

**Language Environment Analysis (LENA) -menetelmän
validiteetti keskosvauvojen ääniympäristön arvioinnissa**

Kati Siirilä

Pro gradu -tutkielma

Turun yliopisto

Psykologian ja logopedian laitos

Psykologia

29.5.2019

TURUN YLIOPISTO Psykologian ja logopedian laitos / Yhteiskuntatieteellinen tiedekunta

SIIRILÄ, KATI: Language Environment Analysis (LENA) -menetelmän validiteetti keskosvauvojen ääniympäristön arvioinnissa

Pro gradu -tutkielma, 45 s

Psykologia

Toukokuu, 2019

Keskosvauvojen ääniympäristöllä vastasyntyneiden teho-osastolla on suuri merkitys heidän kielenkehitykselleen, koska ympäristö eroaa huomattavasti siitä, millainen ääniympäristö normaalisti kehittyvällä sikiöllä on kohdussa. Vastasyntyneiden teho-osaston ääniympäristön tutkiminen on tärkeää, jotta siitä voidaan rakentaa mahdollisimman hyvin keskosten kehitystä tukeva. Lasten ääniympäristön tutkimiseen on kehitetty Language Environment Analysis (LENA) -menetelmä, joka ääniympäristön äänittämisen lisäksi myös automaattisesti analysoi äänitteessä kuuluvaa ääntä, eli esimerkiksi erottelee sitä, paljonko äänitteellä kuuluu puhetta, lapsen ääntelyä tai hiljaisuutta. LENA-menetelmän validiteettia ei ole kattavasti tutkittu vastasyntyneiden teho-osaston ympäristössä ja vain kerran aikaisemmin suomen kielellä. Tämän pro gradu -tutkielman tarkoituksena oli arvioida LENA-menetelmän validiteettia vastasyntyneiden teho-osastolla suomenkielisten perheiden ja keskosvauvojen parissa.

Pro gradu -tutkielman aineisto koostuu kahdeksan satunnaisesti valitun perheen LENA-äänitteistä. Perheet olivat osa laajempaa tutkimusprojektia, jossa tutkitaan vanhemman läheisyyden ja läsnäolon sekä ääniympäristön vaikutusta lapsen kehitykseen. Tähän pro gradu -tutkielmaan valitut perheet olivat suomenkielisiä perheitä, joiden lapsi oli hoidettavana Turun yliopistollisen sairaalan vastasyntyneiden teho-osastolla. Mukaan valittujen perheiden äänitteistä kuunneltiin jokaisesta kuusi viiden minuutin jaksoa. Viisiminuuttisista kaksi valittiin äänitteestä kohdasta, jossa oli runsaasti puhetta, kaksi kohdasta, jossa puhetta oli keskimääräisesti ja vielä kaksi kohdasta, jossa puhetta oli vähän. Kaksi arvioijaa kuuntelivat valitut kohdat ja merkitsivät samalla ylös, oliko LENA-menetelmän äänelle antama koodi yhtenevä heidän havaintonsa kanssa. Koodien yhtenevyys laskettiin tarkastelemalla sitä, kuinka monta prosenttia koodeista LENA-menetelmä oli manuaaliseen koodaukseen verrattuna tunnistanut yhtenevästi.

LENA-menetelmän antamien koodien yhtenevyys arvioitsijoiden havaintojen kanssa aineistossa oli noin 70 prosenttia. Tulokset tukivat pääosin aikaisempia tutkimustuloksia. Aikuisen naisen koodin luotettavuus oli erittäin korkea ja jopa 90 prosenttia koodeista oli yhteneviä arvioitsijan havaintojen kanssa. Aikuisen miehen koodien luotettavuus jäi alhaisemmaksi, mikä voi johtua esimerkiksi siitä, että vauvoille puhuttaessa äänenkorkeus on usein normaalia puhetta korkeampi. Lapsen koodien luotettavuus jäi alhaisemmaksi kuin aiemmissa tutkimuksissa, mikä ei ollut kovin yllättävää ottaen huomioon keskosten keuhkojen kehityksen keskeneräisyyden. Hiljaisuuden koodin luotettavuus oli tarkastelluista huonoin. Melun koodin luotettavuus vastasi kaikkien koodien keskimääräistä luotettavuutta. Tulokset viittaavat siihen, että LENA-menetelmä on luotettava apuväline naisen puheen tunnistamisessa, mutta varsinkin keskosten oman ääntelyn ja hiljaisuuden määrän arviointiin tulee suhtautua varauksella.

Avainsanat: keskonen, vastasyntyneiden teho-osasto, ääniympäristö, Language Environment Analysis (LENA), validiteetti

Sisällysluettelo

1. Johdanto	1
1.1. Puheen merkitys lapsen kehitykselle	2
1.1.1. Puheen merkitys lapsen kielenkehitykselle raskausaikana	2
1.1.2. Varhaisen puheen merkitys lapsen kielenkehitykselle	3
1.2. Vastasyntyneiden teho-osasto ääniympäristönä ja sen tutkiminen	4
1.3. Language Environment Analysis (LENA) -menetelmän validiteetti	5
1.4. Tutkimuskysymykset	8
2. Menetelmät	9
2.1. Aineisto	9
2.2. Tutkimuksen kulku	9
2.3. Mittarit	10
2.3.1. Language Environment Analysis (LENA) -menetelmä ja Transcriber-ohjelma	10
2.3.2. Läheisyyspäiväkirja ja taustatietolomake	13
2.4. Aineiston analysointi	13
3. Tulokset	17
3.1. Perheiden taustatiedot ja äänitteiden kuvaus	17
3.2. LENA-menetelmän validiteetti	21
3.3. Jatkovertailut koodien arvion yhtenevyudessa puheen määrän mukaisissa jaotteluissa ja ihmisiäntä sisältävien koodien välillä	24
4. Pohdinta	25
4.1. Tutkimustulosten merkitys	29
4.2. Tutkimuksen vahvuudet ja rajoitukset	30
4.3. Jatkotutkimukset ja sovellukset	31
4.4. Johtopäätökset	32
Lähteet	33
Liitteet	38

1. Johdanto

Raskaus kestää yleensä 40 viikkoa, jonka jälkeen vauva syntyy maailmaan valmiina kohtaamaan sen ärsykkeet ja kasvamaan ja kehittymään lisää. Osa vauvoista syntyy kuitenkin enneaikaisesti eivätkä he vielä ole täysin valmiita kohtaamaan maailmaa sen monine ärsykkeineen (Parikka & Lehtonen, 2017). Suomessa ennen raskausviikkoa 37 eli keskosina syntyy vuosittain noin 3000 vauvaa. Tämä tarkoittaa, että yksi 20 lapsesta syntyy Suomessa keskosena. Pikkukeskosia Suomessa syntyy vuosittain noin 500-600, joka on alle 1 % kaikista vastasyntyneistä. Pikkukeskosella tarkoitetaan vauvaa, joka on syntynyt ennen raskausviikkoa 32+0 tai jonka syntymäpaino on ollut alle 1500 kilogrammaa. Vauvan syntymäajan merkinnässä ensimmäinen numero viittaa raskausviikkoihin ja toinen numero päiviin.

Raskausviikoilla 26-31 syntyneet vauvat viettävät vastasyntyneiden teho-osastolla keskimäärin 5-12 viikkoa ja ennen raskausviikkoa 26 syntyneet keskimäärin vielä pidempään (Parikka & Leino, 2017). Ottaen huomioon kehityksen hurjan tahdin raskausaikana ja lapsen ensimmäisinä elinvuosina, tällä teho-osastolla vietetyllä ajalla on suuri merkitys lapsen myöhemmällekin kehitykselle. Esimerkiksi aivojen kasvu ja poimuuntuminen, joka yleensä tapahtuu raskauden loppuvaiheessa, tapahtuu pikkukeskosilla suurelta osin osastohoidon aikana (Lehtonen, 2017). Pikkukeskosina syntyneiden vauvojen ääntelyssä on myös havaittu laadullisia eroja täysiaikaisiin vauvoihin verrattuna (Stolt & Yliherva, 2017A). Pikkukeskosten varhainen ymmärretty sanasto myös kehittyy hitaammin kuin täysiaikaisena syntyneiden vauvojen ja lisäksi ilmaistu sanavarasto voi olla pikkukeskosina syntyneillä suppeampi kahden vuoden kehitysiässä. Tämän takia on tärkeää löytää luotettavia välineitä pikkukeskosten ääniympäristön tutkimiseen, jotta elämän alun poikkeavasta (ääni)ympäristöstä huolimatta pikkukeskosten kehitykselle onnistutaan luomaan mahdollisimman optimaalinen ympäristö.

Tämän pro gradu -tutkielman tarkoituksena on selvittää, miten luotettavasti lapsen ääniympäristön tarkasteluun kehitetty Language Environment Analysis (LENA) -menetelmä toimii vastasyntyneiden teho-osastolla pikkukeskosina syntyneiden vauvojen ja heidän vanhempiansa kanssa. Tutkielmani aluksi käyn läpi puheen merkitystä lapsen kielenkehitykselle. Sen jälkeen käsittelen vastasyntyneiden teho-osastoa ääniympäristönä

ja sitä, miten tällainen ympäristö mahdollisesti vaikuttaa sekä vauvan kuuleman puheen määrään että laatuun ja tätä kautta vauvan kehitykseen. Lisäksi esittelen sitä, miten vastasyntyneiden teho-osastoa ääniympäristönä on tähän mennessä tutkittu. Johdannon lopuksi avaan vielä validiteetin käsitettä, koska se on keskeisessä osassa, kun etsitään luotettavia tapoja ääniympäristön tarkasteluun.

1.1. Puheen merkitys lapsen kehitykselle

Puheen merkitystä lapsen kehityksen eri osa-alueille on tutkittu laajasti. Puheen suuri merkitys esimerkiksi itsesäätelyn kehityksessä on tunnistettu ja lasten, joilla on puheen ja kommunikaation vaikeuksia, ajatellaan olevan heikommassa asemassa itsesäätelyn kehitystä ajatellen (Binns, Hutchinson & Cardy, 2019). Kuitenkin pelkästään pikkukeskosiin rajatuissa tutkimuksissa puheen merkitystä on tarkasteltu lähinnä suhteessa kielenkehitykseen, joten tässä yhteydessä käsittelen ainoastaan puheen merkitystä lapsen kielenkehitykselle. Aloitan tarkastelun raskausajasta ja sen jälkeen siirryn tarkastelemaan varhaisen puheen merkitystä lapsen kielenkehitykselle.

1.1.1. Puheen merkitys lapsen kielenkehitykselle raskausaikana

Vauvan altistuminen puheelle alkaa jo sikiöaikana. Tutkimusten mukaan äidin ääni on jo raskausaikana vallitseva piirre kohdun ääniympäristössä ja tutkimuksissa on havaittu, että sikiö on myös kykenevä reagoimaan siihen ainakin raskausviikolta 24 eteenpäin (Moon, 2017). Sikiöt oppivat kielestä ja sen piirteistä jo kohdussa ja he suosivatkin heti syntyessään äitinsä ääntä (DeCasper & Fifer, 1980) ja äidinkieltään (Mehler ym., 1988; Moon, Cooper & Fifer, 1993). Uudemmat tutkimukset kuitenkin esittävät, että tämä raskauden aikainen oppiminen ulottuu pidemmällekin kuin tällaiseen pintapuolisten piirteiden suosimiseen (Gervain, 2018). Lisääntyvä tutkimus viittaa siihen, että sikiö oppii jo raskausaikana kielen prosodiikasta, eli esimerkiksi sanojen kestosta ja painotuksesta (mm. Kisilevsky ym., 2009).

Tämän hetkisen tutkimustiedon valossa yksi hypoteesi on, että vauvan sikiöaikana opittu tieto oman äidinkieltensä prosodiikasta toimii pohjana myöhemmälle kielenkehitykselle

(Gervain, 2018). Pikkukeskoset kohtaavat ennenaikaisen syntymänsä vuoksi kovin erilaisen ääniympäristön varhaisen kehityksensä aikana. Ääniympäristö on usein monimutkaisempi eikä äidin ääni ole samalla tavalla dominoiva kuin kohdussa, joka normaalisti vaimentaa ulkoa tulevia muita ääniä ja niiden yksityiskohtia.

Näin ollen on helppo ymmärtää, miksi pikkukeskosten ääniympäristön tutkiminen on tärkeää. Pikkukeskoset eivät olekaan kohdun ääniympäristössä, kun heidän kielenkehityksensä pitäisi alkaa kielen prosodisten piirteiden oppimisella. He ovat vastasyntyneiden teho-osastolla, jossa kohtu ei olekaan suodattamassa ja vaimentamassa ympäriltä kuuluvia ääniä ja pelkkien kielen prosodisten piirteiden sijaan he saavatkin kaiken mahdollisen ääni-informaation jo kielenkehityksensä alusta lähtien. Mitä aikaisemmin vauva syntyy, sitä vähemmän hän on harjoitellut pelkkien kielen prosodisten piirteiden tunnistamista ja sitä heikompi saattaa olla hänen pohjansa tulevalle kehitykselle. Tämä saattaa osaltaan selittää myös, miksi pikkukeskosilla on suurempi riski kielenkehityksen ongelmille.

1.1.2. Varhaisen puheen merkitys lapsen kielenkehitykselle

Kieli kehittyy edelleen syntymän jälkeen vuorovaikutuksessa ympäröivän maailman kanssa. Vanhemmilla on suuri rooli kielen mallintamisessa asioita nimeämällä ja selittämällä (Stolt & Yliherva, 2017B). Syntymänsä jälkeen vauva kuulee kahdenlaista puhetta. Vauva kuulee puhetta, joka kohdistetaan hänelle ja puhetta, jota puhutaan hänen lähetyvillään. Lapselle kohdistettu puhe (child-directed speech) eroaa monin tavoin aikuiselle kohdistetusta puheesta. Sille tyypillistä on korkeampi sävelkorkeus, yksinkertaisempi sanasto, toistojen suurempi määrä ja liioiteltu intonaatio (Garnica, 1977 viitattu Pine, 1994; Snow, 1972). Useissa tutkimuksissa on löydetty positiivinen yhteys sen väliltä, miten paljon lapselle on puhuttu ja miten hänen sanavarastonsa on kehittynyt (mm. Hoff & Naigles, 2002; Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Weizman & Snow, 2001). Onkin tärkeää, että myös vastasyntyneiden teho-osastolla luodaan ympäristö, jossa vanhemmat pystyvät ja uskaltavat puhua lapsensa kanssa heti tämän syntymästä lähtien.

