



## **AKUSTINEN ÄÄNENLAATUINDEKSI (AVQI) ÄÄNEN ARVIOINNISSA: ALUSTAVA MONITAPAUS- TUTKIMUS ÄÄNITYSTASON, -TILAN JA ÄÄNENTUOTTOTAVAN VAIKUTUKSISTA**

Anne-Maria Laukkanen, Puheen- ja äänentutkimuksen laboratorio, Yhteiskuntatieteiden tiedekunta, Tampereen yliopisto

Tero Ikävalko, Puheen- ja äänentutkimuksen laboratorio, Yhteiskuntatieteiden tiedekunta, Tampereen yliopisto

Leena Rantala, Puheen- ja äänentutkimuksen laboratorio, Logopedia, Yhteiskuntatieteiden tiedekunta, Tampereen yliopisto

Eliina Kankare, Foniatrian poliklinikka, Tampereen yliopistollinen sairaala

Kuudesta parametrilla muodostuvaa akustista äänenlaatuindeksiä (AVQI) käytetään äänihäiriöiden objektiiviseen arviointiin. Tämä tutkimus on alustava selvitys siitä, paljonko AVQI:in vaikuttavat äänitystaso, ympäristön taustakohina, puhujan puhetapa ja sukupuoli silloin, kun käytetään suositeltua äänityslaitteistoa.

Näytteet tallennettiin kahdelta terveääniseltä (mies ja nainen) sekä äänihuulihalvauksesta (mies) ja toiminnallisesta äänihäiriöstä (nainen) kärsineeltä henkilöltä. Terveäänisten puhujien näytteet äänitettiin vaimennetussa studiossa. He lukivat lyhyen tekstin ja äänsivät [a:]-vokaalia 5 sekuntia itselleen tavanomaisella, sekä hiljaisella, voimakkaalla ja korkealla äänellä. Huonon äänitystason ja -ympäristön vaikutusten tarkastelua varten ääninäytteet soitettiin yksitiekaiuttimen kautta ja äänitettiin

---

Kirjoittajan yhteystiedot:

Anne-Maria Laukkanen

anne-maria.laukkanen@tuni.fi

uudelleen studiossa (taustakohinataso 17,3 dBA) ja toimisto-olosuhteissa (taustakohinataso 34,4 dBA) kahdella äänitystasolla. Kaiuttimen kautta soitetuista näytteistä laskettiin AVQI.

Tulosten mukaan äänitystaso, -tila ja sukupuoli vaikuttivat vain vähän AVQI-lukuun ja sen parametreihin. Eniten AVQI-lukua nosti häiriöinen äänenlaatu. Myös puhetapa vaikutti: hiljaisesti tuotetuissa ääninäytteissä terveääniset saivat äänihäiriöille tyypillisiä arvoja. AVQI:a käytettäessä on tärkeää kontrolloida, että puhujat tuottavat ääninäytteet itselleen tavanomaisella puhevoimakkuudella ja -korkeudella. Tämä on oleellista etenkin toistomittauksissa.

**Avainsanat:** akustinen analyysi, huoneakustiikka, äänihäiriöt, äänitys

## 1 JOHDANTO

Kun ääntä arvioidaan, on tärkeää, että tallennettu näyte vastaa mahdollisimman tarkasti alkuperäistä ääntä. Tämä edellyttää sitä, että äänityslaitteisto ja -olosuhteet ovat hyvät (Barsties v. Latoszek & De Bodt, 2015; Deliyki, Shaw, Evans & Vesselinov, 2006; Svec & Granqvist, 2010; Titze & Winholtz, 1993). Esimerkiksi signaalista mitattavaan perturbaatioon ja hälypitoisuuteen vaikuttaa äänityslaitteiston (mikrofoni, esivahvistin, nauhuri) oma kohina (Titze & Winholtz, 1993). Nykyään käytettävissä digitaalisissa äänityksissä audiotallennin ei merkitsevästi lisää perturbaatiota toisin kuin aiemmin käytetyissä analogisissa äänityksissä (Doherty & Shipp, 1988). Mikrofonin tulee olla riittävän herkkä, mahdollisimman kohinatonta, taajuus- ja dynamiikka-alueeltaan riittävän laaja ja taajuusvasteeltaan mahdollisimman suora. Sähkövarauksinen kondensaattorimikrofoni on herkempi kuin dynaaminen mikrofoni (Titze & Winholtz, 1993). Svecin ja Granqvistin (2010) suosituksen mukaan mikrofonin oman kohinatason tulee olla vähintään 15 dB pienempi kuin äänitettävän näytteen hiljaisin ääni, ja dynamiikka-alueen täytyy ylittää vähintään samaan tai korkeampaan dB-tasoon kuin äänitettävän näytteen voimakkain ääni. Lisäksi mikrofonin taajuusalueen on ulotuttava vähintään äänitettävän näytteen matalimmasta perustaajuudesta yli sen taajuuden, mikä on äänen spektritutkimuksen kannalta oleellista (8000 Hz on ANSI-standardin mukaisen LS1-laboratoriomikrofonin ylärajataajuus, mutta äänen tutkimuksessa suositeltavaa on, että mikrofoni tallentaa 20 Hz -20 kHz alueen, ks Barsties v. Latoszek & De Bodt, 2015). Taajuusvasteessa saisi olla korkeintaan vain 2 dB:n poikkeamia koko taajuusalueella.

Myös äänitysetäisyys vaikuttaa signaaliin. Signaalin äänenpainetaso luonnollisesti vaimenee etäisyyden kasvaessa (kaiuttomassa

tilassa vaimenemista tapahtuu 6 dB etäisyyden kahdentumista kohden). Niin sanottu lähiefekti vahvistuu eli signaalin matalimmat osasävelet voimistuvat, kun mikrofonin etäisyys puhujan huulista on alle 30 cm (Leino & Laukkanen, 1993; Svec & Granqvist, 2010). Jos äänitysetäisyys on yli 40 cm, näytteen signaali-kohinasuhde pienenee eli kohina lisääntyy ja huoneakustiikan vaikutus kasvaa. Lähiefekti koskee suuntaavia mikrofoneja (esim. mikroфонia, jonka suuntakuvio on hertta eli kardioidi ja joka ottaa vahvimmin vastaan suoraan edestä tulevia ääniä), ja niiden välillä on lähiefektin suuruudessa eroja, jotka liittyvät mikrofonin herkkyyteen ja taajuusvasteeseen. Lähiefekti ei koske niinkään mikroфонia, jonka suuntakuvio on pallon muotoinen (engl. omnidirectional) eli mikrofoni poimii samanarvoisesti joka suunnasta tulevia ääniä. Huoneakustiikka vaikuttaa erityisesti silloin, kun äänitysetäisyys on suurempi kuin niin sanottu kaikusäde eli se etäisyys, jolla suoraan mikrofoniiin tuleva ja ympäristöstä heijastuva ääni ovat yhtä voimakkaita. Kaikusäteen suuruuteen vaikuttaa huoneen jälkikaiunta-aika: kun jälkikaiunta-aika on pitkä, on kaikusäde lyhyt. Tavallisesti kaikusäde on 50 cm – 4 m (ks. esim. Granqvist, 2020; Granqvist & Svec, 2020; Svec & Granqvist, 2010).

Jos halutaan tarkastella ääninäytteen perturbaatiota ja hälypitoisuutta, kannattaa valita lyhyt äänitysetäisyys (esim. 4-5 cm) ja siten pääpantamikrofoni (esim. Titze & Winholtz, 1993). Pääpantamikrofonin ilmeinen etu on siinä, että äänitysetäisyys pysyy helpoimmin muuttumattomana koko äänityksen ajan. Spektrirakenteen tarkasteluun kannattaa käyttää kuitenkin pitempää, esimerkiksi 30 cm tallennusetäisyyttä. Suositeltavin on yleensä pallokuvioinen kondensaattorimikrofoni, ellei äänitysympäristö ole kohinainen ja kaikuinen (ks. Svec & Granqvist, 2010).

Äänen digitoinnissa käytettävä näytteenottotaajuus vaikuttaa perturbaatioon ja siihen, miten laajalla taajuuskaistalla spektriä on mahdollisuus tarkastella. Mitä suurempi näytteenottotaajuus, sitä tarkemmin voidaan perturbaatiota mitata. Yleisin näytteenottotaajuus on tällä hetkellä 44,1 kHz, joka tarjoaa mahdollisuuden tarkastella spektriä 22050 Hz:iin saakka. Digitoinnissa käytetty bittisyvyys vaikuttaa näytteen dynamiikka-alueeseen ja signaali-kohinasuhteeseen. Tavallisimmin käytetty bittisyvyys on 16 bittiä, joka mahdollistaa 93 dB:n dynamiikka-alueen, kun taas 24 bitin syvyys ylittää 141 dB:n dynamiikkaan (Svec & Granqvist, 2010).

