

Kõne analüüs ja variatiivsuse mudelid

Einar Meister

TTÜ Küberneetika Instituut

Ülevaade ettekandest

- Eesmärgid
- Olulisus
- Varasemad uuringud
- Senised tulemused:
 - Mikroprosoodia
 - Makroprosoodia
 - Vene aktsent
 - Automaatne kvaliteedi hindamine
- Edasine töö
- Tegijad, finantseerimine

Eesmärgid

Uurida ja arendada kõne akustilise/foneetilise analüüsi meetodeid ning luua erinevate kõnevariatsioonide foneetilised kirjeldused ja kõnetehnoloogilisteks rakendusteks sobivad mudelid.

Projekti alamülesanded:

1. uurida ja arendada kõnesignaalide analüüsi, müratötluse ja akustilise keskkonna modelleerimise meetodeid
2. uurida kõneparameetrite varieeruvust erinevate kõnelejate vahel ja töötada välja vastavad normaliseerimismeetodid kõnetuvastuse jaoks ning leida sobivad tunnused/meetodid kõneleja identifitseerimiseks tema häälenäidete alusel
3. luua eesti keelele omaste kvantiteedinähtuste mudelid kõnetuvastuse jaoks, samuti leida meetodid akustilise informatsiooni kasutamiseks liitsõnapiiride tuvastamiseks
4. uurida aktsendiga kõne akustilisi omadusi, luua akustilised mudelid aktsendiga kõne tuvastuseks
5. arendada signaalitötlusmeetodeid aktsendi automaatseks hindamiseks, luua e-õppe tarkvara eesti keele häälduse paremaks omandamiseks

Olulisus

- Kõne loomulikkus ⇒ variatiivsus
- Kõne variatiivsusest tingitud probleemid:
 - Kõnetuvastus – erinevad kõnelejad, kõnestiilid, akustilised tingimused jne
⇒ häda suure variatiivsusega!
 - Kõnesüntees – monotoonne kõne, piiratud kõnestiilid, emotsioonide puudumine, jne
⇒ häda variatiivsuse puudumisega!
- Lahendused kõnetuvastuse (spontaanse kõne) puhul:
 - akustilised mudelid ja keelemudel treenitud spontaanse kõne korpusel
 - spontaanse kõne struktuuri modelleerimine (prosoodilised ja leksikaalsed tunnused; akustilis-foneetilised variatsioonid)
- Lahendused kõnesünteesi puhul:
 - korpuse-põhine kõnesüntees
 - sünteesitava kõnestiili prosoodilise (mikro- ja makro) struktuuri modelleerimine

Varasemad uuringud

- SÕNAPROSOODIA:
 - Häälikukestused ja põhitoon erinevates kontekstides/väldetes – Lehiste, Liiv, Eek, Meister, Mihkla ...
 - Mikroprosoodia:
 - Häälikute omakestused ja põhitoon – Lehiste, Liiv, Eek, ...
 - Vene vokaalide omakestused ja põhitoon – Ott, Kuznetsov
 - Makroprosoodia:
 - Ajaline struktuur (sh pausid) ja intonatsioonimallid loetud kõnes – Mihkla, Pajupuu, Kerge, Asu, Vende, ...
 - Eesti ja vene keele kestus- ja intonatsioonimudelid kõnesünteesiks – Mihkla, Eek, Meister, Siil, Ott, ...
 - Spontaanse kõne uuringud – TÜ suulise kõne töörihm:
 - vestlusanalüüs
 - dialoogi ja suhtlusstrateegiate modelleerimine
- Hennoste, Koit, Gerassimenko, Rääbis, Kasterpalu, Strandson, Valdisoo, Treumuth, ...

Varasemad uuringud

- Aktsendiga kõne akustika – L.Meister 2005
- Kõnekvaliteedi automaatne hindamine – mitmed teadaolevad algoritmid, rakendatakse telefoniside ja kõnekodeerimisseadmete kvaliteedi hindamiseks
- Aktsendi automaatne hindamine – üksikud töö, kus automaatseid meetodeid on kasutatud aktsendi suuruse mõõtmiseks

Mikroprosoodia akustiline analüüs

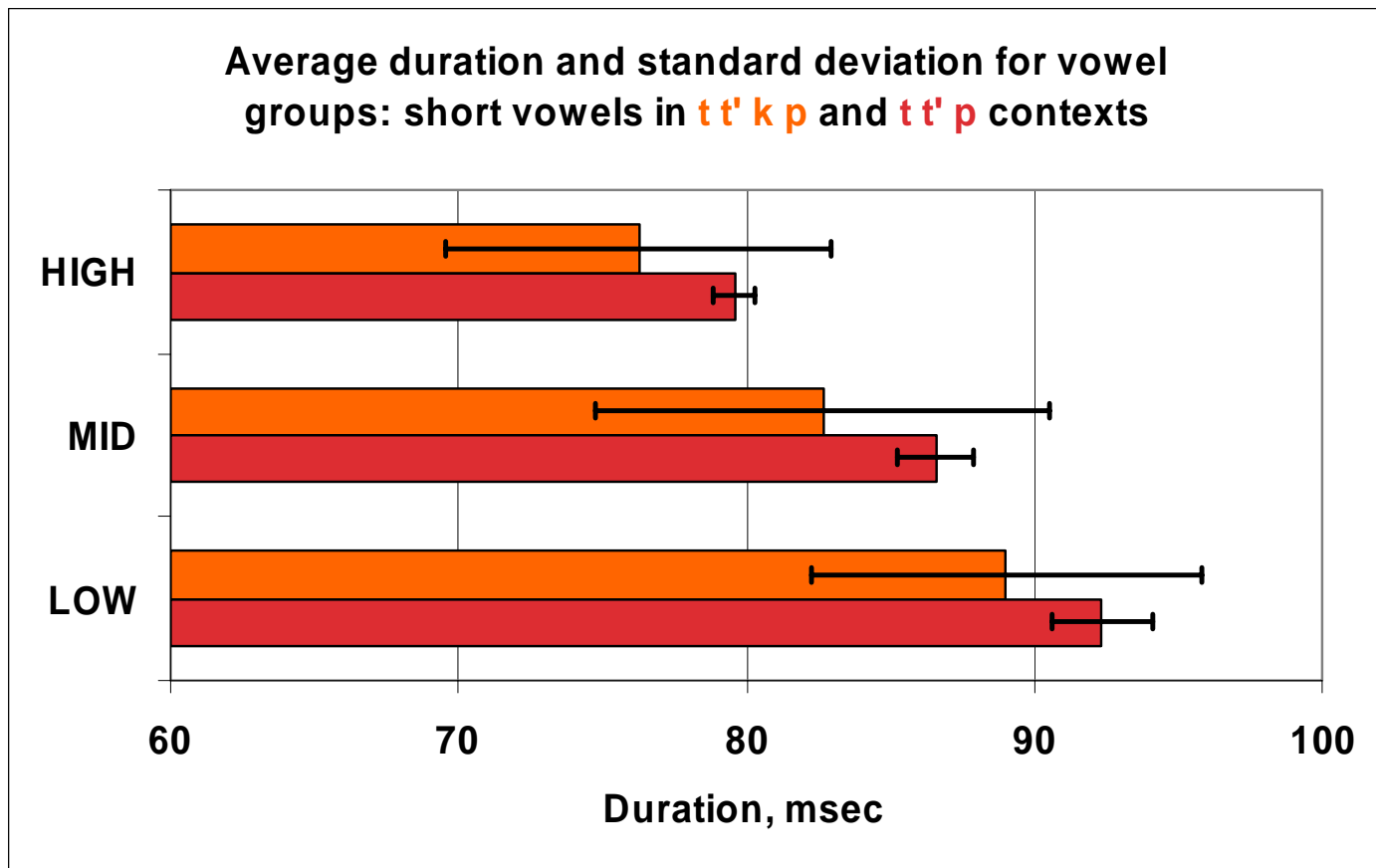
Mikroprosoodilised universaalid:

- Madalate vokaalide omakestus on pikem kui kõrgetel vokaalidel
- Madalate vokaalide põhitoon madalam kui kõrgetel vokaalidel
- Madalate vokaalide intensiivsus on suurem kui kõrgetel vokaalidel

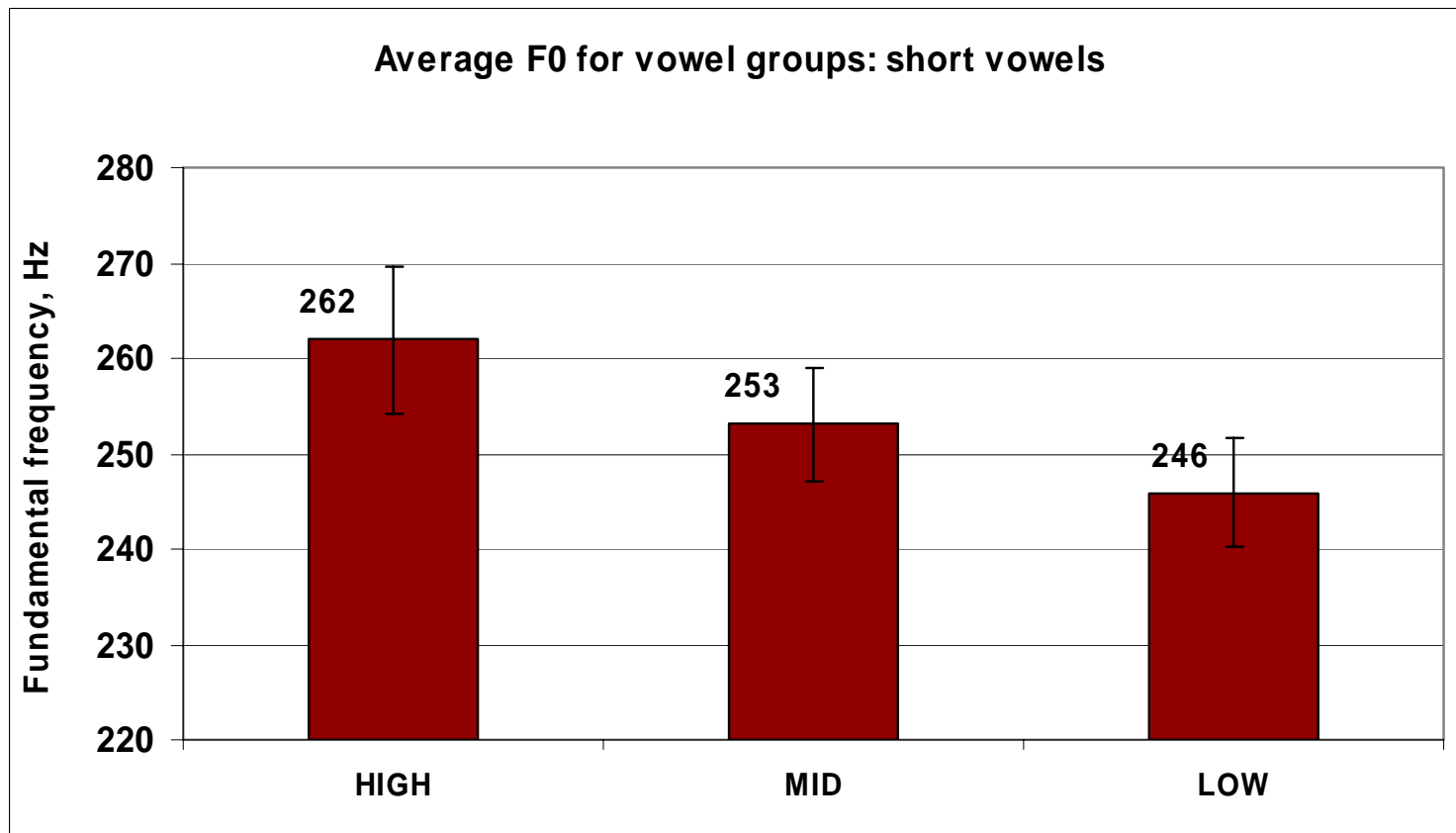
Füsioloogilised põhjused:

- “*Tongue-pull*” hüpotees
- Häälekurdude-alune rõhk
- Kõnetrakti resonantside kompenseerimine
- jm

Vokaalgruppide omakestused



Vokaaligruppide omane põhitoon



Akustiliste mõõtmiste kokkuvõte

- Madalate vokaalide omakestus on 10-15 ms pikem kui kõrgetel vokaalidel
- Madalate vokaalide põhitoon on 10-15 Hz madalam kui kõrgetel vokaalidel
- Individuaalsed erinevused
- Konteksti mõju

Meister, Einar; Werner, Stefan (2006). Intrinsic microprosodic variations in Estonian and Finnish: acoustic analysis. In: Fonetikan Päivät 2006 = The Phonetics Symposium 2006: (Toim.) Aulanko, R.; Wahlberg, L; Vainio, M.. Helsinki: University of Helsinki, 2006, (Publications of the Department of Speech Sciences, University of Helsinki), 103 - 112.

Omakestuste roll kõnetajus

- V. Kouznetsov 2001: Peceptual Role of Inherent Vowel Duration as Distinctive Feature in Russian. In: Proceedings of the XI Session of the Russian Acoustical Society, Moscow, November 19-23, 2001, pp.443-447
- **Põhitulemus:** kui vokaalikvaliteet on ebamäärane, siis lühema vokaalikestuse korral tajutakse kõrgemat vokaali
- Kas omakestusel on samasugune roll ka kvantiteedikeeltes?

Test 1: Vokaalikategooria piiri määramine

- Vokaalide /i/ ja /e/ piiri leidmine:
 - 20 sünteesstiimulit võrdsete F1/F2/F3-intervallidega /i/ ja /e/ prototüüpide vahel; kestus 100 ms, F0 = const
 - 6 kuulajat
 - Sundvalikuga test – kas kuulete /i/-d või /e/-d
 - Leiti kaks kõige ebamäärasema kvaliteediga stiimulit:

X1:

F1 = 351 Hz

F2 = 1933 Hz

F3 = 2616 Hz

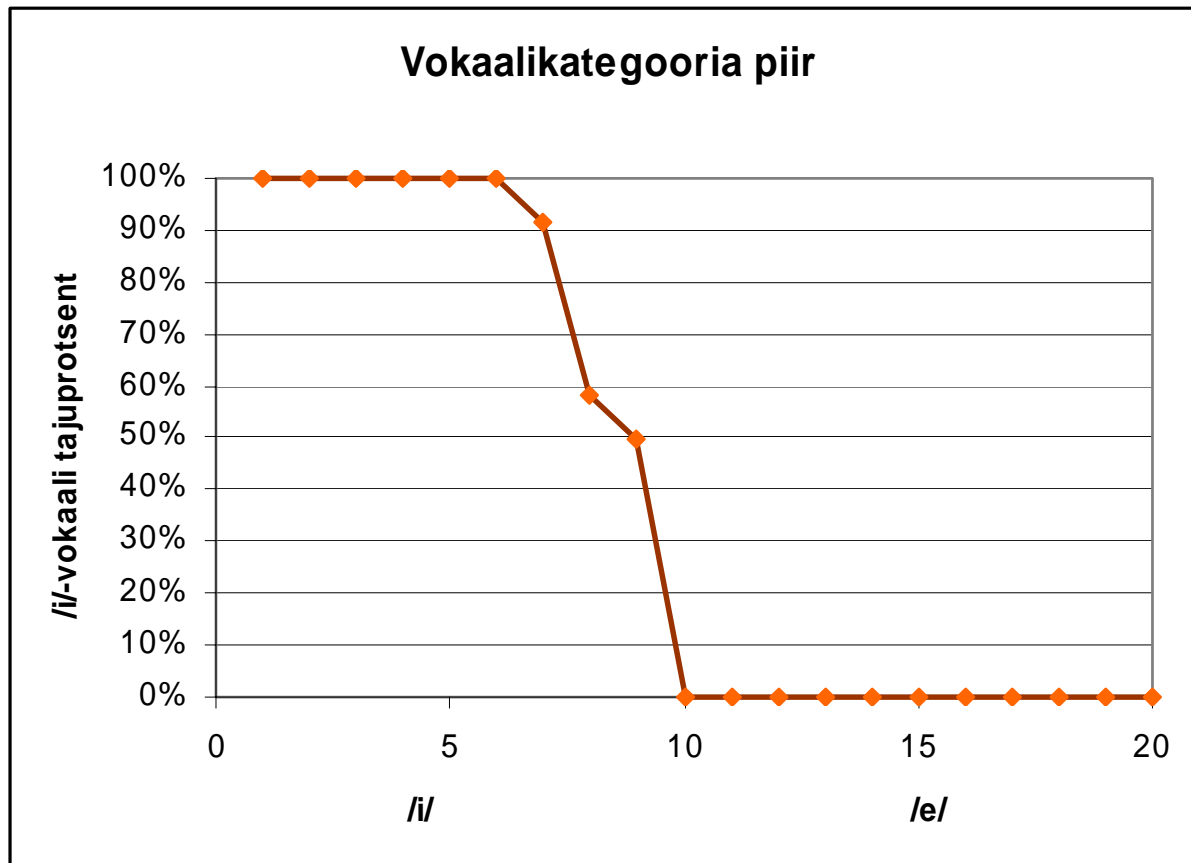
X2:

F1 = 364 Hz

F2 = 1914 Hz

F3 = 2605 Hz

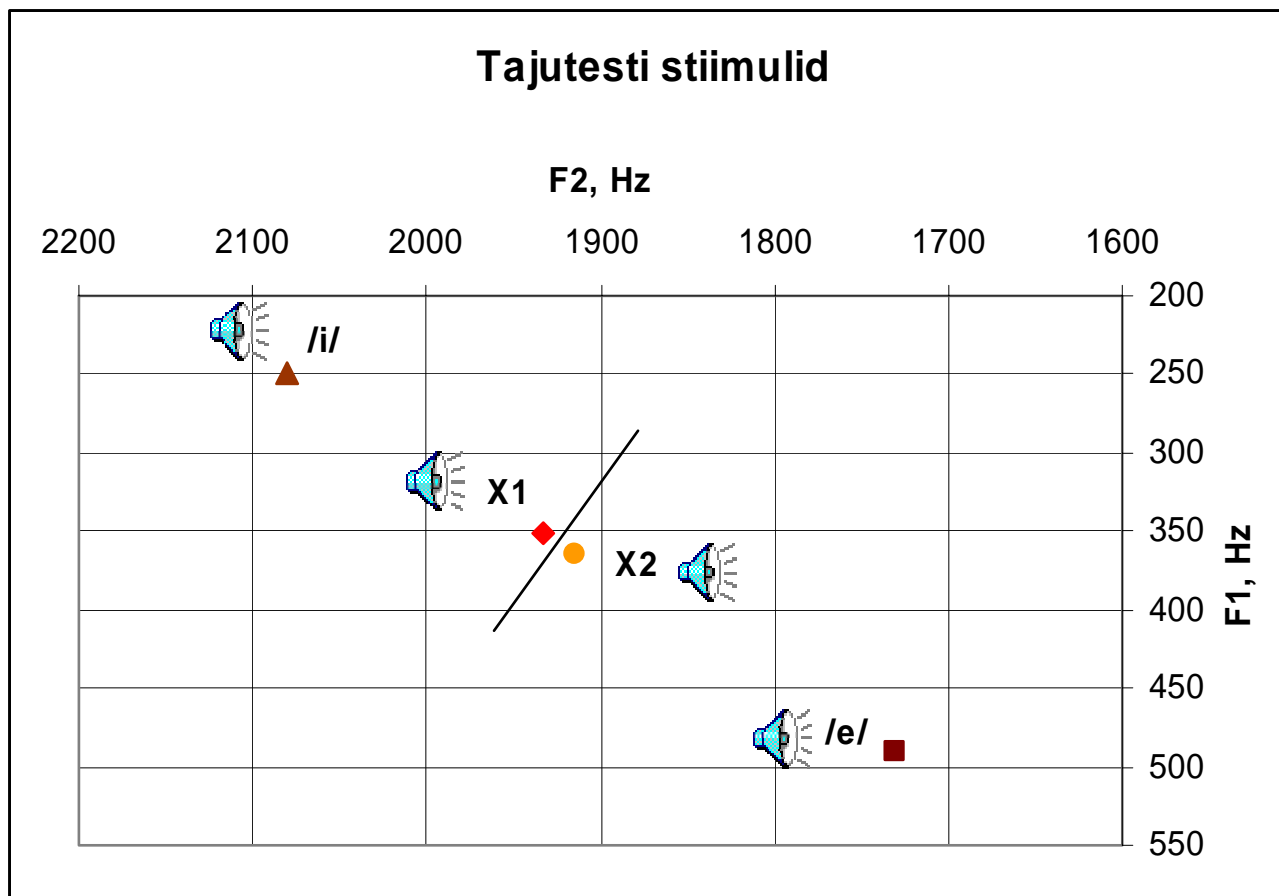
[i]-[e] piir



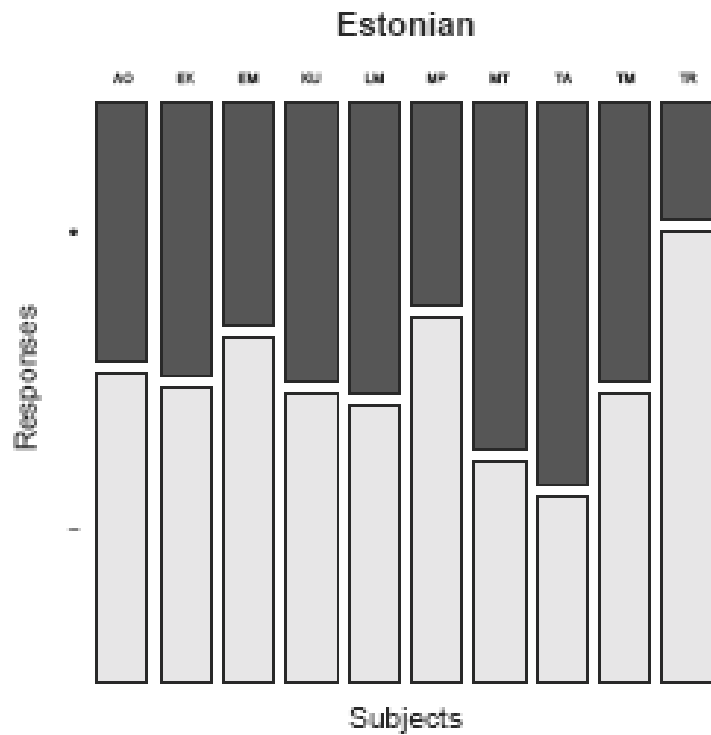
Test 2: Vokaalikvaliteedi ja omakestuse seos

- Vokaalide /i ja e/ taju sõltuvalt vokaali kestusest:
 - 10 sünteesstiimulit /i/ ja /e/ kategooria piiril, $F_0 = \text{const}$
 - Vokaalikestust muudeti 20 ms kaupa alates 60 ms kuni 140 ms
 - 10 kuulajat (5 meest, 5 naist), iga kuulaja tegi testi 2 korda
 - ABX-test, 40 ABX stiimulit + 40 BAX stiimulit ühes seerias
 - Juhuslik stiimulite järjekord
 - Sundvalikuga test – kas kuulete /i/-d või /e/-d