Tutkimuksissa on löydetty positiivinen yhteys 7 ja 18 kuukauden ikäisenä toteutetun lasten kehitystasoa mittaavan Bailey-3 testin pistemäärien ja sen välillä, kuinka suurelle

sanamäärälle pikkukeskonen on altistunut vastasyntyneiden teho-osastolla (Caskey, Stephens, Tucker & Vohr, 2014). Tutkimuksessa ei eritelty sitä, kuinka suuri osa sanamäärästä oli varsinaisesti vauvalle suunnattua, vaan sanamäärään sisältyi kaikki puhe, mitä vauvan ympärillä kuului. Caskey ja kumppanit kuitenkin huomasivat, että sanamäärä oli huomattavasti suurempi silloin, kun vanhemmat olivat paikalla, mikä todennäköisesti lisää myös suoraan vauvalle puhutun puheen määrää. Raskausviikolla 32 äänitettyjen äänitteiden aikuisen puheen sanamäärä per tunti selitti jopa 12 % kielellisen kehitystason kokonaispisteiden vaihtelusta ja jopa 20 % ilmaistun puheen pistemäärän vaihtelusta 18 kuukauden iässä, kun syntymäpainon vaikutus oli vakioitu. Raskausviikolla 36 toteutetun äänityksen aikuisen puheen sanamäärä per tunti selitti jopa 26 % kognitiivisen kehitystason kokonaispisteiden vaihtelusta 7 kuukauden iässä.

Lisäksi Caskey ja kumppanit (2011) huomasivat, että pikkukeskosina syntyneet vauvat alkoivat äännellä vähintään kahdeksan viikkoa ennen laskettua aikaa. Ääntelyiden määrä myös lisääntyi ajan myötä voimakkaasti. Vanhemman puheelle altistuminen oli huomattavasti voimakkaampi ennustaja ääntelyiden määrälle raskausviikolla 32 ja protokeskusteluiden määrälle raskausviikoilla 32 ja 36, kuin altistuminen muiden aikuisten puheelle. Nämä tulokset edelleen alleviivaavat sitä, kuinka tärkeää vastasyntyneiden teho-osastolla on luoda sellainen ympäristö, jossa nimenomaan vanhemmat haluavat ja uskaltavat puhua lapselleen.

1.2. Vastasyntyneiden teho-osasto ääniympäristönä ja sen tutkiminen

Puheella on siis suuri merkitys lapsen kehitykselle jo raskausajasta alkaen. Vastasyntyneiden teho-osasto tuo ympäristönä omat haasteensa siihen, miten paljon vauva kuulee ääntä ja puhetta ja paljonko vanhemmat puhuvat vauvallensa. Vastasyntyneiden teho-osaston fyysiseen suunnitteluun on tutkimusten pohjalta kehitetty ohjeita, joiden avulla pyritään muun muassa minimoimaan mahdolliset kehitykselle haitalliset äänenvoimakkuudet ja pyritään takaamaan, että teho-osastollakin perheillä olisi tarpeeksi yksityisyyttä vauvansa kanssa puhumiseen (Philbin & Evans, 2006). Kaikkiin vastasyntyneiden teho-osaston suunnittelua koskeviin kriteereihin voi tutustua artikkelissa ”Recommended standards for newborn ICU design, eighth edition” (White, Smith & Shepley, 2013). Esimerkiksi keskosten huoneissa ja aikuisten nukkumistiloissa

äänen taso ei saisi ylittää jatkuvasti 45 desibeliä. Jotta nämä äänentasoille annetut standardit täytetään, tulee esimerkiksi myös pintamateriaalien, huonekalujen ja muiden laitteiden täyttää niille annetut kriteerit esimerkiksi siitä, paljonko ne heijastavat ääntä ja siitä, paljonko niistä ylipäänsä lähtee melua käytettäessä.

Vastasyntyneiden teho-osastoa on tutkittu sekä siitä näkökulmasta, kuinka paljon ja kuinka kovalle äänelle pikkukeskoset ja heidän vanhempansa siellä ollessaan altistuvat, mutta myös tarkemmin siitä näkökulmasta, kuinka paljon pikkukeskoset kuulevat nimenomaan puhetta eivätkä vain yleisesti ääntä. Melun määrää erityyppisillä vastasyntyneiden teho-osastoilla on tutkittu esimerkiksi hyödyntämällä video-EEG tallennuksia, joista on laskettu äänen voimakkuuden tasoja (Szymczak & Shellhaas, 2014). Vastasyntyneiden teho-osaston melun on osoitettu aiheuttavan pikkukeskosille fysiologista stressiä ja häiritsevän heidän untaan (Philbin, 2000). Kuhn ja kumppanit (2012) ovat tutkineet äänen voimakkuuden vaihtelun vaikutuksia esimerkiksi pikkukeskosten sydämen sykkeeseen, hengitystiheyteen ja aivojen happisaturaatiotasoihin. Jo 10-15 desibelin yhtäkkinen äänen voimakkuuden tason nousu pikkukeskosten nukkuessa nosti näiden keskimääräistä sydämen sykettä ja laski hengitystiheyttä ja aivojen happisaturaatiotasoa verrattaessa perustasoon.

Keskosten puheelle altistumisen määrää ja sen vaikutuksia kehitykselle on alettu tutkia yhä enemmän käyttämällä hyödyksi Language Environment Analysis (LENA) -menetelmää, joka ääniympäristön tallentamisen lisäksi automaattisesti myös analysoi äänitteestä muun muassa aikuisen ja lapsen puhumaa sanamäärää ja erittelee myös sitä, paljonko melua äänitteellä kuuluu (tarkempi kuvaus kts. kpl 2.3.1.). LENA-menetelmä nopeuttaa ääniympäristön analysointia verrattuna ihmiskoodajaan ja lisäksi se tarjoaa ei-intrusiivisen metodin ääniympäristön analysointiin, kun ohjelmaa käytettäessä tutkijan ei välttämättä ole tarpeen kuunnella äänitteen sisältöä (Caskey & Vohr, 2013).

1.3. Language Environment Analysis (LENA) -menetelmän validiteetti

Menetelmän validiteetti viittaa siihen, mittaako menetelmä sitä, mitä sen halutaan ja ajatellaan mittaavan (Murphy & Davidshofer, 2005). Aiemmin validiteetin tutkimuksessa erotettiin sisältö-, rakenne-, ennustus- ja rinnakkaisvaliditeetti täysin erillisiksi alueiksi, mutta nykyään nämä neljä eri tapaa tutkia validiteettia nähdään vain validiteetin

arvioimisen eri puolina. Tässä tutkimuksessa tutkimme erityisesti LENA-menettelmän sisältövaliditeettia eli kuinka luotettavasti LENA-menettelmän äänen lähteiden koodit vastaavat vastasyntyneiden teho-osaston äänen lähteitä. Menettelmän validiteettia tutkitaan usein vertaamalla sitä johonkin jo olemassa olevaan laajasti käytettyyn ja validiksi todettuun menettelmään. Koska lasten ääniympäristön analysointiin ei ole tällä hetkellä käytössä toista automatisoitua menettelmää, valittiin tässä tutkimuksessa vertailukohdaksi ihmisarvioijat.

Tässä tutkimuksessa arvioidaan LENA-menettelmän validiteettia ääniympäristön tutkimisen välineenä vastasyntyneiden teho-osastolla pikkukeskosten parissa ja suomen kielellä. Hyvä validiteetti kertoisi siitä, että LENA-menettelmä tunnistaa äänen lähteen ihmisarvioijan kanssa yhtenevästi. Äänen lähteille annettuja koodeja ovat esimerkiksi FAN (aikuinen nainen, joka on läsnä huoneessa) ja CHN (lapsi, jonka yllä äänentallennin on) (kaikki koodit katso kpl 2.3.1.). Mahdollisia virhelähteitä LENA-menettelmälle on vastasyntyneiden teho-osaston erikoinen ja monimutkainen ääniympäristö, jossa on esimerkiksi paljon ääntä pitäviä elektronisia laitteita. Lisäksi pikkukeskosten ääntelyn poikkeavuus täysiaikaisista vauvoista ja varsinkin isommista lapsista voi aiheuttaa LENA-menettelmälle virheitä erityisesti lapsen ääntelyn tunnistuksessa. Pikkukeskosten keuhkojen kehitys on usein ennenaikaisen syntymän vuoksi kesken, mikä vaikuttaa esimerkiksi heidän ääntelynsä voimakkuuteen (Parikka, 2017). Myös suomen kieli saattaa itsessään tuottaa ohjelmalle hankaluuksia äänen lähteiden nimeämiseen, sillä esimerkiksi suomen kielen puheen äänenkorkeus eroaa englannin kielen korkeudesta ja ohjelma on alun perin kehitetty englanninkielisen väestön parissa. Suomenkielisten opiskelijoiden parissa toteutetussa tutkimuksessa naisten puheen korkeus oli keskimäärin 194,0 Hz (Leino, Laukkanen, Ilomäki & Mäki, 2008). Englanninkielisiä naisia tutkittaessa puheen korkeus vaihteli 199,7 ja 224,3 Hz:n välillä (Stoicheff, 1981, viitattu Guimarães & Gouveia, 2007). Tämä viittaa siihen, että suomen kielessä naisten puheen korkeus saattaa keskimäärin olla hieman matalampi kuin englannissa.

Tuoreesta kirjallisuuskatsauksesta (Ganek & Eriks-Brophy, 2018A) huomataan, että LENA-menettelmän validiteettia on jo tutkittu suhteellisen paljon englanniksi ja sen on todettu analysoivan äänitettyä puhetta hyvin, jopa 82 prosentin luotettavuudella. Ganek ja Eriks-Brophy huomauttavat kuitenkin kirjallisuuskatsauksessaan, että LENA-menettelmän validiteetin lisätutkiminen muilla kielillä kuin englannilla olisi tärkeää, jotta voisimme entistä paremmin ymmärtää kielen kehittymistä universaalisti. Tähän

mennessä englannin kielen lisäksi LENA-menettelmän validiteettia on tarkasteltu ranskaksi (Canault, Le Normand, Foudil, Loundon & Thai-Van, 2016), kiinaksi (Gilkerson ym., 2015), ruotsiksi (Schwarz ym., 2017), koreaksi (Pae ym., 2016), espanjaksi (Weisleder & Fernald, 2013) ja vietnamiksi (Ganek & Eriks-Brophy, 2018B). Kaikissa näissä tutkimuksissa LENA-menettelmän validiteetti on todettu hyväksi englannista poikkeavasta kielestä huolimatta. Tutkimuksissa LENA-menettelmän validiteettia on usein arvioitu laskemalla, kuinka hyvin LENA-menettelmän laskema aikuisen puheen sanamäärä ja manuaalisesti laskettu sanamäärä vastaavat toisiaan. Korrelaatiot vaihtelivat noin .65 ja .80 välillä, pienimmän korrelaation ollessa ranskan kielessä ja suurimman espanjassa. Näiden tutkimusten välillä on ristiriitaa käsitteiden suhteen, sillä osa puhuu menettelmän reliabiliteetin (kts. esim. Canault ym., 2016) ja osa menettelmän validiteetin (kts. esim. Ganek & Eriks-Brophy, 2018B) tutkimisesta.

Suomeksi LENA-menettelmän validiteettia on tähän mennessä tarkasteltu ainoastaan Hanna Elon (2016) väitöskirjassa. Elon väitöskirjassa tutkimuksen kohteena olivat jo 6 kuukauden ikäiset lapset, eivät keskoset. Elo tutki omassa väitöskirjassaan sekä sitä, kuinka hyvin LENA-menettelmä osasi nimetä ihmisäänien lähteen arvioijan kanssa yhtenevästi ja myös esimerkiksi sitä, kuinka hyvin LENA-menettelmän laskema aikuisen puheen sanamäärä vastasi manuaalisesti laskettua sanamäärää. Elo ei tutkinut väitöskirjassaan muiden kuin ihmisäänien koodien validiteettia. Elon aineistossa LENA-menettelmä nimetä luotettavasti 6 kuukauden ikäisten lasten ja aikuisen naisen äänet, mutta jonkin verran heikommin aikuisen miehen. Manuaalisen koodaukseen verrattaessa LENA-menettelmän antamien lapsen ja aikuisen naisen äänen koodien yhtenevyys oli noin 80-90 prosenttia ja aikuisen miehen noin 50-75 prosenttia.

Lena-menettelmää on käytetty keskosvauvoihin kohdistuneissa tutkimuksissa, mutta sen validiteettia keskosvauvojen parissa ei kuitenkaan ole kattavasti tutkittu (mm. Caskey ym., 2011; Caskey ym., 2014). Tällä hetkellä ei siis vielä tiedetä, kuinka luotettavasti LENA-menettelmä tässä populaatiossa ja kontekstissa toimii ja mitkä ovat sen mahdolliset rajoitukset. Myös suomen kielellä menettelmän luotettavuustutkimusta on vasta rajoitetusti.

1.4. Tutkimuskysymykset

Tässä tutkimuksessa tarkastellaan LENA-menetelmän luotettavuutta vastasyntyneiden teho-osastolla keskoslasten ja heidän vanhempiansa kanssa. Ohjelman validiteettia tarkastellaan vertaamalla LENA-ohjelman nimeämiä äänen lähteitä ihmisarvioijien arvioon saman äänen lähteen alkuperästä. Oletuksena oli, että ihmisarvioija pystyy luotettavammin arvioimaan äänen alkuperän.

Tätä tutkimusta varten muodostimme kolme tutkimuskysymystä:

1. Kuinka luotettavasti LENA-menetelmä tunnistaa ja nimeää eri äänilähteistä tulevat äänet?
2. Kuinka luotettavasti LENA-menetelmä tunnistaa ja nimeää aikuisen naisen (FAN/FAF), aikuisen miehen (MAN/MAF) ja lapsen (CHN/CHF) äänet?
3. Kuinka luotettavasti LENA-menetelmä tunnistaa ja nimeää hiljaisuuden (SIL) ja melun (NON/NOF)?