Parhaat äänitystilat tutkimusäänityksiä varten ovat kaiuton huone tai akustisesti vaimennettu studio. Jos äänitystilassa on taustakohinaa, kaikuisuutta (lyhyt kaikusäde) tai näytteen sisäänmenotaso on alhainen, etenkin puhe- tai lauluäänen hälyisyyttä mitattaavien akustisten parametrien arvot voivat vinoutua. Huoneen taustakohinaa voivat aiheuttaa esimerkiksi ilmastointi, laitteiden käyntiäänet, ulkoa tulevat äänet tai rakennuksen runkoäänet (Rantala & Sala, 2017). Tallennetun puhesignaalin ja huoneessa olevan taustakohinan eron tulisi olla vähintään 30 dB (Deliyski, Shaw & Evans, 2005).

Puhujan äänen perustaajuus ja voimakkuus vaikuttavat äänisignaalin perturbaatioon ja hälypitoisuuteen siten, että korkeammalla perustaajuudella ja voimakkaammin äännetyissä näytteissä on yleensä vähemmän perturbaatiota ja hälyä (Orlikoff & Baken, 1990; Orlikoff & Kahane, 1991). Näin ollen puhujan äänentuottotavalla ja myös sukupuolella on vaikutusta perturbaatioon ja hälyn määrään äänessä. Sukupuoli voi vaikuttaa ei ainoastaan perustaajuuden kautta, vaan myös sen tähden, että kurkunpään rakenteesta johtuen naisten äänirako jää helpommin takaosastaan auki (Sodersten & Lindestad, 1990; Titze, 1989),

mikä lisää todennäköisyyttä turbulenssihälyn esiintymiselle.

Tämän tutkimuksen tarkoituksena oli antaa alustavia tuloksia siitä, missä määrin Akustisen äänenlaatuindeksin AVQI:n (ks. Kankare, Rantala, Ikävalko, Barsties v. Latoszek & Laukkanen, 2020) arvoon voi vaikuttaa (1) digitaalisesti äänitetyn ääninäytteen liian alhainen äänitystaso ja (2) äänitysympäristö. Asiaa tarkasteltiin muutamien potentiaalisesti toisistaan eroavien näytteiden kautta. Mukaan otettiin näytteitä terveääniseltä mies- ja naispuhujalta, tuotettuina eri korkeuksilla ja voimakkuuksilla, ja myös näytteet äänihäiriöiseltä mies- ja naispuhujalta heidän tavanomaisella puhetavallaan tuotettuina. Tutkimuksessa käytettiin sekä AVQI:n versiota 02.02 että 03.01. Äänitykset tehtiin hyvätaoisella mutta huokealla laitteistolla, jollaista AVQI:n kehittäjät ovat suositelleet kliiniseen käyttöön (Barsties & Maryn, 2014). Tällä tutkimuksella saatava tieto on hyödyllistä käytännön kliinisessä työssä, jossa äänityksiä ei useinkaan ole mahdollista tehdä akustisesti vaimennetussa studiossa eikä käytettävissä ole paljon aikaa äänitystason säätämiseen.

## 2 MENETELMÄT

### 2.1 Äänimateriaali ja ääninäytteiden antajat

Äänimateriaalina oli näytteitä terveääniseltä mies- ja naispuhujalta sekä äänihäiriöiseltä mies- ja naispuhujalta. Terveäänisten tutkimushenkilöiden kurkunpää ja kuulo olivat normaalit, mikä oli todettu vähän aiemmin tehdyssä tutkimuksessa. Äänihäiriöisen miehen diagnoosina oli äänihuulihalvaus ja naisen toiminnallinen äänihäiriö. Viisi äänihäiriöpotilaisiin erikoistunutta puheterapeuttia arvioi potilaiden äänenlaadun kuulonvaraisesti GRBAS-asteikon äänen häiriöisyyden

astetta kuvaavalla G:llä (Grade, asteikon vaihteluväli 0–3, jossa 0 = terve ääni, 3 = vaikeasti häiriintynyt ääni) (Hirano, 1981). Naispuhujan äänelle annettu G:n keskiarvo oli 2,2 ja miespuhujan 2,8; molemmilla arvot vaihtelivat välillä 2–3.

Ääninäytteinä oli *Pohjantuuli ja aurinko* -tekstin luenta ja viiden sekunnin pituinen [a:]-vokaalin ääntö tuotettuna tutkittavalle tyypillisellä puheäänienkorkeudella ja -voimakkuudella. Lisäksi terveääniset koehenkilöt tuottivat ääninäytteet (1) hiljaisella äänellä, (2) voimistetulla äänellä ja (3) tavallista korkeammalla puhekorkeudella. Tutkittavat seisoivat tehtävien aikana.

## 2.2 Aineiston muodostaminen

### *Ensimmäinen vaihe: ääninäytteiden tallennus*

Ensimmäisessä vaiheessa terveäänisten henkilöiden ääninäytteet tallennettiin vaimennetussa studiossa pääpantamikrofonilla (AKG C 544), RME Babyface USB -äänikortilla ja Sound Forge Pro 11.0 -äänitysohjelmalla. Mikrofonin oli 4 cm päässä huulikulmasta 45°:een kulmassa. Huoneen taustakohina oli 17,3 dBA. Koska näytteiden oli tarkoitus toimia tutkimuksen toisessa vaiheessa herätesignaaleina, ne pyrittiin tallentamaan mahdollisimman hyvällä signaali-kohinasuhteella. Näytteenottotaajuus oli 44100 Hz, ja hyvän äänityslaadun varmistamiseksi äänityksessä käytettiin 24 bitin bittisyvyyttä sekä mahdollisimman suurta sisäänottotasoa ilman signaalin yliohjautumista. Tällä menetelmällä kaikkein hiljaisimpiinkin puhenäytteisiin saatiin vähintään 35 dB:n ero taustakohinan ja puhenäytteen keskimääräisen äänenpaine-tason (SPL) välille. Ero saatiin mittaamalla SPL ääninäytteen kohdista, joissa ei ollut pu-hetta ja joissa sitä oli.

Äänihäiriöisen mies- ja naispuhujan luenta-näytteet oli äänitetty Tampereen yliopistollisen sairaalan foniatrian poliklinikalla tavallisessa puheterapeutin vastaanottohuoneessa (taustakohina alle 34 dBA). Mikrofonin oli sama merkinen ja sen asema huulikulmasta oli sama kuin studiossa. Äänikorttina oli Focusrite iTrack Solo, ja äänityksessä käytettiin samaa näytteenottotaajuutta kuin studiossa, mutta alhaisempaa bittisyvyyttä (16 bittia).

### *Toinen vaihe: äänitystason ja -ympäristön vaikutusten selvittäminen*

Äänitystason ja -ympäristön vaikutuksia selvitettiin kahdella tavalla. Ensimmäisessä asetelmassa manipuloitiin herätesignaalien (alkuperäisten ääninäytteiden) signaali-kohinasuhdetta digitaalisesti siten, että niiden huippuäänitasoa alennettiin 50 dB ja 60 dB SoundForge Pro 11-ohjelmalla.

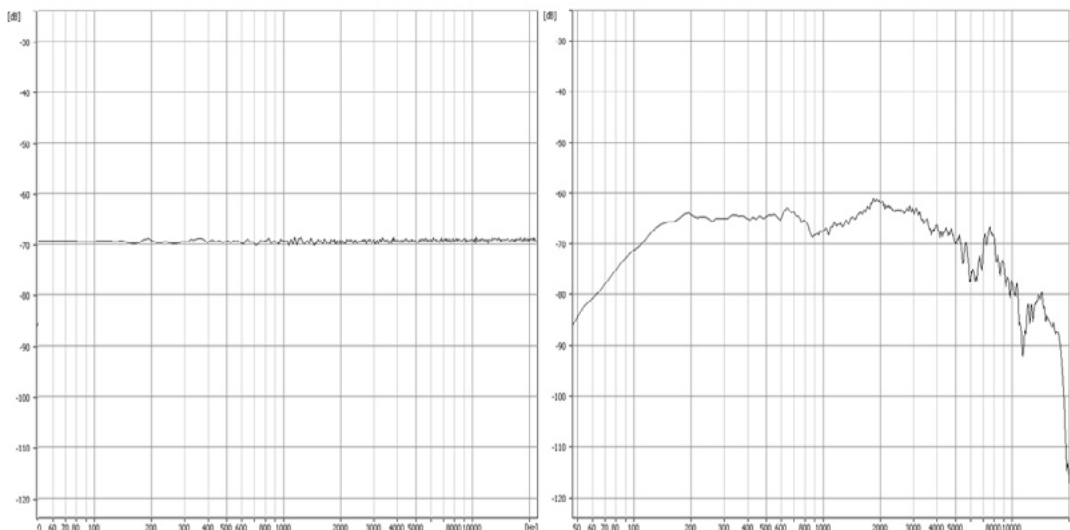
Toisessa asetelmassa herätesignaalit toistettiin kaiuttimen kautta (Auratone 5s Full-Range System –yksitiekaiutin) ja äänitettiin uudelleen kahdella eri sisäänmenotasolla (optimaalinen ja pienin mahdollinen) kahdessa eri ympäristössä (studiossa ja toimistossa). Ainoastaan kaiutintoiston avulla oli mahdollista vakiodia näytteet samoiksi, jotta äänitystason ja äänitystilän vaikutuksia pääpantamikrofonin ja äänikortin kautta äänitettyyn signaaliin voitaisiin tutkia. Kaiuttimesta syötetyn signaalin äänenvoimakkuus säädettiin kuulonvaraisesti sellaiseksi, että se vastaisi mahdollisimman hyvin alkuperäisen puhenäytteen voimakkuutta. Näin saadut uudet ääninäytteet tallennettiin AVQI-analyyssissa tavanomaisesti käytetyillä, AVQI:n kehittäjien kliiniseen käyttöön suosittelemilla laitteilla eli Focusrite iTrack Solo –äänikortilla ja AKG C 544 –pääpantamikrofonilla. Mikrofonin kiinnitettiin 45°:een kulmaan ja 4 cm:n etäisyydelle kaiutinelementin reunasta,

mikä vastasi mikrofonin sijaintia tutkimushenkilöiden äänityksissä (kuva 1). Kaiutin oli asetettu jalustalle, jonka korkeus säädettiin tutkimushenkilöiden pituutta vastaavaksi. Näytteet tallennettiin Sound Forge Pro 11.0 -ohjelmalla. Äänityksessä käytettiin 16 bitin bittisyvyyttä ja 44100 Hertsin näytteenotto-taajuutta.