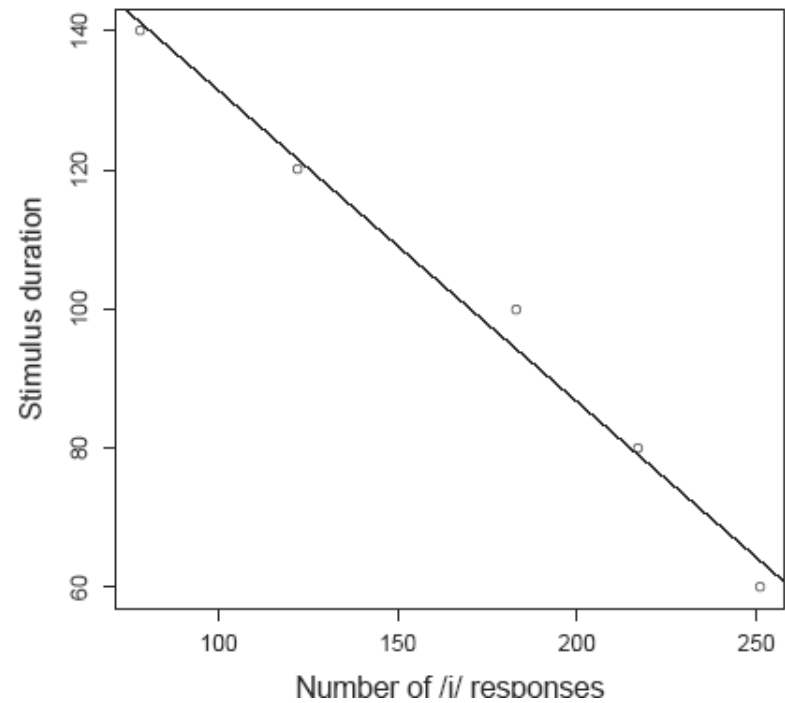
Test 2 stiimulid



Test 2: Tulemused



Linear regression: duration vs. /i/ response count



Test 2: Kokkuvõte

- Vokaalikvaliteedi tajus sõltub vokaali kestusest – pikemad vokaalid tajutakse /e/-na, lühemad /i/-na

- Werner, Stefan; Meister, Einar (2007). Microdurational influences on perceived vowel quality. The Third Baltic Conference on Human Language Technologies, Kaunas, October 4-5, 2007.

Spontaanse kõne makroprosoodia

□ Spontaanset kõne iseloomustavad:

- kordused
- parandused
- lõpetamata lausungid
- venitused
- partiklid, üneemid
- mitut liiki pausid
- sageli ei lange pausid kokku grammatiliste üksuste piiridega
- varieeruv kõnetempo
- foneetiline variatiivsus
- ...

□ Mittesoravuse põhjused:

- kõne planeerimise probleemid – kognitiivne
- suhtlussituatsioonist tingitud – kommunikatiivne
- hingamine – füsioloogiline
- ...

Uuritav kõnematerjal

- Konverentsiettekande kõnesalvestus
 - Kas see on (piisavalt) spontaanne kõne?
 - Spontaanse kõne alaliik – loengukõne?
- Keelejuht:
 - mees, eestlane, kõnehäireid pole diagnoositud
- Salvestus:
 - Edirol R1 (16 bitti, 44,1 kHz, lineaarne, mono)
 - mikrofon AKG C 444L, eelvõimendi AKG B29L
 - kestus ca 45 min

Teoreetiline raamistik

- Konversatsioonianalüüs (Chafe, Biber, ...)
- Intonatsioonianalüüs (Pierrehumbert, Beckman, ...)
- Üksused:
 - hingamisgrupp (-rühm)
 - intonatsiooniüksus (-grupp, -fraas, ...) – lähtub eelkõige intonatsioonist
 - prosoodiline üksus (-rühm, -grupp, -fraas) – lisaks F0-le hõlmab ka muid prosoodilisi tunnuseid
 - ...

Hierarhilised temporaalsed üksused

Teema – ühe slaidi esitamiseks kulunud aeg

Teemapaus (TP) – slaidivahetusega kaasnev paus

Hingamisrühm (HR) – kahe järjestikuse sissehingamise vaheline kõneintervall

Hingamispaus (HP)

Prosoodiline rühm (PR) – vaikse või täidetud pausiga eraldatud HR-sisene kõneintervall

Prosoodiliste rühmade vaheline paus (PRP)

Loengukõne temporaalse struktuuri mudel

PROSOODILINE RÜHM

PR $\approx 1,2$ s
1-3 sõna
PRP
 $\approx 0,4$ s
...

HINGAMIS- RÜHM

HR ≈ 5 s
2-4 PR
HP
 ≈ 1 s
...

TEEMA

TEEMA
 ≈ 70 s
TP
 ≈ 3 s
...

Kokkuvõte

- Sarnased tulemused teiste uurimustega (GROG-projekt):
 - 2-4 sõnalised üksused raadiointervjuudes (Strangert 2004)
 - 2-2.5 s pikkused kõnesegmendid (Horne 2006)
 - ca 2 s ajapiirangud kõnetöötuses (mälu-uuringud, Baddeley 1997)

- Esmane mudel põhiliste temporaalsete üksuste tuvastuseks

- Meister, Einar; Lippus, Pärtel (2007). On temporal organization of spontaneous Estonian: preliminary analysis results of lecture speech. The Third Baltic Conference on Human Language Technologies, Kaunas, October 4-5, 2007

Võõrkeelse aktsendi uurimine

Aktsendiga kõne:

- võõrkeelt (L2/K2) kõneleva inimese kõne
- hälbib süstemaatiliselt emakeele (L1/K1) kõnelejate hääldest
- mitte-patoloogiline
- hälbed on kuuldeliselt tajutavad
- eelkõige täiskasvanud keeleõppija “probleem”

Aktsendinähtuste klassifikatsioon

- Vene aktsendi tüüpilisemad ilmingud:
 - sõnarõhu asukoht
 - temporaalne struktuur
 - palatalisatsioon
 - lühike/pikk vastandus
 - vokaali kvaliteet
 - diftongide kvaliteet
 - klusiilide helilisus
 - intonatsioonimallid

Vene aktsent eesti keeles

Vene aktsent eesti keeles avaldub kombinatsioonina erinevatest akustilistest tunnustest, millest olulisemad on

