

**NÄYTETAJUUDEN JA BITTISYVYYDEN VAIKUTUS
SUBJEKTIIVISEEN ÄÄNENLAATUEROJEN
TUNNISTAMISEEN**

Heikki Kivelä
Kandidaatintutkielma
Musiikkitiede
Kevät 2012
Jyväskylän yliopisto

JYVÄSKYLÄN YLIOPISTO

Tiedekunta – Faculty Humanistinen	Laitos – Department Musiikin laitos
Tekijä – Author Kivelä Heikki Joonas Veijonpoika	
Työn nimi – Title Näytetaajuuden ja bittisyvyyden vaikutus subjektiiviseen äänenlaatuerojen tunnistamiseen	
Oppiaine – Subject Musiikkitiede	Työn laji – Level Kandidaatintutkielma
Aika – Month and year 5/2012	Sivumäärä – Number of pages 27
Tiivistelmä – Abstract <p>Äänenlaatu muodostaa yhden olennaisimmista tallennettujen äänten ominaisuuksista. Äänenlaatuerojen tunnistaminen taas liittyy keskeisesti ihmisen kuulokykyyn ja sen rajoituksiin. Tässä tutkimuksessa selvitettiin kuuntelukokeen avulla erilaatuisten äänten keskinäisiä eroavaisuuksia. Työn teoreettinen osuus keskittyy äänenlaadun historiaan, ihmisen kuulojärjestelmään ja digitaaliseen äänenkäsittelyyn jälkimmäisen osan sisältäessä kuuntelukoeasetelman esittelyn ja kokeesta saatujen tulosten esittämisen.</p> <p>Tutkimuksen tulokset osoittavat, että toisin kuin yleisesti luullaan, koehenkilöt kokivat eroavaisuuksia esimerkiksi CD-äänenlaatua korkeammilla näytetaajuuksilla. Tämä tieto hyödyttänee niitä, jotka ovat pohtineet korkeampien tallennusresoluutioiden käyttöä nykyaikaisessa äänityöskentelyssä. Samalla tulokset tuovat näkökulman siihen, rajoittuuko ihmisen kuulokyky todellisuudessa yleisesti rajana pidettyyn noin 20 kilohertsiin.</p>	
Asiasanat – Keywords Äänenlaatu näytetaajuus bittisyvyys musiikkiteknologia kuuntelukoe	
Säilytyspaikka – Depository JYX	
Muita tietoja – Additional information –	

SISÄLLYS

1 JOHDANTO	4
2 ÄÄNENLAADUN MERKITYS MUSIIKIN TALLENTAMISESSA	6
2.1 Äänenlaadun kehitys tallentamisen ja toistamisen historiassa	6
2.2 Äänenlaadun musiikintuotannollinen näkökulma	8
3 ÄÄNENLAATUEROJEN KUULEMINEN	11
3.1 Kuulemisprosessi	11
3.2 Kuuloalueen rajoituksista	12
3.3 Kuulojärjestelmän herkkyys	13
4 DIGITAALISEN ÄÄNENTALLENNUKSEN PERUSTEET	14
4.1 Näytteistysperiaate	14
4.2 Bittisyvyys	15
5 TUTKIMUKSEN TOTEUTUS	16
5.1 Tutkimuskysymykset	16
5.2 Tutkimusmenetelmän kuvaus	16
5.3 Tutkimuksessa käytetty laitteisto	20
6 TUTKIMUKSEN TULOKSET	21
7 PÄÄTÄNTÖ	24
LÄHTEET	26

1 JOHDANTO

Valtaosa kuulemistamme tallennetuista äänistä aina musiikillisista teoksista matkapuhelinten soittoaäniin käsitellään vähintään jossain tuotantovaiheessa digitaalisesti. Tämän huomioiden digitaalisen äänenkäsittelyn voidaan sanoa olevan keskeisimpiä musiikkituotannollisia elementtejä nykyaikana. Täten on syytä huomioida myös se, millä tavalla digitaalisen äänenkäsittelyn erilaiset tekniikat vaikuttavat tallennettavan äänen laatuun ja autenttisuustekijöihin.

Tuon tässä kandidaatintutkielmassa esiin näkökulman niihin vaikutuksiin, joita kahden tärkeän mallinnusfaktorin, näytetaajuuden ja bittisyvyyden, varioiminen aiheuttaa subjektiivisessa äänenlaatuerojen tunnistamisessa. Se, miten laadukkaiksi tallennettuja ääniä arvioidaan, liittyy ensisijaisesti kuulijan subjektiivisiin preferenssitekijöihin, mutta myös kuulemisprosessiin sekä kuunteluolosuhteisiin. Tutkimustehtäväni on selvittää missä määrin kuulija erottaa taltioitun äänen hienovaraisimpia nyansseja laadullisesti toisistaan eroavissa tallennuskäytännöissä.

Keskityn työni teoreettisessa osuudessa digitaalisen näytteistysperiaatteen tarkempaan selittämiseen tai muihin teknisiin seikkoihin vain yleisluontoisesti, sillä jo olemassaoleva kirjallisuus tarjoaa varsin hyvän katsauksen näihin teorioihin. Tarkemmin pyrin tuomaan esille aiheeni keskeisimpien tekijöiden, eli näytetaajuuksien ja bittisyvyyden variaatioiden peruspiirteet ja niiden vaikutukset äänentallennusprosessiin, sekä osoittamaan tekijäkohtaisten valintojen suomien mahdollisuuksien hyödyntämisen äänenlaadultaan mahdollisimman hyvien tallenteiden aikaansaamisessa. Lisäksi esittelen lyhyesti ihmisen kuulemisprosessiin liittyviä tekijöitä ja pyrin selvittämään onko yleisesti kuuloalueenamme pidetty 20–20 000 hertsiä se skaala, jonka rajoissa tapahtuvat muutokset pelkästään määrittävät valintojamme, vai vaikuttavatko preferensseihin myös 20 000 hertsin yläpuolella olevat taajuudet.

Työni tarkoitus on osoittaa kliinisissä olosuhteissa järjestetyn kuuntelukokeen avulla miten herkästi ihmisen kuuloaisti erottaa erilaisten tallennustekniikoiden väliset eroavaisuudet. Koemateriaalina toimivat lyömäsoitinnäytteet, joiden keskeinen taajuussisältö ulottuu yli 20 000 hertsin taajuuksille. Tutkimukseni tarjoaa täten näkökulman tallennusmetodien variaatiomahdollisuuksien hyödyntämiseen, joista etenkin musiikkiteknologiaharrastajien ja -ammattilaisten keskuudessa käydään paikoin hyvinkin monipuolista keskustelua ja ajatustenvaihtoa. Osalle alalla toimivista esimerkiksi CD-tasoista äänenlaatua korkeampien tallennusresoluutioiden käyttö tuntuu olevan yhdentekevää kovalevytilan haaskausta joka ei mainittavasti paranna lopullisen tuotteen laatua. Toisille taas pyrkimys autenttisempien äänitteiden luomiseen lähtee perinteisten tallennusresoluution rajojen ylittämisestä. Yhtä kaikki toistaiseksi

mitään merkittävää ja tieteellisesti vedenpitävää näyttöä subjektiivisten äänenlaatuerojen arvottamisesta näytetaajuuksien ja bittisyvyysvariaatioiden perusteella ei kuitenkaan vielä ole.

2 ÄÄNENLAADUN MERKITYS MUSIIKIN TALLENNUKSESSA

Äänentallennus on nykyisin eräs keskeisemmistä audiotekniikan käyttömuodoista tuottaessa ääni- ja/tai kuvataallenteita kulttuuriviestinnän, kuten musiikin, elokuvan, videon, radion ja television puitteissa, sekä teatterissa ja muussa esitystoiminnassa tuottaessa äänitehosteita ja äänitaustoja (Blomberg & Lepoluoto 2005, 23). Jotta tallennustapahtuma ja tallennetut äänet palvelisivat parhaiten kuulijoita ja yleistä käyttötarkoitustaan, on prosessin laadullisiin ominaisuuksiin vaikuttaviin valintoihin kiinnitettävä huomioita jo varhaisimmissa tuotantovaiheissa. Tämä tarkoittaa käytännössä sopivimpien saatavilla olevien musiikkiteknologisten laitteistojen ja sovellusten oikeanlaista käyttöä.

Äänenlaatuun ja sen maksimointiin liittyvät tekijät korostuvat musiikillisten tallennusten laatimisessa. Musiikillista informaatiota sisältävät äänitteet tulevat usein kuunnelluiksi ja toistetuiksi hyvin erilaisissa olosuhteissa hyvin erilaisten äänentoistojärjestelmien kautta, jolloin keskeisimmän musiikillisen sisällön tulisi säilyä mahdollisimman häiriöttömänä ja laadultaan tasapainoisena formaatista riippumatta. On syytä kiinnittää huomiota myös äänien, kuten myös niiden havainnoinnin moniulotteisuuteen. Käytettävissä olevien tallennus- ja toistolaitteiden ominaisuuksien mittaaminen ja äänen fysikaalisten ominaisuuksien tarkastelu ei useinkaan kerro sitä, miten, tai millä tavalla äänen laadullisia tekijöitä arvioidaan (Bech & Zacharov 2006, 1). Täten huomion kohdistaminen nimenomaan ihmisen kuulemiskykyyn on ensiarvoisen tärkeää. Jos kuulemamme signaali toistetaan kuuloaistimme herkkyyttä karkeammalla resoluutiolla, tunnistamme ja koemme eron alkuperäiseen mielikuvaan verrattuna. Jos taas käytössä oleva toistojärjestelmä on kuuloaistiamme herkempi, voi äänenlaatu vaikuttaa *täydelliseltä*, vaikkei se välttämättä sitä olekaan (Watkinson 2001, 2).