Laajemman tutkimuksen kannalta tärkeimpiä ovat koodit, jotka sisältävät puhetta ja siksi tällaisia koodeja tarkasteltiin erikseen. Lisäksi, koska melulla on todettu olevan haitallisia vaikutuksia keskosille, tarkastelemme myös sitä, kuinka luotettava melun koodi on. Toisaalta, koska puheen kehittymiselle on tärkeää, että lapsi kuulee puhetta eikä ympärillä kuulu pelkästään hiljaisuutta, tarkastelemme erikseen myös hiljaisuuden koodin luotettavuutta. Jatkovertailuina tarkastelemme sitä, eroaako koodien luotettavuus tilastollisesti merkitsevästi riippuen siitä, paljonko kohdassa on puhetta. Tarkastelemme jatkovertailuina myös sitä, eroaako ihmisääntä sisältävien koodien luotettavuus tilastollisesti merkitsevästi toisistaan.

2. Menetelmät

2.1. Aineisto

Tutkielma on osa Turun yliopistollisen sairaalan (TYKS) vastasyntyneiden teho-osaston ja Tallinnan lastensairaalan (Tallinn Children's Hospital) vastasyntyneiden teho-osaston yhteistä tutkimusprojektia, jossa tutkitaan vanhempien läheisyyden ja kommunikaation vaikutusta lapsen kehitykseen (Effects of Parent-Infant Closeness and Communication on Child Development). Projektissa on tarkoitus tarkastella muun muassa sitä, miten vanhempien läsnäolon määrä vastasyntyneiden teho-osastolla vaikuttaa vauvan ääniympäristöön ja miten nämä tekijät yhdessä vaikuttavat lapsen tarkkaavaisuuteen 7 kk:n iässä tai lapsen kielen kehitykseen 12 ja 24 kk:n iässä.

Laajemman projektin aineisto koostuu 50 perheestä TYKS:stä ja 50 perheestä Tallinnan lastensairaalaan, joihin oli syntynyt lapsi ennen raskausviikkoa 32+0. Tutkimuksen poissulkukriteerit olivat: 1) lapsen henkeä uhkaava sairaus, 2) synnynnäinen poikkeama, 3) kromosomipoikkeama tai 4) kliinisesti merkittävä syndrooma. Mukaan otettiin suomen-, ruotsin-, eestin- ja venäjänkielisiä perheitä.

Tämän pro gradu -tutkielman aineistoon valittiin mukaan 8 perhettä Turun yliopistollisesta keskussairaalaan kerätystä aineistosta. Kaikki mukaan valitut perheet olivat suomenkielisiä. Ruotsinkieliset perheet jätettiin pois siksi, että niiden osuus aineistosta oli niin pieni. Perheet valittiin aineistosta satunnaisesti niin, että satunnaisen numeron (1-50) valinnan jälkeen tarkistettiin perheen käyttämä kieli, jotta varmistettiin se, että tähän tutkimukseen mukaan tuli ainoastaan suomenkielisiä perheitä. Tallinnan lastensairaalan aineisto jätettiin kokonaan pois tästä tutkimuksesta aikataulusyistä.

2.2. Tutkimuksen kulku

Tutkimuksen perheiden rekrytointi aloitettiin Turussa maaliskuussa 2017 ja Tallinnassa huhtikuussa 2017. Rekrytoinnin hoitivat Turussa vastasyntyneiden teho-osaston tutkimushoitaja Minna Paaso ja neonatologi Anette Aija. Tallinnassa rekrytoinnista vastasi ainoastaan Anette Aija. Rekrytoijan tehtävänä oli kertoa perheille tutkimuksesta

ja selvittää myös, että perhe täytti tutkimukseen mukaan pääsemisen ehdot (kts. kappale 2.1.). Tutkimukseen suostumisen jälkeen mukaan valittuja perheitä äänitettiin yhtäjaksoisesti 16 tunnin ajan käyttäen LENA-menetelmään kuuluvaa äänentallenninta. Äänentallennin asetetaan liiviin, joka puetaan lapselle päälle ja se tallentaa kaiken lapsen lähellä kuuluvan äänen. Äänentallentimen asettamisen ja käynnistämisen hoiti joko tutkimushoitaja Minna Paaso tai neonatologi Anette Aija, jotka myös hakivat äänentallentimen pois äänityksen päätyttyä. Äänitykset toteutettiin raskausviikoilla 32+0-33+6 ja äänityspäivää ympäröivältä kahdelta viikolta vanhemmat täyttivät myös läheisyyspäiväkirjaa siitä, milloin he olivat paikalla vastasyntyneiden teho-osastolla ja koska lapsi oli sylissä tai kenguruhoivossa. Tutkimus jatkuu tarkkaavaisuuden kehityksen tarkastelulla lapsen ollessa 7 kuukautta (+- 7 päivää) ja lapsen kielen kehityksen arvioinnilla lapsen ollessa 12 kuukautta (+- 7 päivää).

2.3. Mittarit

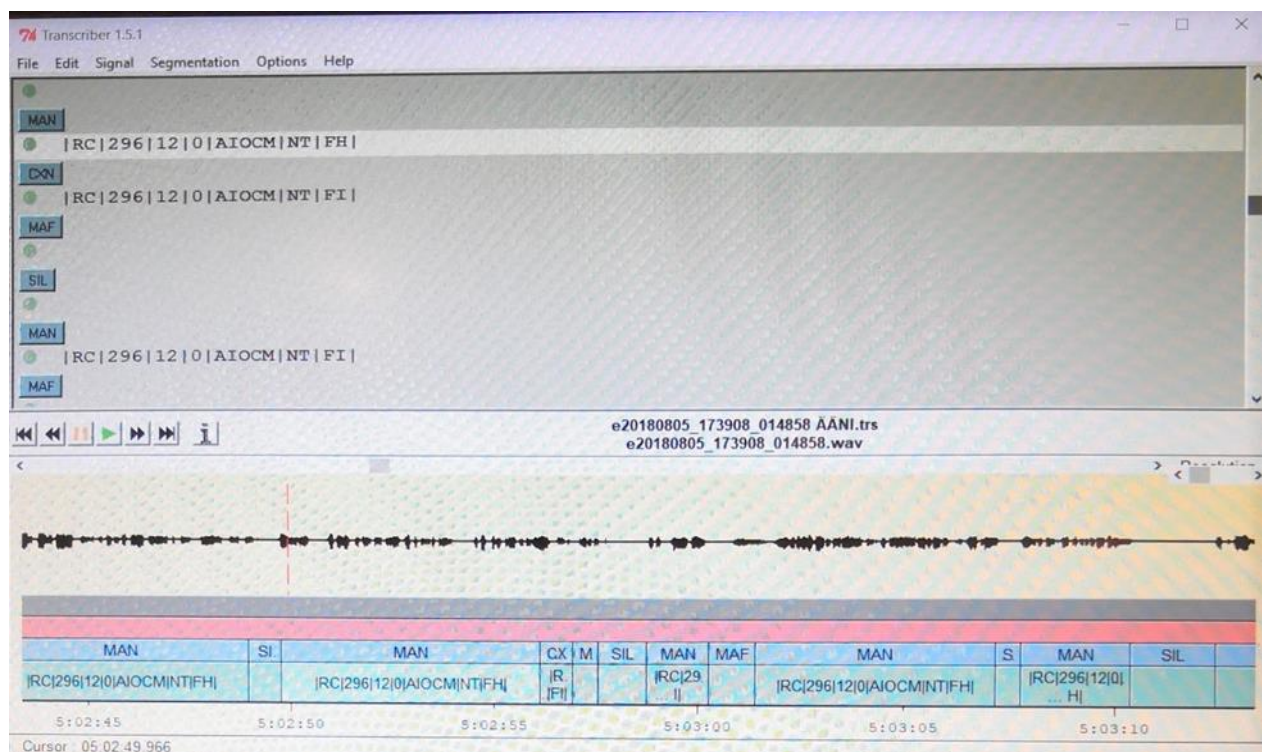
2.3.1. Language Environment Analysis (LENA) -menetelmä ja Transcriber-ohjelma

Tässä tutkimuksessa äänen tallentamiseen ja äänitteiden analysointiin käytettiin Language Environment Analysis (LENA) -menetelmää, jonka ovat kehittäneet Terrance ja Judith Paul. LENA-menetelmä on lasten kieliympäristön analysointiin kehitetty ohjelma ("What is LENA?", n.d.). LENA-menetelmä on äänen tallennus- ja analysointipaketti, jonka avulla lasta ympäröivää ääniympäristöä voidaan helposti tallentaa, käsitellä ja analysoida. LENA-menetelmä kehitettiin alun perin 2-48 kuukauden ikäisille lapsille ja heidän vanhemmilleen. Alun perin LENA-menetelmä kehitettiin avuksi palautteen antamiseen vanhemmille siitä, paljonko he puhuvat lastensa kanssa ja toiveena oli palautteen kautta lisätä lasten kuuleman puheen määrää.

LENA-menetelmä koostuu kevyestä äänentallentimesta, joka asetetaan sitä varten suunnitellun liivin sisään ("What is LENA?", n.d.). Liivi on suunniteltu niin, että äänentallennin pystyisi mahdollisimman tarkasti äänittämään kaikki ympärillä kuuluvat äänet. Äänentallentimen ja liivin lisäksi LENA-menetelmään kuuluu myös tietokoneohjelmisto, jonka avulla äänitteitä voidaan kuunnella. Ohjelmisto antaa myös erilaisia tietoja äänitteistä, kuten esimerkiksi, kuinka monta aikuisten puhumaa sanaa

äänitteessä on tai kuinka monta lapsen puhumaa sanaa. LENA-menetelmä myös erittelee esimerkiksi sitä, paljonko melua äänitteellä kuuluu.

Tässä pro gradu -tutkielmassa tarkempaa äänitteiden tarkastelua varten käytimme erillistä Transcriber-ohjelmistoa, jonka avulla äänitteitä pystyi kuuntelemaan kokonaisuudessaan. Transcriber-ohjelmisto on puheen segmentointiin, nimeämiseen ja litterointiin kehitetty työkalu (Barras, Geoffrois, Wu & Liberman, 1998). Transcriber-ohjelma avaa äänitteen niin, että kuunneltaessa äänitettä se näyttää samalla, minkä koodin LENA on siinä kohdassa kuuluvalla äänelle antanut. Kuvassa 1 on nähtävillä, miltä Transcriber-ohjelma näyttää ja miten siinä ovat näkyvillä puheelle annetut koodit.



Kuva 1. Transcriber-ohjelman ulkoasu. Kuvassa näkyy kuinka LENA-menetelmän antamat koodit näkyvät Transcriber-ohjelmassa.

Kaikki LENA-menetelmän ja Transcriber-ohjelman käyttämät koodit ja niiden selitykset ovat esitettyinä taulukossa 1. Koodit voidaan jakaa pääluokkiin äänen lähteen mukaan, eli esimerkiksi aikuinen nainen tai lapsi ("LENA Speaker labels in exported data", n.d.). Sen jälkeen, muut ääniluokat paitsi hiljaisuus, jakautuvat vielä kahtia sen mukaan, onko

äänen lähde lähellä vai kaukana. Äänen lähteen etäisyys selviää koodin viimeisestä kirjaimesta; viimeinen kirjain N kertoo, että äänen lähde on lähellä ja viimeinen kirjain F taas siitä, että äänen lähde on kauempana lapsen päällä olevasta äänentallentimesta, se on hiljainen tai että ääni eroaa jonkin verran siitä mallista, johon LENA-menetelmä sen ääniluokan ääniä vertaa. Koodeja tarkastellessa ei ole otettu huomioon niiden pituutta, joten koodien pituudet voivat erota toisistaan huomattavasti. Esimerkiksi sekä yhtäjaksoinen viiden sekunnin pituinen hiljaisuus että minuutin kestävä yhtäjaksoinen hiljaisuus saavat molemmat vain yhden koodin.

Taulukko 1. *Taulukossa on esiteltyinä LENA-menetelmän sekä LENA-äänitteiden kuunteluun käytetyn Transcriber-ohjelman käyttämät äänikoodit ja niiden merkitykset.*

Koodi	Koodin kuvaus
CHN	Kohde lapsi -eli lapsi, jolla liivi on
FAN	Aikuinen nainen, läsnä huoneessa
MAN	Aikuinen mies, läsnä huoneessa
CXN	Joku muu lapsi, läsnä huoneessa
TVN	TV/elektroniikka
NON	Melu -esim. liivin hankautuminen tai live-musiikki
OLN	Päällekkäisyys (ihmisen ääni + mikä tahansa muu ääni)
CHF	Mahdollisesti kohde lapsi, mutta kaukana/hiljainen/eroaa mallista
FAF	Mahdollisesti aikuinen nainen, mutta kaukana/hiljainen/eroaa mallista
MAF	Mahdollisesti aikuinen mies, mutta kaukana/hiljainen/eroaa mallista
CXF	Mahdollisesti joku muu lapsi, mutta kaukana/hiljainen/eroaa mallista
OLF	Mahdollisesti päällekkäisyyttä, mutta kaukana/hiljainen/eroaa mallista
TVF	Mahdollisesti TV/elektroniikka, mutta kaukana/hiljainen/eroaa mallista
NOF	Mahdollisesti melua, mutta kaukana/hiljainen/eroaa mallista
SIL	Hiljaisuus

2.3.2. Läheisyyspäiväkirja ja taustatietolomake

Äänitteiden kuuntelukohtien valinnassa käytettiin apuna myös vanhempien täyttämää läheisyyspäiväkirjaa, josta tarkistettiin, milloin vanhemmat olivat paikalla vastasyntyneiden teho-osastolla. Molemmat vanhemmat merkitsivät erikseen päiväkirjaan oman paikallaolonsa vastasyntyneiden teho-osastolla. Päiväkirjaan merkittiin myös, koska lapsi oli vanhemman sylissä tai kenguruhoiossa. Paikallaololla tarkoitettiin sitä, että vanhempi oli jossain sisällä teho-osastolla (Raiskila ym., 2017). Kenguruhoito erosi sylissä pitämisestä siinä, että kenguruhoion aikana vauva ja vanhempi olivat iho ihoa vasten. Lisäksi vanhemmat täyttivät taustatietolomakkeen, jossa kysyttiin tietoja perheestä, vanhemmista sekä keskosen voinnista ja tarvituista tukimenetelmistä.