- **temporaalse struktuuri**
- **rõhuparameetrite**
- **häälikukvaliteediga**

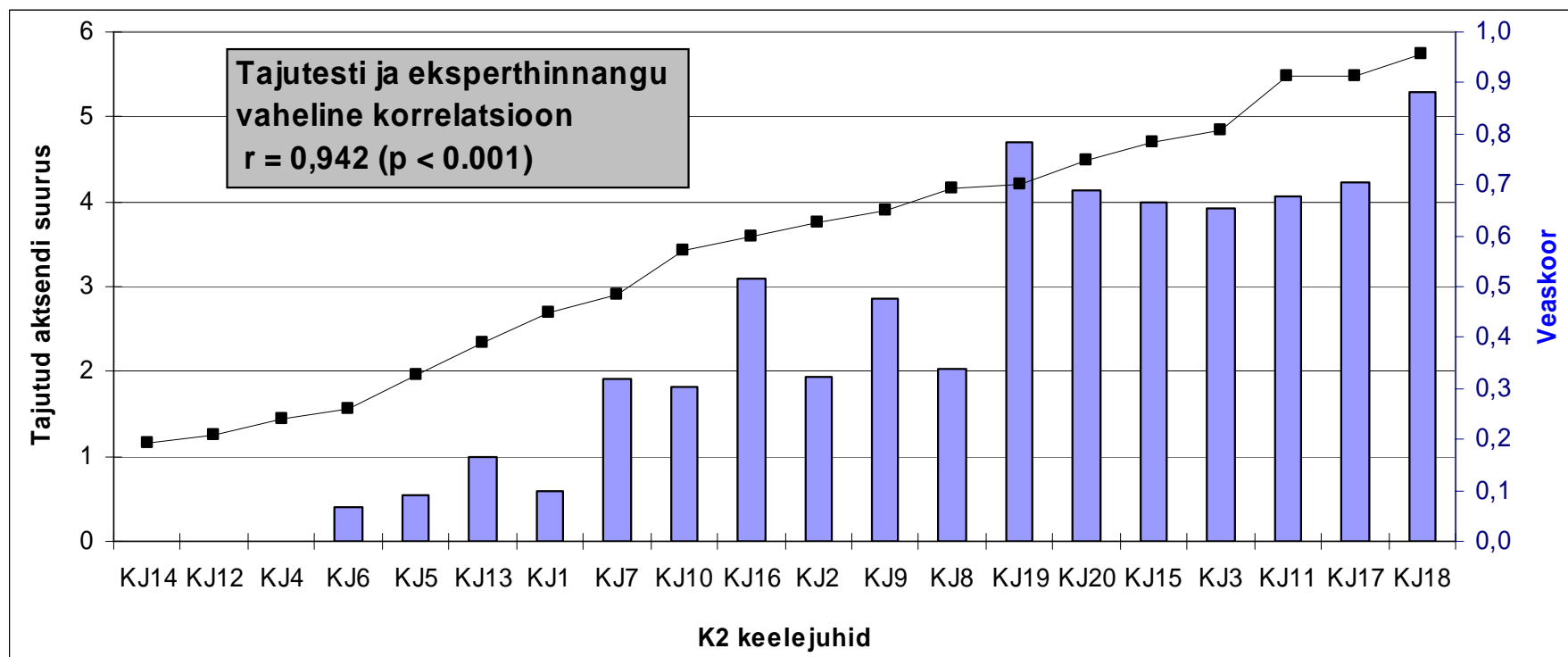
seotud tunnuste kõrvalekalded eesti normkõnest

Meister, Lya (2006). Vene aktsent eesti keele : akustilise analüüsi tulemusi. Metslang, H.; Langemets, M. (Toim.). Eesti Rakenduslingvistika Ühingu aastaraamat 2 = Estonian papers in applied linguistics 2 (131 - 152). Tallinn: Eesti Keele Sihtasutus.