2.1 Äänenlaadun kehitys tallentamisen ja toistamisen historiassa

Äänten tallennuksen voidaan katsoa saaneen alkunsa vuonna 1877, jolloin yhdysvaltalainen keksijä ja liikemies Thomas Alva Edison kehitti fonografin ja lausui sen suppiloon sanat ”Mary had a little lamb”. Nämä siirtyivät neulan kautta kojeessa olleelle tinapaperille, jota uudelleenliikuttamalla suppilosta kuului epäselvästi mutta kuitenkin ymmärrettävästi, mitä siihen oli puhuttu. (Gronow & Saunio 1990, 19.) Vastaavia keksintöjä syntyi pitkin 1800-luvun loppua muuallakin Yhdysvalloissa ja Euroopassa, mutta niiden yksityiskohtaisempi hyödyntäminen musiikin tallennuksessa ja

toistossa jäi muiden käyttötarkoitusten varjoon.

Äänilevyn varsinainen aika koitti vuonna 1889, kun saksalainen Kämmerer & Reinhardtin lelutehdas toi markkinoille ensimmäiset viralliset kumista tai selluloidista valmistetut levyt sekä niiden toistamiseen sopivat soittimet. Näiden laitteiden ja levyjen lisenssit omisti tuolloin gramofonia kehittänyt Emil Berliner, joka toi markkinoille myös ensimmäiset 7-tuumaiset äänilevyt. Vaikka äänenlaadussa ei näidenkään laitteiden yhteydessä ollut kehumista, äänilevyteollisuus oli syntynyt. Tästä kertoo Berliner-yhtiön toimintakertomus tilivuodelta 1897–98, jolloin yhtiö valmisti 11 211 gramofonia ja 408 195 äänilevyä. Äänilevyille tallennetun materiaalin kirjo sinänsä oli hyvin monipuolinen ja musiikkihistoriallisesti tärkeä. (Gronow & Saunio, 36–37.)

Seuraavina vuosikymmeninä aina toiseen maailmansotaan saakka markkinoille tuotiin käytettävyydeltään ja toiminnaltaan hyvin erityyppisiä äänen tallennukseen ja toistoon soveltuvia laitteita. Näistä mainittakoon rakenteeltaan magnetoituvan teräslangan käyttöön perustuvat kojeet. Erityisesti Saksassa 1930-luvulle tultaessa magneettilankaperiaatteella toimivien laitteiden suosio oli huomattavaa varsinkin radioalalla. Kun kaupallisesti potentiaalisten tuotteiden kehittäminen monipuolistui, kiinnostivat saksalaiskeksijät ensimmäisten joukossa huomiota myös laitteiden ja niissä käytettävien medioiden äänenlaatuun. Esimerkiksi Kurt Stille, joka oli jo pitkään toiminut alalla, toi 1930 markkinoille ensimmäisen tallennuslaitteen, jonka taajuusvaste pysyi tasaisena aina 250 hertsistä 3 000 hertsiin. Vuonna 1935 taajuusvaste oli jo 70–5 500 hertsiä, joka laatunsa puolesta riitti jo nauhoitetun puheen toistoon AM-verkossa ilman kuultavissa olevaa häiriötä. (Daniel, Mee & Clark 1999, 30–33.)

Ajan kuluessa ja äänien tallentamiseen ja toistamiseen soveltuvien laitteiden yleistyessä levy-yhtiöt valtasivat kaupallista alaa tuomalla erilaisia ratkaisuja kuluttajien saataville. Vuosina 1967–69 Japanin NHK-lähetysverkon insinöörit suunnittelivat ensimmäisen PCM (pulssikoodaus)-modulaatiota hyödyntävän täysin digitaalisen kaksiraitatallentimen, jonka näytetaajuus oli 32 kilohertsiä ja resoluutio 13 bittiä. Samoihin aikoihin Nippon Columbia -yhtiön insinöörit pyrkivät kehittämään analogista LP-äänenlaatua parempia menetelmiä kohdistuen kritiikkinsä nauhatallentimien aiheuttamaan säröytymään. Vuosina 1970–72 valmistui 8-kanavainen äänitysjärjestelmä, jonka näytetaajuus oli jo 47,25 kilohertsiä bittisyvyyden pysyessä kolmessatoista. (Fine 2008, 2–3.)

1970-luvun lopussa digitaaliset äänitysjärjestelmät olivat tavoittaneet jo lukuisia huippustudioita ympäri maailman. Digitaalisesti toteutettujen tallenteiden äänenlaadun katsottiin korvanneen analogiset vastineensa, mutta varsinainen standardi haki vielä paikkaansa. Varsinainen

käänteeseen äänilevyn historiassa tapahtui 1980-luvun alussa, kun Philips- ja Sony-yhtiöiden yhteistyössä kehitetty CD-levy tuli markkinoille suuren suosion saattamana. Yhtiöt olivat kehittäneet ensimmäisenä standardin, joka yhdistäisi sekä tallennusmedian (CD-levy), että toistolaitteiden ominaisuudet valtavaan kaupalliseen potentiaaliin. CD-levyn suosiosta kertovat myyntiluvut; vinyylimuotoisten levyjen ja kasettien ollessa vielä yleisesti markkinoilla CD-levyjä myytiin Yhdysvalloissa julkaisuvuotenaan 1983 800 000 kappaletta ja 80-luvun loppuun mennessä maailmanlaajuisesti yli miljardi kappaletta (Pohlmann 1992, 12).

Historiallinen kehitys Edisonista ja analogisen tallennuksen erilaisista muodoista aina digitaalisen äänen yleistymiseen ja CD-levyyn kesti noin sata vuotta, mutta ajallisesti huomattavasti lyhyempi kehitys CD-levystä nykypäivään pitää sisällään useita tärkeitä vaiheita, jotka vaikuttavat ratkaisevasti musiikin tekemiseen ja kuluttamiseen. Mielestäni on kuitenkin olennaista pohtia nykyaikaisia kehityssuuntia esimerkiksi tiedonvälityksen kautta. Yleistäen voidaan sanoa, että analogisten ja digitaalisten äänentallennuslaitteistojen ja -formaattien kehitystä on ohjannut aina tietty pyrkimys kaupallisen potentiaalin maksimointiin. Kaikkien audioteknologisten keksintöjen pyrkimys on ollut tuottaa uudelleen tallennettu ääni jossain toisaalla. Kuluttajan kannalta mielenkiinto kohdistuu kohtuullisesti hinnoiteltuihin laitteisiin jotka pystyvät enemmän toistamaan tai vastaanottamaan signaaleja, kun taas levy-yhtiöiden ja median pyrkimyksenä on muovata signaaleista jollain tasolla vetoavia erilaisten kaupallisten ohjelmien yhteydessä (Watkinson 2001, 1). Täten teknologioiden joustavuus ja toiminnallinen nopeus usein syrjäyttää muita arvoja, joista äänenlaatu on yksi hinnan ohella. Ja koska CD-levyn iltapuoli on tiettyssä mielessä jo käsillä (Fine, 2) ja häkellyttävän suuri osa musiikista kuunnellaan ja hankitaan internetin avulla (vuodesta 1999), levymyynti Yhdysvalloissa on laskenut 47 % 14,6 miljardista 7,7 miljardiin dollariin, ja vuosien 2004 ja 2009 välillä noin 30 miljardia musiikkikappaletta ladattiin laittomasti (RIAA 2012) jää äänenlaadun nousujohteisempi kehitys näillä näkymin toisarvoiseen asemaan.

2.2 Äänenlaadun musiikintuotannollinen näkökulma

Musiikintuotannon kannalta pyrkimys mahdollisimman hyvään äänenlaatuun on ollut hyvin vahva tekijä. Digitaalisen äänenkäsittelyn yleistyttyä studioissa maailmanlaajuisesti erilaiset tuotekehittelyt ja tuotantotavat ovat alkaneet edesauttaa äänenlaadultaan entistä parempien tallennusten aikaansaantia, joihin liittyy paitsi analoginen, lämmin ja akustinen äänimaailma myös virtuaalisten tilojen ja synteettisten äänien käyttö. Red Book -standardin myötä digitaalisen

äänenkäsittelyn perusta on yleensä mielletty korreloivaksi tallennusmedian eli CD-levyn kanssa. Täten näytetaajuutena on käytetty 44,1 kilohertsiä ja resoluutiona kuudentoista bitin syvyyttä. Näiden arvojen ominaisuuksista kerron tarkemmin luvussa 4.

Ajatus 44,1 kHz / 16-bit -resoluutiosta perustuu Philipsin ja Sonyn neuvotteluihin sopivimpien arvojen löytämiseen. Alun perin Philipsin toive näytetaajuudeksi ja bittiresoluutioksi oli 44,050 kilohertsiä / 14 bittiä, kun taas Sonyn edustajat ehdottivat 50 000 kilohertsiä / 16 bittiä. Pitkien kriittisten kuuntelukokeiden jälkeen standardiksi valittiin 44,1 kilohertsiä ja 16 bittiä (Pohlmann 1992, 11). Korkeammista tallennustarkkuuksista keskusteltiin kuitenkin heti 80-luvun alussa, sillä monet audiofiilit kokivat CD-levyn äänenlaadun hirveäksi (Downes 2010, 305). Mahdollista on, että juuri markkinoille tuodun formaatin ominaisuuksia tuomittiin liian nopeasti, sillä keskeisimpänä vertailukohtena toimi analoginen LP-toisto, jossa taajuusvaste ja toistotarkkuus saattoi yltää tavanomaisellakin hifi-äänirasiolla kymmenestä hertsistä 25 kilohertsiin (Hansen 2002, 358; Roads 1996, 31).