2.4. Aineiston analysointi

Jokaisen kahdeksan perheen äänitteistä kuunneltiin 30 minuuttia puhetta. Koodien yhtenevyyden tarkasteluun valittiin kuusi viisiminuuttista LENA-menetelmän kehittäneen LENA Foundationin konsultaation jälkeen. LENA Foundation suositteli koodien luotettavuuden tarkastelua erikseen puheen määrän mukaan. Viisiminuuttisia valittiin kaksi puheen määrän mukaan jaotelluista tunneista niin, että sekä runsaan, keskimääräisen että vähäisen puheen määrän kohdat tulivat edustetuiksi. Puheen määrällä viitataan siihen, paljonko LENA-menetelmä oli ilmoittanut kohdassa olevan ihmisääntä. Tämän puheen määrän mukaisen tarkastelun perusteena on se, että sen kautta voidaan saada kiinni mahdolliset systemaattiset virheet, joita äänitteessä saattaisi olla, jos tarkasteluun ei otettaisi puheen määrän koko laajuutta (Bland & Altman, 2011). Systemaattisia virheitä voisi aiheutua esimerkiksi siitä, että runsaan puheen kohdassa paikalla voi olla enemmän henkilöitä, joista yhden äänen LENA-menetelmä nimeää jatkuvasti väärin. Ottamalla tarkasteluun kohtia niin runsaan, keskimääräisen kuin vähäisenkin puheen määrän kohdista voidaan tällaisia virheitä saada kiinni.

Näin ollen, jokaisen perheen äänitteestä etsittiin kuunneltavat viisiminuuttiset niin, että kaksi viisiminuuttista edustivat runsaan puheen kohtia, kaksi keskimääräisen puheen kohtia ja kaksi vähäisen puheen kohtia. Jokaisen perheen äänitettä kuunneltiin yhteensä

30 minuuttia. Kuunneltavat kohdat valittiin äänitteistä kohdista, joissa vähintään toinen vanhemmista oli paikalla vastasyntyneiden teho-osastolla. Kuuntelukohtien valinta aloitettiin siis tarkastelemalla läheisyyspäiväkirjan avulla sitä, koska vanhemmat olivat paikalla. Tämän jälkeen äänitteestä etsittiin se tunti, jolla oli LENA-menettelmän mukaan ollut prosentuaalisesti eniten ihmisääntä ja se tunti, jolla ihmisääntä oli ollut menetelmän mukaan prosentuaalisesti vähiten. Näiden avulla äänitteelle laskettiin ihmisäänen mukaan ylimmän 10 prosentin, keskimmäisen 20 prosentin ja alimman 10 prosentin luokat. Tämän jälkeen äänitteestä etsittiin jokaiseen luokkaan yksi sopiva tunti. Jos useampi tunti olisi sopinut yhteen luokkaan, valinta tehtiin sattumanvaraisesti. Tämän jälkeen jokaista valittua tuntia tarkasteltiin vielä tarkemmin ja sen sisältä valittiin kaksi viisiminuuttista, jotka sisälsivät LENA-menettelmän mukaan eniten ihmisääntä. Tällä tavalla varmistettiin, että kaikki valitut viisiminuuttiset sisälsivät myös ihmisääntä. Jos useampi kuin kaksi viisiminuuttista olisi tullut valituksi, tehtiin valinta tässäkin tapauksessa sattumanvaraisesti. Tämä sama valinta toteutettiin kaikille tutkimuksessa mukana olleille äänitteille erikseen.

Esimerkiksi äänitteestä yksi kuunneltavat kohdat etsittiin seuraavalla tavalla: Ensin läheisyyspäiväkirjan avulla tarkasteltiin, koska vanhemmat olivat äänityspäivänä paikalla vastasyntyneiden teho-osastolla. Läheisyyspäiväkirjasta nähtiin, että vähintään toinen vanhemmista oli paikalla 10.00-13.00, 15.00-18.00 ja 20.00-24.00 välisinä aikoina. Sen jälkeen LENA-menettelmästä tarkastettiin, montako prosenttia ihmisääntä oli ollut korkeimmillaan näinä tunteina ja montako prosenttia sitä oli ollut alhaisimmillaan. Tämän äänitteen tapauksessa korkeimmillaan ihmisääntä oli ollut 10 prosenttia ja alhaisimmillaan 1 prosentti. Näiden avulla laskettiin äänitteelle ihmisäänen mukaan alimman 10 prosentin luokka, joka oli 1-1,9 prosenttia, keskimmäisen 20 prosentin luokka, joka oli 4,6-6,4 prosenttia sekä ylimmän 10 prosentin luokka, joka oli 9,1-10 prosenttia. Luokkien laskemisen jälkeen jokaiseen luokkaan valittiin äänitteestä vanhempien paikallaolo tunneista yksi tunti, jonka ihmisäänen prosentuaalinen määrä sopi siihen luokkaan. Tämän äänitteen tapauksessa alimman 10 prosentin luokkaan sopi aikaväli 15.00-16.00, keskimmäisen 20 prosentin luokkaan 10.00-11.00 ja ylimmän 10 prosentin luokkaan 11.00-12.00. Tämän jälkeen kuunneltavat viisiminuuttiset etsittiin tarkastelemalla erikseen jokaista tuntia. Esimerkiksi aikaväliltä 15.00-16.00 valittiin kuunneltavaksi ne viisiminuuttiset, jotka sisälsivät LENA-menettelmän mukaan eniten ihmisääntä. Äänitteen yksi tapauksessa ne olivat 15.25-15.30 ja 15.30-15.35.

Ennen kuin äänitteitä lähdettiin kuuntelemaan ja koodaamaan manuaalisesti, pidettiin arvioijien kesken palaveri, jossa sovittiin yhteisistä kriteereistä koskien kuuntelua. Ensimmäisen äänitteen kuuntelun jälkeen pidettiin uusi palaveri, jossa käsiteltiin esille nousseita ongelmakohtia ja tarkennettiin koodaamisen sääntöjä. Molemmat arvioijat kuuntelivat äänitteet kuulokkeilla, jotta voitiin varmistua siitä, että molemmat kuulisivat äänitteet mahdollisimman samankaltaisesti. Lisäksi sovittiin, että jos kuunneltavan kohdan aloitus tai lopetus osui keskelle ohjelman antamaa koodia, aloitettiin kuuntelu koodin alusta ja lopetettiin kuuntelu koodin loppuun. Tällöin valittu viisi minuuttinen saattoi käytännössä olla pidempikin kuin viisi minuuttia. Äänitteitä kuunnellessa arvioitiin, vastaako LENA-menetelmän antama äänen lähteen pääluokka arvioijan antamaa pääluokkaa, mutta ei arvioitu LENA-menetelmän ja arvioijan välistä vastaavuutta äänen lähteen etäisyyden osalta. Äänitteissä esille tulevien ongelmakohtien ratkaisuksi sovittiin myös yhteisistä säännöistä:

- Jos arvioijan mukaan kohdassa on selkeää puhetta, täytyy se tulla ilmi myös LENA-menetelmän antamassa koodissa tai koodi arvioidaan ei-yhteneväksi.
- Jos kohta on epäselvä tai LENA-menetelmä koodaa suurimman osan yhtenevästi arvioijan kanssa, koodi arvioidaan yhteneväksi. Eli jos esimerkiksi minuutin kestävä hiljaisuuden koodi sisältää oven sulkeutumisesta aiheutuvan äänen, on LENA:n antama hiljaisuuden koodi yhtenevä arvioijan kanssa.
- Jos kohdassa on sekä miehen että naisen puhetta yhtä paljon, mutta LENA-menetelmä koodaa vain toisen, arvioidaan koodi ei-yhteneväksi.
 - Mutta jos kohdassa on selkeä pääpuhuja ja LENA-menetelmä on tunnistanut tämän, koodi arvioidaan yhteneväksi.
- Jos LENA-menetelmän antama pääkategoria on yhtenevä arvioijan arvion kanssa (esimerkiksi mies tai nainen), ei arvioida, onko LENA-menetelmän antaman koodin etäisyys yhtenevä arvioijan kanssa (koodin viimeinen kirjain N tai F).

- Jos kaksoset ovat samassa sängyssä, merkitään kaikki lapsen puheen koodit yhteneviksi, silloin kun kohdassa oikeasti kuuluu lapsen ääntä.

Äänitteiden yhtenevyyden analysointi toteutettiin Transcriber-ohjelmalla. Excel-
taulukkoon merkittiin äänitteitä kuunneltaessa, minkä koodin LENA-menetelmä oli
äänelle antanut ja viereen se, oliko koodi arvioijan mielestä yhtenevä hänen havaintonsa
kanssa äänen alkuperästä (koodi=1) vai ei-yhtenevä (koodi=0). Äänitteitä arvioi 2
henkilöä, joista toinen toimi ensisijaisena arvioijana ja toinen niin sanottuna
reliabiliteetti-arvioijana. Perheiden äänitteistä kaksi kuunneltiin päällekkäin niin, että
ensisijainen arvioija kuunteli yhteensä 7 äänitettä ja reliabiliteetti-arvioija 3. Arvioijien
välistä reliabiliteettia arvioitiin kahden äänitteen eli yhteensä 60 minuutin avulla niin, että
molemmat kuuntelivat niistä valitut kohdat (runsaat, keskimääräiset ja vähäiset). Toisessa
äänitteessä arvioijat olivat noin 74-prosenttisesti samaa mieltä ja toisessa 87-
prosenttisesti. Arvioijien välille laskettiin myös Cohenin kappa, joka oli ensimmäisen
äänitteen kohdalla .442, $p < .001$ ja toisen .699, $p < .001$. Yli .4 kappa-kerrointa voidaan
pitää tyydyttävänä ja yli .6 kerrointa hyvänä (McHugh, 2012).

Ensin kaikista äänitteistä laskettiin Excel-ohjelman avulla, kuinka monta prosenttia
kaikista koodeista oli yhteneviä LENA-menetelmän ja arvioijien välillä. Tämän jälkeen
tarkasteltiin, kuinka monta prosenttia tarkempaan tarkasteluun valituista aikuisen naisen
(FAN/FAF), aikuisen miehen (MAN/MAF), lapsen (CHN/CHF), melun (NON/NOF) ja
hiljaisuuden (SIL) koodeista oli yhteneviä LENA-menetelmän ja arvioijien välillä.
Tämän jälkeen tarkasteltiin vielä erikseen sitä, miten suuri osa kaikista koodeista oli
yhteneviä LENA-menetelmän ja arvioijien välillä runsaan, keskimääräisen ja vähäisen
puheen osuuksissa. Myös aikuisen naisen, aikuisen miehen, lapsen, melun ja hiljaisuuden
koodien yhtenevyyttä LENA-menetelmän ja arvioijien välillä arvioitiin erikseen runsaan,
keskimääräisen ja vähäisen puheen osioissa. Nämä kaikki vertailut tehtiin sekä jokaiselle
äänitteelle erikseen että myös koko datalle yhdessä. Jatkovertailuina tarkasteltiin SPSS-
ohjelmassa khiin neliö -testillä vielä sitä, erosivatko vähäisen, keskimääräisen ja runsaan
puheen kohdat toisistaan koodien yhtenevyydessä tilastollisesti merkitsevästi. Lisäksi
jatkovertailuina tarkasteltiin khiin neliö -testillä myös sitä, erosivatko ihmisäänien
koodien yhtenevyydet toisistaan. Jatkovertailut toteutettiin khiin neliö -testillä, koska
tarkasteltava muuttuja oli kategorinen eli yhtenevä – ei-yhtenevä. Vertailuille laskettiin
efektikoon estimaatti Cramerin V. Efektikokojen tulkinnassa käytettiin Cohenin (1988)

antamia suuntaviivoja. Arvioijien välistä reliabiliteettia tarkasteltiin sekä Excel-ohjelmassa lasketun prosenttien kautta että SPSS-ohjelmistolla lasketun kappa-arvon avulla.

3. Tulokset

3.1. Perheiden taustatiedot ja äänitteiden kuvaus

Tutkimukseen mukaan valitut perheet sisälsivät laajasti eri painoisina sekä erilaisilla raskausviikoilla syntyneitä keskosia. Mukana oli sekä tyttöjä että poikia ja myös sekä yksösinä että kaksosina syntyneitä. Vanhempien iät ja koulutusasteet vaihtelivat myös perheiden välillä. Lämmönsäätelyyn ja hengitykseen tarvittavan tuen määrä vaihteli vauvojen välillä. Taulukossa 2 on esiteltyinä tarkemmin jokaiseen perheeseen liittyvät taustatiedot.

Taulukko 2. Perheiden taustatiedot.

Äänite	Rv	Paino (g)	Sp	Lasten lukumäärä	Äidin ikä	Isän ikä	Koulutus (äiti)	Koulutus (isä)	LENA-nauhoituspäivä	Lämpö	Hengitystuki
1	30+6	1640	1	1	26	25	toinen aste	toinen aste	7.3.2017	2	0
2	26+5	600	1	1	32	34	korkeakoulu	korkeakoulu	21.4.2017	2	4
3	31+6	770	2	1	33	41	-	toinen aste	18.8.2017	2	1
4	28+4	880	1	2	30	30	toinen aste	korkeakoulu	19.9.2017	2	3
5	29+2	1300	1	2	25	26	toinen aste	toinen aste	4.11.2017	2	3
6	26+4	860	1	1	21	21	toinen aste	toinen aste	25.1.2018	2	3
7	28+3	1080	1	2	38	36	korkeakoulu	korkeakoulu	25.7.2018	0	0
8	30+4	1410	2	2	28	30	toinen aste	peruskoulu	23.9.2018	0	3

Huom.

