

Spontaanipuheen fonologian numeeriset menetelmät 1–5- vuotiaiden tutkimuksissa

Sofia Grundström, Lotta Hurskainen

Kandidaatintutkielma

Ohjaajat: Elina Mainela-Arnold, Minna Vanhala

Yhteiskuntatieteellinen tiedekunta, psykologian ja logopedian laitos, logopedia

Toukokuu 2025

Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisu alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.

TURUN YLIOPISTO

Psykologian ja logopedian laitos/Yhteiskuntatieteellinen tiedekunta

GRUNDSTRÖM, SOFIA:

HURSKAINEN, LOTTA:

Spontaanipuheen fonologian

numeeriset menetelmät 1–5-vuotiaiden
tutkimuksessa

Kandidaatintutkielma, 37 s., 2 liitettä

Logopedia

Toukokuu 2025

Tämän tutkimuksen tavoitteena oli tarkastella, millaisia numeerisia menetelmiä on käytetty 1–5-vuotiaiden lasten spontaanin puheen tutkimuksissa sekä kuinka luotettaviksi nämä menetelmät on arvioitu. Spontaanipuheen havainnointi antaa keskeistä tietoa siitä, miten lapsi hyödyntää fonologisia taitojaan luonnollisessa kommunikaatiossaan, ja sen numeerinen analyysi mahdollistaa mitattavan ja objektiivisen taitojen tarkastelun.

Toteutimme tutkimuksen systemaattisena kirjallisuuskatsauksena, ja tutkimusaineisto haettiin seuraavista tietokannoista: PsychINFO- ja PubMed. Aineistoon valikoitiin vertaisarvioituja tutkimuksia, jossa hyödynnettiin numeerisia menetelmiä osana spontaanipuheen fonologian tutkimusta. Lopulliseen analyysiin valikoitui yhteensä 10 artikkelia.

Tulosten perusteella yleisimmin käytettyjä numeerisia menetelmiä olivat konsonanttien oikeellisuusprosentti (PCC) ja ilmauksien fonologinen keskipituus (pMLU). Molemmissa menetelmissä todettiin korkeita luotettavuusarvoja. Menetelmien käyttöön liittyy kuitenkin eroavaisuuksia reliabiliteetti-arviot tehneissä arvioijissa, arvioitujen näytteiden määrässä sekä litteroinnin tasossa. Nämä kaikki voivat vaikuttaa reliabiliteetin arviointiin ja tulosten toistettavuuteen.

Katsauksen perusteella voidaan todeta, että PCC- ja pMLU-menetelmät ovat käyttökelpoisia ja luotettavia välineitä lasten spontaanipuheen fonologian arviointiin. Rajoitteena katsauksessa olivat pieni otoskoko ja vaihtelevat tutkimusasetelmat, jotka vaikeuttivat tulosten arviointia sekä yleistettävyyttä. Tulevaisuuden tutkimuksessa huomiota tulisi kiinnittää reliabiliteetti-arvojen systemaattiseen raportointiin. Jatkossa spontaanin puheen analysointia voidaan mahdollisesti hyödyntää nykyistä enemmän ja systemaattisemmin fonologisten taitojen kliinisessä arvioinnissa.

Asiasanat: fonologinen kehitys, numeeriset menetelmät, spontaanipuhe, taaperoikäiset, leikki-ikäiset

Sisällys

1 Johdanto	3
1.1 Varhainen fonologinen kehitys	4
1.2 Fonologian tutkimuksen menetelmät.....	6
1.2.1 Laadulliset menetelmät fonologian tutkimuksissa	7
1.2.2 Numeeriset menetelmät fonologian tutkimuksissa	7
1.3 Reliabiliteetti.....	9
1.4 Tutkimuskysymykset.....	10
2 Menetelmät.....	11
2.1 Aineiston hankinta.....	11
2.2 Aineiston luokittelu	13
3 Tulokset.....	14
3.1 Tutkimusaineisto.....	14
3.2 Spontaanipuheen fonologian numeeriset menetelmät.....	14
3.3 Numeeristen menetelmien reliabiliteetti.....	19
3.3.1 Arvioidut näytteet.....	19
3.3.2 Arviointitapa	20
3.3.3 Reliabiliteettiarvot	20
4 Pohdinta.....	25
4.1 Tulosten tulkinta	25
4.1.1 Arvioijat.....	25
4.1.2 Arvioidujen näytteiden määrä	26
4.1.3 Litteroinnin taso - foneeminen vs. foneettinen	26
4.2 Vahvuudet ja rajoitukset.....	27
4.3 Johtopäätökset ja jatkotutkimusehdotukset.....	28
Lähteet.....	29

1 Johdanto

Valmiudet äänteelliseen kehitykseen rakentuvat useista toisiaan tukevista taidoista. Kuulonvaraisen erottelukyvyn ja motorisen valmiuden lisäksi lapselta edellytetään myös monipuolisia puheen tuottamiseen, havainnointiin ja käsittelyyn liittyviä taitoja sekä kiinnostusta vuorovaikutukseen (Kunnari & Savinainen-Makkonen, 2009). Lapsen fonologis-foneettista kehitystä voidaan tarkastella monesta eri näkökulmasta, kuten vertaamalla lapsen puhetta aikuisen puheeseen tai tarkastella niitä itsenäisenä järjestelmänä sekä rinnakkain muiden kielellisen kehityksen osa-alueiden kanssa (Kunnari ym., 2006). Puheen ja kielen häiriöt usein liittyvät äänteiden tuottoon ja äännejärjestelmän poikkeavuuksiin, minkä takia fonologian tutkimus on keskeinen osa logopedian alaa (Owens ym., 2015)

Fonologia käsitteenä viittaa kielen äänteiden ja niiden järjestelmän sääntöjen tutkimukseen, ja tämä liittyy olennaisesti siihen, miten lapset oppivat erottamaan sekä tuottamaan kielellisiä yksiköitä (Bernthal ym., 2022; Vihman, 2014). Lapsen kyky havaita ja tuottaa äänteitä oikein on olennainen edellytys puheen ymmärrettävyydelle sekä tehokkaalle viestinnälle (Rvachew & Brosseau-Lapré, 2018). Fonologian tutkimus mahdollistaa myös kielenkehityksen häiriöiden, kuten fonologisten viivästymien tai puheen tuoton häiriöiden, varhaisen tunnistamisen (Bishop & Leonard, 2014). Tämän avulla voidaan suunnitella yksilöllisiä tukitoimia, jotka edistävät lapsen kielellistä ja sosiaalista osallistumista.