Äänenlaadun ja erilaisten tallennusresoluutioiden merkityksen erittäin subjektiivisesta luonteesta kertoo alan arvostettujen ammattilaisten Rupert Neven ja Dan Lavryn näkemysten välille muodostuva kahtiajako. Neve on tuonut esiin korkeiden tallennusresoluutioiden merkityksen suunnittelemissaan musiikkiteknologisissa laitteistoissa ja kertonut mm. 1998 julkaistussa AudioTechnology Magazinen haastattelussa, että vasta noin 150 kilohertsin tarkkuus pystyy tallentamaan ”taiteesta” kaiken riittävän, pohjaten lausuntonsa useiden ihmisten kuulohavaintoihin (Simmons 1998). Lavryn kanta asiaan on päinvastainen; tutkimuksensa *Sampling Theory for Digital Audio* (2004) mukaan kaikki argumentit ratkaisevasti yli ihmisen kuulokynnyksen ylittyvistä näytetaajuuksista osoittavat vajaata ymmärrystä näytteistysperiaatteesta. Lisäksi Lavry tuo esiin näkökulmia, joiden mukaan korkeampien resoluutioiden käyttö rasittaa laitteistoa ja aiheuttaa signaalivirheitä (Lavry 2004, 1–2).

Väitteitä, mielipiteitä ja näkökulmia erilaisten korkealaatuisten näytetaajuuksien ja korkeamman bittisyvyyden käytöstä sisältyy myös musiikkiteknologia-alan peruskirjallisuuteen. Yhteistä näille näkökulmille on niiden tietynlainen tulkinnanvaraisuus ja kirjoittajan/tekijän omat tottumukset. Esimerkiksi Bobby Owsinski tiivistää vertailun CD-laadun ja 96:n kilohertsin näytetaajuuden välillä yksinkertaisesti muotoon ”korkeampi näytetaajuus, parempi äänen representaatio ja laajempi käyttökelpoisen äänen kaistanleveys” (Owsinski 1999, 87). Vastaavasti John Watkinsonin esittää teoksessaan *The Art of Digital Audio*, että kaikki 96:n tai 192 kilohertsin näytetaajuuksilla tehdyt kokeet sisältävät väärinymmärryksiä, ja ettei mainittujen näytetaajuuksien käyttämisellä saavuteta minkäänlaista kuultavissa olevaa hyötyä (Watkinson 2001, 729–730).

Neutraalimman näkökulman asiaan tuo Bob Katz teoksessaan *Mastering Audio* ehdottaen, että tutkimusten perusteella toistaiseksi ihmiskorvalle parhaimman ja vähiten transientti- ja terävyyshäviötä aiheuttavia tai äänen selkeyttä heikentäviä faktoreita sisältävä näytetaajuus on noin 50 kilohertsiä, jolloin CD-äänenlaatu jää tämän alapuolelle 44.1 kilohertsin tarkkuudella ja 48 kilohertsin näytetaajuus on melkein riittävä (Katz 2007, 252–253).

Yhteistä äänenlaatua koskevissa vertailussa on ensiksi niiden perustellusti subjektiiviseen havaintoihin liittyvä luonne. Toiseksi havaintojen mittausongelmat ja niiden vaikutus yhteisymmärryksen löytämiseen. Kolmanneksi ongelmakohdat laitteistojen todellisen tallennuksen ja toiston kontrolloinnissa. Aihetta on kuitenkin perusteltua tarkastella käynnissä olevan formaattimuutoksen vuoksi, sillä tallennusstandardina pidetty Red Book CD-standardi alkaa CD-levyformaatin myötä käydä tarpeettomaksi, koska äänitetty musiikki voidaan tallentaa, säilöä ja toistaa joko pakatussa ja hävikillisessä muodossa tai CD-resoluutiota korkeammalla näytetaajuudella ja bittisyvyydellä riippuen kuluttajan henkilökohtaisista preferenssitekijöistä. Äänenlaatu tulee kuitenkin säilyttämään paikkansa yhtenä musiikin kulutuskäyttäytymistä ohjaavana faktorina. Tämä on äärimmäisen tärkeää sen vuoksi, että tallennustilanteen äänenlaadun asettaminen tiettyyn muotoon vaikuttaa erittäin keskeisesti siihen, miltä musiikki kuulostaa eri kuunteluolosuhteissa. Mikäli näytetaajuuden ja bittisyvyyden arvot eivät tue alkuperäisen tallennustilanteen äänimaailmaa, ei mielestäni ole perusteltua maksaa riittämättömällä laadulla tehdystä työstä tai tuotteesta. Vastaavasti teknisten äänenlaatuarvojen ja tallennusmetodien korostaminen asiayhteyksissä, joissa niistä ei ole lopputuloksen kannalta selkeää hyötyä (kuten äänilevyjen remasterointikulttuuri, joissa raidat käsitellään vain päivitettyjen laitteiden kautta tarjoten lopputuloksena vain kasvanutta äänenvoimakkuutta) ei myöskään tulisi olla liian keskeisessä arvossa musiikin markkinoinnissa.

3 ÄÄNENLAATUEROJEN KUULEMINEN

Äänenlaadun subjektiivisen havaitsemisen kannalta ensimmäisen osatekijän muodostavat kuuloaistin kautta hahmottuvat havainnot. Esittelen tässä luvussa tutkimukseni kannalta vain olennaisimpia elementtejä kuulemisprosessissa, sillä ihmisen kuuloaistia, sen kautta muodostuvia havaintoja sekä äänen yleistä olemusta on hankalaa esitellä yksiselitteisesti vain yhden tieteenalan kautta. Musiikkiteknologian ohella tärkeän näkökulman äänihavaintojen käsittelyyn tässä yhteydessä tarjoaa psykoakustiikka. Psykoakustiikassa tutkitaan, miten havaitsemme ääniä ja erityisesti miten erilaiset fysikaaliset muutokset äänessä muuttavat ihmisten kuuloaikutelmaa (Järveläinen 2010, 33).

3.1 Kuulemisprosessi

Kuuloaistin kannalta välttämättömiä ärsykeitä ovat äänet. Ääni on väliaineessa etenevää paineenvaihtelua. Se etenee ali- ja yli- paineisina vyöhykkeinä ja on siis pitkittäissuuntaista aaltoliikettä – värähtelyä (Blomberg & Lepoluoto 2005, 27). Ihmisen kuulojärjestelmän saavuttaessaan paineenvaihtelut muuntuvat ensin mekaaniseksi värähtelyksi ulko-, väli- ja sisäkorvassa, ja sitten sähköisiksi impulsseiksi sisäkorvassa. Nämä sähköiset impulssit kulkevat hermorojoja pitkin aivoihin. Sekä matkalla aivoihin että perillä aivoissa ääni-informaatiota käsitellään, ja lopulta informaatio voidaan aistia äänenä (Huotilainen 2010, 47). Vaikka kuuloelimiemme perustoiminnan kuvaaminen on mahdollista näin lyhyesti, on kuitenkin syytä huomioida prosessin muuttuminen verrattain monimutkaiseksi siinä vaiheessa, kun ääni-informaation käsittely muuttuu kuulluista havainnoista subjektiiviseksi kokemusmaailman tuotteiksi. Kuvaamassani vaiheessa kuulemisen fysiologinen puoli väistyy ja käsittely muuttuu enemmän psykologiseksi tavaksi hahmottaa maailmaa. Esimerkiksi tarkempi taajuussisällön analyysi ja eteenpäin välittäminen sekä äänen tarkka ajallinen analyysi tapahtuvat kuuloaistin simpukkatumakkeen osassa. Tämän prosessin yhteydessä hahmottuvat myös äänen tulosuunta sekä äänen alkamis- ja päättymisominaisuudet (Huotilainen 2010, 53–55). Tarkemmin nimenomaan äänenlaatuerojen havainnointiin keskittyvistä kuulemisprosessin fysiologisista osatekijöistä keskeisimmäksi nousevat kuuloalueen rajoitukset ja kuulojärjestelmän herkkyyys.

3.2 Kuuloalueen rajoituksista

Korkealaatuisten äänen tallennus- ja toistokäytäntöjen yhteydessä nostetaan usein esiin ihmisen kuuloalueen rajoitukset. Taajuusalueensa puolesta ihmisen kuulon katsotaan yleisesti vastaanottavan ääniä 20–20 000 hertsin kaistalta (Järveläinen 2010, 34; Blomberg & Lepoluoto 2005, 27). Musiikkiteknologia-alan kirjallisuudessa yläraja on toisinaan jopa matalampi, kuten 15 000 hertsiä (Katz 2007, 249) tai jotain 16 000:n ja 20 000 hertsin välillä (Watkinson 2001, 27).

Oletus 20000 hertsin ylärajasta on määrittänyt musiikkiteknologisten laitteiden kehitystä ja toimintaa jo vuosikymmeniä. Koska ihmisen ei oleteta havaitsevan yli 20000 hertsin taajuuksia, esimerkiksi monien mikrofoniin taajuuskaista rajoittuu myös kyseisen lukeman alapuolelle (Lavry 2004, 1). Toisaalta lukuisat tuntemistamme instrumenteista tuottavat runsaasti energiaa myös korkeilla, yli 20 kilohertsin taajuusalueilla, kuten Boyk tutkimuksessaan (1997) esittää.