Rv= raskausviikot, joilla vauva syntynyt

Sp= sukupuoli: 1=tyttö, 2=poika

Lasten lukumäärä: 1=yksönen, 2=kaksonen

Lämpö= lämmönsäätelyn tuki: 0=ei mitään, 1=keskoskaappi, 2=lämpöpöytä

Hengitystuki: 0=ei mitään, 1=hengityskone, 2=CPAP (ylipainehoito), 3=high flow-viikset, 4=CPAP/high flow-viikset

Taulukosta 3 nähdään kategorioihin jaettujen äänen lähteiden prosenttiosuudet jokaisessa äänitteessä sekä näiden keskiarvot kaikissa äänitteissä. Näissä ei ole otettu huomioon LENA-menettelyn mahdollisia vääriä tunnistuksia. Merkityksellisen puheen luokka sisältää kaikki ne äänet, jotka LENA-menettely on nimennyt ihmisääniksi. Eniten äänitteissä oli LENA-menettelyn mukaan hiljaisuutta, jonka keskimääräinen osuus oli noin 65 prosenttia äänitteistä.

Taulukko 3. Merkityksellisen puheen, hiljaisuuden, melun, tv:n ja elektroniikan sekä kaukaisen äänen prosenttiosuuksien keskiarvot sekä osuudet jokaisessa äänitteessä erikseen.

	Merkityksellinen puhe (%)	Hiljaisuus (%)	Melu (%)	Tv ja elektroniikka (%)	Kaukainen ääni (%)
Koko data ka	7.5	65.1	17	2.6	7.8
1	4	86	1	5	4
2	7	60	23	1	9
3	7	75	8	1	9
4	7	70	11	3	9
5	16	43	30	1	10
6	6	41	45	1	7
7	6	68	17	1	8
8	7	78	1	8	6

Taulukoissa 4-6 nähdään merkityksellisen puheen, hiljaisuuden ja melun prosenttiosuudet erikseen runsaan, keskimääräisen ja vähäisen puheen määrän mukaan jaoteltuina. Taulukoissa on nähtävillä sekä prosenttiosuuksien keskiarvo kaikissa äänitteissä että kaikkien kategorioiden prosenttiosuudet kaikissa äänitteissä erikseen. Taulukoista voidaan huomata, että merkityksellisen puheen osuus äänitteissä keskimäärin vähenee suhteellisen voimakkaasti, kun siirrytään runsaan puheen määrän kohdasta vähäisen puheen määrän kohtaan. Toisaalta taas hiljaisuuden ja melun määrät kasvavat samanaikaisesti.

Taulukko 4. Merkityksellisen puheen määrän keskiarvo ja prosenttiosuudet jokaisessa äänitteessä jaoteltuna runsaan, keskimääräisen ja vähäisen puheen määrän mukaan.

	Merkityksellinen puhe (%) -runsaan puheen kohdat	Merkityksellinen puhe (%) -keskimääräisen puheen kohdat	Merkityksellinen puhe (%) -vähäisen puheen kohdat
Koko data ka	48.8	27.6	10.1
1	57.5	21.5	1.5
2	45.5	40.5	20.5
3	47.5	27	24.5
4	30.5	21	4
5	70	42	9
6	38.5	21	4
7	48	31	16
8	53	17	1.5

Taulukko 5. Hiljaisuuden määrän keskiarvo ja prosenttiosuudet jokaisessa äänitteessä jaoteltuna runsaan, keskimääräisen ja vähäisen puheen määrän mukaan.

	Hiljaisuus (%) -runsaan puheen kohdat	Hiljaisuus (%) -keskimääräisen puheen kohdat	Hiljaisuus (%) -vähäisen puheen kohdat
Koko data ka	33.8	44.5	71.2
1	29	65	97
2	27.5	33	50.5
3	28	40	52
4	55	55,5	50.5
5	24.5	41,5	84
6	30.5	13,5	91
7	43	46	48
8	32.5	61,5	96.5

Taulukko 6. Melun määrän keskiarvo ja prosenttiosuudet jokaisessa äänitteessä jaoteltuna runsaan, keskimääräisen ja vähäisen puheen määrän mukaan.

	Melu (%) -runsaan puheen kohdat	Melu (%) -keskimääräisen puheen kohdat	Melu (%) -vähäisen puheen kohdat
Koko data ka	0.8	5.8	7.8
1	0	0.5	0
2	0	9.5	6.5
3	3	0.5	6.5
4	0	0	41
5	0	0	0
6	3.5	34	0.5
7	0	1	6.5
8	0	0.5	1.5

3.2. LENA-menetelmän validiteetti

Koko aineistoa tarkastellessa kaikki koodit olivat LENA-menetelmän ja arvioijien välillä yhteneviä noin 66-prosenttisesti. Kaikkien koodien arvion yhtenevyyden kohdalla oli suhteellisen vähän hajontaa, kun koodeja tarkasteltiin sen mukaan, oliko kohdassa runsaasti, keskimääräisesti vai vähäisesti puhetta. LENA-menetelmän ja arvioijien välillä yhtenevien koodien määrät vaihtelivat noin 62 ja 69 prosentin välillä. Koodien yhtenevyys oli matalin keskimääräisen puheen määrän koodeissa ja korkein vähäisen puheen määrän koodeissa.

Tutkimuskysymyksen kaksi mukaisesti ihmisääntä sisältäviä aikuisen naisen, aikuisen miehen ja lapsen koodeja tarkasteltiin vielä erikseen. Aikuisen naisen (FAN/FAF) -koodit olivat LENA-menetelmän ja arvioijien välillä yhteneviä koko aineistossa lähes 89-prosenttisesti. Puheen määrän mukaan jaoteltuna aikuisen naisen koodin arvion yhtenevyydessä oli myös suhteellisen vähän vaihtelua: koodin arvion yhtenevyys vaihteli noin 86 ja 90 prosentin välillä. Yhtenevyys oli matalin vähäisen puheen määrän koodeissa ja korkein runsaan puheen määrän koodeissa. Aikuisen miehen (MAN/MAF) -koodit olivat LENA-menetelmän ja arvioijien välillä yhteneviä koko aineistossa lähes 60-

prosenttisesti. Puheen määrän mukaan jaoteltuna koodin arvion yhtenevydessä löytyi hieman aikuisen naisen -koodia enemmän vaihtelevuutta koodin arvion yhtenevyyden vaihdellessa 55-65 prosentin välillä. Yhtenevyys oli matalin runsaan puheen määrän koodeissa ja korkein vähäisen puheen määrän koodeissa. Koko datassa lapsen (CHN/CHF) -koodit olivat LENA-menetelmän ja arvioijien välillä yhteneviä 61-prosenttisesti. Tämän koodin kohdalla arvioiden yhtenevydessä löytyi ihmisäänistä kaikkein eniten vaihtelua, kun sitä tutkittiin erikseen runsaan, keskimääräisen ja vähäisen puheen kohdissa. Koodin arvioiden yhtenevyys vaihteli näissä eri tilanteissa noin 55 ja 70 prosentin välillä. Yhtenevyys oli matalin runsaan puheen määrän koodeissa ja korkein vähäisen puheen määrän koodeissa.

Tutkimuskysymyksen kolme mukaisesti myös melun ja hiljaisuuden koodeja tarkasteltiin lähemmin. Melun koodit olivat LENA-menetelmän ja arvioijien välillä yhteneviä koko aineistossa 63-prosenttisesti. Puheen määrän mukaan tehdyissä jaotteluissa tämänkin koodin kohdalla löytyi jonkin verran vaihtelevuutta. Koodin arvioiden yhtenevyys vaihteli 53-71 prosentin välillä. Matalin yhtenevyys oli keskimääräisen puheen määrän koodeissa ja korkein runsaan puheen määrän koodeissa. Hiljaisuuden koodit olivat LENA-menetelmän ja arvioijien välillä yhteneviä koko aineistossa noin 53-prosenttisesti. Tämän koodin kohdalla ei ollut suurta vaihtelua, kun sitä tarkasteltiin erikseen runsaan, keskimääräisen ja vähäisen puheen kohdissa. Vaihteluväli tämän koodin arvioiden yhtenevydessä oli noin 49-57 prosenttia. Arvioiden yhtenevyys oli matalin keskimääräisen puheen määrän koodeissa ja korkein vähäisen puheen määrän koodeissa. Taulukossa 7 on nähtävillä eri tilanteissa koodien lukumäärät sekä kuinka monta kappaletta ja prosenttia koodeista oli yhteneviä LENA-menetelmän ja arvioijien välillä.

Liitteenä olevista taulukoista 1-8 on nähtävillä erikseen jokaisen äänitteen koodien määrä ja arvioiden yhtenevyys eri tilanteissa. Taulukoista nähdään, että koodien lukumäärissä oli jonkin verran vaihtelua kaiken kaikkiaan ja että lähemmin tarkastellut aikuisen naisen (FAN/FAF), aikuisen miehen (MAN/MAF), lapsen (CHN/CHF), melun (NON/NOF) ja hiljaisuuden (SIL) koodien määrät vaihtelivat myös äänitteiden välillä. Lisäksi erityisesti esille nousee lasten koodien pieni määrä, joka voi vaikuttaa tämän koodin validiteetin arvion luotettavuuteen. Myös aikuisen miehen koodeja oli keskimäärin huomattavasti vähemmän kuin aikuisen naisen koodeja.

Taulukko 7. Taulukossa 7 on esiteltyinä koko datasta koodien lukumäärä, kuinka monta kappaletta koodeista oli yhteneviä LENA:n ja arvioijien välillä ja kuinka monta prosenttia koodeista oli yhteneviä LENA:n ja arvioijien välillä eri tilanteissa.

Koko data	N	Yhteneviä (N)	Yhteneviä (%)
Kaikki koodit	6920	4557	65.9
Kaikki koodit_R	2650	1780	67.2
Kaikki koodit_K	2630	1645	62.5
Kaikki koodit_V	1640	1132	69.0
FAN/FAF	1813	1611	88.9
FAN/FAF_R	850	764	89.9
FAN/FAF_K	707	626	88.5
FAN/FAF_V	256	221	86.3
MAN/MAF	890	525	59.0
MAN/MAF_R	481	269	55.9
MAN/MAF_K	246	150	61.0
MAN/MAF_V	163	106	65.0
CHN/CHF	159	97	61.0
CHN/CHF_R	40	22	55.0
CHN/CHF_K	68	39	57.4
CHN/CHF_V	51	36	70.6
NON/NOF	1634	1030	63.0
NON/NOF_R	447	318	71.1
NON/NOF_K	659	349	53.0
NON/NOF_V	528	363	68.8
SIL	1081	575	53.2
SIL_R	419	229	54.7
SIL_K	384	189	49.2
SIL_V	278	157	56.5

Huom.

Kaikki koodit sisältävät kaikki LENA-menetelmässä käytössä olevat koodit

FAN/FAF=aikuinen nainen, MAN/MAF=aikuinen mies, CHN/CHF=lapsi, NON/NOF=melu, SIL=hiljaisuus

R, K ja V viittaavat puheen määrään: R=runsaan puheen kohta, K=keskimääräisen ja V=vähäisen.

3.3. Jatkovertailut koodien arvion yhtenevydessä puheen määrän mukaisissa jaotteluissa ja ihmisääntä sisältävien koodien välillä

Jatkovertailuina tarkasteltiin khiin neliö -testillä sitä, erosiko koodien yhtenevyys LENA-menetelmän ja arvioijien välillä vähäisen, keskimääräisen ja runsaan puheen kohdissa tilastollisesti merkitsevästi. Koko datassa kaikkien koodien kohdalla koodien yhtenevydet LENA-menetelmän ja arvioijien välillä vähäisen, keskimääräisen ja runsaan puheen kohdissa erosivat khiin neliö -testin mukaan toisistaan tilastollisesti merkitsevästi, $\chi^2(2)=21.87$, $p<.001$, *Cramerin V*=.04. Lähemmässä tarkastelussa huomattiin, että toisistaan erosivat runsaan ja keskimääräisen puheen kohtien arvioiden yhtenevydet, $\chi^2(1)=12.38$, $p<.001$, *Cramerin V*=.05, niin että runsaan puheen kohtien arvioiden yhtenevyys oli korkeampi. Myös keskimääräisen ja vähäisen puheen kohtien arvioiden yhtenevydet erosivat toisistaan, $\chi^2(1)=18.28$, $p<.001$, *Cramerin V*=.07. Vähäisen puheen kohtien koodien arvioiden yhtenevyys oli korkeampi. Runsaan ja vähäisen puheen kohdissa koodien arvioiden yhtenevydet eivät eronneet toisistaan tilastollisesti merkitsevästi, $\chi^2(1)=1.94$, $p=.222$, *Cramerin V*=.02.

Myös melun koodissa havaittiin jatkovertailuissa khiin neliö -testillä tilastollisesti merkitsevä ero koodin arvioiden yhtenevydessä LENA-menetelmän ja arvioijien välillä vähäisen, keskimääräisen ja runsaan puheen kohdissa, $\chi^2(2)=48.72$, $p<.001$, *Cramerin V*=.12. Parittaisissa tarkasteluissa toisistaan erosivat runsaan ja keskimääräisen puheen kohtien arvioiden yhtenevyys, $\chi^2(1)=36.78$, $p<.001$, *Cramerin V*=.18, joista runsaan puheen kohtien arvioiden yhtenevyys oli parempi. Keskimääräisen ja vähäisen puheen kohtien arvioiden yhtenevydet erosivat myös toisistaan tilastollisesti merkitsevästi, $\chi^2(1)=30.45$, $p<.001$, *Cramerin V*=.16. Näistä vähäisen puheen määrän kohtien arvioiden yhtenevyys oli korkeampi. Koodien arvioiden yhtenevyys runsaan ja vähäisen puheen kohdissa ei eronnut tilastollisesti merkitsevästi, $\chi^2(1)=.66$, $p=.418$, *Cramerin V*=.03.

Aikuisen naisen koodin arvioiden yhtenevyys LENA-menetelmän ja arvioijien välillä ei koko datassa eronnut puheen määrän mukaisissa jaotteluissa tilastollisesti merkitsevästi, $\chi^2(2)=2.63$, $p=.269$, *Cramerin V*=.03. Myöskään aikuisen miehen ($\chi^2(2)=4.73$, $p=.094$, *Cramerin V*=.05), lapsen ($\chi^2(2)=2.96$, $p=.228$, *Cramerin V*=.09) ja hiljaisuuden ($\chi^2(2)=3.40$, $p=.135$, *Cramerin V*=.04) koodien arvioiden yhtenevyys ei eronnut tilastollisesti merkitsevästi, kun arvioiden yhtenevyyttä tarkasteltiin vähäisen, keskimääräisen ja runsaan puheen kohdissa.