Kuva 1. Äänitys vaimennetussa studiossa. Kuvassa Auratone -yksitiekaiutin ja pääpantamikrofonin sijoittaminen.

Toimistotilassa oli äänityksen aikana päällä kaksi pöytätietokonetta ja yksi kannettava tietokone, ja taustakohinasoksi mitattiin 34,4 dBA. Toimistossa tehtyjen äänitysten oli tarkoitus simuloida taustahälyltään puheterapeutin luonnollista työympäristöä. Tutkimushenkilöiden ääninäytteiden lisäksi samalla periaatteella äänitettiin äänihäiriöisen mies- ja naispuhujan ääninäytteet. Kuva 2 havainnollistaa, miten äänittämismenetelmä vaikutti alkuperäiseen signaaliin. Vaikka alkuperäiset testiäännet muuttuivatkin kaiutintois- tossa ja näytteiden uudelleen äänittämisessä, tämä menettely kuitenkin teki mahdolliseksi äänitystason ja huoneakustiikan vaikutusten mittaamisen suuntaa-antavien tulosten saamiseksi, koska koeasetelmissamme kaikkia signaaleja käsiteltiin samalla tavalla.



Kuva 2. Vasemmalla valkoisen kohinan spektri, oikealla sama kohina äänitettynä yksitiekaiuttimen kautta pääpantamikrofonilla.

### 2.3 Akustinen analyysi

Ääninäytteet analysoitiin Praat (5.3.57) -ohjelmaan (Boersma & Weenink, 2013) asennetuilla AVQI-skripteillä (sekä versiolla 02.02 että 03.01; Barsties & Maryn, 2016; Maryn & Weenink, 2015). AVQI 02.02:lla tehtävään analyysiin leikattiin tekstistä ensimmäiset 23 tavua eli sanat *Pohjantuuli ja aurinko väittelivät kummalla olisi enemmän voimaa*, ja AVQI 03.01:ella tehtävään analyysiin otettiin mukaan 8 tavua lisää eli sanat *kun he samalla näkivät*. Viiden sekunnin mittaisen [a:]-vokaalin keskikohdasta otettiin analyysiin kolmen sekunnin mittainen osa. Analyysi yhdistää ääninäytteet ja antaa tulokseksi yhden AVQI-luvun. Poikkeavaa äänenlaatua kuvaa yli 3,09 olevat arvot analysoituna skriptin 02.02 mukaan ja yli 1,83 arvot skriptin 03.01 mukaan (Kankare ym., tässä numerossa). Analyysi antaa arvon myös jokaiselle kuudelle parametrille, joista indeksi rakentuu: CPPS, HNR, shim%, shim dB, slope ja tilt (ks. Kankare ym. tässä numerossa). Myös näytteiden perustaajuuskeskiarvot ja SPL mitattiin Praatilla (ks Liite 1).

### 2.4 Tilastollinen analyysi

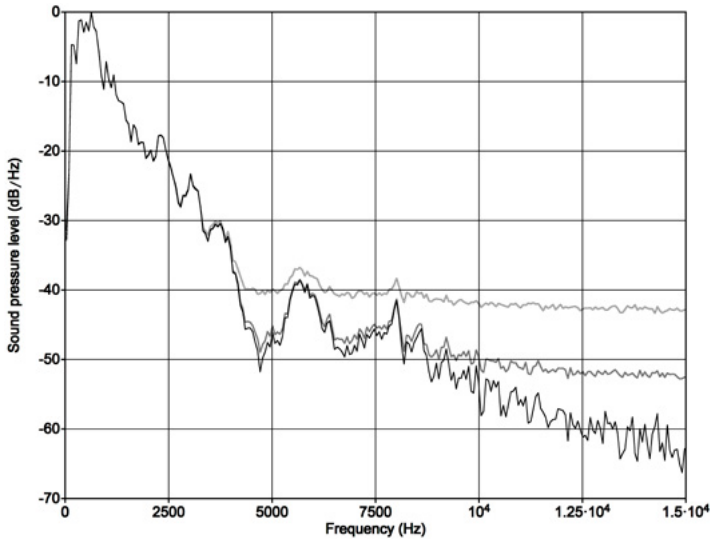
Tämän alustavan tutkimuksen näytemäärä oli kokonaisuudessaan 40: Neljältä henkilöltä tallennettiin näyte tavanomaisella voimakkuudella ja kahdelta lisäksi hiljaisella, voimakkaalla ja tavallista korkeammalla puheäänellä. Kaikki nämä näytteet (yhteensä 10) äänitettiin vielä kahdella sisäänmenotasolla ja kahdessa eri huoneessa. Kahden äänitystason ja kahden äänitystilan vaikutusta AVQI-lukuihin (AVQI 02.02 ja AVQI 03.01) tarkasteltiin pareittaisilla testeillä (Studentin t-testi), koska niiden todettiin noudattavan normaalijakaumaa.

## 3 TULOKSET JA POHDINTA

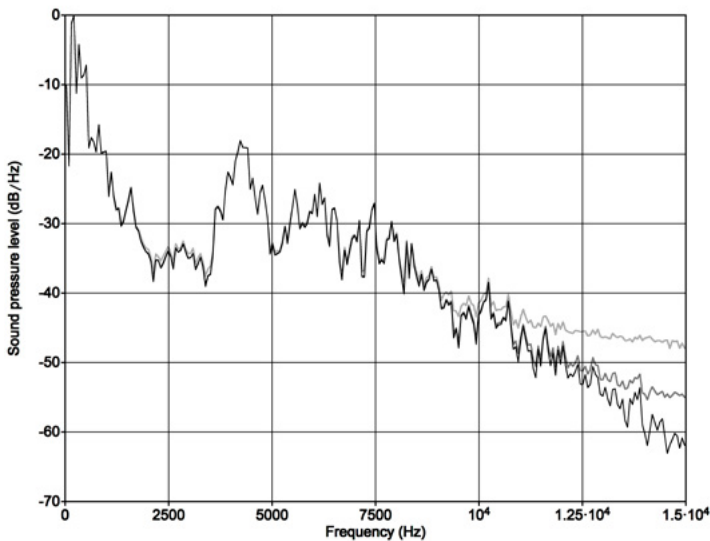
### 3.1 Signaali-kohinasuhteen digitaalisen heikentämisen vaikutukset

Kuva 3 havainnollistaa, miten signaali-kohinasuhteen pienentämisestä seurannut signaalin dynamiikka-alueen kaventuminen vaikutti spektriin siten, että korkeiden taajuusalueiden energiamäärä suhteellisesti voimistui. Dynamiikka-alueen kaventuminen näkyi selvimmin voimakkaan miesäänen spektrissä alueella 5-10 kHz.

Myös AVQI:n (03.01) ja sen parametrien (ks. Kankare ym. tässä numerossa) arvot reagoivat muutokseen, mikä näkyi selvimmin silloin, kun dynamiikka-aluetta kavennettiin todella huomattavasti, 50 dB ja 60 dB: hälypitoisuutta mittaavien CPPS:n ja HNR:n arvot alenivat, eli hälyn määrä lisääntyi, spektrin kaltevuutta kuvaavan tiltin arvo kasvoi (kaltevuus loiveni), ja myös AVQI-luku kasvoi, mikä viittasi äänenlaadun huonontumiseen (taulukko 1). Näin suuria muutoksia dynamiikka-alueeseen ei ole mahdollista saada aikaan käyttämällä äänityksessä hyvätasoista pääpantamikrofonia ja äänikorttia. Taulukossa on merkitty harmaalla ne AVQI-tulokset, joita saatiin silloin, kun dynamiikka-aluetta kavennettiin sen verran kuin oli mahdollista äänitettäessä näytteet kaiuttimen kautta huonolla äänitystasolla. Muutokset ovat vähemmän systemaattisia, HNR pieneni ja slope ja tilt kasvoivat hieman, ja AVQI-luku joko pieneni tai kasvoi hieman.