Japanilaisen Tsutomu Oohashin (2000) työryhmineen suorittaman tutkimuksen mukaan 20 000 hertsin yläraja on kuitenkin syytä kyseenalaistaa. Tutkimuksesta selviää, että ihmisäivot reagoivat eri tavalla ääninäytteisiin, joihin sisältyy tutkimuksessa käytettyjä ns. korkeataajuuskomponentteja. Korkeataajuuskomponenttien katsotaan vaikuttavan kuuntelutilanteeseen tallennuksen autenttisuutta lisäävällä tavalla; onhan akustisessa ääniympäristössämme todellisuudessa hyvinkin korkeita ääniä, esimerkiksi sademetsäolosuhteissa jopa yli 100 000 hertsiin saakka. Tutkimukseen liittyneessä koeasetelmassa koehenkilöt myös arvioivat näytteiden subjektiivisia äänenlaatueroja. Tässä kokeessa korkeataajuuskomponentteja sisältäneet näytteet erottuivat tilastollisesti merkitsevästi niistä näytteistä, joissa kyseisiä komponentteja ei ollut. (Oohashi et al. 2000, 3548–3553.)

Vaikka Oohashin ja hänen työryhmänsä kokeet tuovatkin uuden näkökulman korkeita taajuuksia sisältävien ääninäytteiden äänenlaatuerojen arviointiin, tutkimusta ei sinällään voi vielä arvioida tieteellisesti käänteentekevänä lähinnä siitä syystä, ettei siinä käytettyjä metodeja ja asetelmia ole tietävästi uusittu alkuperäisen kokeen jälkeen. Tämä kertoo osaltaan siitä, ettei tarvittavaa laitteistoa ja tietotaitoa välttämättä löydy muualta, mutta myös siitä, ettei kyseisten taajuusalueiden sisällyttäminen kaupallisiin äänentallennus- ja toistolaitteisiin tai tallennusmedioihin ole tarpeellista tai kannattavaa.

3.3 Kuulojärjestelmän herkkyys

Äänenlaatuerojen arvioinnissa tässä yhteydessä kuuloalueen taajuusrajoihin liittyvän näytetaajuuden ja sen variaatioiden lisäksi toinen keskeinen osatekijä on bittisyvyys ja sen variaatiot. Bittisyvyyden teoriasta kerron tarkemmin luvussa 4.2, keskittyen tässä kuulojärjestelmän herkkyuden periaatteiden lyhyehköön esittelyyn.

Korva kestää vaurioitumatta noin 140 desibelin hetkellisen huippuäänipainetason ja kuulee herkimällä taajuusalueella jopa alle 0 desibelin tasoisia ääniä, kun taas hyvin matalilla ja korkeilla taajuuksilla vasta kymmenien desibelien ääni tulee kuuluviin (Järveläinen 2010, 34). Tarkemmin kuulojärjestelmän toiminnasta kertoo ns. Fletcher-Munsonin kuulokäyrä.

Fletcher-Munsonin kuulokäyrän perusta on jaettavissa käytännössä viiteen kohtaan: 1. Ihmiskorva ei aisti kuuloalueen kaikki taajuuksia tasaisen herkästi. 2. Hetkellisesti kaikista herkimmin kuultavan taajuuden suhde muihin taajuuksiin korreloi äänenvoimakkuuden muutosten kanssa. 3. Korvan herkkyuden erilaisuudet korostuvat matalilla äänenvoimakkuuden tasoilla. 4. Korvan herkkyuden erilaisuudet korostuvat vähiten korkeilla äänenvoimakkuuden tasoilla. 5. Korva on yleisesti ottaen kaikista herkin keskiäänitaajuuksille ja vähiten herkkä matalille taajuuksille. Herkkyys korkeille taajuuksille sijoittuu matalien ja keskiäänien herkkyystaajuuksien välille. (Gottlieb 2007, 42.)

Esimerkkinä kuulojärjestelmän herkkyuden maksimaalisesta huomioimisesta ja hyödyntämisestä äänityöskentelyssä ja musiikin tuotantoprosesseissa voidaan mainita lähestulkoon välttämättömäksi toimitukseksi muodostunut taajuusvasteen korjaus (ekvalisoiminen), joka perustuu erilaisten taajuuskomponenttien äänikuvaan asetteluun monesti juuri ihmiskorvalle herkintä taajuusaluetta silmälläpitäen. Kuulomme pystyy havaitsemaan jopa vain 0,1–0,2 % muutoksen värähtelytaajuudessa herkimällä 1–2 kilohertsin taajuusalueella (Järveläinen 2010, 37). Täten esimerkiksi mainitulla herkkyyalueella esiintyvä epäviireisyys tai erilaiset häiriötekijät voivat vaikuttaa negatiivisesti soittosuorituksen äänenlaadun subjektiiviseen arviointiin. Musiikillisen informaation välityksessä myös nk. peittoilmiön vaikutus on huomattava. Peittoilmiössä ääni häipyä kokonaan tai osittain kuulumattomiin jonkin toisen äänen, kuten kohinan vaikutuksen vuoksi (Järveläinen 2010, 34).

4 DIGITAALISEN ÄÄNENTALLENNUKSEN PERUSTEET

Digitaalisen äänien tallennuksen ydinkäsite, näytteistäminen (*sampling*) tarkoittaa käytännössä jatkuvien analogisten signaalien (esimerkiksi mikrofonista tulevien) muuntamista erillisiksi aikänäyte-signaaleiksi (Roads 1996, 9). Tällaiset näytteet muodostavat tallennuttuaan äärimmäisen lyhyitä impulsseja. Tietokonejärjestelmille ja tietojenkäsittelyyn pohjautuville laitteistolle tämä sopii hyvin, sillä kyseiset kojeet perustuvat binääriseen järjestelmään missä kaikella informaatiolla voi olla vain arvo ”kyllä” tai ”ei”, ”päälle” tai ”pois päältä” (Davies 1996, 3). Pitkästä impulssijonosta muodostuva signaali muuttuu ääneksi vasta, kun se muunnetaan digitaalisesta muodosta takaisin analogiseksi ja johdetaan äänentoistojärjestelmään. Seuraavassa esittelen tarkemmin näytteistysperiaateja ja toiminnallisia periaatteita, joista keskeisimpiä ovat näytteistysperiaate ja näytetaajuudet sekä bittisyvyys.

4.1 Näytteistysperiaate

Keskeisin näytteistysperiaate pohjaa vuonna ruotsalaisen Harry Nyquistin vuonna 1928 kehittämään informaatioteknologiseen teoreemaan. Nyquist kehitti teoriansa työskennellessään Bellin laboratorioissa ja saavutti sen myötä suurta kunnioitusta mm. informaatioteorian isältä Claude Shannonilta (Lavry 2004, 1). Shannon todisti teoreeman paikkansapitävyyden vuonna 1949 (Gallagher 2008, 141).

Nyquistin (myös Nyquist-Shannonin) teoreema perustuu yksinkertaiseen sääntöön: signaali voidaan tuottaa uudelleen täydellisesti vain, jos tallennusprosessin näytetaajuus on vähintään kaksinkertainen korkeimpaan näytteistettävän signaalin taajuuteen verrattuna. Täten esimerkiksi 10 kilohertsin signaalin täydelliseen tallentamiseen ja uudelleentoistamiseen käytetyn näytetaajuuden tulee olla vähintään 20 kilohertsiä. (Gallagher 2008, 141.) Tätä periaatetta noudattaen Red Book -standardin mukainen CD-äänenlaatu, jonka näytetaajuus on 44,1 kilohertsiä, voi ongelmitta näytteistää maksimissaan 22,05 kilohertsin signaalin. Tätä 22,05 kilohertsin signaalia voidaan kutsua Red Book -näytetaajuudella operoitaessa ns. Nyquistin taajuudeksi (Gallagher 2008, 140), sillä se on korkein mahdollinen näytteistyskelpoinen signaali, ja siitä voidaan 44,1 kilohertsin näytetaajuudella saada kaksi näytettä; positiivinen ja negatiivinen huippuarvo. Mikäli signaali ylittää Nyquist-teoreeman säännön ja on edellisen esimerkitapauksen olosuhteissa yli 22,05 kilohertsiä, muodostuu signaaliin yhteyteen ilmiö, jota kutsutaan laskostumiseksi. Teoriassa

esimerkiksi voimakas 28 kilohertsin signaali voi 44,1 kilohertsin näytetaajuudella tallennettaessa aiheuttaa huomattavaa tunnistettavaa säröä 16,1 kilohertsin taajuudella.

Käytännössä laskostumisen estävä laskostumisenestosuodatin (*anti-alias filter*) tai jokin vastaava teknologia kuitenkin suodattaa lähes kaikissa digitaali–analogi-muuntimissa näytetaajuuden asettaman nyquist-taajuuden yli menevien signaalien energian. Samalla tämä tarkoittaa sitä, että sellaisten instrumenttien, joiden energiasta keskeinen taajuussisältö on esimerkiksi yli 20 kilohertsin taajuuksilla, puolikkaan näytetaajuuden ylimenevät taajuusosat suodattuvat pois.

4.2 Bittisyvyys

Näytetaajuuden ohella tallennusprosessin toinen merkittävä osatekijä on bittisyvyys (*bit depth*). Bitti tässä yhteydessä on lyhenne termistä binäärinen luku (Roads 1996, 22). Siinä missä näytteistysperiaate keskittyy näytekohtaisesti tallennettavien äänten taajuussisältöihin, toimii bittisyvyys äänen dynaamisten ominaisuuksien määrittäjänä. Käytännössä, menemättä yksityiskohtaisempiin tietoteknologisiin selitystapoihin, bittisyvyys tai -resoluutio tarkoittaa käytettävissä olevien bittien lukumäärää äänen dynaamisen skaalan tallennuksessa, jolloin yksi bitti vastaa noin kuuden desibelin skaalan tallennusmahdollisuutta (Gallagher 2008, 21). Täten esimerkiksi Red Book -standardissa toteutuva kuudentoista bitin syvyys mahdollistaa äänien dynaamisten ominaisuuksien tallentamisen 0–96 desibelin välillä (Roads 1996, 40). Samalla periaatteella esimerkiksi DVD-audion mahdollistama 24-bittinen tallennusresoluutio kasvattaa dynaamista asteikkoa 144 desibeliin saakka.