Jatkovertailuina tarkasteltiin myös, erosivatko aikuisen naisen, aikuisen miehen ja lapsen koodien arvioiden yhtenevyudet LENA-menetelmän ja arvioijien välillä toisistaan tilastollisesti merkitsevästi. Koodien arvioiden yhtenevyudessa oli tilastollisesti merkitsevä ero, $\chi^2(2)=339.02$, $p<.001$, *Cramerin V*=.24. Parittaisissa tarkasteluissa toisistaan erosivat tilastollisesti merkitsevästi aikuisen naisen koodin arvioiden yhtenevyys ja aikuisen miehen koodin arvioiden yhtenevyys, $\chi^2(1)=321.30$, $p<.001$, *Cramerin V*=.35. Naisen koodin arvioiden yhtenevyys oli korkeampi. Myös aikuisen naisen ja lapsen koodien arvioiden yhtenevyudet erosivat toisistaan tilastollisesti merkitsevästi, $\chi^2(1)=97.80$, $p<.001$, *Cramerin V*=.22. Aikuisen naisen koodin arvion yhtenevyys oli tässäkin korkeampi. Parittaisissa vertailuissa toisistaan eivät eronneet tilastollisesti merkitsevästi aikuisen miehen ja lapsen koodien arvioiden yhtenevyys, $\chi^2(1)=.23$, $p=.633$, *Cramerin V*=.01.

4. Pohdinta

Tämän pro gradu -tutkielman tarkoituksena oli selvittää LENA-menetelmän validiteettia ääniympäristön mittaamisessa vastasyntyneiden teho-osastolla suomenkielisten keskosten ja heidän vanhempiansa kanssa. Validiteettia arvioitiin vertaamalla LENA-menetelmän äänen lähteille antamien koodien ja ihmisarvioijien arvion yhtenevyyttä. Tässä pro gradu -tutkielmassa tutkittiin LENA-menetelmän sisältövaliditeettia vastasyntyneiden teho-osastolla. Sisältövaliditeetin käsite on erittäin lähellä reliabiliteetin käsitettä ja myös esimerkiksi Murphy ja Davidshofer (2005) nostavat esiin sisältövaliditeetin ja reliabiliteetin käsitteellisen vahvan samankaltaisuuden. Tämä samankaltaisuus on nähtävissä myös aiemmissa LENA-menetelmän luotettavuutta koskevilla tutkimuksissa, joissa puhutaan vaihdellen reliabiliteetista ja validiteetista (kts. esim. Elo, 2016; Schwarz ym., 2017; Ganek & Eriks-Brophy, 2018B).

Tämän tutkimuksen tulosten mukaan LENA-menetelmän kokonaisvaliditeetti vastasyntyneiden teho-osastoympäristössä oli kohtalaisen heikko, koska vain noin 66 prosenttia kaikista LENA-menetelmän antamista koodeista oli yhteneviä ihmisarvioijan antaman arvion kanssa äänen alkuperästä. Eri perheiden äänitteiden väliltä löytyi suhteellisen paljon vaihtelua sekä koodien määrässä että koodien arvion yhtenevyudessa LENA-menetelmän ja arvioijien välillä. Kaikki koodit huomioon otettaessa niiden

luotettavuudessa oli tilastollisesti merkitsevä ero, kun verrattiin runsaan ja keskimääräisen sekä keskimääräisen ja vähäisen puheen kohtia. Runsaan ja vähäisen puheen kohdissa koodien luotettavuudessa ei löytynyt tilastollisesti merkitsevää eroa. Vähäisen puheen kohdassa koodien luotettavuus oli korkein. Käytännössä erot runsaan ja keskimääräisen sekä keskimääräisen ja vähäisen puheen kohtien välillä eivät olleet kuitenkaan kovin suuria. Suurimmillaankin ero oli vain noin 7 prosenttiyksikköä ja voi olla, että koodien suuri määrä vaikutti siihen, että eroista tuli tilastollisesti merkitseviä. Myös näiden vertailujen välille lasketut todella pienet efektikoot viittaavat siihen, että ero ei käytännössä ole kovin merkittävä. Vaikka tilastollisia vertailuja ei tehty, niin näyttää siltä, että satunnaisesti mukaan valitut perheet edustavat taustatietojen osalta laajemmassa tutkimuksessa olevia perheitä hyvin. Eri äänityyppien prosenttiosuudet vaihtelivat tarkastelluissa runsaan, keskimääräisen ja vähäisen puheen määrän kohdissa, joten kohtien tarkastelu erikseen vaikuttaa myös perustellulta.

Tulosten mukaan aikuisen naisen koodien luotettavuus oli keskimäärin erittäin hyvä. Tulos vastasi aiempien tutkimusten tuloksia (esim. Elo 2016). Aikuisen naisen koodin luotettavuus vaihteli vain vähän prosentuaalisesti eikä tilastollisesti merkitsevästi sen mukaan, oliko kohdassa runsaasti, keskimääräisesti vai vähäisesti puhetta. Myös efektikoko tässä vertailussa tukee sitä, että koodin luotettavuudessa ei ole vaihtelua riippuen siitä, kuinka paljon siinä kohdassa on puhetta. Toisin sanoen tämän tutkimuksen tulos antaa tukea aikaisemmille tutkimuksille siitä, että LENA-menetelmä nimeää aikuisen naisen puheen luotettavasti myös suomen kielellä, riippumatta siitä, paljonko kyseisessä kohdassa on puhetta.

Tämän tutkimuksen tulosten perusteella miesten koodien luotettavuus oli suhteellisen heikko. Myös tämä tulos oli linjassa tuloksiin, joita Elo (2016) sai väitöskirjatutkimuksessaan. Vaikka aikuisen miehen koodien luotettavuudessa vaihtelua puheen määrän mukaisissa jaotteluissa löytyi prosentuaalisesti enemmän kuin esimerkiksi aikuisen naisen koodeissa, eivät nämä erot kuitenkaan olleet tilastollisesti merkitseviä ja niiden efektikoot jäivät myös pieniksi. Huomiota herättävää on se, että miesten koodien luotettavuus oli huomattavasti alhaisempi kuin naisten koodien. Miesten koodeja oli myös yli puolet vähemmän kuin naisten, mikä voi tietenkin osaltaan vaikuttaa sen luotettavuuteen. Toinen syy, joka saattaa vaikuttaa miesten koodien arvion luotettavuuteen, on LENA-menetelmän tapa kategorisoida äänet. LENA-menetelmä luokittelee äänet eri luokkiin vertaamalla ääntä menetelmässä olevaan malliin. Varsinkin

pienille lapsille kohdistetulle puheelle on tyypillistä korkeamman sävelkorkeuden käyttö puheessa (Snow, 1972), mikä voi aiheuttaa sen, että miesten puhe vastaa korkeudeltaan huomattavasti LENA-menetelmässä olevaa mallia ja sitä kautta vaikeuttaa LENA-menetelmän miesten äänien tunnistamista. Toisaalta voi olla, että suomenkielisten miesten äänenkorkeus eroaa ylipäättään englanninkielisten miesten äänenkorkeudesta ja suomenkielisten miesten tunnistus jää tämän vuoksi alhaisemmaksi. Ainakin naisten parissa äänenkorkeuden eroja suomenkielisten ja englanninkielisten naisten välillä on löydetty (Leino ym., 2008; Stoicheff, 1981, viitattu Guimarães ja Gouveia, 2007).

Lapsen äänen koodien luotettavuus jäi tässä tutkimuksessa alhaiseksi. Luotettavuus jäi myös huomattavasti alhaisemmaksi kuin aikaisemmissa tutkimuksissa vanhemmilla lapsilla (kts. esim. Elo, 2016), mikä viittaa siihen, että keskosten ääntely todella on poikkeavaa ja lisäksi sitä on suhteellisen vähän. Sekä ääntelyn määrällinen että laadullinen ero vanhempiin lapsiin verrattaessa johtuu varmasti suurelta osin keskosten keuhkojen kehittymättömyydestä (Parikka, 2017). Lapsen koodien luotettavuudessa vaihtelua puheen määrän mukaisissa jaotteluissa löytyi prosentuaalisesti enemmän kuin muissa ihmisäänen koodeissa. Nämä erot eivät kuitenkaan olleet tilastollisesti merkitseviä. Kuitenkin efektikoko lapsen koodien luotettavuuden vertailuissa saattaisi osoittaa siihen suuntaan, että tilanteiden välillä voisi olla eroja. Lapsen koodeja oli kokonaisuudessaan huomattavasti vähemmän kuin muita koodeja ja voikin olla, että tämä koodien pieni lukumäärä vaikutti myös siihen, että sen luotettavuuden vaihtelu on suurinta, vaikka tilanteiden väliset erot eivät koodien pienen määrän vuoksi nousseetkaan tilastollisesti merkitseviksi.

Jatkovertailuina tarkasteltiin vielä sitä, erosivatko ihmisääntä sisältävät koodit luotettavuudessaan myös tilastollisesti merkitsevästi. Jatkovertailut tukivat silmämääräistä arviota, eli naisten koodien luotettavuus erosi sekä miesten koodien että lapsen koodien luotettavuudesta. Miesten koodien ja lapsen koodien luotettavuudet eivät eronneet toisistaan tilastollisesti merkitsevästi. Miesten ja lapsen koodien luotettavuuden eroavaisuus naisten koodien luotettavuudesta selittyy varmasti ainakin osittain jo aiemmin esiin nostetuilla syillä. Miesten puheen äänenkorkeus saattaa erota LENA-menetelmän käyttämästä mallista ja keskosten keuhkojen kehittymättömyys saattaa vaikuttaa heidän ääntelyynsä niin voimakkaasti, että LENA-menetelmä ei pysty heidän ääntelyään luokittelemaan yhtä tarkasti kuin esimerkiksi Elon (2016) tutkimuksessa mukana olleiden jo kuuden kuukauden ikäisten lasten ääniä.

Melun ja varsinkin hiljaisuuden koodien luotettavuus oli tässä tutkimuksessa heikko. Hiljaisuuden ja melun koodien luotettavuutta ei ole aikaisemmissa tutkimuksissa tarkasteltu, mutta niitä tarkasteltiin tässä tutkimuksessa, koska niillä molemmilla on oma vaikutuksensa keskosen kehitykseen. Hiljaisuuden koodin luotettavuus oli tässä tutkimuksessa tarkastelluista koodeista kaikkein alhaisin ja LENA-menetelmä oli tunnistanut keskimäärin ainoastaan hieman yli puolet koodeista yhtenevästi ihmisarvioijan kanssa. Äänitteitä kuunnellessa syntynyt laadullinen arvio oli, että monet hiljaisuudeksi luokitellut kohdat sisälsivät todellisuudessa erilaisia häiriöääniä, joiden olisi oikeasti kuulunut saada koodikseen melu tai erilaisten laitteiden ääniä, mikä viittaa siihen, että LENA-menetelmä saattaa aliarvioida vastasyntyneiden teho-osastolla kuuluvaa häiriöäänien määrää. Hiljaisuuden koodin luotettavuudessa ei löytynyt tilastollisesti merkitsevää eroa vähäisen, keskimääräisen ja runsaan puheen kohdissa ja myös vertailun efektikoko jäi pieneksi, mikä viittaa siihen, että se toimii samankaltaisesti siitä huolimatta, paljonko kohdassa on puhetta. Hiljaisuuden määrä lisääntyi odotusten mukaan prosentuaalisesti, kun tarkasteltiin runsaan, keskimääräisen ja vähäisen puheen määrän kohtia. Hiljaisuuden koodien lukumäärä taas pieneni, mikä johtuu siitä, että yksittäisten koodien kesto kasvaa ja koodien lukumäärien laskemisessa ei koodin kesto otettu huomioon. Hiljaisuuden koodien huono luotettavuus viittaa siihen, että jos tutkimuskysymys kohdistuu hiljaisuuteen, tarvitaan sen arviointiin muita menetelmiä avuksi.

Myös melun koodin luotettavuus jäi tässä tutkimuksessa kohtalaisen heikoksi. Luotettavuus oli parhain runsaan puheen kohdissa ja heikoin keskimääräisen puheen kohdissa. Jatkovertailut paljastivat, että koodin luotettavuudessa oli tilastollisesti merkitsevä ero runsaan ja keskimääräisen sekä keskimääräisen ja vähäisen puheen kohdissa, mutta eroa ei ollut runsaan ja vähäisen puheen kohdissa. Erot olivat myös prosentuaalisesti suhteellisen isoja: noin 18 ja 13 prosenttiyksikköä ja näiden vertailujen efektikoot tukevat myös sitä, että tämän koodin luotettavuudessa oikeasti on vaihtelua riippuen siitä, paljonko kohdassa on puhetta. Melun koodin luotettavuudessa havaitulle näinkin suurelle vaihtelulle on hankala löytää syytä. Melun koodit sisälsivät pääasiassa suhteellisen hiljaisia ääniä. Jonkin verran ongelmia melun koodin nimeämisessä syntyi siitä, että LENA-menetelmä luokitteli esimerkiksi hengityskoneiden tasaisen kohinan meluksi eikä elektronisiksi laitteiksi. Toisaalta esimerkiksi hengityskoneista lähtevä tasainen kohina oli läsnä koko äänitteen ajan ja voi olla, että jako runsaan, keskimääräisen

ja vähäisen puheen määrän kohtiin ei ole tämän koodin kohdalla yhtä hyvin perusteltu kuin muiden koodien.

4.1. Tutkimustulosten merkitys

Tutkimustulokset osoittavat, että LENA-menetelmä on naisen äänen osalta luotettava apuväline lapsen ääniympäristön arvioinnissa, mutta laajemmin menetelmän keskosen ääniympäristöstä antamien tulosten tulkinnassa vastasyntyneiden teho-osastolla tulee käyttää harkintaa. LENA-menetelmä tarjoaa naisen puheen määrän tarkasteluun luotettavan työkalun. Nämä tutkimustulokset tukevat pääosin Elon (2016) tuloksia siitä, että suomenkielisessä väestössä LENA-menetelmä toimii samankaltaisesti kuin muillakin englannista poikkeavilla kielillä.