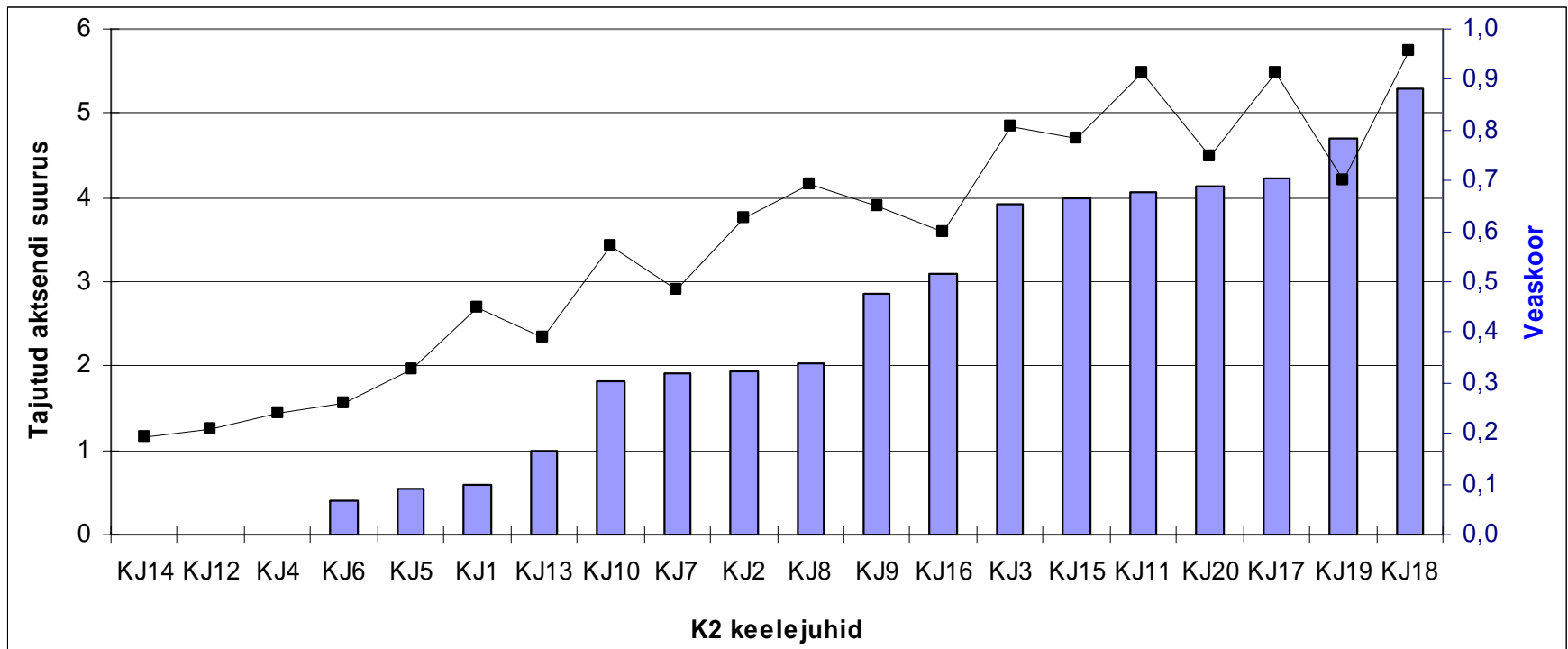
Aktsendi suuruse hindamine

- Kõnematerjal sisaldab 20 K2 keelejuhi eksamisalvestusi
- Igalt keelejuhilt 20 sek kõnelõik, stiimulite vaheline paus 5 sek
- 6-palline intervallskaala:
 - “1” - aktsendita kõne
 - ...
 - “6” - maksimaalne aktsent
- Kuulajad: 20 eesti emakeelega isikut (10 naist, 10 meest) + 2 eksperti-foneetikut

Koondtulemused: järjestatud aktsendi suuruse järgi



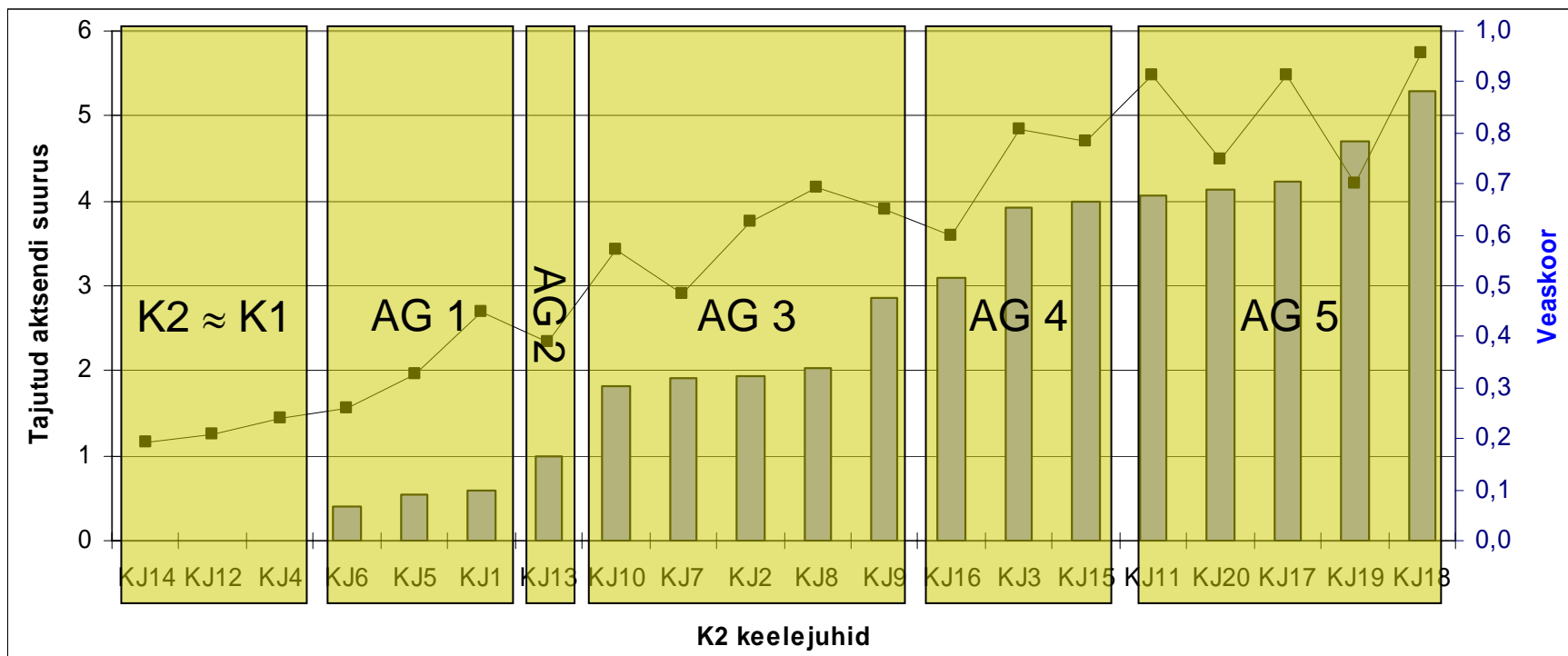
Koondtulemused: järjestatud veaskoori järgi