3a



3b

Kuva 3. Signaali-kohinasuhteen digitaalisen manipuloinnin vaikutus (a) miespuhujan luentaan voimakkaalla äänellä ja (b) naispuhujan luentaan hiljaisella äänellä. Musta käyrä on alkuperäisen (optimaalisella sisäänottotasolla äänitetyn) äänen keskiarvospektri, tummanharmaa käyrä kuvaa keskiarvospektriä, kun signaalia vaimennettiin 50 dB ja vaaleanharmaa käyrä, kun vaimennus oli 60 dB.



Taulukko 1. Äänitystason digitaalisen alentamisen vaikutukset AVQI:in (03.01) ja sen parametrien arvoihin kuvassa 3 näkyville signaaleille. Miespuhujan voimakas luentanäyte, naispuhujan hiljainen näyte. Harmaalla merkittynä on huonoin taso, jolla pystyttiin äänittämään kaiuttimen kautta toistettuja näytteitä.

Miespuhujaja	CPPS	HNR	Shim. (%)	Shim. (dB)	Slope LTAS	Tilt LTAS	AVQI 03.01
opt. taso	19,5	19,98	4,19	0,49	-16,01	-13,72	-1,59
-13 dB	19,56	19,86	4,67	0,57	-15,97	-13,73	-1,44
-30 dB	19,22	20	4,5	0,48	-16	-13,74	-1,5
-40 dB	18,59	19,93	4,28	0,48	-15,96	-13,41	-1,07
-50 dB	17,64	19,68	4,08	0,46	-15,98	-12,81	-0,46
-60 dB	15,96	18,11	4,38	0,5	-15,96	-11,06	0,97
<b>Naispuhujaja</b>							
opt. taso	12,19	17,74	6,14	0,62	-20,45	-10,93	2,96
-29 dB	12,16	17,66	6,06	0,62	-20,43	-10,88	2,99
40 dB	12,18	17,57	6,33	0,66	-20,42	-10,86	3,07
-50 dB	11,89	17,61	6,02	0,57	-20,47	-10,57	3,08
-60 dB	11,33	16,58	6,13	0,62	-20,29	-9,29	3,87

### 3.2 Äänitystason ja -tilan vaikutukset AVQI-lukuun

Äänitystason ja -ympäristön vaikutukset äänihäiriönäytteiden AVQI-lukuihin olivat käytännöllisesti katsoen hyvin pieniä, 0–0,61 (taulukko 2). Se, että äänitystason alentaminen vaikutti jokseenkin vähän AVQI-lukuihin, johtunee erityisesti siitä, että käytetyllä laitteistolla ei saada syntymään tarpeeksi matalaa signaali-kohinasuhdetta. (Liitteessä 2 näkyvät toimistossa tehtyjen äänitysten signaali-kohinasuhteet.) Myöskin käytetty lyhyt äänitetyisyys minimoi äänitisympäristön mahdolliset vaikutukset. Kun ääninäytteen sisäänmenotaso oli matala, tavallista hiljaisemmalla äänellä tuotetut näytteet saattoivat saada jopa pienempiä AVQI-lukuja eli parempaa äänenlaatua kuvaavia arvoja. Tätä selittää se, että osa ääninäytteen sisältämästä ja periaatteessa AVQI-lukua kasvattavasta hälystä jäi äänityksen ulkopuolelle, kun ääni-

kortin sisäänmenotaso oli minimissään. Toisaalta myöskin tilt-luku saattoi hieman kasvaa (spektrin kaltevuuden jyrkkyys vähentyä), koska laitteiston oma kvantisaatiokohina nosti suhteellisesti ottaen äänienergian määrää alueella 5–10 kHz.

Huonetilan vaikutus oli yleisesti ottaen hieman suurempi kuin äänitystason vaikutus ja edelleen suurempi silloin, kun äänitystaso oli hyvä (taulukko 3). Äänitystason ollessa hyvä AVQI 02.02-lukujen keskimääräinen ero (erotuksen itseisarvo) hyvän ja huonon äänitystilan välillä oli 0,226 ja AVQI 03.01-lukujen keskimääräinen ero oli 0,252 (ks. taulukko 3). Erot eivät olleet tilastollisesti merkitseviä (AVQI 02.02:lle  $t(9) = -0,018$ ,  $p = 0,986$  ja AVQI 03.01:lle  $t(9) = -1,383$ ,  $p = 0,200$ ). Kun äänitystaso oli huono, keskimääräinen ero (erotuksen itseisarvo) oli 0,185 AVQI 02.02:lle ja 0,153 AVQI 03.01:lle. (erot  $t(9) = -1,036$ ,  $p = 0,327$  ja  $t(9) = 0,120$ ,  $p = 0,907$ ). Huonommassa akustisessa ympä-



Taulukko 3. Äänitystilän vaikutus AVQI-lukuun. Taulukossa olevat lukuarvot ovat studiossa ja toimistossa mitattujen AVQI-lukujen erotuksia (ks. alkuperäiset arvot taulukosta 2).

<b>AVQI 02.02</b>	<b>Hyvä äänitystaso</b>	<b>Huono äänitystaso</b>	<b>AVQI 03.01</b>	<b>Hyvä äänitystaso</b>	<b>Huono äänitystaso</b>
<b>Mies</b>			<b>Mies</b>		
Tavallinen	0,09	-0,24	Tavallinen	-0,2	-0,09
Hiljainen	-0,03	0,04	Hiljainen	0,65	0,33
Korkea	0,22	0,19	Korkea	0,27	0,16
Voimakas	0,01	0,05	Voimakas	-0,03	-0,16
Äänihäiriö	0,75	0,2	Äänihäiriö	0,3	0
<b>Nainen</b>			<b>Nainen</b>		
Tavallinen	0,02	-0,37	Tavallinen	-0,08	-0,15
Hiljainen	-0,42	-0,22	Hiljainen	0,39	0,17
Korkea	-0,53	-0,41	Korkea	-0,06	-0,09
Voimakas	-0,16	-0,05	Voimakas	-0,25	-0,24
Äänihäiriö	0,03	0,08	Äänihäiriö	0,29	0,14
Itseis- arvon keskiarvo	0,226	0,185		0,252	0,153

Taulukko 4. AVQI 02.02:n ja AVQI 03.01:n parametrit eri tavoin tuotetuille näytteille.

Äänitystaso 1 = optimaalinen, 0 = huono.