1.1 Varhainen fonologinen kehitys

Varhainen fonologinen kehitys on keskeinen osa lapsen kielellistä kehitystä. Fonologinen kehitys alkaa jo varhaislapsuudessa ja liittyy tiiviisti muihin kielellisiin osa-alueisiin, kuten sanavaraston karttumiseen ja kieliopillisten rakenteiden omaksumiseen (Kunnari ym., 2006; Storkel & Morrisette, 2002; Torvelainen, 2007; Turunen, 2003; Yang, 2024). Fonologinen kehitys etenee vaiheittain ja tutkijat ovat jaotelleet fonologian kehityksen vaiheita eri tavalla. Esimerkiksi Ingram (1976) on jakanut kehityksen viiteen vaiheeseen: esikielellinen ääntely, yksisanavaihe, systemaattinen vaihe, inventaarin viimeistely ja morfofonologinen kehitys.

Ingram (1976) mukaan kehitys alkaa esileksikaalisella kaudella, jossa lapsi harjoittelee äänteiden tuottamista ilman merkityksellisiä sanoja, esimerkiksi jokeltelun avulla. Tätä seuraa ensisanojen kausi, jolloin lapsi alkaa omaksua ja tuottaa ensimmäisiä sanoja. Samalla äänteellinen kehitys jatkuu, vaikka lapsen tuotokset ovat edelleen yksinkertaisia. Äänteiden hallinta alkaa vähitellen vakiintua

systemaattisessa vaiheessa, jossa lapselle kehittyy johdonmukaisia sääntöjä ja käyttää näitä aktiivisesti, vaikka puheessa esiintyykin esimerkiksi inventaaripuutteesta sekä äänteiden yhdistelyn rajoitteista aiheutuvia fonologisia prosesseja (Ingram, 1976; Kunnari & Savinainen-Makkonen, 2012). Fonologiset prosessit ovat systemaattisia aikuiskielen yksikertaistuksia, kuten äänteiden korvaamista helpommilla äänneillä tai äänteiden poisjättämistä, jotka helpottavat puheen tuottamista äännejärjestelmän vielä kehittyessä (Kunnari & Savinainen-Makkonen, 2012; Yang, 2024). Näiden prosessien pitäisi tyypillisesti hävitä neljänteen ikävuoteen mennessä. Inventaarin viimeistelyvaiheessa lapsen äännejärjestelmä laajenee ja lähenee äänneellisesti aikuisen puhetta (Kunnari & Savinainen-Makkonen, 2012). Tässä vaiheessa jo useimmat äänneet tuotetaan oikein. Äänneellisen kehityksen päättää morfofonologinen vaihe, jossa lapsi oppii hallitsemaan sanojen taivutuksissa tapahtuvat äännevaihtelut (Ingram, 1976; Kunnari & Savinainen-Makkonen, 2012).

Fonologinen kehitys on tiiviisti yhteydessä muihin kielen osa-alueisiin. Esimerkiksi sanavaraston laajentuminen edellyttää kykyä erottaa ja tuottaa uusia äänneitä sekä äänneyhdistelmiä (Edwards ym., 2004). Myös kieliopillisten rakenteiden omaksuminen vaatii fonologisten sääntöjen ymmärtämistä ja soveltamista. Storkel ja Morrisette (2002) havaitsivat tutkimuksessaan, että fonologinen kehitys ja sanaston kasvu tukevat toisiaan: kun lapsi oppii enemmän sanoja, hänen kykynsä käsitellä ja tuottaa äänneitä paranee, mikä helpottaa uusien sanojen oppimista.

Fonologisen kehityksen häiriöt voivat ilmetä vaikeuksina äänteiden tuottamisessa, erottelussa tai järjestelmässä (Ikonen & Ström, 2009). Tällaiset haasteet voivat heijastua puheen epäselvyytenä ja vaikuttaa negatiivisesti esimerkiksi sanaston ja kieliopin kehittymiseen. Fonologisten häiriöiden on todettu myös vaikuttavan muuhun oppimiseen, kuten lukemis- ja kirjoitustaitojen omaksumiseen sekä kehittymiseen (Ikonen & Ström, 2009; Turunen, 2003).

Fonologisen kehityksen arviointi ja tukeminen ovat tärkeitä erityisesti varhaislapsuudessa, jolloin kielen kehitys on nopeaa ja herkkyyskaudet oppimiselle aktiiviset. Varhainen puuttuminen fonologisiin vaikeuksiin voi ehkäistä myöhempiä kielellisiä ja oppimiseen liittyviä haasteita (Torvelainen, 2007; Yang, 2024). Spontaanipuheen analyysi tarjoaa puolestaan arvokasta tietoa lapsen fonologisesta kehityksestä ja mahdollisista häiriöistä, ja sen avulla voidaan suunnitella yksilöllisiä tukitoimia kielen kehityksen edistämiseksi.

1.2 Fonologian tutkimuksen menetelmät

Lasten spontaaneja fonologis-foneettisia tuotoksia ja niiden kehitystä voidaan tarkastella kvalitatiivisin eli laadullisin sekä kvantitatiivisin eli numeerisin menetelmin. Pääsääntöisesti kuvaukset puhetuotoksista tehdään vertaamalla lapsen tuotoksia aikuisen puheeseen (eng. *relational analysis*), itsenäisenä kuvausmallina, jossa lapsen puhetuotosta kuvataan itsenäisesti ilman vertailuja aikuisen malliin (eng. *independent analysis*) tai molempia malleja yhdistelemällä (Kunnari ym., 2006; Stoel-Gammon & Dunn, 1985). Aikuisen tuotokseen vertaavan mallien analyysissä keskitytään muun muassa oikein tuotettujen konsonanttien ja vokaalien prosentuaalisiin osuuksiin, fonologisiin prosesseihin sekä lapsen foneemiseen inventaariin (Kunnari ym., 2006; Stokes ym., 2005). Itsenäisessä mallissa lasten puhetuotosten kokonaisuutta kuvataan puolestaan kehittyvänä ja monipuolistuvana järjestelmänä (Kunnari ym., 2006). Tämän mallin analyysissä tarkastellaan muun muassa äänneitä, jotka kuuluvat lapsen foneettiseen inventaariin tai niitä tuotoksia, jotka lapsi hallitsee, vaikka äänneet eivät vastaisikaan täysin aikuisen kieltä (Stokes ym., 2005)

Fonologian arvioinnissa on myös oleellista tarkastella yksittäisten sanojen tai lauseiden tuoton fonologian lisäksi myös spontaanipuheen fonologiaa. Spontaanipuheen havainnointi antaa merkityksellistä tietoa siitä, miten lapsi hyödyntää fonologisia taitojaan luonnollisessa kommunikaatiossa. Spontaanissa puheessa voi ilmetä myös sellaista vaihtelua ja virheitä, joita strukturoidut testit eivät välttämättä paljasta. Esimerkiksi Cray (1983) tarkasteli tutkimuksessaan, miten spontaanin puheen näytteen koko vaikuttaa fonologisten prosessien analyysiin, ja havaitsi, että laajemmat puhenäytteet mahdollistivat kattavamman ja luotettavamman analyysin lapsen fonologisista prosesseista. Näin ollen voidaan olettaa, että spontaaneista puhenäytteistä saadaan realistisempi kuva lapsen puheesta sekä puhetuotosten fonologisesta rakenteesta. Spontaanipuheen analysointia kliinisessä työssä toteutetaan usein laadullisin havainnoin, eikä systemaattisia ja numeerisia menetelmiä arviointiin juurikaan ole tai ne ovat harvemmin käytössä. Spontaanipuheen sisällyttäminen osaksi kliinistä arviointia ja sen tarkastelu numeerisin menetelmin voi tarjota olennaista lisätietoa esimerkiksi kuntoutuksen suunnittelun tueksi.