Mikäli Red Book -standardin mukaista CD-äänenlaatua korkeamman bittisyvyyden hyödyntäminen on mahdollista, on sen käytössä kuitenkin huomioitava myös tallennusympäristö ja sen vaikutus lopputulokseen. Russ Hainesin teoksessaan *Digital Audio* (2001) esittelemät kriteerit ansaitsevat huomion: Äänitystilän ambienssin on oltava mahdollisimman lähellä nollaa, mikä myös on ihmisen kuuloherkkyyden raja, ja tallennettavien äänien dynamiikka-alueen tulee kattaa olennaisesti määritetyn bittisyvyyden ala- ja yläraja-arvojen alueen (Haines 2001, 275). Käytännössä tarve esiintyy esimerkiksi ihmisäänen tallennuksen kohdalla, mikäli kaikki signaalit hiljaisimmasta kuiskauksesta kovimpaan huutoon tulisi saada tallennettua mahdollisimman puhtaasti (Haines 2001, 276).

5 TUTKIMUKSEN TOTEUTUS

Koska näkökulmia ja perusteita äänenlaadultaan erilaisten taltiointimenetelmien käytöstä on esiintynyt käytännössä 1800-luvun lopulta saakka ja yleisellä tasolla äänenlaadun paremmuuden teoreettiset lähtökohdat esiintyvät enenevässä määrin erilaisten musiikkiteknologisten laitteistojen, työskentelytapojen ja tiedonvälityksen yhteydessä, on mielestäni syytä selvittää, missä kohtaa kuulijan subjektiivinen raja-arvo äänenlaatuerojen tunnistamisessa tulee vastaan.

Tämän työn yhteydessä toteuttamani kuuntelukokeen perusta on yhtäällä kaikkien parametrien kontrolloimisessa optimaalisten tuloksien saavuttamiseksi, mutta myös kuulijan henkilökohtaisten mieltymystekijöiden huomioonottaminen. Empiiristen havaintojeni perusteella merkittävä osa tutkimusaiheeni tapaisista epävirallisista äänenlaadua koskevista kuuntelukoetyylisistä arviointiprosesseista tapahtuu yleensä jokseenkin epäselvissä olosuhteissa testaamattomilla laitteilla ilman jokaisen osatekijän ominaisuuksien lähempää tarkastelua. Lisäksi koetilanteista harvemmin jää mitään systemaattisesti käsiteltyä dataa tarkasteltavaksi ja keskusteltavaksi.

5.1 Tutkimuskysymykset

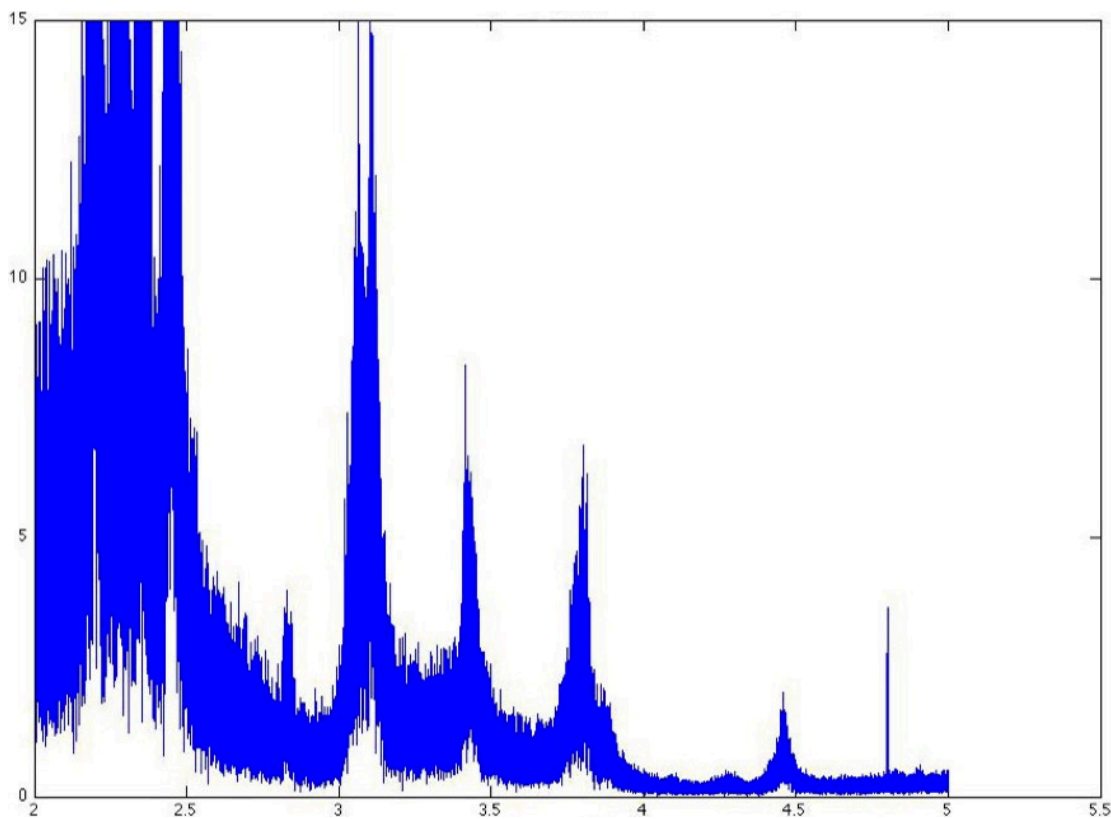
Tehtävänäni on selvittää kuuntelukokeen avulla missä määrin koehenkilöt kuulevat eroavaisuuksia kaksiosaisien näytteiden keskenään äänenlaadullisesti eroavien osien välillä. Kokeesta saaduilla tuloksilla voidaan siten antaa näkökulma esimerkiksi siihen, miten paljon ihmisen oletetun kuuloalueen ylärajan ylittävät (yli 20 kilohertsiset) taajuussisällöt vaikuttavat eroavaisuuksien tunnistamiseen suhteessa kuuloalueen rajoissa pysyviin taajuussisältöihin (alle 20 kilohertsiä). Lisäksi tarkoituksena on esittää näyttöä siitä, rajoittuuko eroavaisuuksien tunnistaminen nimenomaan esimerkiksi CD-resoluutiona pidetyn 44,1 kilohertsin rajaan, vai pystyvätkö koehenkilöt erottamaan edellä mainitun tallennustarkkuuden äänenlaadun ylittäviä näytteitä.

5.2 Tutkimusmenetelmän kuvaus

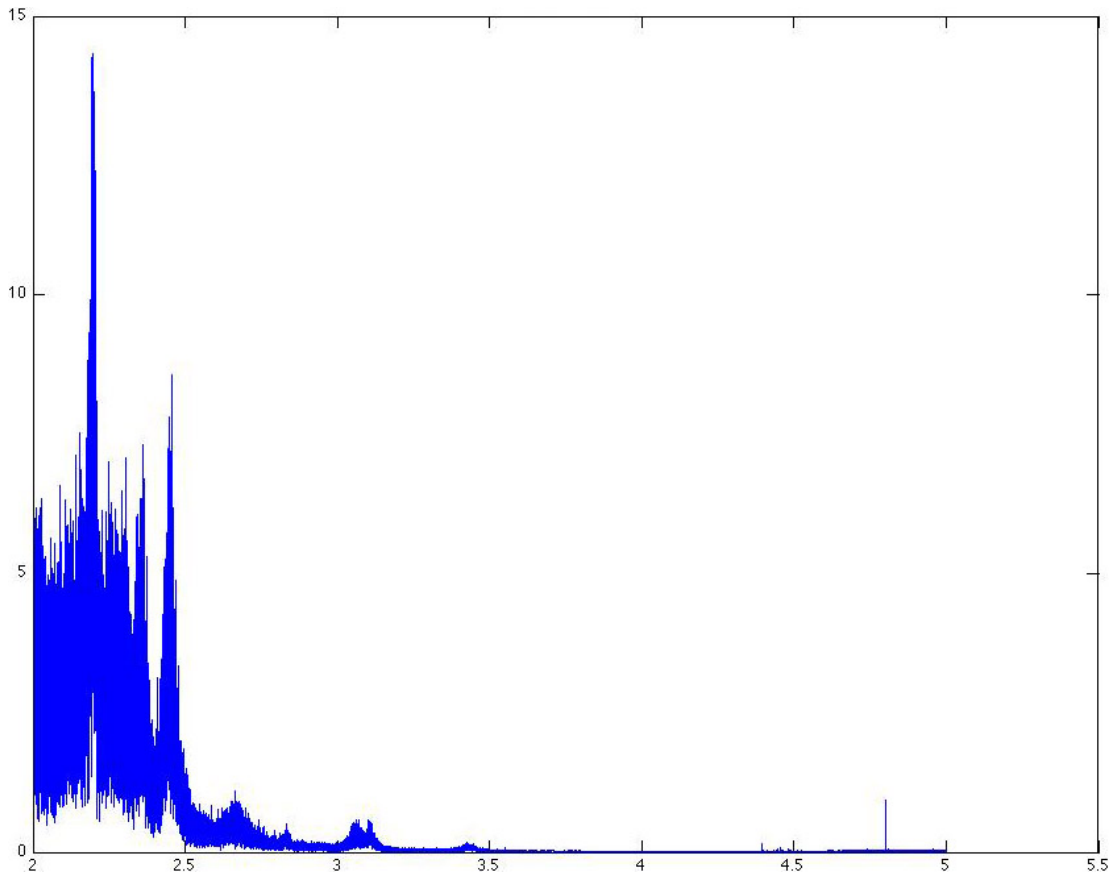
Koemateriaali muodostuu kuudesta erilaisesta metalliseoslyömäsoitinnäytteestä, jotka ovat; kilkka,