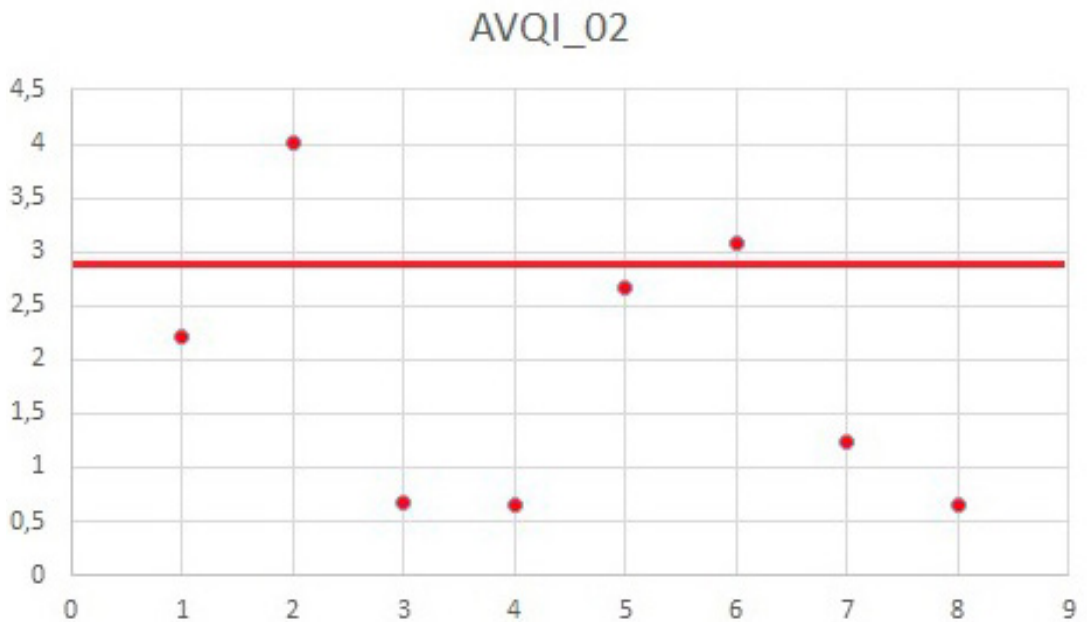
		AVQI 03.01						AVQI 02.02					
	Äänitys taso	CPPS _03	HNR _03	Shim _03	ShimdB _03	Slope _03	Tilt _03	CPPS _02	HNR _02	Shim _02	ShimdB _02	Slope _02	Tilt _02
<b>Studio</b>													
<b>Mies</b>													
tavallinen	1	16,3	16,83	7,2	0,71	-20,53	-14,5	16,63	17,43	7,05	0,71	-20,64	-14,71
tavallinen	0	16,26	16,96	7,04	0,7	-20,52	-14,03	16,61	17,64	6,72	0,63	-20,64	-14,51
hiljaa	1	10,16	16,06	7,08	0,63	-26,73	-12,92	10,32	16,42	6,88	0,6	-26,59	-12,94
hiljaa	0	10,12	15,86	6,87	0,65	-26,79	-13,09	10,24	16,17	6,71	0,62	-26,62	-12,98
korkea	1	16,91	22,8	3,8	0,42	-20,45	-13,74	17,15	23,49	3,54	0,35	-20,28	-13,68
korkea	0	16,88	22,79	3,82	0,42	-20,4	-13,75	17,12	23,48	3,55	0,35	-20,22	-13,7
voimakas	1	18,97	20,09	6,01	0,57	-17,83	-14,62	19,48	21,09	3,76	0,45	-17,9	-14,57
voimakas	0	18,96	20,07	5,97	0,57	-17,79	-14,63	19,47	21,06	3,8	0,45	-17,86	-14,59
äänihäiriö	1	3,91	5,54	14,84	1,51	-25,14	-11,87	4,05	5,65	15,07	1,53	-24,9	-11,78
äänihäiriö	0	3,96	5,5	15,23	1,45	-25,25	-11,84	4,08	5,75	14,91	1,42	-25,17	-11,75
<b>Nainen</b>													
tavallinen	1	13,8	17,73	6,61	0,63	-24,04	-13,68	13,87	18,02	5,92	0,53	-24,33	-13,48
tavallinen	0	13,78	17,75	6,58	0,63	-24,06	-13,59	13,81	18,03	6,06	0,53	-24,39	-13,43
hiljaa	1	11,9	18,31	5,79	0,63	-21,93	-12,63	12,18	18,52	5,22	0,5	-21,47	-12,75
hiljaa	0	11,91	18,33	5,5	0,58	-22,02	-12,54	12,17	18,48	5,21	0,48	-21,56	-12,65
korkea	1	14,86	21,09	3,97	0,48	-17,95	-14,18	15,23	21,74	3,76	0,38	-17,71	-14,3
korkea	0	14,87	21,13	3,95	0,48	-18,04	-14,22	15,23	21,78	3,69	0,38	-17,79	-14,3
voimakas	1	17,11	19,72	3,97	0,5	-11,28	-15,09	17,48	20,6	3,38	0,41	-10,91	-14,9
voimakas	0	17,11	19,75	3,92	0,49	-11,32	-15,06	17,44	20,57	3,53	0,44	-10,96	-15,06
äänihäiriö	1	5,22	11,72	11,23	1,11	-31,51	-13,05	5,01	12,66	10,49	1,01	-31,55	-12,8
äänihäiriö	0	5,29	11,86	11,12	1,06	-31,56	-12,96	5,1	12,86	10,41	1,01	-31,61	-12,71
<b>Toimisto</b>													
<b>Mies</b>													
tavallinen	1	16,15	16,99	7,38	0,73	-20,4	-14,17	16,5	17,67	6,92	0,67	-20,5	-14,14
tavallinen	0	16,1	16,98	7,39	0,73	-20,4	-14,17	16,42	17,63	6,9	0,67	-20,51	-14,13
hiljaa	1	11,13	16,32	6,38	0,57	-27,28	-13,12	11,26	16,61	6,14	0,59	-27,25	-13,37
hiljaa	0	10,93	16,23	7,18	0,62	-27,28	-12,36	11,06	16,52	6,96	0,65	-27,26	-12,37
korkea	1	17,02	22,84	4,01	0,36	-20,46	-13,9	17,33	23,53	3,4	0,33	-20,22	-14,19
korkea	0	16,77	22,78	4	0,36	-20,48	-13,88	17,06	23,45	3,39	0,32	-20,23	-14,17
voimakas	1	18,95	20,22	4,46	0,51	-17,95	-14,47	19,44	21,23	3,97	0,46	-18,03	-14,44
voimakas	0	18,6	20,19	4,35	0,48	-18,01	-14,48	19,05	21,21	3,89	0,43	-18,09	-14,45
äänihäiriö	1	4,05	5,73	15,01	1,42	-24,96	-11,76	4,18	5,95	14,82	1,4	-25,03	-11,71
äänihäiriö	0	4,02	5,91	16,03	1,49	-24,96	-11,8	4,16	6,05	15,89	1,47	-25,05	-11,72
<b>Nainen</b>													
tavallinen	1	13,8	17,65	6,58	0,65	-23,95	-13,63	13,93	17,94	6,13	0,55	-24,33	-13,72
tavallinen	0	13,6	17,59	6,6	0,65	-23,95	-13,62	13,72	17,88	6,22	0,61	-24,33	-13,7
hiljaa	1	12,06	18,54	5,88	0,53	-22,09	-12,74	12,34	18,7	5,33	0,59	-21,78	-12,91
hiljaa	0	11,9	18,31	5,71	0,55	-22,07	-12,83	12,14	18,54	5,78	0,56	-21,8	-12,84
korkea	1	14,82	21,05	3,8	0,47	-17,91	-14	15,16	21,5	3,44	0,43	-17,67	-14,1
korkea	0	14,74	21,05	3,63	0,45	-17,94	-13,98	15,05	21,49	3,36	0,4	-17,71	-14,1
voimakas	1	17,01	19,44	4,13	0,49	-11,29	-14,29	17,39	20,27	3,62	0,43	-10,93	-14,35
voimakas	0	16,98	19,46	4,05	0,48	-11,37	-14,3	17,36	20,29	3,61	0,43	-11,02	-14,37
äänihäiriö	1	5,35	11,96	10,11	0,95	-31,33	-12,69	5,07	13,08	10,39	1	-31,4	-12,5
äänihäiriö	0	5,29	11,93	10,22	0,95	-31,3	-12,7	5,01	13,14	10,42	1	-31,36	-12,52

Äänitystaso vaikutti varsin vähän yksittäisiin parametreihin (taulukko 4) ja näytteiden välillä oli eroa siinä, mitkä parametrit muuttuivat ja mihin suuntaan. Äänitystila vaikutti niinkään eri näytteissä erisuuntaisesti.

### 3.3 Äänentuottotavan, sukupuolen ja äänenlaadun vaikutus

Mittausten mukaan äänentuottotapa vaikuttaa AVQI-lukuun. Varsinkin hiljaisella äänellä tuotetut näytteet poikkesivat tavanomaisesti tuotetuista. Poikkeamaa oli jopa niin paljon, että hiljaa tuotettujen näytteiden AVQI-arvot ylittivät äänihäiriölle asetetun raja-arvon (taulukko 2, kuva 4; suomalaiset viitearvot: ks. Kankare ym. tässä numerossa). AVQI-luvut puolestaan pienenevät eli äänenlaatu ikään kuin parani niissä näytteissä,

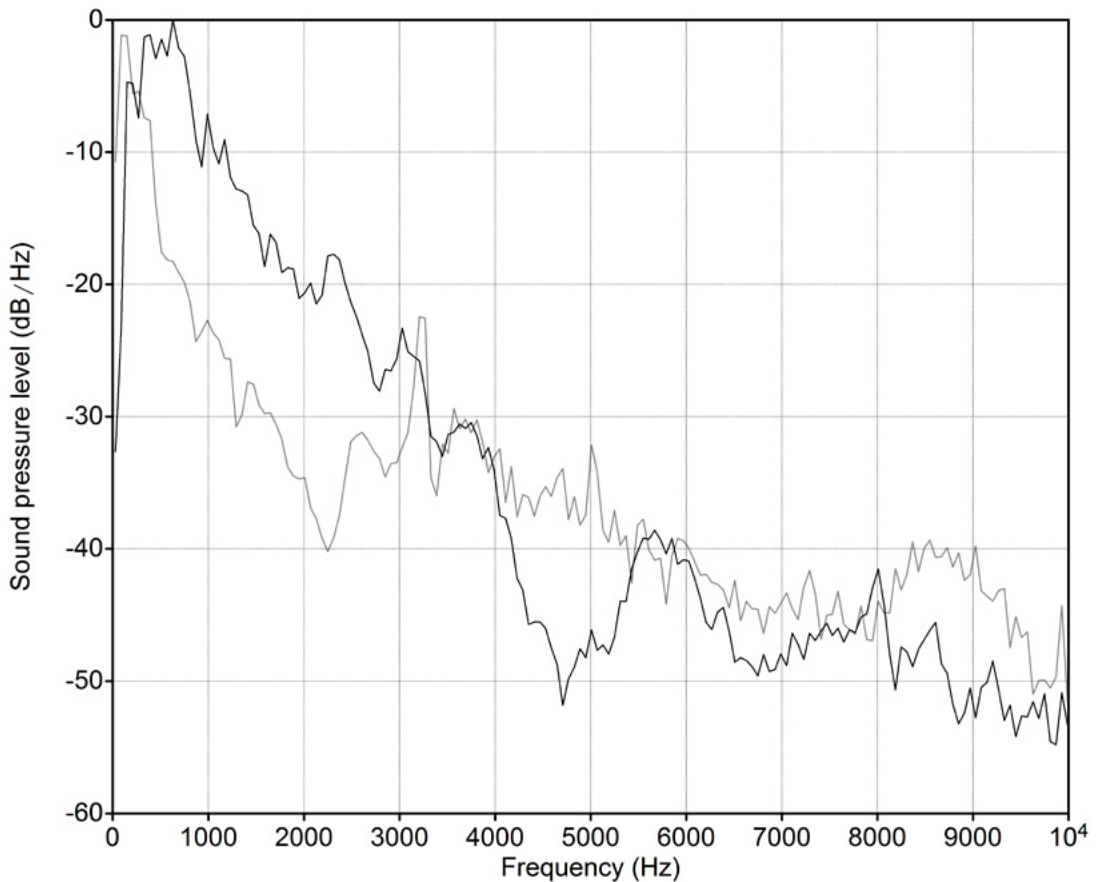
jotka oli tuotettu tavallista voimakkaammalla äänellä. Taulukossa 5 ovat eri tavalla tuotettujen näytteiden SPL ja fo. Äänentuottotapa vaikutti selvästi myös joihinkin AVQI:n parametreihin. Sekä CPPS:n että HNR:n arvot laskivat hiljaisessa ja kasvoivat voimakkaassa ja tavallista korkeammassa äänentuotossa. Shimmerin arvo pieneni kaikissa äänentuoton variaatioissa tavanomaiseen verrattuna. Osaksi muutokset voivat liittyä siihen, että kun puhuja on tietoisesti tuottanut erilaisia äänivariaatioita, hänen ääntökontrollinsa on lisääntynyt. Toisaalta aiempien tutkimusten perusteella myös tiedetään, että äänenvoimakkuuden lisääminen ja sävelkorkeuden nostaminen vähentävät shimmerin määrää (Glaze, Bless & Susser, 1990; Koike, 1973).