1.2.1 Laadulliset menetelmät fonologian tutkimuksissa

Perinteisesti logopedian alalla fonologian ilmiöitä on analysoitu laadullisin menetelmin. Laadullinen tutkimus fonologian alueella keskittyy puheen ja äänteiden tuottamisen ilmiöiden merkitykselliseen tulkintaan ja kontekstualisointiin (Ball ym., 2014). Ne eivät itsessään mittaa tutkittavaa ilmiötä, vaan ennemminkin pyrkivät ymmärtämään itse ilmiön luonnetta, kuten millaisia äänteitä, äänneyhdistelmiä ja virheitä puhetuotoksessa havaitaan sekä mitä nämä kertovat fonologisesta järjestelmästä tai kehityksestä. Kliinisessä työssä arvioidaan ja havainnoidaan laadullisesti fonologisia taitoja esimerkiksi strukturoiduilla testeillä sekä erilaisilla toistotehtävillä. Tämänkaltaisista menetelmistä keskeinen esimerkki on 2–6-vuotiaiden lasten fonologisen kehityksen arviointiin ja seurantaan tarkoitettu Fonologiatesti (Kunnari ym., 2012). Testissä lapsi nimeää erilaisia kuvia, joiden äänteiden tuottamisen pohjalta tarkastellaan lapsen fonotaktisia ja paradigmaattisia taitoja laadullisesti tyypillisten ja epätyypillisten kehityspiirteiden kautta. Fonotaktiset taidot tarkoittavat kykyä tunnistaa ja tuottaa kielen äänne- ja tavarakenteita koskevia sääntöjä, kun taas paradigmaattiset taidot liittyvät yksittäisten puheäänteiden hallintaan kielessä. Laadulliset menetelmät tarjoavat arvokasta tietoa lapsen fonologisesta kehityksestä, sillä niiden avulla voidaan tarkastella puheen äänteellisen kehityksen piirteitä yksityiskohtaisesti ja määritellyssä kontekstissa.

1.2.2 Numeeriset menetelmät fonologian tutkimuksissa

Laadullisten menetelmien rinnalle on kuitenkin noussut korostunut tarve kehittää ja hyödyntää systemaattisia sekä toistettavia mittareita, erityisesti kliinisessä työssä, jossa arviointien on oltava luotettavia, vertailukelpoisia ja ajallisesti mitattavissa (Peña-Brooks & Hegde, 2023; Ramig & Fox, 2012). Numeeriset menetelmät tarjoavat tähän ratkaisuja, sillä ne mahdollistavat lapsen fonologisen kehityksen määrällisen ja objektiivisen arvioinnin. Näitä menetelmiä voidaan hyödyntää esimerkiksi puheen tarkkuuden, ymmärrettävyyden tai kehitysvaiheen mittaamiseen (Ingram, 2002). Numeeriset mittarit ovat erityisen hyödyllisiä spontaanipuheen analyysissä, jossa lapsen kielellistä ilmaisuun tarkastellaan luonnollisessa tilanteessa (Watson & Terrell, 2012).

Numeeriset menetelmät perustuvat yksinkertaisuudessaan tuotetun puheen yksiköiden mittaamiseen ja siitä saatavan numeerisen aineiston tarkasteluun. Ingram (2002) ehdotti, että kliininen tutkimus ja arviointi keskittyisivät yksittäisten äänteiden sijaan lapsen koko-sanan tuotoksiin (eng. *whole-word productions*). Hän esitti, että lapsen puheen arviointia voisi tehostaa mittaamalla koko-sanan oikeellisuutta, moninaisuutta, ymmärrettävyyttä sekä vaihtelevuutta. Tällä lähestymistavalla pyritään

saamaan kokonaisvaltaisempi kuva lapsen fonologisista taidoista ja vertaamaan lapsen kokosanan tuotosta aikuisen puheeseen (Ingram & Ingram, 2001). Ingram (2002) esitti neljä samankaltaista, mutta toisistaan erillistä mittaria analysoimaan numeerisesti lasten kokosanan tuottoa. Nämä ovat ilmauksien fonologinen keskipituus eli pMLU (eng. *Phonological Mean Length of Utterance*) sekä siihen liittyvät kolme dimensiota: ilmaisun suhteellinen tarkkuus eli PWP (eng. *Proportion of Whole-word Proximity*), täysin oikein tuotettujen sanojen osuus eli PWC (eng. *Proportion of Whole-word Correctness*) sekä sanojen vaihtelun suhteellinen osuus eli PWV (eng. *Proportion of Whole-word Variability*).

Keskeisin mittareista on pMLU. Sitä hyödynnetään lapsen äänteellisen kehityksen vaiheen määrittämiseen, sen arviointiin, kuinka lähellä lapsen tuottamat sanat ovat aikuisen tuotosta sekä sanojen monimutkaisuuden mittaamiseen (Ingram, 2002). pMLU mittaa MLU:n (eng. *Mean Length of Utterance*) tavoin lapsen tuotoksen pituutta, mutta eroaa ottamalla kantaa myös oikein tuotettujen konsonanttien määrään (Tomas & Dorofeeva, 2019). Menetelmää varten on kehitetty säännöt, jotka käsittelevät otoskoon määrittelyä, sanavalinnan periaatteita sekä pisteyttämistä konsonanttien ja vokaalisen esiintymisestä (Ingram, 2002). Näiden sääntöjen tavoitteena on varmistaa menetelmän korkea validiteetti ja reliabiliteetti. pMLU:n laskeminen on merkitty liitteeseen 1.