tingsha-symbaalit, rin gong -kulho, kellosymbaali, crash-symbaali ja tamburiini. Yhteistä näille näytteille on niiden äänienergian ulottuminen yli 20 kilohertsin. Kokeessa käytetyt näytteet taltioitiin Pro Tools 9 HD -ohjelmiston avulla (tarkempi laitteistokuvaus luvussa 5.3). Myös kuuntelukoe toteutettiin standardeista hieman poiketen Pro Toolsilla. Syynä tähän oli varmistua siitä, että oletetut erot säilyvät kuultavissa laitteistojen ja ohjelmiston toimiessa varmuudella annetulla tallennus- ja toistotarkkuudella, joka tässä tapauksessa oli maksimissaan 192 kilohertsin näytetaajuus sekä 24 bitin syvyys.

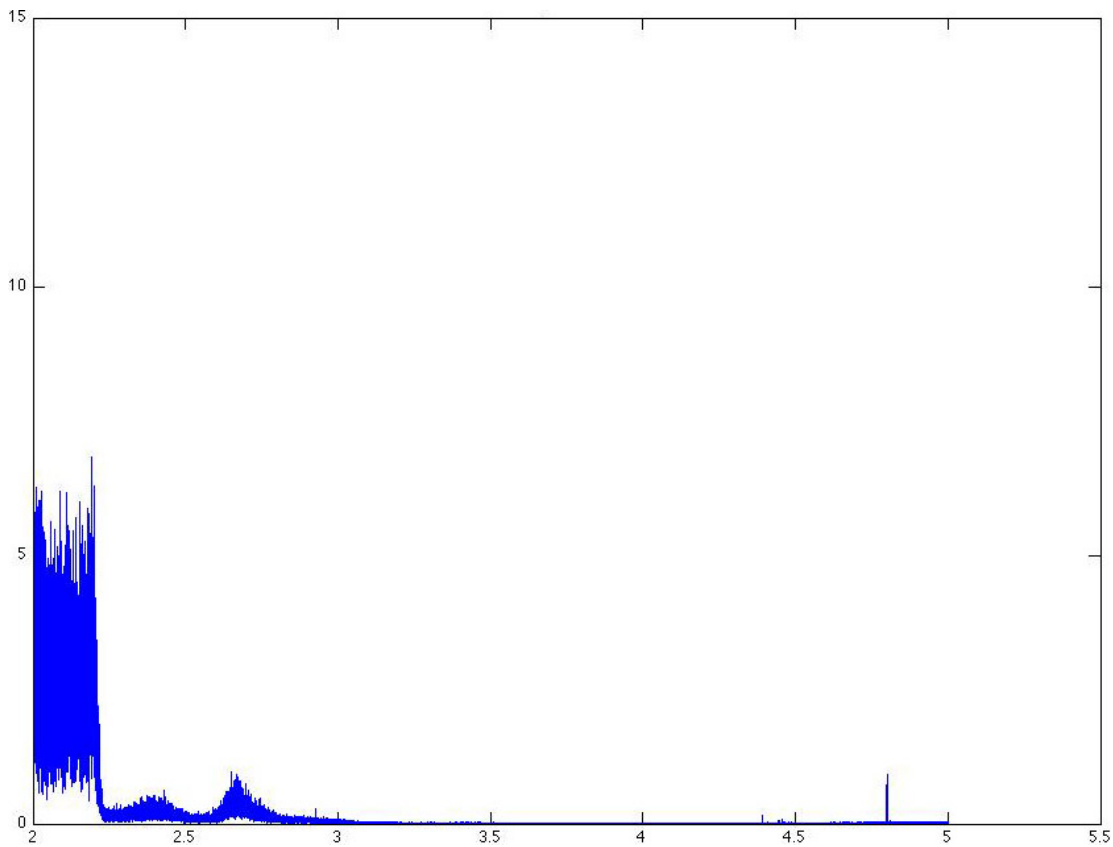
Todisteeksi äänenlaatuerojen todellisesta esiintymisestä näyteäänissä tallennus- ja toistojärjestelmä mitattiin ja tulokset analysoitiin MirToolboxilla. Ohessa tamburiininäytteen taajuussisältöä kuvaavat havainnekuvat (kuvat 1, 2 ja 3, joiden Y-akselilla suhteellinen äänenvoimakkuus, X-akselilla taajuus 20–55 kilohertsiä):



Kuva 1. Tamburiininäytteen (192 kHz / 24-bit) datan taajuussisältö



Kuva 2. Kaiuttimesta toistetun tamburiininäytteen (192 kHz / 24-bit) datan taajuussisältö



Kuva 3. Kaiuttimesta toistetun tamburiininäytteen (44,1 kHz / 16-bit) datan taajuussisältö

Kuvissa 1. ja 2. näkyvät 192 kilohertsin näytetaajuudella 24 bitin syvyydellä näytteistetyt tamburiininäytteen taajuussisällöt. Ylempi kuva osoittaa näytteen taajuussisällön ulottuvan merkitsevästi 45 kilohertsiin saakka, mutta kaiuttimesta toistettuna (ja siitä samalla laitteistolla uudelleentaltioituna) korkeimmat taajuudet häviävät 25 kilohertsin jälkeen. Alempi kuva osoittaa kuitenkin selvästi, että korkeammalla näytetaajuudella näytteistetyt signaalien cd-resoluutiota korkeammat taajuudet tallentuvat ja toistuvat tässä koeasetelmassa. Vertailun vuoksi kuvassa 3 44,1 kilohertsin ja 16 bitin syvyydellä näytteistetyt tamburiininäytteen kaiuttimesta uudelleentaltioitujen signaalien taajuussisältökuvaaja ja ilmiselvä taajuuksien poisleikkautuminen noin 22 kilohertsin kohdalla.

Oheiset kuvaesimerkit osoittavat selvästi, että korkeampien näytetaajuuksien käyttö vaikuttaa äänen korkeataajuisen osien tallentumiseen. Huomattavaa kuitenkin on, että melkoinen osa datan sisältämästä taajuussisällöstä ei kuitenkaan koeolosuhteissa toistunut kaiuttimista. Kuvan 1 sisältö on alkuperäinen soitinnäyte ja kuvat 2 ja 3 ovat kaiuttimista samoilla mikrofoneilla, etuasteilla ja tallentimilla tallennetut mittauskuvat.

Kuuntelukoe

Kuuntelukoe koostui 27 näytteestä, joista jokaisessa verrattiin toisiinsa kahta saman- tai erilaatuista ääntä näytetaajuuksilla 22,05–192 kilohertsiä ja bittisyvyyksillä 16 tai 24. Peruste heikkolaatuisempien näytetaajuuksien (<44,1 kHz) käytölle oli kuuntelukoetilanteen mielekkyyden ylläpitämisellä. Oletettavasti pelkkien korkealaatuisten (>44,1 kHz) näytteiden vertailu toisiinsa olisi voinut synnyttää keskihajonnaltaan käyttökelvottomia tuloksia. Lisäksi teoriassa kuultavissa olevien eroavaisuuksien mittaaminen toimii perusteena kokeen koherenssille. Koehenkilöt arvioivat näytteiden äänenlaatueroja arvoilla 0, 1, 2 tai 3, 0:n kuvatessa tilaa jossa eroa ei koeta ja 3:n kuvatessa tilaa jossa ero oli selvä.

Tutkimukseen osallistuneet koehenkilöt (N=12) olivat kaikki musiikkiteknologisessa mielessä vähintään valveutuneita harrastajia, mutta osittain myös sen parissa ammatikseen työskenteleviä. Peruste koehenkilöiden tietotaidon ja asiantuntijuuden huomioimiseen valintakriteerinä oli saada luotettavia tuloksia yksilöiltä, jotka ovat mahdollisesti arvioineet korkeamman äänenlaadun vaikutuksia työskennellessään äänten taltioinnin parissa. Mainittakoon kuitenkin, että esimerkiksi mahdollisen jatkotutkimuksen kannalta erilaisista koehenkilöistä

muodostuvan kontrolliryhmän sisällyttäminen tutkimukseen olisi suotavaa.

Kokeeseen tullessaan koehenkilöiden vireystilaa tiedusteltiin suullisesti, jonka jälkeen kokeen rakenne kerrottiin lyhyesti samalla, kun vastauslomake ja koeasetelmanäkymä Pro Toolsissa käytiin läpi väärinkäsitysten ja virheiden välttämiseksi. Koehenkilöiden oli mahdollista säätää optimaalinen kuunteluäänenvoimakkuus itse lähtötasosta 70 dB. Koehenkilöt suorittivat 27-vaiheisen kokeen yksin studiotarkkaamossa omaan vapaaseen tahtiin vapaavalintaisessa järjestyksessä. Kokeen kesto vaihteli koehenkilöstä riippuen 17–47 minuutin välillä. Kokeen päätteeksi jokaisen koehenkilön kanssa käytiin lyhyt keskustelu.