Kuva 4. Kahden terveäänisen henkilön ääninäytteiden AVQI 02.02-luvut, kun ääni on tuotettu tavanomaisesti, hiljaa, korkeasti ja voimakkaasti ja äänitetty vaimennetussa studiossa optimaalisella äänitystasolla. Punainen viiva osoittaa raja-arvon (2.87, Kankare ym., 2019), jonka ylittäviä arvoja mitataan äänihäiriöissä. Raja-arvon ylittävät arvot ovat hiljaisesti tuotetuista näytteistä. AVQI:n lukuarvot löytyvät taulukosta 3.

Mies- ja naisäänien välillä ei ollut suuria eroja missään AVQI:n parametreissa. Ne pienet erot, joita yksittäisissä parametreissa esiintyi, selittyivät äänentuottotapojen yksilöllisillä eroilla. Ainoastaan shimmeriä oli systemaattisesti hieman vähemmän naisella, mihin oli todennäköisesti syynä naisen korkeampi puhekorkeus. Aiemman kirjallisuuden perusteella tiedetään, että naisilla on yleensä miehiä pienemmät perturbaatioarvot (Teixeira & Fernandes, 2014). Lisäksi miehen ja naisen näytteiden spektri-arvoissa oli havaittavissa eroja. Miehen näytteissä spektrin yleinen kaltevuus (*slope*) jyrkkeni hiljaisessa ja loiveni korkeassa ja voimakkaassa äänessä (taulukko 4). Naisella puolestaan *slope* loiveni kaikissa äänentuottotavan muutoksissa. Spektrin yleisen kaltevuuden tiedetään loiventuvan voimakkuuden ja puristeisuuden lisääntyessä (Sundberg & Nordenberg, 2006). Se, että naisäänessä myös hiljaisessa äänessä *slope* arvo hieman kasvoi, johtuu todennäköisesti hiljaiseen äänentuottoon liittyneestä korkeuden noususta (tässä tutkimuksessa naisääni oli 26 Hz korkeampi hiljaisessa kuin tavanomaisella tavalla tuotetussa vokaaliäännessä, ks Liite 1). Spektrin toinen kaltevuusluku, *tilt*, käyttäytyi eri tavoin kuin *slope*, mikä johtuu siitä, että *tilt* algoritmi ottaa huomioon spektrin kaltevuuden muutoksen nopeuden taajuuden kasvaessa. Sekä miehellä että naisella *tilt* loiveni hiljaisessa ja jyrkkeni voimakkaassa äänessä. Korkeassa äänessä *tilt* loiveni miehellä, jyrkkeni naisella.

Eniten AVQI:n arvoon vaikutti puhujan äänenlaatu: äänihäiriöisillä puhujilla AVQI-luvut olivat selvästi suuremmat kuin terveäänisillä tutkimushenkilöillä (taulukko 2). Äänen häiriöisyys pienensi huomattavasti CPPs:n ja HNR:n arvoja, lisäsi shimmeriä, jyrkkensi *slope* ja loivensi *tilt*. Syy eroon näiden kahden spektrin kaltevuutta kuvaavan luvun muutoksessa johtuu erosta niiden laskentatavassa. *Slope* lasketaan erotuksena 1 kHz:n taajuudella ja 10 kHz:n taajuudella olevasta äänienergiasta. *Tilt* taas kuvaa, miten jyrkkä spektrin läpi kulkevan regressiosuoran kaltevuus on. Kuva 5 havainnollistaa terveestä ja häiriöisestä (syynä äänihuulihalvaus) äänestä piirrettyjen spektrien eroa. Häiriöisessä äänessä spektrin kaltevuus oli jyrkempi 3000 Hz:iin saakka kuin terveessä äänessä, mutta hälyn aiheuttamaa äänienergiaa oli suhteellisesti enemmän alueella 4-10 kHz.



Kuva 5. Terveen (paksu käyrä) ja dysfonisen (syynä äänihuulihalvaus) miesäänien keskiarvospektrit. Vaaka-akselilla taajuus, pystyakselilla äänenpainetaso.

### 3.4 Käytetyn tutkimusmenetelmän vaikutus tuloksiin

Missään analysoiduista näytteistä ei ollut merkittävän matalaa signaali-kohinasuhdetta (alle 30 dB). Tämä johtuu siitä, että Focusrite iTrack Solo -äänikortilla ei saa tallennetuksi kovin matalaa signaalitasoa silloin, kun signaali äänitetään AKG C544 -pääpantamikrofonilla. Matalimmalla mahdollisella äänitystasolla äänitetyn hiljaisimman näytteen (naispuhujan hiljainen näyte) signaali-kohinasuhde pysyi hieman 30 dB:n yläpuolella myös silloin, kun näyte oli äänitetty toimistossa (Liite 2). Matalimmalla mahdollisella tasolla äänitetyssä kaikkein voimakkaimmassa näytteessä (miespuhujan voimakas näyte)

signaali-kohinasuhde oli yli 50 dB sekä studioäänityksessä että toimistoäänityksessä. Optimaalisen äänitystason ja matalimman mahdollisen äänitystason välille syntyi eroa kaikkein voimakkaimman puhenäytteen kohdalla noin 12 dB ja kaikkein hiljaisimman kohdalla noin 28 dB. Näin ollen voimakkain näyte oli mahdollista äänittää ainoastaan noin 12 dB äänikortin matalinta mahdollista äänitystasoa voimakkaammalla tasolla. Tätä suurempi äänitystaso olisi aiheuttanut signaalin ylioheutuksen. Koska käytetyllä laitteistolla ei voinut äänittää kovin matalatasoista signaalia, signaalitason digitaalisella alentamisella (-50 dB ja -60 dB) saadut vaihtokäytökset ovat lähinnä teoreettisia.

Herätesignaalin toistaminen kaiuttimen kautta ja sen äänittäminen uudelleen pääpantamikrofonilla vaikutti näyttöiden spektrissä voimakkaimmin korkeilla taajuuksilla ( $> 8 < 10$  kHz), jotka vaimenivat noin 10 dB kuten kuvasta 2 on nähtävissä. Kaiutintoisto oli kuitenkin ainoa tapa vakioida näyttöet samoiksi, jotta äänitystason ja -tilan vaikutusta voitiin tarkastella. Mikäli olisimme äänittäneet samoja tutkimushenkilöitä heidän toistaessaan näyttöä useampia kertoja, emme voisi tietää, mikä vaikutus kerrasta toiseen tapahtuvalla puhettavan pienellä muutoksella olisi ollut tuloksiin. Niinikään käyttämämme kaiutin oli ainoa käytettävissämme oleva, tähän tarkoitukseen soveltuva yksitiekkaiutin. Kaiutinmenetelmän johdosta korkeilla taajuuksilla tapahtunut vaimeneminen vaikutti todennäköisimmin mittaustuloksiin niiden parametrien kohdalla, jotka reagoivat herkimmin muutoksiin spektrin kaltevuudessa ja korkeataajuuksisen hälyn määrässä (*tilt*, *slope*, *HNR*). Periaatteessa *tilt* ja *slope* voisivat tällöin saada hieman todellista pienempiä arvoja ja *HNR* hieman todellista suurempia arvoja näyttöissä, jotka ovat lähtötilanteessa kohinaisempia, ja päinvastoin lähtötilanteessa varsin kohinatomiin näyttöisiin voisi tulla mukaan laitteiston kvantisaatiokohinaa, joka kasvattaa *tilt*- ja *slope*-lukuja ja pienentää *HNR*:ää. Tällaisten muutosten myötä hälypitoisten näyttöiden AVQI-luku voisi hieman pienentyä (parantua) ja vähähälyisten näyttöiden AVQI-luku voi puolestaan kasvaa (huonontua). Liitteessä 3 on esimerkki siitä, miten käyttämämme kaiutinmenetelmä on vaikuttanut kahteen kohinaisimpaan tutkituista näyttöistä (naispuhujan hiljainen luentanäyte ja äänihuilihalvauksesta kärsivän miehen näyte) sekä terveäänisten mies- ja naispuhujan tavallisella voimakkuudella tuottamiin luentanäyttöisiin

ja miespuhujan voimakkaaseen näyttöeseen. Kaikissa esimerkinnäytteissä oli havaittavissa, että kaiuttimen kautta uudelleen äänittäminen pienensi hieman *tilt*-lukua ja *slopea* ja tuotti hieman pienemmän (paremman) AVQI-luvun.