Loput kolme mittaria liittyvät pMLU-pisteytyksiin ja voidaan katsoa pMLU:n alamittareiksi (Ingram, 2002). PWP on luotu eräänlaiseksi epäsuoraksi mittariksi sanan ymmärrettävyydelle. Siinä tarkastellaan, kuinka hyvin lapsen yrittämä tuotos täsmää aikuisen vastineeseen. PWC puolestaan kertoo, kuinka suuri osuus lapsen tuottamista sanoista on tuotettu oikein tämän koko puhenäytteestä. PWV perustuu taas lapsen koko-sanatuotoksen yhteneväisyyteen ja siihen, että lapset eivät tuota sanoja aina johdonmukaisesti samalla tavalla eri kerroilla. Tämän yhteneväisyyden vaihtelun mittaaminen ei ole kuitenkaan yksinkertaista, eikä eroja saada välttämättä tyypillisten artikulaatiotestien yhteydessä selville. (Ingram, 2002; Peña-Brooks & Hegde, 2023)

Fonologisia tuotoksia voidaan tarkastella myös Shribergin ja Kwiatkowskiin (1982) mukaan konsonanttien oikeellisuusprosentti (eng. *Percentage of correct consonants*) indeksin eli PCC:n avulla. PCC ilmoittaa keskustelupuheessa oikeintuotettujen konsonanttien prosenttiosuuden (Shriberg ym., 1997). Menetelmä soveltuu näin ollen erityisesti spontaanipuheen ja puheen ymmärrettävyyden tutkimukseen. PCC lasketaan jakamalla oikein tuotetut konsonantit oikein- sekä

väärintuotettujen konsonanttien summalla eli toisin sanoen kaikilla tuotetuilla konsonanteilla. PCC:n laskukaava on liitteessä 1.

Tällä hetkellä käytössä on jokseenkin vähän numeerisia menetelmiä, mukaan lukien aiemmin mainitut menetelmät, analysoimaan spontaanipuhetta kliinisessä työssä. Perinteisesti käytetyillä fonologian arviointimenetelmillä, kuten fonologiatestillä, ei välttämättä saada riittävää kokonaiskuvaa lapsen äänteellisestä kehityksestä, jonka vuoksi arvioon on oleellista sisällyttää myös lapsen spontaanipuheen tarkastelu. Tämän systemaattisen katsauksen tarkoituksena onkin kartoittaa, millaisia analyysimenetelmiä spontaanipuheen fonologian tutkimuksessa on käytetty 1–5-vuotiaiden lasten tutkimuksessa.

1.3 Reliabiliteetti

Reliabiliteetilla tarkoitetaan mittarin kykyä tuottaa johdonmukaisia ja toistettavia tuloksia saman ilmiön mittaamisessa (Heikkilä, 2008). Fonologian tutkimuksessa tämä tarkoittaa esimerkiksi sitä, että eri arvioijat päätyvät samoihin tuloksiin (interrater-reliabiliteetti) tai että sama arvioija tuottaa saman tuloksen eri aikoina (intrarater-reliabiliteetti). Numeeristen menetelmien, kuten pMLU:n tai PCC:n käyttö edellyttää selkeitä pisteytysohjeita ja luotettavaa datan käsittelyä, jotta tulokset olisivat vertailukelpoisia ja toistettavia (Ingram, 2002; Shriberg ym., 1997).

Ingram (2002) korostaa, että pMLU:n käyttöön liittyy tarkasti määritellyt säännöt sanavalintaan, pisteytykseen ja otoskoon juuri reliabiliteetin varmistamiseksi. Vastaavasti Shriberg ja Kwiatkowski (1982) ovat osoittaneet PCC-mittarin olevan luotettava työkalu, kun arvioijat koulutetaan yhtenevästi. Myöhemmin Shriberg kolleegoineen (1997) ovat laajentaneet PCC-mittarin soveltuvuutta ja arvioineet sen reliabiliteettia eri populaatioissa raportoiden hyviä interrater- ja intrarater-arvoja. Reliabiliteettia parantaa tarkat ja yhtenäiset ohjeet, aineiston analysointi uudelleen joko saman tai eri arvioijan toimesta sekä arvioijien asianmukainen koulutus.

Reliabiliteetin arviointi on tärkeä osa tutkimusten laatua. Systemaattisessa katsauksessa kiinnitämme huomiota siihen, onko tutkimuksissa raportoitu käytettyjen analyysien reliabiliteettia ja millä tavoin sitä on mitattu.

1.4 Tutkimuskysymykset

Tämän tutkimuksen tavoitteena on tarkastella, millaisia numeerisia menetelmiä on käytetty 1–5-vuotiaiden spontaanipuheen fonologian tutkimuksessa. Tarkastelemme myös, kuinka luotettavia tuloksia numeerisilla menetelmillä on saatu. Toteutamme tutkimuksen systemaattisena kirjallisuuskatsauksena, jonka tavoitteena on vastata seuraaviin kysymyksiin:

1. Mitä spontaanipuheen numeerisia menetelmiä on 1–5-vuotiaiden fonologian tutkimuksessa käytetty?
2. Kuinka luotettavia tuloksia numeerisilla menetelmillä on saatu?

2 Menetelmät

2.1 Aineiston hankinta

Toteutimme tutkimuksen systemaattisena kirjallisuuskatsauksena. Teimme tutkimusaineiston haun 18.2.2025 kahdesta tietokannasta, jotka olivat PubMed ja PsycINFO. Valitsimme hakuun keskeisimmät logopedian ja psykologian alan tietokannat. Hakulauseke tietokannoissa oli seuraava:

("phonological dev" OR "phonology" OR "phonetic analysis" OR "phonological complexity" OR "speech analysis" OR "speech processing") AND ("child*" OR "preschool*" OR "toddler*" OR "infant*") AND ("spontaneous speech" OR "spontaneous speaking" OR "spontaneous talk*" OR "spontaneous language" OR "spontaneous speech samples" OR "natural* speech" OR "unscripted speech").*

Liitteessä 2 on tarkemmat tietokantakohtaiset hakulausekkeet.