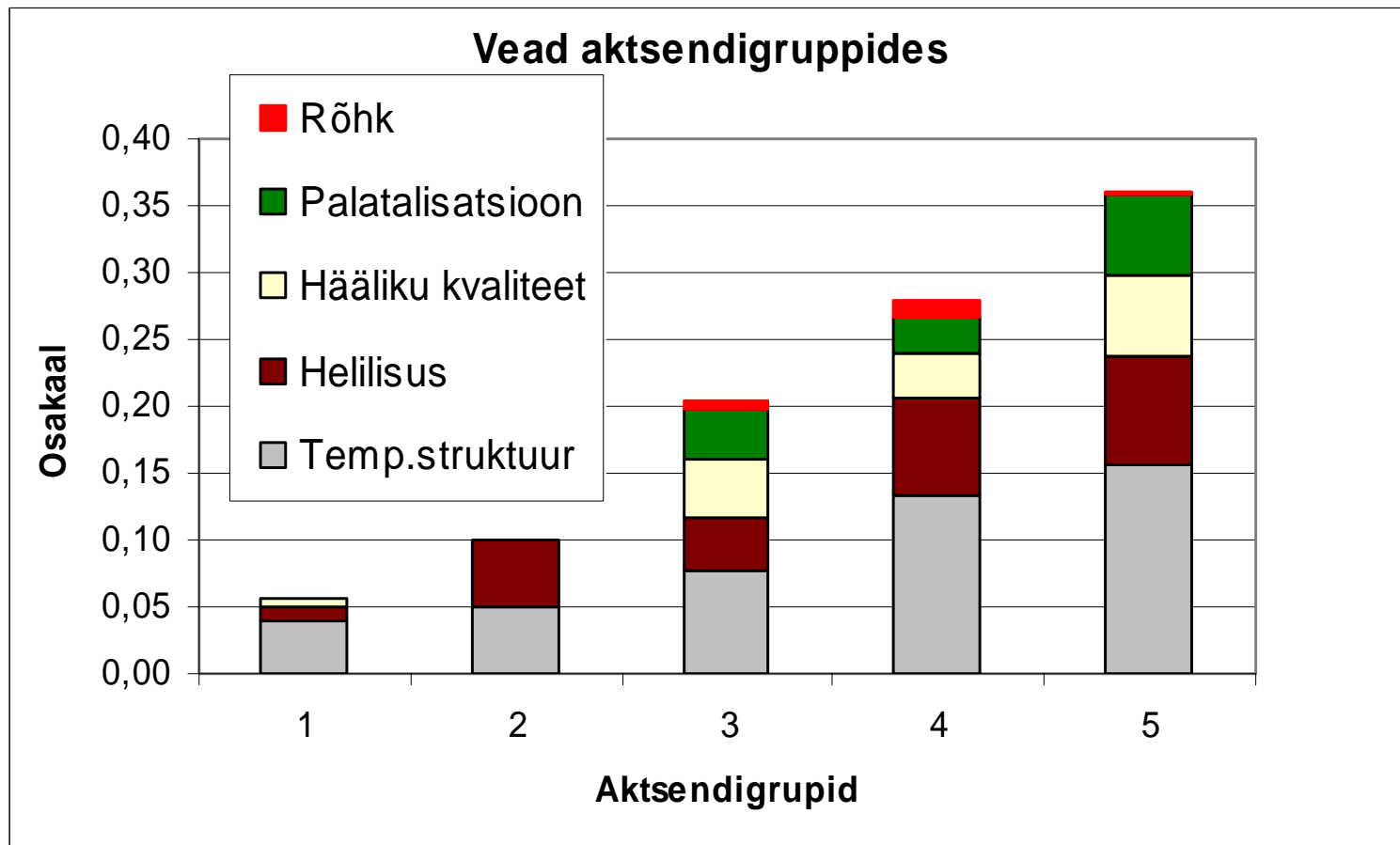
Aktsendigruppide defineerimine

Aktsendigrupp (AG)	Aktsendi suurus	Veamäär: <u>sõnade arv</u> vigade arv
1	Väga nõrk aktsent	10 - 20
2	Nõrk aktsent	4 - 10
3	Keskmine aktsent	2 - 4
4	Tugev aktsent	1,5 - 2
5	Väga tugev aktsent	1 - 1,5

Aktsendigrupid



Vead aktsendigruppides



Kokkuvõtteks

- K1 kuulajate hinnangute korrelatsioon on kõrge: 0,85 – tajuhinnangud on usaldusväärsed
- Tavakuulajate ja foneetikute hinnangute vahel on tugev positiivne korrelatsioon ($r = 0,942$) – mida kõrgem on tajutud aktsendi suurus, seda rohkem on kõnes hääldusvigu
- Sagedasemad hääldusvead (esinemissageduse järgi): temporaalstruktuur, klusiilide helilisus, häälikute kvaliteet, palatalisatsioon, rõhk.

- Meister, Lya (2007). Perceptual assessment of the degree of Russian accent. In: Proceedings of the 16th Nordic Conference of Computational Linguistics NODALIDA-2007: (Toim.) Nivre, J. et al.. Tartu: University of Tartu, 2007, 345 – 348.;
- Meister, Lya; Meister, Einar (2007). Perceptual assessment of Russian-accented Estonian. In: ICPhS XVI : Proceedings of the 16th International Congress of Phonetic Sciences, 6-10 August 2007, Saarbrücken Germany: Saarbrücken: Universität des Saarlandes, 2007, 1717 - 1720.

Kõnekvaliteedi automaatne hindamine

- PESQ – Perceptual Evaluation of Speech Quality – algoritmi omaduste testimine
- PESQ algoritmi võrdlus 3SQM (Single Sided Speech Quality Measure) algoritmiga
- Aktsent on normist hälbiv kõne → aktsendi suurust saab hinnata automaatse meetodiga, näiteks 3SQM algoritmiga (?)
- Muskat, Margus; Meister, Einar (2007). Quality estimation of time-scale modified signals. The Third Baltic Conference on Human Language Technologies, Kaunas, October 4-5, 2007.

Edasine töö

□ Mikroprosoodia:

- Tajutestid omakeestuse ja -põhitooni rolli selgitamiseks sõna kontekstis
- ...
- Võimalikud rakendused – kõne ajalise struktuuri modelleerimine kõnesünteesil

□ Spontaanse kõne makroprosoodia:

- Erinevate üksuste akustilised, leksikaalsed, süntaktilised tunnused
- Uute kõnelejate analüüs
- ...
- Võimalikud rakendused – kõne ajalise struktuuri modelleerimine spontaanse kõne sünteesil ja tuvastamisel

Edasine töö

- Aktsendiga kõne uuringud:
 - Muudest keeltest (soome, inglise, jne) tingitud aktsendinähtuste uurimine
 - Aktsendi automaatse hindamise meetodite uurimine/arendamine
 - ...
 - Võimalikud rakendused:
 - Eesti keele kui võõrkeele õpetamise metoodika täiustamine
 - Häädustreeningu programmid aktsendivaba hääduse saavutamiseks
 - ...



Tegijad, finantseerimine

Põhitäitjad:

Einar Meister

Lya Meister

Margus Muskat

Koostöö:

TÜ – Pärtel Lippus

Joensuu Ülikool – Stefan Werner

NordForsk VISPP võrgustik

EKI (ETF grant) – H.Pajupuu

Finantseerimine:

2006 – 420 000 EEK

2007 – 420 000 EEK