#### 4 JOHTOPÄÄTÖKSET

Äänitutkimuksissa on pidetty tärkeänä, että äänitystasot ovat sopivat ja äänitysympäristö on vaimennettu studio. Tässä tutkimuksessa tehtyjen havaintojen perusteella äänitystaso ja -tila näyttäisivät vaikuttavan varsin vähän AVQI:n ja sen parametrien arvoihin, kunhan äänitys tehdään riittävän hyvällä laitteistolla. Kliinisessä työssä käytössä oleva työhuone on siis aivan riittävä ääninäytteen tallentamiseen. Myös AVQI-analyysiin suositeltu hyväntasoinen pääpantamikrofoni (esim. AKG C544) ja iFocusrite-äänikortti täyttävät onnistuneen tallennuksen vaatimukset. Suurempi merkitys AVQI:lla analysoituihin tuloksiin on sillä, tuottaako puhuja ääninäytteet omalla tavanomaisella puhevoimakkuudellaan ja -korkeudellaan. Hiljaisella äänellä tuotetuissa näyttöissä voidaan helposti saada äänihäiriöille tyypillisiä arvoja. Luonnollisesti AVQI:n antamien tulosten yhteydessä on aina ilmoitettava, millä skriptiversiolla analyysit on tehty, koska siitä määräytyy terveen ja sairaan äänen kynnyksarvo. Tässä tutkimuksessa tehtyjen havaintojen perusteella on suositeltavaa, että – mikäli mahdollista – näyttöiden sävelkorkeus ja äänenpainetaso vakioidaan samoiksi silloin, kun esimerkiksi arvioidaan ääniterapiassa tai interventiotutkimuksissa tapahtuneita muutoksia.



## 5 LÄHTEET

- ANSI S1.15-1997/Part 1 (R2006) American National Standard Measurement Microphones. Part 1: Specifications for Laboratory Standard Microphones, 2006.
- Awan, S. N. & Roy, N. (2009). Outcomes measurement in voice disorders: Application of an acoustic index of dysphonia severity. *Journal of Speech, Language, and Hearing Research*, 52, 482–499.
- Barsties v. Latoszek, B. & De Bodt, M. (2015) Assessment of voice quality: Current state-of-the-art. *Auris Nasus Larynx*, 42, 183–188.
- Barsties, B. & Maryn, Y. How to use the Acoustic Voice Quality Index? An objective approach measuring overall voice quality. Presentation in 11th International Voice Symposium Salzburg, August 23-24.2014.
- Barsties, B. & Maryn, Y. (2016). External validation of the Acoustic Voice Quality Index version 03.01 with extended representativity. *Annals of Otolaryngology & Rhinology*, 125, 571–583.
- Boersma, P. & Weenink, D. (2013). Praat [computer program]: *Doing Phonetics by Computer (Version 5.3.57)*. Amsterdam, The Netherlands: Institute of Phonetic Sciences.
- Bottalico, P., Codino, J., Cantor-Cutiva, L. C., Marks, K., Nudelman, C. J., Skeffington, J., Shrivastav, R. ... Rubin, A.D. (2020). Reproducibility of voice parameters: The effect of room acoustics and microphones. *Journal of Voice*, 34, 320–334.
- Buder, E. H. (2000). Acoustic analysis of voice quality: A tabulation of algorithms 1902–1990. Teoksessa R. D. Kent & M. J. Ball (toim.), *Voice quality measurement* (s. 119–244). San Diego, CA: Singular Publishing Group.
- De Bodt, M. S., Wuylts, F. L., Van de Heyning, P. H. & Croux, C. (1997). Test-retest study of the GRBAS scale: Influence of experience and professional background on perceptual rating of voice quality. *Journal of Voice*, 11, 74–80.
- Deliyski, D. D., Shaw, H. S. & Evans, M. K. (2005). Adverse effects of environmental noise on acoustic voice quality measurements. *Journal of Voice*, 19, 15–28.
- Deliyski, D. D., Shaw, H. S., Evans, M. K. & Vesselinov, R. (2006). Regression tree approach to studying factors influencing acoustic voice analysis. *Folia Phoniatrica et Logopaedica*, 58, 274–288.
- Doherty, E. T. & Shipp, T. (1988). Tape recorder effects on jitter and shimmer extraction. *Journal of Speech and Hearing Research*, 31, 485–490.
- Eadie, T. L. & Baylor, C. R. (2006). The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *Journal of Voice*, 20, 527–544.
- Granqvist, S. A tutorial on acoustic measurements for the non-technician. <https://slideplayer.com/slide/4594335/> Haettu 7.1.2020.
- Granqvist, S. & Svec, J. Microphones and room acoustics and their influence of voice. <https://slideplayer.com/slide/4812542/> Haettu 7.1.2020.
- Glaze, L. E., Bless, D. M. & Susser, R. D. (1990). Acoustic analysis of vowel and loudness differences in children's voice. *Journal of Voice*, 4, 37–44.
- Hammarberg, B., Fritzell, B., Gaufin, J., Sundberg, J. & Wedin, L. (1980). Perceptual and acoustic correlates of abnormal voice qualities. *Acta Oto-Laryngologica*, 90, 441–451.
- Hillenbrand, J., Cleveland, R. A. & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech, Language, and Hearing Research*, 37, 769–778.
- Hirano, M. (1981). Psycho-acoustic evaluation of voice. Teoksessa G. E. Arnold, F. Winckel & B. D. Wyke (toim.), *Disorders of human communication 5. Clinical examination of voice* (s. 81–84). Vienna: Springer-Verlag.
- Hosokawa, K., Barsties, B., Iwahashi, T., Iwahashi, M., Kato, C., Iwaki, S., Sasai, H. ... & Maryn, Y. (2017). Validation of the acoustic voice quality index in the Japanese language. *Journal of Voice*, 31, 260-e1. doi: 10.1016/j.jvoice.2016.05.010.
- Kankare, E., Barsties v. Latoszek, B., Maryn, Y., Asikainen, M., Rorarius, E., ym. (2019). The Acoustic Voice Quality Index Version 02.02 in the Finnish-speaking Population. *Logopedics Phoniatrics Vocology*, 30, 1–8.