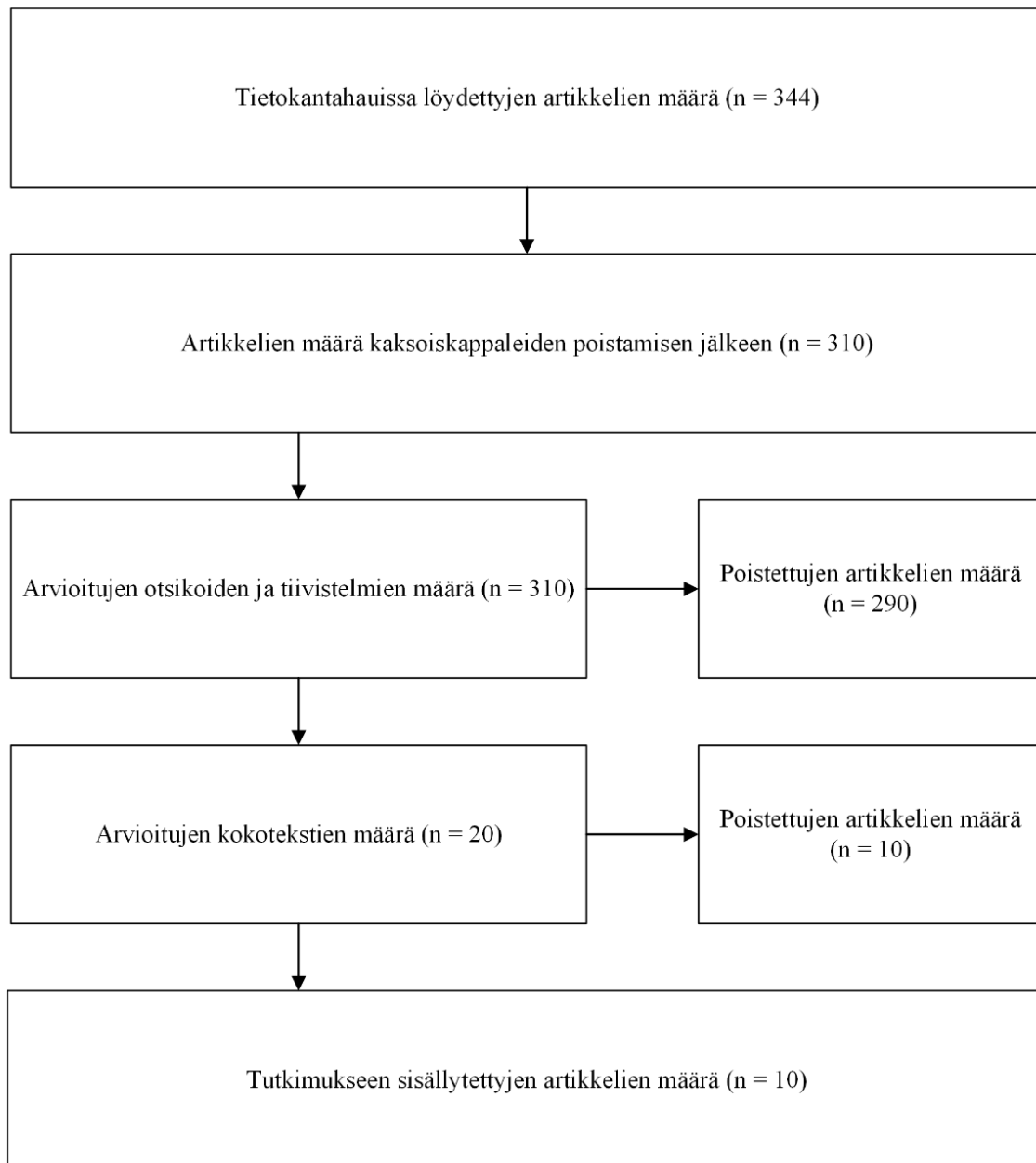
Katsauksen tavoitteena oli kerätä tietoa spontaanipuheen fonologian tutkimuksissa käytetyistä numeerisista menetelmistä sekä niiden luotettavuudesta arvioimalla menetelmien reliabiliteettia. Tämän vuoksi sisäänottokriteereinä olivat, että 1) tutkimuksessa on osana analysoitu fonologista kehitystä spontaanipuheen avulla, 2) tutkimuksessa on käytetty numeerisia menetelmiä spontaanipuheen fonologian tutkimiseen, 3) käytettyjen menetelmien reliabiliteettia voidaan analysoida, 4) tutkittavat ovat 1–5-vuotiaita lapsia ja 5) artikkelit ovat vertaisarvioitu.

Poissulkukriteereinä olivat seuraavat: 1) spontaanipuheen fonologian tutkimiseen on käytetty vain laadullisia menetelmiä ja 2) artikkelin kokotekstiä ei ole saatavilla englanniksi tai suomeksi.

Kuva 1 havainnollistaa artikkelien sisällyttämistä katsaukseen vuokaaviona. Hausta saatiin tuloksena 344 artikkelia, joista kaksoiskappaleiden poiston jälkeen jäljelle jäi 310. Kävimme artikkelit läpi sokkoutettuna. Rayyan-ohjelmalla huomioiden asetetut sisäänotto- ja poissulkukriteerit. Ensin tarkastelimme artikkeleiden otsikot ja tiivistelmät itsenäisesti, minkä jälkeen valitsimme yhdessä pohtien 20 artikkelia kokotekstiarvioon. Huomioitavaa oli, että sisäänottokriteereistä menetelmien reliabiliteetin analyysia voitiin tarkastella tarkemmin vasta kokotekstitarkastelussa. Yhteen artikkeliin kahdestakymmenestä ei ollut pääsyä, joten sitä ei voitu katsaukseen sisällyttää. Kokotekstitarkastelu toteutettiin samalla tavalla kuin otsikoiden ja tiivistelmien arviointi, jonka seurauksena katsaukseen valikoitui tämän perusteella 10 artikkelia.

Kuva 1.

Vuokaavio tutkimusten sisällyttämisestä katsaukseen.



Seuraavassa vaiheessa katsaukseen mukaan otetut artikkelit luettiin tarkemmin läpi ja samalla luotiin työskentelytaulukot, johon kerättiin systemaattisen katsauksen kannalta olennaisimmat tiedot: tutkimuksen tekijät ja julkaisuvuosi, tutkimukseen osallistuneiden lasten otoskoko, ikä ja

spontaanipuheen fonologian numeeriset menetelmät, artikkelien tutkimuskohteet sekä menetelmien reliabiliteettia koskevat tiedot.

2.2 Aineiston luokittelu

Sisäännotetuista artikkeleista etsimme tämän systemaattisen katsauksen tutkimuskysymysten kannalta olennaiset tiedot, ja ne koottiin kahteen työskentelytaulukoon tarkempaan analyysiin. Taulukkoon 1 on koottu keskeiset tiedot valikoitujen artikkelien menetelmistä, joita ovat tutkimuksen tekijät ja julkaisuvuosi, otoskoko, koehenkilöiden ikä, artikkelien tutkimuskohteet sekä käytetyt numeeriset menetelmät spontaanipuheen fonologian tutkimiseen. Taulukkoon 2 kerättiin olennaiset tiedot tutkimusten numeerisien menetelmien reliabiliteetista.

3 Tulokset

3.1 Tutkimusaineisto

Tämän systemaattisen kirjallisuuskatsauksen tarkoituksena on selvittää, 1) millaisia spontaanipuheen fonologian numeerisia menetelmiä on käytetty 1–5-vuotiaiden lasten fonologian tutkimuksessa ja 2) tarkastella näiden menetelmien luotettavuutta reliabiliteettianalyysien avulla. Katsauksen aineisto on esitetty taulukossa 1. Systemaattiseen katsaukseen valikoitui yhteensä 10 tutkimusta, jotka ovat julkaistu vuosina 1985–2020. Tutkimuksiin osallistuneiden lasten ikä vaihteli 1 vuoden 1 kuukauden iästä 5 vuoden 4 kuukauden ikään. Tutkimusten otoskoot vaihtelivat 5 osallistujasta 64 osallistujaan. Suurimmassa osassa (7) tutkimuksista oli kuitenkin vähintään 20 osallistujaa. Tutkimukset olivat joko pitkittäis- tai poikittaistutkimuksia tai yhdistelmiä molemmista. Huomioitavaa on, että osana yhtä tutkimusta toteutettiin kaksi erilaista koeasetelmasta ja tarkasteltiin näistä saatuja tuloksia erikseen (Camarata & Schwartz, 1985). Molemmissa asetelmissä tarkasteltiin kuitenkin semanttis-fonologisen tason yhteyttä lapsilla, joiden puhe oli rajoittunut yksittäisiin sanoihin.