Tulosten perusteella voidaan todeta, että LENA-menetelmän luotettavuudessa vastasyntyneiden teho-osastolla on omat haasteensa, jotka tulee ottaa tutkimuksissa ja sen johtopäätöksissä huomioon. Keskosten ääntelyn tunnistamisen luotettavuus jäi alhaisemmaksi kuin esimerkiksi vanhemmilla lapsilla toteutetussa Elon (2016) tutkimuksessa, joten keskosten ääntelyn määrän arvioinnissa tulee käyttää harkintaa. Myös varsinkin hiljaisuuden koodin luotettavuuden heikkous tulee ottaa huomioon, kun tarkastellaan vastasyntyneiden teho-osaston ääniympäristön sisältämää äänen määrää tai jos tutkimuskysymys liittyy hiljaisuuteen.

Menetelmän vaivattomuus mahdollistaa keskosten ääniympäristön tarkastelun helposti ja voi nostaa esille mahdollisia ongelmia ääniympäristössä, vaikka sen antamiin tuloksiin tuleekin suhtautua varauksella. Menetelmä voi nostaa esille ongelmia vastasyntyneiden teho-osaston ääniympäristössä ja kun nämä ongelmat ovat tiedossa, voidaan niihin puuttua ja kehittää vastasyntyneiden teho-osaston ääniympäristöä siihen suuntaan, että se tukee parhaalla mahdollisella tavalla keskosten kokonaisvaltaista kehitystä eli esimerkiksi keskosten puheen kehitystä tai vanhempien kiintymyksen muodostumista. Tällainen automatisoitu menetelmä voi myös helpottaa perheiden osallistumista tutkimukseen, kun heidän ei tarvitse huolehtia siitä, että joku ihminen kuuntelee äänitteet, vaan heille voidaan kertoa, että äänitteitä analysoidaan ilman sisällön kuuntelua.

4.2. Tutkimuksen vahvuudet ja rajoitukset

Tällä tutkimuksella oli sekä vahvuuksia että rajoituksia. Tutkimuksen vahvuutena voidaan pitää sitä, että äänitteet valittiin sattumanvaraisesti laajemman tutkimuksen aineistosta. Kokonaisuudessaan koodien lukumäärä nousi korkeaksi, mikä myös auttaa tarkentamaan koodien luotettavuuden arviota. Yhtenä vahvuutena voidaan pitää myös sitä, että arvioijien välinen reliabiliteetti oli suhteellisen hyvä, mikä viittaa siihen, että molemmat arvioijat tekivät arvionsa samankaltaisesti. Vaikka arvioijien välinen reliabiliteetti nousi vähintään kohtalaiseksi, olisi sitä voinut mahdollisesti parantaa käyttämällä suurempaa pilottiaineistoa.

Yhtenä tämän tutkimuksen heikkoutena voidaan pitää sitä, että jos kaksoset olivat sijoitettuina samaan sänkyyn, emme voineet koodien yhtenevyyttä arvioidessa tunnistaa, kumpi kaksosista äänitelee, koska arvio tehtiin ainoastaan äänitteen perusteella ilman esimerkiksi videokuvaa. Tämä johti siihen, että molemmat lapsen koodien pääluokat arvioitiin yhteneviksi, jos kohdassa kuului lapsen ääntä. Kuitenkaan kaksosia ei haluttu rajata tämän tutkimuksen ulkopuolelle, koska pikkukeskosina syntyneistä noin 30 prosenttia on kaksosia (Parikka & Lehtonen, 2017).

Koodien yhtenevyyden arvioinnista jouduttiin myös poistamaan etäisyyden arvio kokonaan. Etäisyyden arvion poistamisella, eli koodien pääluokkien yhdistämisellä, pystyttiin koodien yhtenevyydestä tekemään luotettavampia arvioita, koska etäisyyden arvioiminen äänitteitä kuunneltaessa todettiin monissa tilanteissa mahdottomaksi, kun arvio täytyi tehdä ainoastaan äänityksestä ilman esimerkiksi videokuvaa. Toisaalta etäisyyden arviota hankaloitti myös ihmisen ja LENA-menetelmän eroavaisuus äänen tunnistamisessa ja nimeämisessä. LENA-menetelmä vertaa ääntä vain yhteen malliin, kun taas ihmiskuuntelijalla on käytössään useita erilaisia malleja esimerkiksi aikuisen miehen äänestä ja lisäksi ihminen voi käyttää puheessa olevia sisällöllisiä vihjeitä hyväkseen tunnistessaan puhujaa ja tämän etäisyyttä äänentallentimeen. Etäisyyden arvio koettiin myös tarpeettomammaksi tiedoksi, kuin se, kuinka luotettavasti LENA-menetelmä osasi tunnistaa äänen lähteen pääluokan.

4.3. Jatkotutkimukset ja sovellukset

Vaikka LENA-menetelmää on keskoslasten parissa käytetty tätä ennen jo kolmessa tutkimuksessa maailmalla (Caskey ym., 2011; Caskey ym., 2013; Caskey ym., 2014), ei sen luotettavuutta keskosten ääntelyn tunnistamisessa ole kuitenkaan kattavasti tutkittu ja jatkossa tulisikin tutkia vielä tarkemmin sitä, kuinka luotettavasti LENA-menetelmä tunnistaa keskosien ääntelyn määrän ja sen, minkälaista ääntelyä nauhalla kuuluu eli onko ääntely esimerkiksi itkua. Keskosien ääntelyn tunnistuksen luotettavuus jäi alhaisemmaksi kuin vanhemmilla lapsilla toteutetuissa tutkimuksissa ja tämä todennäköisesti vaikuttaa myös esimerkiksi LENA-menetelmän arvioihin protokeskusteluiden määrästä vanhempien ja keskosien välillä. Keskosten ääntelyn tunnistuksen alhaisen luotettavuuden vuoksi siihen liittyvien tulosten johtopäätösten kanssa tulee olla varovainen. Jos tutkimuskysymys kohdistuu jollain tavalla keskosien ääntelyyn ja sen määrään, olisi hyvä arvioida aineistoa myös manuaalisesti. Menetelmän luotettavuutta keskosten ääntelyn tunnistamiseen voisi myös kehittää, tarkentamalla menetelmän käytössä olevia malleja lapsen ääntelystä vastaamaan paremmin keskosille tyypillistä ääntelyä.

Myös aikuisen miehen koodin luotettavuus jäi huomattavasti aikuisen naisen koodin luotettavuutta heikommaksi ja tämän tuloksen jatkotutkimus olisi ehdottomasti tarpeen. Toisaalta laadullinen arvio äänitteitä kuunnellessa oli, että aikuisen miehen koodin luotettavuutta heikensi se, että aikuisen naisen eli usein äidin ääni joutui luokitelluksi tähän luokkaan. Tämä pro gradu -tutkielma on osa Turun yliopistollisen sairaalan ja Tallinnan lastensairaalan vastasyntyneiden teho-osastojen tutkimusta, jossa tarkastellaan esimerkiksi sitä, miten vanhempien läsnäolon määrä vaikuttaa keskosien ääniympäristöön ja miten nämä asiat ovat yhteydessä keskosien kielenkehitykseen myöhemmin. Tähän tarkoitukseen LENA-menetelmä tuntuu tämän tutkimuksen perusteella soveltuvan vähintäänkin kohtalaisesti, koska vanhempien puheen määrää ei ole tarkoitus eritellä erikseen äitiin ja isään. Yleisemminkin keskosten kielenkehityksen tutkimuksen näkökulmasta olisi tarpeen vielä tarkastella lähemmin, kuinka luotettavasti naisen ja miehen äänien koodit yhdessä toimivat. Kielenkehityksen kannalta tärkeää on keskosien kuuleman ja toisaalta suoraan hänelle osoitetun puheen määrä, ei se kumpi vanhemmista puhuu.

Lisäksi tämä tutkimus nostaa esiin vanhempien puheen määrän jakautumisen eroavaisuuden: äidit puhuivat huomattavasti enemmän kuin isät. Voi olla, että ilmiölle on täysin luonnollinen selitys, kuten se, että isät pystyvät olemaan vähemmän paikalla vastasyntyneiden teho-osastolla. Vaikka tämä kysymys ei liity menetelmän kehittämiseen, ilmiötä ja sen syitä tulisi kuitenkin jatkossa tutkia, jotta tarvittaessa voitaisiin miettiä, miksi isät puhuvat vähemmän tai miksi he mahdollisesti viettävät vähemmän aikaa osastolla.

4.4. Johtopäätökset

Tämän tutkimuksen perusteella LENA-menetelmä tunnistaa naisen äänen luotettavasti vastasyntyneiden teho-osaston ääniympäristössä ja suomen kielellä. Laajemman tutkimuksen, jonka osa tämä pro gradu -tutkielma on, näkökulmasta tämä tarkoittaa sitä, että LENA-menetelmä tunnistaa luotettavasti äidin äänen ja arvioi luotettavasti äidin puheen määrää. Muiden äänien tunnistuksessa LENA-menetelmän validiteetti jää tämän tutkimuksen perusteella kohtalaisen heikoksi. LENA-menetelmä voi kuitenkin antaa arvokasta tietoa siitä, millainen ääniympäristö vastasyntyneiden teho-osastolla olevilla keskoslapsilla ja heidän vanhemmillaan on, kunhan tulosten tulkinnassa muistetaan käyttää harkintaa. Tulokset viittaavat siihen, että LENA-menetelmällä voidaan luotettavasti arvioida keskosen kuulemaa naisen puheen määrää vastasyntyneiden teho-osastolla, mutta huonommin keskosten omaa ääntelyn määrää ja muita ääniä.

Lähteet

- Barras, C., Geoffrois, E., Wu, Z. ja Liberman, M. (1998). Transcriber: a Free Tool for Segmenting, Labeling, and Transcribing Speech. *First International Conference on Language Resources and Evaluation (LREC)*, 1373–1376.
- Binns, A., Hutchinson, L. ja Cardy, O. (2019). The speech-language pathologist's role in supporting the development of self-regulation: A review and tutorial. *Journal of Communication Disorders*, 78, 1–17.
<https://doi.org/10.1016/j.jcomdis.2018.12.005>
- Bland, M. ja Altman, D. (2011) Correlation in restricted ranges of data. *BMJ*, 342,
<https://doi.org/10.1136/bmj.d556>
- Canault, M., Le Normand, M., Foudil, S., Loundon, N. ja Thai-Van, H. (2016). Reliability of the Language ENvironment Analysis system (LENA™) in European French. *Behavior Research Methods*, 48(3), 1109–1124.
<https://doi.org/10.3758/s13428-015-0634-8>
- Caskey, M., Stephens, B., Tucker, R. ja Vohr, B. (2011). Importance of Parent Talk on the Development of Preterm Infant Vocalizations. *Pediatrics*, 128(5), 910–916.
<https://doi.org/10.1542/peds.2011-0609>
- Caskey, M., Stephens, B., Tucker, R. ja Vohr, B. (2014). Adult Talk in the NICU With Preterm Infants and Developmental Outcomes. *Pediatrics*, 133(3), e578–e584.
<https://doi.org/10.1542/peds.2013-0104>
- Caskey, M. ja Vohr, B. (2013). Assessing language and language environment of high-risk infants and children: a new approach. *Acta Paediatrica*, 102(5), 451–461.
<https://doi.org/10.1111/apa.12195>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Hillsdale.
- DeCasper, A. ja Fifer, W. (1980). Of human bonding: newborns prefer their mothers' voices. *Science*, 208. <https://doi.org/10.1126/science.7375928>
- Elo, H. (2016). *Acquiring language as a twin*. Tampere: University Press.

- Ganek, H. ja Eriks-Brophy, A. (2018A). Language ENvironment analysis (LENA) system investigation of day long recordings in children : A literature review. *Journal of Communication Disorders*, 72, 77–85.
<https://doi.org/10.1016/j.jcomdis.2017.12.005>
- Ganek, H. ja Eriks-Brophy, A. (2018B). A Concise Protocol for the Validation of Language ENvironment Analysis (LENA) Conversational Turn Counts in Vietnamese. *Communication Disorders Quarterly*, 39(2), 371–380.
<https://doi.org/10.1177/1525740117705094>
- Gervain, J. (2018). The role of prenatal experience in language development. *Current Opinion in Behavioral Sciences*, 21, 62–67.
<https://doi.org/10.1016/j.cobeha.2018.02.004>
- Gilkerson, J., Zhang, Y., Xu, D., Richards, J., Xu, X., Jiang, F., Harnsberger, J. ja Topping, K. (2015). Evaluating language environment analysis system performance for Chinese: a pilot study in Shanghai. *Journal of Speech, Language, and Hearing Research*, 58(2), 445–452. https://doi.org/10.1044/2015_JSLHR-L-14-0014
- Guimarães, I. ja Gouveia, S. (2007). Do Portuguese women have a lower pitch than other female subjects. Saatavissa:
http://www.sld.cu/galerias/pdf/sitios/rehabilitacion-logo/mujer_portuguesa.pdf
- Hoff, E. ja Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, 73(2), 418–433. <https://doi.org/10.1111/1467-8624.00415>
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. ja Lyons, T. (1991). Early vocabulary Growth: Relation to Language Input and Gender. *Developmental Psychology*, 27(1), 1236–1248. <https://doi.org/10.1037/0012-1649.27.2.236>
- Kisilevsky, B., Hains, S., Brown, C., Lee, C., Cowperthwaite, B., Stutzman, S., Swansburg, M., Lee, K., Xie, X., Huang, H., Ye, H., Zhang, K. ja Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, 32(1), 59–71. <https://doi.org/10.1016/j.infbeh.2008.10.002>
- Kuhn, P., Zores, C., Pebayle, T., Hoeft, A., Langlet, C., Escande, B., Astruc, D. ja Dufour, A. (2012). Infants born very preterm react to variations of the acoustic environment in their incubator from a minimum signal-to-noise ratio threshold of 5

- to 10 dBA. *Pediatric Research*, 71(4), 386-392. <https://doi.org/10.1038/pr.2011.76>
- Lehtonen, L. (2017). Keskosen kehitys ja sen tukeminen sairaalahoidon aikana. Teoksessa Stolt, S., Yliherva, A., Parikka, V., Haataja, L. ja Lehtonen, L. (toim.) *Keskosen hoito ja kehitys* (115-122). Helsinki: Duodecim.
- Leino, T., Laukkanen, A., Ilomäki, I. ja Mäki, E. (2008). Assessment of vocal capacity of Finnish university students. *Folia Phoniatrica et Logopaedica*, 60(4), 199–209. <https://doi.org/10.1159/000133651>
- LENA speaker labels in exported data. (n.d.) Saatavissa: <https://lenafoundation.screensteps.com/s/pro-support/m/18913/1/453824-lena-speaker-labels-in-exported-data>
- McHugh, M. (2012). Interrater reliability : the kappa statistic. *Biochemia Medica*, 22(3), 276–282. <https://doi.org/10.11613/BM.2012.031>
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertocini, J. ja Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2), 143–178. [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2)
- Moon, C. (2017). Prenatal experience with the maternal voice. Teoksessa Filippa, M., Kuhn, P. ja Westrup, B. (toim.) *Early vocal contact and preterm infant brain development* (25-37). Sveitsi: Springer.
- Moon, C., Cooper, R., ja Fifer, W. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495–500. [https://doi.org/10.1016/0163-6383\(93\)80007-U](https://doi.org/10.1016/0163-6383(93)80007-U)
- Murphy, K. ja Davidshofer, C. (2005). *Psychological testing: Principles and applications*. New Jersey: Pearson.
- Pae, S., Yoon, H., Seol, A., Gilkerson, J., Richards, J., Ma, L. ja Topping, K. (2016). Effects of feedback on parent-child language with infants and toddlers in Korea. *First Language*, 36(6), 549–569. <https://doi.org/10.1177/0142723716649273>
- Parikka, V. (2017). Keskosen keuhkot ja hengityksen tukeminen. Teoksessa Stolt, S., Yliherva, A., Parikka, V., Haataja, L. ja Lehtonen, L. (toim.) *Keskosen hoito ja kehitys* (42-54). Helsinki: Duodecim.