5.3 Tutkimuksessa käytetty laitteisto

Instrumentit: kilkka, tingsha-symbaalit, rin gong, kellosymbaali, crash-symbaali ja tamburiini.

Mikrofonit: 2x DPA-4006A, joissa UA0777 Nose Cone -pallokuviokapselit. (Mikrofonien taajuusvaste valmistajan ilmoituksen mukaan on 10–20000 hertsiä, mutta tämän tutkimuksen kannalta tärkeisiin yli 20 kilohertsin taajuuksiin mikrofonit reagoivat luotettavasti vielä yli 25 kilohertsissä).

Etuaste: GML 8304 (taajuusvaste 1,7–260000 hertsiä).

Ensisijainen tallennuslaite: Korg MR-2000S -DSD-tallennin tarkkuudella 5,6 Mhz.

Lopullinen tallennus- ja toistolaitte: Digidesign Pro Tools|HD 2 Accel / 192 I|O audio Interface

Kaiuttimet: Adam S3A (taajuusvaste 32–35000 hertsiä.)

6 TUTKIMUKSEN TULOKSET

Tutkimukseni on osoittanut tähän mennessä todennetuksi ainakin sen, että erilaisten näytetaajuuksien ja bittisyvyyden käyttö vaikuttaa digitaalisen äänen ominaisuuksiin hyvin paljon sen ”varastoinnissa”, mutta todistetusti myös sen toistoon koelolosuhteissa. Kuten seuraavalla sivulla olevasta kaaviosta 1 voidaan huomata, koehenkilöiden mukaan jokaisessa näyteparissa oli ainakin jossain määrin kuultavia eroja. Taulukkoon on listattu näyteparin järjestysnumero, näyteääni, näyteparin ensimmäisen osan tallennustarkkuus, näyteparin toisen osan tallennustarkkuus, havaittujen erojen keskiarvo (asteikolla 0-3) sekä havaittujen erojen keskihajonta (asteikolla 0-3).

Merkitsevimpiä eroavaisuuksia taulukon mukaan on oletetusti sellaisten näyteparien kohdalla, joiden verrattavana oleva näytetaajuus on heikkolaatuisempi (< 44100 hertsiä) ja joiden taajuussisältöjen vajavaisuus verrattuna korkealaatuisempiin näytteisiin (> 44100 hertsiä) on teoriassa kuultavissa eroavaisuuksien keskittyessä ratkaisevasti alle 20 kilohertsin kaistalle. Näistä näytepareista viisi yli eroavaisuusarvo 2:n ylittävää vertailuparia olivat näyteparit 24, 6, 7, 23 ja 5.

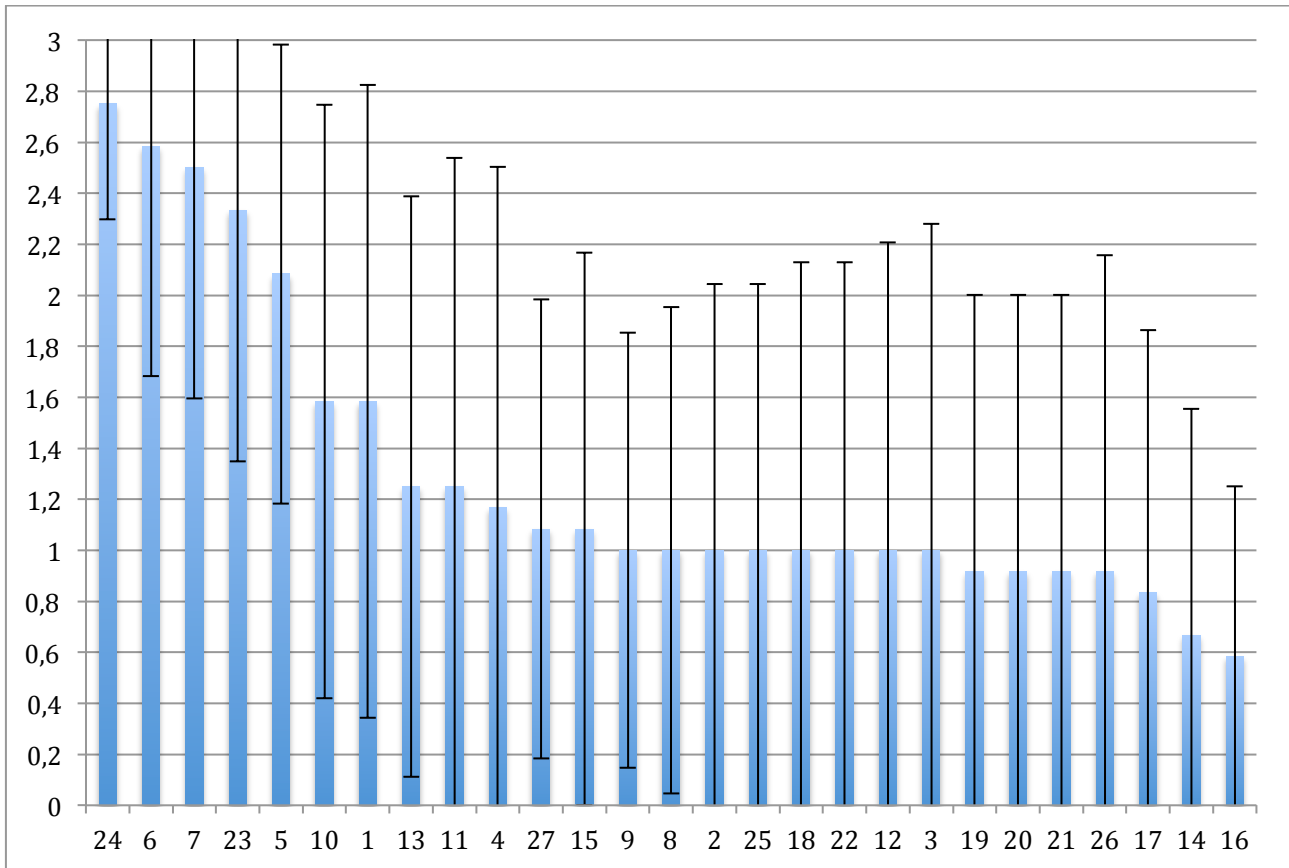
Vastaavasti äänenlaadullisesti samankaltaisimmiksi koettuja näytepareja 19, 20, 21 ja 26 yhdistää niiden näytetaajuuden nosto suhteellisesti pienemmällä porrasasteikolla, mistä johtuen kyseisten näyteparien eroavaisuuksien havaitseminen on ollut huomattavasti epäselvempää. Lisäksi näytteissä 17 ja 16, joissa CD-äänenlaatua verrataan korkeampiin näytetaajuuksiin eroavaisuuksien havaitseminen on ollut hankalampaa. Toisaalta on huomioitava, että esimerkiksi näytteet 8, 9, 15 ja 18, joissa CD-äänenlaatua (44,1 kilohertsiä) verrataan sitä korkeampiin näytetaajuuksiin, on koettu näytepareja 17 ja 16 erottuvammaksi. Toiseksi vähiten eroavaisuutta sisältävä näytepari 14 löytyy taulukon lopusta hyvästä syystä, koska sen osat olivat äänenlaadultaan identtiset.

Bittisyvyysvertailua (näytetaajuuden pysyessä samana) koskevat näyteparit 1, 2 ja 3 sijoittuvat taulukon keskikohtiin.

Kaiken kaikkiaan tulokset antavat hyvän kuvan siitä, että äänenlaadullisesti eroavien tallennustekniikoiden käytölle on kuultavissa olevia perusteita. Huomattavaa kuitenkin on, että erotuskyky laskee selvästi, jos näytteiden äänenlaatuerot ovat verrattain pieniä, eli toisin sanoen näytetaajuudet ovat lähellä toisiaan. Lisäksi erotuskyvyn laskuun vaikuttaa tietyissä tapauksissa CD-äänenlaatua korkeampien näytetaajuuksien käyttäminen, muttei kuitenkaan niin yksiselitteisesti, että sen perusteella olisi syytä hylätä korkeampien näytetaajuuksien käyttö vedoten vaikutusten kuulumattomuuteen.