3.2 Spontaanipuheen fonologian numeeriset menetelmät

Katsaukseen valikoituneissa tutkimuksissa hyödynnettiin keskenään erityyppisiä numeerisia menetelmiä spontaanipuheen fonologian tutkimiseen. Taulukossa 1 on esitetty käytetyt menetelmät tutkimuskohtaisesti. Käytetyin (7) numeerinen menetelmä oli PCC, jota hyödynnettiin joko yksinään tai toisen numeerisen menetelmän kanssa. PCC tarjoaa numeerisen arvon tuotettujen konsonanttien tarkkuudesta ja soveltuu hyvin virheiden kokonaismäärän arviointiin.

Tutkimuksissa PCC laskettiin lapsen spontaanista puheesta laaditun litteraatin perusteella. Useimmiten materiaali litterointiin lapsen ja tutkijan välisestä vuorovaikutuksesta (Camarata & Schwartz, 1985; Martikainen ym., 2020; McLeod ym., 2001; Morris & Ozanne, 2003; Smith ym., 2006). Tutkija tarvittaessa ohjasi vuorovaikutusta, mutta kerätty materiaali sisälsi vain lapsen spontaaneja tuotoksia. Kolmessa tutkimuksessa tarkasteltiin lapsen ja vanhemman välistä spontaania vuorovaikutustilannetta, jossa tutkija toimi vain tarkkailijana (Camarata & Schwartz, 1985; Nelson & Bauer, 1991; Roberts ym., 1998). Nämä tilanteet olivat puolistrukturoituja, eli toteutettiin kontrolloidussa ympäristössä, mutta tilanteeseen oli tuotu erilaisia virikkeitä, jotta se olisi lapsille mahdollisimman luonnollinen. Tähän hyödynnettiin muun muassa leluja, kirjoja ja ruokaa. Kaikissa tutkimuksissa vuorovaikutustilanteet videoitiin ja äänitettiin myöhempää litteraatin tarkistusta varten.

Kolmessa tutkimuksessa spontaanista puheesta toteutettiin analyysina pMLU, joka toimii kokosanamittarina fonologiselle sujuvuudelle. Jokaisessa pMLU:n sisältävässä tutkimuksessa videointi ja litterointi toteutettiin vanhemman ja lapsen välisestä vuorovaikutustilanteesta (Beers ym., 2019; Van Noort-Van Der Spek ym., 2010; Watson & Terrell, 2012). Vuorovaikutustilanteissa oli käytössä joko vanhemman itse tuomia tai tutkijoiden tarjoamia leluja tai kirjoja.

Mainittujen menetelmien lisäksi tutkimuksissa hyödynnettiin myös muita yksittäisiä numeerisia spontaanipuheen fonologian menetelmiä (Martikainen ym., 2020; Morris & Ozanne, 2003; Nelson & Bauer, 1991; Smith ym., 2006; Watson & Terrell, 2012). Kyseiset menetelmät esiintyivät aina joko PCC:n tai pMLU:n kanssa. Muita menetelmiä olivat muun muassa vokaalien oikeellisuusprosentti indeksi eli PVC (eng. *Percentage of Vowels Correct*) sekä tuotettujen konsonanttien määrää ja oikeellisuutta tarkastelevia analyyseja. Yhdessä tutkimuksessa tarkasteltiin syvemmin pMLU:n osal alueita, PWC:tä ja PWP:tä (Watson & Terrell, 2012).

Taulukko 1.

Tutkimuksen aineisto.

Tekijät ja vuosi	Otoskoko	Ikä	Artikkelien tutkimuskohteet	Numeeriset menetelmät
Beers ym., 2019	n = 47	1;3–4;0 v	Tutkimuksessa tavoitteena oli selvittää hollantilaisten lasten fonologisen kehityksen vaiheita ja kuinka tämä kehitys heijastuu pMLU-mittarin pisteisiin iän myötä.	pMLU ¹
Camarata & Schwartz, 1985	n = 27*	1;5–3;4 v	Tutkimuksen tarkoituksena oli tarkastella mahdollista semanttis-fonologisen tason yhteyttä puheessa lapsilla, joiden puhe oli rajoittunut yksittäisiin sanoihin.	PCC ²
Martikainen ym., 2020	n = 64	3;2–5;4 v	Tutkimuksessa tutkittiin sanansisäistä tarkkuutta ja johdonmukaisuutta suomea puhuvilla lapsilla, joilla on puhehäiriö (eng. <i>Speech Sound Disorder, SSD</i>) ja näitä verrattiin tyypillisesti kehittyviin ikätovereihinsa.	PCC PVC ³
McLeod ym., 2001	n = 16	2;0–2;11 v	Tutkimuksessa tarkasteltiin sananalkuisten ja -loppuisten konsonanttiyhdistelmien (eng. <i>consonant cluster</i>) tuottoa keskittyen erityisesti yleisten kehityssuuntausten ja yksilöllisten erojen väliseen suhteeseen analysoimalla pitkittäis- ja poikittaisdataa tyypillisesti kehittyneillä lapsilla.	PCC

Tekijät ja vuosi	Otoskoko	Ikä	Artikkelien tutkimuskohteet	Numeeriset menetelmät
Morris & Ozanne, 2003	n = 20	1;10–2;2 v	Pitkittäistutkimuksessa tarkasteltiin kahden suulakihalkion kanssa syntyneen lapsiryhmän kielellisiä, foneettisia ja fonologisia taitoja 3-vuoden iässä.	PCC
Nelson & Bauer, 1991	n = 5	1;1–2;0 v	Tutkimuksessa tutkittiin, kuinka 2-vuotiaat lapset hallitsevat äänteiden tuottamisen ja sanayhdistelmien muodostamisen suhdetta spontaanissa puheessaan.	PCC
Roberts ym., 1998	n = 48	2;0–2;7 v	Tutkimuksessa tarkasteltiin lapsia, joilla oli diagnosoitu ekspressiivinen kielen kehityksen erityisvaikeus (eng. <i>Specific Language Impairment – Expressive, SLI-E</i>), ja selvitettiin, jatkuivatko puheen ja äänteiden tuottoon liittyvät ongelmat, sekä millainen vaikutus näillä on puheen ymmärrettävyyteen ja kielen kehitykseen.	PCC