- Parikka, V. ja Lehtonen, L. (2017) Keskonen. Teoksessa Stolt, S., Yliherva, A., Parikka, V., Haataja, L. ja Lehtonen, L. (toim.) *Keskosen hoito ja kehitys (9-16)*. Helsinki: Duodecim.
- Parikka, V. ja Leino, M. (2017). Keskovauvan kotiutuminen. Teoksessa Stolt, S., Yliherva, A., Parikka, V., Haataja, L. ja Lehtonen, L. (toim.) *Keskosen hoito ja kehitys (131-142)*. Helsinki: Duodecim.
- Philbin, M. (2000). The Full-Term and Premature Newborn: The Influence of Auditory Experience on the Behavior of Preterm. *Journal of Perinatology*, 20, S76-S86. <https://doi.org/10.1038/sj.jp.7200453>
- Philbin, M. ja Evans, J. (2006). Standards for the acoustic environment of the newborn ICU. *Journal of Perinatology*, 26, S27–S30. <https://doi.org/10.1038/sj.jp.7211585>
- Pine, J. (1994). The language of primary caregivers. Teoksessa Gallaway, C. ja Richards, B. (toim.) *Input and interaction in language acquisition (15-37)*. Cambridge: University Press.
- Raiskila, S., Axelin, A., Toome, L., Caballero, S., Tandberg, B. S., Montirosso, R., Normann, E., Hallber, B., Westrup, B., Ewald, U. ja Lehtonen, L. (2017). Parents' presence and parent–infant closeness in 11 neonatal intensive care units in six European countries vary between and within the countries. *Acta Paediatrica, International Journal of Paediatrics*, 106(6), 878–888. <https://doi.org/10.1111/apa.13798>
- Schwarz, I., Botros, N., Lord, A., Marcusson, A., Tideliu, H. ja Marklund, E. (2017). The LENA™ system applied to Swedish : Reliability of the Adult Word Count estimate. *Interspeech*, 2088–2092. <https://doi.org/10.21437/Interspeech.2017-1287>
- Snow, C. (1972). Mothers' speech to children learning language. *Child Development*, 43, 549–565. <https://doi.org/10.2307/1127555>
- Stolt, S. ja Yliherva, A. (2017A). Varhainen kielen ja kommunikaation kehitys. Teoksessa Stolt, S., Yliherva, A., Parikka, V., Haataja, L. ja Lehtonen, L. (toim.) *Keskosen hoito ja kehitys (185-195)*. Helsinki: Duodecim.
- Stolt, S ja Yliherva, A. (2017B) Varhaisen kielellisen kehityksen tukeminen. Teoksessa Stolt, S., Yliherva, A., Parikka, V., Haataja, L. ja Lehtonen, L. (toim.) *Keskosen*

hoito ja kehitys (196-203). Helsinki: Duodecim.

Szymczak, S. ja Shellhaas, R. (2014). Impact of NICU design on environmental noise.

Journal of Neonatal Nursing, 20(2), 77–81.

<https://doi.org/10.1016/j.jnn.2013.07.003>

Weisleder, A. ja Fernald, A. (2013). Talking to Children Matters: Early Language

Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*,

24(11), 2143–2152. <https://doi.org/10.1177/0956797613488145>

Weizman, Z. ja Snow, C. (2001). Lexical input as related to children’s vocabulary acquisition: effects of sophisticated exposure and support for meaning.

Developmental Psychology. <https://doi.org/10.1037/0012-1649.37.2.265>

What is LENA. (n.d.) Saatavissa: [https://lenafoundation.screenstepslive.com/s/pro-](https://lenafoundation.screenstepslive.com/s/pro-support/m/18913/1/291987-what-is-lena)

[support/m/18913/1/291987-what-is-lena](https://lenafoundation.screenstepslive.com/s/pro-support/m/18913/1/291987-what-is-lena)

White, R., Smith, J. ja Shepley, M. (2006). Recommended standards for newborn ICU

design. *Journal of Perinatology*, 26(S1), S2–S18.

<https://doi.org/10.1038/sj.jp.7211587>

Liitteet

Taulukko 1. Äänitteen 1 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

1	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	689	373	54.1
Kaikki koodit_R	287	169	58.9
Kaikki koodit_K	341	178	52.2
Kaikki koodit_V	61	26	42.6
FAN/FAF	231	203	87.9
FAN/FAF_R	132	124	93.9
FAN/FAF_K	94	79	84.0
FAN/FAF_V	5	0	0
MAN/MAF	84	1	1.2
MAN/MAF_R	62	0	0
MAN/MAF_K	20	1	5.0
MAN/MAF_V	2	0	0
CHN/CHF	14	12	85.7
CHN/CHF_R	2	2	100.0
CHN/CHF_K	11	9	81.8
CHN/CHF_V	1	1	100.0
NON/NOF	90	50	55.6
NON/NOF_R	9	5	55.6
NON/NOF_K	68	36	52.9
NON/NOF_V	13	9	69.2
SIL	182	80	44.0
SIL_R	59	28	46.5
SIL_K	95	38	40.0
SIL_V	28	14	50.0

Taulukko 2. Äänitteen 2 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

2	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	1087	715	65.8
Kaikki koodit_R	366	264	72.1
Kaikki koodit_K	352	218	61.9
Kaikki koodit_V	369	233	63.1
FAN/FAF	335	262	78.2
FAN/FAF_R	112	96	85.7
FAN/FAF_K	116	82	70.7
FAN/FAF_V	107	84	78.5
MAN/MAF	181	132	72.5
MAN/MAF_R	114	88	77.2
MAN/MAF_K	41	23	56.1
MAN/MAF_V	26	21	80.8
CHN/CHF	19	6	31.6
CHN/CHF_R	2	0	0
CHN/CHF_K	4	2	50.0
CHN/CHF_V	13	4	30.8
NON/NOF	383	206	53.8
NON/NOF_R	108	62	57.4
NON/NOF_K	127	69	54.3
NON/NOF_V	148	75	50.7
SIL	6	4	66.7
SIL_R	-	-	-
SIL_K	6	4	66.7
SIL_V	-	-	-

Taulukko 3. Äänitteen 3 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

3	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	1083	831	76.7
Kaikki koodit_R	392	284	72.4
Kaikki koodit_K	373	299	79.9
Kaikki koodit_V	317	248	78.2
FAN/FAF	264	253	95.8
FAN/FAF_R	111	105	94.6
FAN/FAF_K	94	90	95.7
FAN/FAF_V	59	58	98.3
MAN/MAF	260	116	44.6
MAN/MAF_R	104	28	26.9
MAN/MAF_K	86	57	66.3
MAN/MAF_V	70	31	44.3
CHN/CHF	6	3	50.0
CHN/CHF_R	3	3	100.0
CHN/CHF_K	2	0	0
CHN/CHF_V	1	0	0
NON/NOF	116	107	92.2
NON/NOF_R	38	32	84.2
NON/NOF_K	27	26	96.3
NON/NOF_V	51	49	96.1
SIL	260	221	85.0
SIL_R	83	78	94.0
SIL_K	102	83	81.4
SIL_V	75	60	80.0

Taulukko 4. Äänitteen 4 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

4	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	746	530	71.0
Kaikki koodit_R	318	223	70.1
Kaikki koodit_K	312	212	67.9
Kaikki koodit_V	116	95	81.9
FAN/FAF	199	169	84.9
FAN/FAF_R	107	87	81.3
FAN/FAF_K	87	77	88.5
FAN/FAF_V	5	5	100.0
MAN/MAF	71	60	84.5
MAN/MAF_R	42	37	88.1
MAN/MAF_K	15	9	60.0
MAN/MAF_V	14	14	100.0
CHN/CHF	25	19	76.0
CHN/CHF_R	3	0	0
CHN/CHF_K	13	10	76.9
CHN/CHF_V	9	9	100.0
NON/NOF	295	167	56.6
NON/NOF_R	113	63	55.8
NON/NOF_K	114	53	46.5
NON/NOF_V	68	51	75.0
SIL	33	24	72.7
SIL_R	32	23	71.9
SIL_K	-	-	-
SIL_V	1	1	100.0

Taulukko 5. Äänitteen 5 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

5	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	800	492	61.5
Kaikki koodit_R	237	138	58.2
Kaikki koodit_K	341	214	62.8
Kaikki koodit_V	222	140	63.1
FAN/FAF	234	229	97.9
FAN/FAF_R	110	110	100.0
FAN/FAF_K	110	107	97.3
FAN/FAF_V	14	12	85.7
MAN/MAF	55	26	47.2
MAN/MAF_R	20	2	10.0
MAN/MAF_K	35	24	68.6
MAN/MAF_V	-	-	-
CHN/CHF	23	17	73.9
CHN/CHF_R	3	0	0
CHN/CHF_K	2	1	50.0
CHN/CHF_V	18	16	88.9
NON/NOF	224	112	50.0
NON/NOF_R	10	9	90.0
NON/NOF_K	130	32	24.6
NON/NOF_V	84	71	84.5
SIL	137	13	9.5
SIL_R	76	5	6.6
SIL_K	-	-	-
SIL_V	61	8	13.1

Taulukko 6. Äänitteen 6 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

6	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	727	522	71.8
Kaikki koodit_R	387	287	74.2
Kaikki koodit_K	186	129	69.4
Kaikki koodit_V	154	106	68.8
FAN/FAF	147	122	83.0
FAN/FAF_R	98	79	80.6
FAN/FAF_K	36	33	91.7
FAN/FAF_V	13	10	76.9
MAN/MAF	47	37	78.7
MAN/MAF_R	36	32	88.9
MAN/MAF_K	-	-	-
MAN/MAF_V	11	5	45.5
CHN/CHF	27	11	40.7
CHN/CHF_R	13	7	53.8
CHN/CHF_K	13	4	30.8
CHN/CHF_V	1	0	0
NON/NOF	217	186	85.7
NON/NOF_R	110	105	95.5
NON/NOF_K	78	55	70.5
NON/NOF_V	29	26	89.7
SIL	78	67	85.9
SIL_R	12	9	75.0
SIL_K	1	0	0
SIL_V	65	58	89.2

Taulukko 7. Äänitteen 7 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

7	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	1005	670	66.7
Kaikki koodit_R	280	171	61.1
Kaikki koodit_K	372	228	61.3
Kaikki koodit_V	353	271	76.8
FAN/FAF	227	224	98.7
FAN/FAF_R	65	65	100.0
FAN/FAF_K	109	107	98.2
FAN/FAF_V	53	52	98.1
MAN/MAF	120	102	85.0
MAN/MAF_R	50	48	96.0
MAN/MAF_K	31	20	64.5
MAN/MAF_V	39	34	87.2
CHN/CHF	27	15	55.6
CHN/CHF_R	-	-	-
CHN/CHF_K	20	10	50.0
CHN/CHF_V	7	5	71.4
NON/NOF	193	125	64.8
NON/NOF_R	4	2	50.0
NON/NOF_K	64	51	79.7
NON/NOF_V	125	72	57.6
SIL	169	87	44.4
SIL_R	81	48	59.3
SIL_K	64	24	37.5
SIL_V	24	15	62.5

Taulukko 8. Äänitteen 8 koodien lukumäärä ja arvion yhtenevyys LENA-menetelmän ja arvioijien välillä.

8	N	Yhtenevyys (N)	Yhtenevyys (%)
Kaikki koodit	783	424	54.2
Kaikki koodit_R	383	244	63.7
Kaikki koodit_K	352	167	47.4
Kaikki koodit_V	48	13	27.1
FAN/FAF	176	149	84.7
FAN/FAF_R	115	98	85.2
FAN/FAF_K	61	51	83.6
FAN/FAF_V	-	-	-
MAN/MAF	72	51	70.8
MAN/MAF_R	53	34	63.2
MAN/MAF_K	18	16	88.9
MAN/MAF_V	1	1	100.0
CHN/CHF	18	14	77.8
CHN/CHF_R	14	10	71.4
CHN/CHF_K	3	3	100.0
CHN/CHF_V	1	1	100.0
NON/NOF	116	77	66.4
NON/NOF_R	55	40	72.7
NON/NOF_K	51	27	52.9
NON/NOF_V	10	10	100.0
SIL	216	79	36.6
SIL_R	76	38	50.0
SIL_K	116	40	34.5
SIL_V	24	1	4.2