- Kankare, E., Rantala, L., Ikävalko, T., Barsties v. Latoszek, B. & Laukkanen, A.-M. (2020). Akustisen äänenlaatuindeksin (AVQI) version 03.01 validointi suomenkielisille puhujille. *Puhe ja kieli*, 40, X-X.
- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A. & Berke, G. S. (1993). Perceptual evaluation of voice quality review, tutorial, and a framework for future research. *Journal of Speech, Language, and Hearing Research*, 36, 21–40.
- Karnell, M. P., Melton, S. D., Childes, J. M., Coleman, T. C., Dailey, S. A. & Hoffman, H. T. (2007). Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *Journal of Voice*, 21, 576–590.
- Koike, Y. (1973). Application of some acoustic measures for the evaluation of laryngeal dysfunction. *Studia Phonologica*, 7, 17–23.
- Leino, T. & Laukkanen, A.-M. (1993). Äänitysetäisyyden vaikutus puheäänen keskiarvospektriin. Teoksessa A. Iivonen & R. Aulanko (toim.), *Papers from the 17th meeting of Finnish Phoneticians* (s. 117–129). Publications of the Department of Phonetics, University of Helsinki, Helsinki.
- Lieberman, P. (1963). Some acoustic measures of the fundamental periodicity of normal and pathologic larynges. *The Journal of the Acoustical Society of America*, 35, 344–353.
- Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N. & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice*, 24, 540–555.
- Maryn, Y., De Bodt, M., Barsties, B. & Roy, N. (2014). The value of the Acoustic Voice Quality Index as a measure of dysphonia severity in subjects speaking different languages. *European Archives of Oto-Rhino-Laryngology*, 271, 1609–1619.
- Maryn, Y., De Bodt, M. & Roy, N. (2010). The Acoustic Voice Quality Index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*, 43, 161–174.
- Maryn, Y., Kim, H. T. & Kim, J. (2016). Auditory-perceptual and acoustic methods in measuring dysphonia severity of Korean speech. *Journal of Voice*, 30, 587–594.
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P. & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis a. *The Journal of the Acoustical Society of America*, 126, 2619–2634.
- Maryn, Y. & Weenink, D. (2015). Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and acoustic voice quality index. *Journal of Voice*, 29, 35–43.
- Orlikoff, R. E. & Baken, R. J. (1990). Consideration of the relationship between the fundamental frequency of phonation and vocal jitter. *Folia Phoniatrica*, 42, 31–40.
- Orlikoff, R. F. & Kahane, J. C. (1991). Influence of mean sound pressure level on jitter and shimmer measures. *Journal of Voice*, 5, 113–119.
- Rabinov, C. R., Kreiman, J., Gerratt, B. R. & Bielamowicz, S. (1995). Comparing reliability of perceptual ratings of roughness and acoustic measures of jitter. *Journal of Speech, Language, and Hearing Research*, 38, 26–32.
- Rantala, L. & Sala, E. (2017). Voice ergonomics for professional speakers: Who needs them? Who provides them? What do they do? Teoksessa A. Khidr (toim.), *Voice Training Programs for Professional Speakers: Global Outcomes* (s. 3–26). San Diego: Plural Publishing.
- Revis, J., Giovanni, A., Wuyts, F., & Triglia, J. M. (1999). Comparison of different voice samples for perceptual analysis. *Folia Phoniatrica et Logopaedica*, 51, 108–116.
- Rosen, C. A., Lombard, L. E. & Murry, T. (2000). Acoustic, aerodynamic, and videostroboscopic features of bilateral vocal fold lesions. *Annals of Otology, Rhinology & Laryngology*, 109, 823–828.
- Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D. & Hillman, R. (2013). Evidence-based clinical voice assessment: A systematic review. *American Journal of Speech-Language Pathology*, 22, 212–226.
- Roy, N., Gouse, M., Mauszycki, S. C., Merrill, R. M. & Smith, M. E. (2005). Task specificity in adductor spasmodic dysphonia versus muscle tension dysphonia. *Laryngoscope*, 115, 311–316.
- Svec, J. G. & Granqvist, S. (2010). Guidelines for selecting microphones for human voice production research. *American Journal of Speech-Language Pathology*, 19, 356–368.

- Sundberg, J. & Nordenberg, M. (2006). Effects of vocal loudness variation on spectrum balance as reflected by the alpha measure of long-term-average spectra of speech. *The Journal of the Acoustical Society of America*, 120, 453–457.
- Södersten, M. & Lindestad, P. A. (1990). Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech, Language, and Hearing Research*, 33, 601–611.
- Teixeira, J. P., Fernandes, P. O. (2014). Jitter, shimmer and HNR classification within gender, tones and vowels in healthy voices. *Procedia Technology*, 16, 1228–1237.
- Titze, I.R. (1989). Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America*, 85, 1699–1707.
- Titze, I.R. & Winholtz, W. S. (1993). The effect of microphone type and placement on voice perturbation measurements. *Journal of Speech and Hearing Research*, 36, 1177–1190.
- Uloza, V., Petrauskas, T., Padervinskis, E., Ulozaitė, N., Barsties, B. & Maryn, Y. (2017). Validation of the acoustic voice quality index in the Lithuanian language. *Journal of Voice*, 31, 257-e1.
- Wuyts, F. L., De Bodt, M. S., Molenberghs, G., Remacle, M., Heylen, L., Millet, B., Van Lierde, K., Raes, J. & Van de Heyning, P. H. (2000). The dysphonia severity index: An objective measure of vocal quality based on a multiparameter approach. *Journal of Speech, Language, and Hearing Research*, 43, 796–809.
- Yu, P., Revis, J., Wuyts, F. L., Zanaret, M. & Giovanni, A. (2002). Correlation of instrumental voice evaluation with perceptual voice analysis using a modified visual analog scale. *Folia Phoniatrica et Logopaedica*, 54, 271–281.
- Zraick, R. I., Wendel, K. & Smith-Olinde, L. (2005). The effect of speaking task on perceptual judgment of the severity of dysphonic voice. *Journal of Voice*, 19, 574–581.

## LIITE 1.

Tutkimuksessa käytettyjen näytteiden kalibroidut äänenpainetasot (SPL, dB) ja keskimääräiset perustaajuudet (fo, Hz). Äänitysetäisyys 40 cm. Äänitykset ja tasomittaukset tehty Bruel & Kjaer 4188-mikrofonilla ja Bruel & Kjaer 2238 Mediatorilla.

Mies	SPL teksti	fo teksti	SPL vokaali	fo vokaali
Tavallinen	75,8	126	83,5	115
Hiljainen	66,3	115	68,4	118
Korkea	76,5	173	75,8	162
Voimakas	83,3	170	87,6	167
Äänihäiriö	65,6	81	66,0	103
Nainen				
Tavallinen	71,3	173	67,9	158
Hiljainen	64,6	173	65,3	184
Korkea	79,6	263	86,1	313
Voimakas	81,2	223	85,1	209
Äänihäiriö	68,8	104-260	61,1	57

## LIITE 2.

Kaiuttimen kautta pääpantamikrofonilla äänitettyjen näytteiden signaali-kohinasuhde. Äänitystilana toimisto.

	Signaali-kohinasuhde (dB) äänityksessä	
Miespuhuja	Hyvä äänitystaso	Huono äänitystaso
Tavallinen	43,63	44,23
Hiljainen	36,5	35,88
Korkea	46,57	45,87
Voimakas	52,29	51,55
Äänihäiriö	36,38	35,33
Naispuhuja		
Tavallinen	42,19	40,89
Hiljainen	32,81	32,54
Korkea	49,62	48,33
Voimakas	50,86	50,08
Äänihäiriö	40,84	39,95

## LIITE 3.

Kaiuttimen kautta äänittämisen vaikutus viiteen erilaiseen näytteeseen.

<b>Naispuhuja</b>	<b>CPPS</b>	<b>HNR</b>	<b>Shim %</b>	<b>Shim dB</b>	<b>Slope</b>	<b>Tilt</b>	<b>AVQI 03.01</b>
hiljainen originaali	12,06	17,54	6,23	0,61	-20,78	-10,83	3,02
kaiuttimesta	12	12,23	5,89	0,61	-21,87	-12,44	2,6
tavallinen originaali	13,99	16,86	7,09	0,67	-22,58	-11,63	1,84
luenta kaiuttimesta	13,65	17,65	6,53	0,62	-23,87	-13,53	1,37
<b>Miespuhuja</b>							
tavallinen originaali	16,53	16,45	7,69	0,75	-17,68	-13,71	0,28
kaiuttimesta	16,35	16,94	7,17	0,71	-20,23	-14,68	-0,01
voimakas originaali	19,5	19,98	4,19	0,49	-16,01	-13,72	-1,59
kaiuttimesta	18,97	20,09	6,01	0,57	-17,83	-14,62	-1,58
<b>Mies, äänihäiriö</b>							
originaali	3,86	6,08	15,65	1,49	-24,73	-9,11	9,14
kaiuttimesta	3,95	5,6	16,71	1,52	-24,92	-12,10	8,25

## **THE ACOUSTIC VOICE QUALITY INDEX (AVQI) IN THE ASSESSMENT OF VOICE QUALITY: A PRELIMINARY MULTIPLE-CASE STUDY ON THE EFFECTS OF RECORDING LEVEL, ROOM NOISE AND PHONATION TYPE**

*Anne-Maria Laukkanen, Speech and Voice Research Laboratory, Faculty of Social Sciences, Tampere University*

*Tero Ikävalko, Speech and Voice Research Laboratory, Faculty of Social Sciences, Tampere University*

*Leena Rantala, Speech and Voice Research Laboratory, Faculty of Social Sciences, Tampere University / Unit of Logopedics, Faculty of Social Sciences, Tampere University*

*Eliina Kankare, Phoniatic Clinic, Tampere University Hospital*

The AVQI is used for the evaluation of dysphonia. This preliminary study investigated how much the index is affected by a low input gain level or background noise during recording, as well as the speaker's gender and way of speaking when the recommended recording device are used.

Samples were obtained from two vocally healthy speakers and two voice patients (male and female in both cases). The healthy participants were recorded in an acoustically treated studio with a head-set microphone while reading a passage and sustaining [a:] in a habitual way, softly, loudly, and in a higher than the habitual pitch. The recordings and the voice patients' samples were re-recorded in a studio and in an ordinary office using an optimal and low gain level. The AVQI was calculated.

The recording level, room noise, and speaker's gender had a minor effect on the AVQI. The dysphonic voices obtained the highest (= worst) AVQI values. The healthy participants' soft samples resulted in AVQI values typical of dysphonia. When the AVQI is utilized for voice evaluation, it is important to ensure that the voice samples are produced at a conversational loudness.

**Keywords:** acoustic analysis, dysphonic voices, recording, room acoustics