Tekijät ja vuosi	Otoskoko	Ikä	Artikkelien tutkimuskohteet	Numeeriset menetelmät
Smith ym. 2006	n = 51	2;0–2;6 v	Tutkimuksessa tutkittiin nuorten lasten sanaston koon ja fonologisten taitojen yhteyttä.	<ol style="list-style-type: none"> 1. Yksittäisten konsonanttien tavoiteltu määrä 2. Yksittäisten konsonanttien määrä, jotka tuotettiin oikein ainakin kerran 3. Konsonanttien määrä, jotka tuotettiin oikein ≥ 75 % kerroista 4. PCC
Van Noort-Van Der Spek ym., 2010	n = 40	1;11–2;2 v	Tutkimuksessa tarkasteltiin nuorten lasten, joilla on erittäin alhainen syntymäpaino (eng. <i>Very Low Birth Weight, VLBW</i>), fonologista kehitystä.	pMLU
Watson & Terrell, 2012	n = 12	2;0–3;0 v	Tutkimuksen tarkoituksena oli kuvata muutoksia lasten kokosana-tuotoksissa 24 kuukauden iästä 36 kuukauden ikään.	pMLU

1 = pMLU, Phonological Mean Length of Utterance; 2 = PCC, Percentage of Consonants Correct; 3 = PVC, Percentage of Vowels Correct

** tutkimus koostui kahdesta erilaisesta tutkimusasetelmästä, jossa ensimmäisessä koehenkilöitä oli 12 ja toisessa 15.*

3.3 Numeeristen menetelmien reliabiliteetti

3.3.1 Arvioidut näytteet

Katsauksen tutkimuksissa oli erilaisia tapoja valikoida litterointien reliabiliteetin arviointiin otetut näytteet. Taulukkoon 2 on koottu kaikki reliabiliteetti-arvioon valikoidut tutkimuskohtaiset näytteet. Kolmessa tutkimuksessa reliabiliteetin arviointiin valitut näytteet valittiin satunnaisesti kaikkien näytteiden määrästä (Martikainen ym., 2020; McLeod ym., 2001; Nelson & Bauer, 1991). Näissä artikkeleissa näytteiden valintaan ei sovellettu erityisiä kriteerejä, vaan näytteet poimittiin satunnaisesti koko aineistosta uudelleenlitterointia varten. Esimerkiksi Martikainen kumppaneiden (2020) tutkimuksessa videoitiin tyypillisesti kehittyneiden ja puhehäiriöisten lasten puhetta. Kaikista näytteistä valikoitiin satunnaisesti 10 % uudelleenlitterointiin, joiden perusteella tarkasteltiin yhteneväisyyttä alkuperäiseen puheterapeutin tekemään litterointiin. Camaratan ja Schwartzin (1985) tutkimus sisälsi puolestaan kaksi erilaista tutkimusasetelmaa, joissa ensimmäisessä näytteiden valikointi tapahtui satunnaisesti, mutta toisessa kaikki näytteet litteroitiin kokonaisuudessaan kahdesti ja verrattiin näiden yhteneväisyyksiä.

Neljässä tutkimuksessa oli päätetty kriteeri, jonka perusteella näytteet valikoitiin uudelleenlitterointiin (Beers ym., 2019; Roberts ym., 1998; Smith ym., 2006; Van Noort-Van Der Spek ym., 2010). Beers kumppaneineen (2019) tarkastelivat tutkimuksessaan kielenkehityksen vaiheita, jonka myötä jokaisesta ikäluokasta valikoitiin satunnaisesti yhden pojan ja yhden tytön ääninäytteet osaksi reliabiliteettitarkastelua. Robertsin ja kumppaneiden (1998) sekä Van Noort-Van Der Spekin ja kumppaneiden (2010) tutkimuksissa valikoitiin tietty osuus verrokkiryhmän sekä kontrolliryhmän näytteistä, jotka laskettiin mukaan reliabiliteetti-arviointiin.

Morris ja Ozanne (2003) hyödynsivät useampaa arvioijaa litteroinnin luotettavuuden tarkastelussa. Yksi puheterapeutti litteroi kokonaan videoidut ja nauhoitetut näytteet kaikilta koehenkilöiltä. Tämän lisäksi kaksi ulkopuolista puheterapeuttia arvioivat uudelleen kuuden koehenkilön puhenäytteistä 30 %, ja näiden yhteneväisyyttä verrattiin alkuperäiseen, eli ensimmäisen puheterapeutin tekemään litterointiin. Huomioitavaa oli se, että näille puheterapeuteille valikoitiin satunnaisesti samat näytteet litteroitavaksi.

3.3.2 Arviointitapa

Artikkeleissa tarkasteltiin pääsääntöisesti eroavaisuuksia ja yhteneväisyyksiä litteroinnin osalta. Kaikissa tutkimuksissa pääsääntöisesti yksi henkilö litteroi kaikki tai suurimman osan tallennetuista puhenäytteistä. Tämän lisäksi yksi tai useampi henkilö uudelleenlitteroi tietyn osuuden samoista näytteistä uudelleen. Litterointien yhteneväisyyksien vertailu perustui näihin uudelleenlitteroituihin osuuksiin, jotka ovat merkitty taulukkoon 2. Useimmiten tutkimuksissa (8) litterointi sekä sen yhteneväisyyden uudelleenarviointi toteutettiin jälkikäteen koetilanteista tallennettujen videoiden ja nauhoitusten perusteella (Beers ym., 2019; Martikainen ym., 2020; Morris & Ozanne, 2003; Nelson & Bauer, 1991; Roberts ym., 1998; Smith ym., 2006; Van Noort-Van Der Spek ym., 2010; Watson & Terrell, 2012). Osassa tutkimuksista (2) jo itse koetilanteen aikana toteutettiin litterointia spontaanipuheesta, ja myöhemmät tarkastelut tehtiin jälkikäteen videon sekä nauhoituksen perusteella (Camarata & Schwartz, 1985; McLeod ym., 2001). Esimerkiksi McLeodin ja kumppanien (2001) tutkimuksessa puheterapeutit tekivät litterointia jo koetilanteissa sekä myöhemmin vielä videoiden ja nauhoitusten perusteella uudelleen varmistaakseen litteroinnin tarkkuuden.