Näytteen nro	Näyteääni	A-osan näytetaajuus ja bittisyvyys	B-osan näytetaajuus ja bittisyvyys	Havaitun eron keskiarvo (0-3)	Havaitun eron keskihajonta (0-3)
1	Kilkka	22,05 kHz / 16-bit	22,05 kHz / 24-bit	1,583	1,240
2	Tamburiini	44,1 kHz / 16-bit	44,1 kHz / 24-bit	1,000	1,044
3	Tingsha-symbaalit	192 kHz / 16-bit	192 kHz / 24-bit	1,000	1,279
4	Rin Gong –kulho	22,05 kHz / 16-bit	22,05 kHz / 16-bit	1,166	1,337
5	Kellosymbaali	22,05 kHz / 16-bit	44,1 kHz / 24-bit	2,083	0,900
6	Crash-symbaali	22,05 kHz / 16-bit	192 kHz / 24-bit	2,583	0,900
7	Kilkka	22,05 kHz / 24-bit	44,1 kHz / 16-bit	2,500	0,904
8	Tamburiini	44,1 kHz / 24-bit	192 kHz / 16-bit	1,000	0,953
9	Tingsha-symbaalit	192 kHz / 24-bit	44,1 kHz / 16-bit	1,000	0,852
10	Rin Gong –kulho	22,05 kHz / 16-bit	44,1 kHz / 16-bit	1,583	1,164
11	Kellosymbaali	24 kHz / 16-bit	44,1 kHz / 16-bit	1,250	1,288
12	Crash-symbaali	32 kHz / 16-bit	44,1 kHz / 16-bit	1,000	1,206
13	Kilkka	40 kHz / 16-bit	44,1 kHz / 16-bit	1,250	1,138
14	Tamburiini	44,1 kHz / 16-bit	44,1 kHz / 16-bit	0,666	0,887
15	Tingsha-symbaalit	44,1 kHz / 16-bit	48 kHz / 16-bit	1,083	1,083
16	Rin Gong –kulho	44,1 kHz / 16-bit	64 kHz / 16-bit	0,583	0,668
17	Kellosymbaali	44,1 kHz / 16-bit	88,2 kHz / 16-bit	0,833	1,029
18	Crash-symbaali	44,1 kHz / 16-bit	192 kHz / 16-bit	1,000	1,128
19	Kilkka	22,05 kHz / 16-bit	24 kHz / 24-bit	0,916	1,083
20	Tamburiini	32 kHz / 16-bit	40 kHz / 24-bit	0,916	1,083
21	Tingsha-symbaalit	44,1 kHz / 16-bit	48 kHz / 24-bit	0,916	1,083
22	Rin Gong –kulho	64 kHz / 16-bit	88,2 kHz / 24-bit	1,000	1,128
23	Kellosymbaali	192 kHz / 16-bit	22,05 kHz / 24-bit	2,333	0,984
24	Crash-symbaali	24 kHz / 16-bit	32 kHz / 24-bit	2,750	0,452
25	Kilkka	40 kHz / 16-bit	44,1 kHz / 24-bit	1,000	1,044
26	Tamburiini	48 kHz / 16-bit	64 kHz / 24-bit	0,916	1,240
27	Tingsha-symbaalit	88,2 kHz / 16-bit	192 kHz / 24-bit	1,083	0,900

Taulukko 1. Kuuntelukokeen rakenne ja tulokset



Kaavio 1. Koettujen eroavaisuuksien keskiarvot ja keskihajonnat (Y-akselilla koetun eron määrä, X-akselilla näytteen numero, ks. Taulukko 1.)

7 PÄÄTÄNTÖ

Tutkimustuloksieni mukaan väite siitä, ettei CD-äänenlaatua korkeampien tallennusresoluutioiden käytöllä ole vaikutusta koettuun äänenlaatuun voidaan kyseenalaistaa tietyin reunaehdoin. Ensinnäkin kaikkien valitsemieni näytteiden kohdalla koetun eron keskiarvo ylitti nollarajan, joten koehenkilöt todella kokivat eroavaisuuksia eri resoluutioilla tallennettuja ääniä vertailtaessa. Täten tulokseni haastavat yleisen oletuksen, jonka mukaan ihmisen äänenlaadullisen erottelukyvyn yläraja olisi välttämättä rajoittunut standardina pidetyn CD-äänenlaadun ja 20 kilohertsin tietämille. Koska tutkimukseni ei pyrkinyt varsinaisesti hakemaan niinkään laatuun keskittyvää arviota, en katso aiheelliseksi esittää väitteitä siitä millä tavalla käytettyjen äänenlaatuojen varsinainen paremmuus tai koehenkilöiden preferenssit esiintyvät tutkimuksessa. Joka tapauksessa taipumus erottaa näytteet toisistaan osoittaa, että äänenlaatuerojen kokeminen tämäntyyppisen vertailun avulla tuottaa tulosta.

Tuloksiin vaikuttavana tekijä on myös ehdottomasti koeympäristön ja tutkimuksessa käytetyn laitteiston toimivuus. Mahdollisesti aiemmat tutkimukset ja kokeet aiheen yhteydessä ovat voineet epäonnistua tai tuottaa käyttökelvottomia tuloksia sen takia, ettei jokin osatekijä laajemmassa ketjussa ole pystynyt tallentamaan tai toistamaan korkeataajuisempia ääniä. Tässä tapauksessa kuitenkin sekä suorittamani koelaitteiston mittaus, että tutkimustulokset osoittavat selvästi, että toistetun äänen CD-resoluutiota korkeammat taajuuskaistat (ainakin 25 kilohertsiin saakka) todella välittyvät kuulijoille vaikuttaen heidän erotuskykyynsä.

Vaikka tutkimukseni erosikin luonteeltaan esimerkiksi Oohashin (2000) ja hänen työryhmänsä kokeesta, tukevat tulokseni samansuuntaista havaintoa ihmisen kuulemiskyvyn konventionaalisten rajojen venyttämisestä. Samalla hieman kyseenalaiseen valoon joutuu esimerkiksi Watkinsonin (2001, 729–730) väite siitä, ettei CD-resoluutiota korkeammilla tallennusresoluutioilla saavuteta kuultavissa olevaa hyötyä, puuttumatta siihen millaista ”hyötyä” Watkinson tarkoittaa. Lavryn (2004) mukaan korkeammilla näytetaajuuksilla esiintyvät signaalivirheet voivat taas olla yksi tämänkin tutkimuksen osatekijöistä äänenlaatuerojen tunnistamisessa, mutta toisaalta koehenkilöiden kommentit esimerkiksi crash-symbaalinäytteiden luonnollisuudesta eivät tue mielikuvaa ”virheiden” perusteella muodostetuista mielipiteistä.

Lopuksi voin todeta, että huolimatta tutkimukseni kohtuullisen hyvästä onnistumisesta jatkotutkimus aiheen piirissä olisi toivottavaa ja suotavaa. Jatkotutkimuksen kannalta sekä käytettävien näyteäänien, että koehenkilöiden ja –ryhmien laajentaminen ja jakaminen olisi kuitenkin tarpeellista entistä täsmällisempien tutkimustulosten saamiseksi. Myös tutkimussuuntia

tarkentamalla äänenlaatuun keskittyvistä kokeista voisi saada tuloksia, jotka vastaavat tarkemmin juuri tiettyihin äänten ominaisuuksiin liittyviin kysymyksiin. Joka tapauksessa jatkotutkimusten laatiminen esimerkiksi käyttämäni tutkimusasetelman avulla tuottaa kohtuullisella todennäköisellä käyttökelpoisia tuloksia.

LÄHTEET

- Bech S. & Zacharov N. (2006). *Perceptual Audio Evaluation – Theory, Method and Application*. West Sussex, England: John Wiley & Sons.
- Daniel E., Mee C.D. & Clark M. (1999). *Magnetic recording: the first 100 years*. West Sussex, England: Wiley-IEEE Press.
- Davies H. (1996). A History of Sampling. *Organised Sound* 1(1), 3–11.
- Downes K. (2010). Perfect Sound Forever. *Technology and Culture*, 15(2), 305–331.
- Fine T. (2008). The Dawn of Commercial Digital Recording. *ARSC Journal*, 39(1), 1–17.
- Gallagher M. (2008). *Music Tech Dictionary: A Glossary of Audio-Related Terms and Technologies*. Boston, MA: Course Technology.
- Gottlieb G. (2007). *Shaping Sound in the Studio and Beyond: Audio Aesthetics and Technology*. Boston, MA: Course Technology.
- Gronow P. & Saunio I. (1990). *Äänilevyn historia*. Helsinki: WSOY.
- Haines R. (2001). *Digital Audio*. Scottsdale, AZ: Coriolis Group, LLC, The.
- Hansen K. F. (2002). The Basics of Scratching. *Journal of New Music Research*, 31(4), 357–365.
- Huotilainen M. (2010). Kuulojärjestelmä. Teoksessa *Musiikkipsykologia* (toim. Louhivuori J. & Saarikallio S.). Jyväskylä: Atena.
- Järveläinen H. (2010). Psykoakustiikka. Teoksessa *Musiikkipsykologia* (toim. Louhivuori J. & Saarikallio S.). Jyväskylä: Atena.
- Katz B. (2007). *Mastering Audio – The Art and the Science, second edition*. Burlington, MA: Focal Press.
- Oohashi T. & al (2000). Inaudible High-Frequency Sounds Affect Brain Activity: Hypersonic Effect. *Journal of Neurophysiology*, 83(6), 3548–3558.
- Owsinski B. (1999). *The Mixing Engineer's Handbook*. Vallejo, CA: Mixbooks.
- Pohlmann K. (1992). *The Compact Disc Handbook - 2nd ed*. Madison, WI: A-R Editions.
- Roads C. (1996). *The Computer Music Tutorial*. Cambridge, MA: The MIT Press.
- Simmons G. (1998). Rupert Neve interview part 1 of 3. *AudioTechnology Magazine*, 1(1), 56–60.
- Watkinson J. (2001). *The Art of Digital Audio*. Burlington, MA: Focal Press.

Verkkolähteet:

- Blomberg E. & Lepoluoto A. (2005). Audiokirja: audiotekniikkaa ammattilaisille ja kehittyneille harrastajille. [WWW-dokumentti]. Tapiolan viestintäsuunnittelu. [Viitattu 20.5.2012]. Saatavissa: <http://ari.lepoluo.to/audiokirja/>
- Boyk J. (1997). There's Life Above 20 KiloHertz! A Survey of Musical Instrument Spectra to 102.4 Khz. [WWW-dokumentti]. California Institute of Technology. [Viitattu 20.5.2012]. Saatavissa: <http://www.cco.caltech.edu/~boyk/spectra/spectra.htm>
- Lavry D. (2004). Sampling Theory For Digital Audio. [WWW-dokumentti]. Lavry Engineering Inc. [Viitattu 20.5.2012]. Saatavissa: http://www.lavryengineering.com/documents/Sampling_Theory.pdf
- RIAA (2012). Student FAQ. [WWW-dokumentti]. Recording Industry Association of America. [Viitattu 20.5.2012]. Saatavissa: http://www.riaa.com/toolsforparents.php?content_selector=resources-for-students