 100
l 60 0 119 100 124	# 100	e 80 0 114 66 117 100 115
l 85 0 124 33 127 100 124	j7 50 0 120 100 121	e: 165 0 115 100 102
h: 205	a 55 0 121 66 122 100 115	s 100
t 100	k 60	
n 55 0 107 100 103	{ 65 0 121 66 125 100 124	
e3 70 0 103 100 100	l 65 0 124 100 121	
# 300	v 60 0 121 100 117	
s1 80	l 80 0 117 100 113	
l 65 0 134 66 138 100 137	t 180	
s 60	a 70 0 103 100 100	
e 100 0 133 100 126		
s 65		
t 100		
a 70 0 115 100 110		

Automaatselt
genereeritud
prosoodilised
parameetrid:
hääliku kestus
põhitooni
kontuur

Joonis 4. Kõne prosoodilise struktuuri modelleerimine.

Eestikeelse kõnesünteesi tarkvara on vabalt saadav Internetist
<http://www.phon.ioc.ee/synt>,
<http://www.eki.ee/keeletehnoloogia/projektid/syntees/tns.html>, veebidemo saab kuulata aadressil <http://kiisu.eki.ee>.

KES KÕNESÜNTEESI KASUTAVAD?

Eestikeelse kõnesünteesi tarkvara on tasuta vabalt kasutatav mitteäriksel ja mittesõjalistel eesmärkidel. Seda on laialdaselt levitatud nägemispuuetega arvutikasutajate hulgas. Eesti Pimedate Liidu ja Põhja-Eesti Pimedate Ühingu kaudu on tasuta jagatud 100 sünteesitarkvaraga CD-ROMi, üle saja korra on seda allalaaditud Internetist (<http://www.phon.ioc.ee/synt>).

Eriti oluline on eestikeelse kõnesünteesi tarkvara kasutuselevõtt nägemispuuetega laste õpetamisel Tartu Emajõe Koolis, sest seni kasutuses olnud soomekeelne kõnesüntees võis negatiivselt mõjutada laste keelelist arengut.

TULEVASED RAKENDUSED

Võimalikud tulevased rakendused on seotud mobiilside teenuste arendamisega, näiteks elektronposti ettelugemine mobiiltelefoni kaudu. Inimese ja arvuti vahelises suhtluses võib kõnesüntesaatorit kasutada mitmesugustes dialoogsüsteemides, väljastamaks informatsiooni sünteeskõne kujul, kuid seni, kui pole loodud eestikeelse kõnetuvastuse tarkvara, jääb sünteeskõne kasutus paratamatult piiratuks.

KAS EESTIKEELNE KÕNESÜNTEES ONGI NÜÜD VALMIS?

Kindlasti mitte. Sünteesikvaliteedi tõstmiseks on vajalik täiustada nii lingvistilise töötluse moodulit (süntaktilise analüüsi lisamine) kui ka lauseprosoodia mudeleid. Need ülesanded nõuavad aga veel põhjalikku uurimistööd.

Küll aga võib täie kindlusega väita, et eestikeelse kõnesünteesi kvaliteet on piisav selle raskusteta mõistmiseks ja sünteesitarkvara on küps laialdaseks kasutuselevõtuks.

KOKKUVÕTTEKS

Eestikeelse kõnesünteesi tarkvara loomine on ilmekas näide keeleteaduslike uuringutulemuste edukast kasutamisest keeletehnoloogilise rakenduse loomisel.

Töö aluseks on suulise kõne akustilis-pertseptiivsetel [Eek, Meister, 1998; 1999a] uuringutel põhinev eesti keele häälikusüsteemi ja prosoodia foneetilis-fonoloogiline kirjeldus [Eek, Meister, 1999b], millest lähtuvalt on koostatud eesti keele difoonide andmebaas [Mihkla jt, 1998]. Täiendades kõnesünteesi prototüüpi [Mihkla jt, 1999a; 1999b] täiuslikumate teksti lingvistilise töötluse ja prosoodia moodulitega [Mihkla jt, 2000a; 2000b; Kaalep, Vaino, 2001] ning testides sünteesikvaliteeti rahvusvaheliselt aktsepteeritud meetodika kohaselt [Mihkla jt, 2001a; 2001b], on saavutatud eestikeelse kõnesünteesi arengus kvalitatiivselt uus tase [Mihkla, Meister, 2002]. SAPI-liidese loomine on taganud sünteesitarkvara ühilduvuse MS Windowsi keskkonnaga ja andnud seega võimaluse sünteesitarkvara laialdaseks rakendamiseks uute mobiilside ning IT-teenuste arendamisel.

Eestikeelse kõnesünteesi tarkvara kasutavad laialdaselt Eesti nägemispuuetega inimesed.

KIRJANDUS:

Eek, A., Meister, E. 1998. Quality of standard Estonian vowels in stressed and unstressed syllables of the feet in three distinctive quantity degrees. Proc. of the Finnic Phonetics Symposium, Pärnu, 11.-14.08.1998. *Linguistica Uralica*, 3, 226-233.

Eek, A., Meister, E. 1999a. The perception of stop consonants: linking the strongest spectral region of the burst to the following vowel. Järvikivi, J., Heikkinen, J. (eds.) *Out Loud: Papers from 19th Meeting of Finnish Phoneticians*. *Studies in Languages*, 33, 10-23.

Eek, A., Meister, E. 1999b. Estonian speech in the BABEL multi-language database: Phonetic-phonological problems revealed in the text corpus. Fujimura et al. (eds.) Proc. of LP'98. Vol. II. Prague: The Karolinum Press, 529-546.

Kaalep, H.-J., Vaino, T. 2001. Complete morphological analysis in the linguist's toolbox. *Congressus Nonus Internationalis Fenno-Ugristarum*. Pars, V. Tartu, 9-16.

Mihkla, M., Eek, A., Meister, E. 1998. Creation of the Estonian diphone database for text-to-speech synthesis. Proc. of the Finnic Phonetics Symposium, Pärnu, 11.-14.08.1998. *Linguistica Uralica*, 3, 334-340.

Mihkla, M., Eek, A., Meister, E. 1999a. Text-to-speech synthesis of Estonian. Proc. of 6th European Conference on Speech Communication and Technology. Vol. 5. Budapest, 2095-2098.

Mihkla, M., Eek, A., Meister, E. 1999b. Diphone synthesis of Estonian. Proc. of the International Workshop on Computational Linguistics and its Applications. Vol. 2. Tarusa, 351-353.

Mihkla, M., Eek, A., Meister, E., Hein, I., Tatter, P. 2000a. Non-words interpreter, prosody generator and screen reader for the Estonian text-to-speech synthesizer. Narin'yan, A.S. (ed.) Proc. of the International Workshop Dialogue'2000. *Computational Linguistics and Its Applications*. Vol. 2 (Applications). Protvino, 399-407.

Mihkla, M., Eek, A., Meister, E. 2000b. Eesti keele tekst-kõne süntees: grafeem-foneem teisendus ja prosoodia modelleerimine. Hennoste, T. (toim.) Tartu Ülikooli üldkeeleteaduse õppetooli toimetised 1. Arvutuslingvistikalt inimesele. Tartu, 309-319.

Mihkla, M., Eek, A., Meister, E., Lasn, J. 2001a. Testing the quality of Estonian text-to-speech synthesis. Ojala, S., Tuomainen, J. (eds.) 21. Fonetikaan Päivät, Turku 4.-5.1.2001. Publications of the Department of Finnish and General Linguistics of the University of Turku. Turku, 40-45.

Mihkla, M., Meister, E., Lasn, J. 2001b. Quality evaluation of Estonian text-to-speech synthesis. Proc. of the International Workshop SPEECH and COMPUTER (SPECOM'2001), Moscow, Russia, 29-31 October, 163-166.

Mihkla, M., Meister, E. 2002. Eesti keele tekst-kõne-süntees. *Keel ja Kirjandus*, 2, 88-97; 3, 173-182.

EESTI VABARIIGI TEADUSPREEMIA

2003

TALLINN, 2003

Raamat sai ilmuda tänu

**EESTI VABARIIGI HARIDUS- JA
TEADUSMINISTEERIUMI**
toetusele

Jüri ENGELBRECHT (vastutav toimetaja)
Riigi teaduspreemiate komisjoni esimees

Helle-Liis HELP, Galina VARLAMOVA

Kaante kujundamisel kasutati 2003. a teaduspreemiate laureaaside
diplomi fotosid (Tõnu KRÜNVALD)

ISSN 1406-2321

© EESTI TEADUSTE AKADEEMIA

OÜ-INFOTRÜKK