Tutkimusten litterointi toteutettiin eri tarkkuudella riippuen siitä, kuinka yksityiskohtaista litteraattia tuotoksista tavoiteltiin. Useimmiten (6) tutkimuksissa toteutettiin foneettinen litterointi, jossa voidaan tarvittaessa kuvata pienetkin ääneen liittyvät ilmiöt (Beers ym., 2019; Camarata & Schwartz, 1985; Morris & Ozanne, 2003; Nelson & Bauer, 1991; Smith ym., 2006; Van Noort-Van Der Spek ym., 2010). Martikainen ja kumppanit (2020) toteuttivat ainoana katsauksen aineistosta näytteilleen foneemisen litteroinnin, eli merkitsivät vain mitä ääniteitä on käytetty ottamatta tarkemmin kantaa, miten ne ääntyvät. McLeod ja kumppanit (2001) toteuttivat tutkimuksessaan niin foneettisen kuin foneemisen litteroinnin. Kahdessa artikkelissa litteroinnin laajuudesta ei ollut saatavilla yksiselitteistä tietoa (Roberts ym., 1998; Watson & Terrell 2012).

3.3.3 Reliabiliteettiarvot

Jokaisessa katsaukseen valikoituneessa artikkeleissa tarkasteltiin arvioijien välistä reliabiliteettia (eng. *Interrater reliability*). Reliabiliteettianalyysi tehtiin eri tarkastelutasoilla, jotka ovat merkitty taulukkoon 2. Katsaukseen valikoituneiden artikkelien litteroinneissa ja niiden tarkemmissa analyyseissä keskityttiin tutkimusasetelmien sekä käytettyjen menetelmien mukaisiin painotuksiin. Pääsääntöisesti tutkimuksissa tarkasteltiin yhteneväisyyksiä kokonaisuudessaan foneettisissa ja

foneemisissa litteroinneissa (4) tai tarkemmin konsonantti- tai vokaalitason oikeanlaisissa tuotoksissa (4). Kokonaisuuden litteroinnissa tarkasteltiin lapsen puhenäytteessä esiintyviä äänteitä, sanoja ja lauseita, kuitenkin ottamatta tarkemmin kantaa siihen, missä kohtaa mahdolliset erot ilmenevät. Kokonaisuuden litterointia käsittelevissä tutkimuksissa tarkasteltiin tutkijoiden tekemiä päätöksiä litteroinnin osalta sekä näiden päätösten yhteneväisyyttä (Camarata & Schwartz, 1985; Nelson & Bauer, 1991; Van Noort-Van Der Spek ym., 2010; Watson & Terrell, 2012). Esimerkiksi Nelson ja Bauer (1991) vertailivat tutkimuksessaan 5% litteroiduista puhenäytteistä kokonaisuuden kannalta ja tarkastelivat litteroijien välisiä yhteneväisyyksiä. Konsonantti- tai vokaalitasolle tarkentuneissa tutkimuksissa keskityttiin litteroinnin näkökulmasta syvällisemmin tiettyihin puheen osa-alueisiin, eli konsonanttien ja vokaalien tuottamiseen (Beers ym., 2019; Martikainen ym., 2020; McLeod ym., 2001; Morris & Ozanne, 2003).

Kahdessa tutkimuksessa tarkasteltiin vielä erikseen, miten eroavaisuudet litteroinnissa vaikuttavat käytetyn numeerisen menetelmän tulokseen. Esimerkiksi Smithin ja kumppanien (2006) tutkimuksessa keskityttiin PCC-menetelmän käsittelyyn ja oikeiden konsonanttien prosentuaaliseen tuottoon sanapaikoissa. Puolestaan Robertsien ja kumppaneiden (1998) tutkimuksessa vertailtiin yhteneväisyyksiä litteraattikokonaisuudesta, vokaali- ja konsonanttitasosta sekä tutkimuksessa käytetystä PCC-menetelmästä.

Kolmessa artikkelissa tarkasteltiin myös arvioijan sisäistä yhteneväisyyttä (eng. *Intrarater reliability*) (Martikainen ym., 2020; Smith ym., 2006; Van Noort-Van Der Spek ym. 2010). Martikaisen ja kumppanien (2020) tutkimuksessa arvioijien sisäistä reliabiliteettia tarkasteltiin valikoimalla aikaisemmin litteroiduista näytteistä 10 % uudelleenlitteroitavaksi. Tämän pohjalta havaittiin, että litteroinneissa oli 93 % arvioijan sisäinen yhteneväisyys konsonanteille ja 97 % vokaaleille. Smithin ja kumppaneiden (2006) tutkimuksessa vertailtiin neljän koehenkilön PCC:ta sanan eri asemissa yhden arvioijan ensimmäisen ja toisen litteroinnin välillä. Litterointikertojen välillä oli aina muutama kuukausi. Ero litterointikerroilla oli keskimäärin 7.4 % eli arvioijan sisäiseksi reliabiliteetiksi määritettiin noin 92.6 %. Van Noort-Van Der Spekin ja kolleegoiden (2010) tutkimuksessa puheterapeutti, joka oli toteuttanut kaikkien näytteiden litteroinnit, arvioi uudelleen viisi videonäytettä. Näiden litterointien välillä todettiin 92 % sisäinen yhteneväisyys sanoissa ja 90 % sisäinen yhteneväisyys äänteiden osalta.

Taulukko 2

Tutkimusten litteroinnin reliabiliteetin arviointi.

Tekijät, julkaisuvuosi	Kaikki litteroijat	Arvioitujen näytteiden määrä kaikista näytteistä (%) [*]	Reliabiliteetin arviointitapa	Reliabiliteettiarvo, yhteneväisyys (%)		
				Koko foneeminen ja foneettinen litterointi	Vokaali- ja konsonanttitaso	Numeerinen menetelmä
Beers ym., 2019	2 asiantuntijaa	20	Yhteneväisyys konsonanttien, vokaalien ja konsonanttiyhdistelmien välillä	-	Konsonanttiyhdistelmät 85 konsonantit 78 vokaalit 90	-
Camarata & Schwartz, 1985	2 asiantuntijaa	ei selkeästi määritettävissä ¹	Oikein tuotetut konsonantit	1. koeasetelma: 95 2. koeasetelma: 97	-	-
Martikainen ym., 2020	2 puheterapeuttia	10	Oikein tuotetut konsonantit ja vokaalit	-	konsonantit: 88 vokaaleille: 94	-

Tekijät, julkaisuvuosi	Kaikki litteroijat	Arvioitujen näytteiden määrä kaikista näytteistä (%)*	Reliabiliteetin arviointitapa	Reliabiliteetti-arvo, yhteneväisyys (%)		
				Koko foneeminen ja foneettinen litterointi	Vokaali- ja konsonanttitaso	Numeerinen menetelmä
McLeod ym., 2001	2 puheterapeuttia	10	Oikein tuotetut konsonanttiyhdistelmät	-	laaja litterointi: 90.8 yksityiskohtainen litterointi: 84.5	-
Morris & Ozanne, 2003	3 puheterapeuttia	ei tarkasti määriteltävissä ²	Oikein tuotetut konsonantit	-	konsonantit 86	-
Nelson & Bauer, 1991	2 jatko-opiskelijaa	10	Oikein tuotetut konsonantit	72	-	-
Roberts ym., 1998	Pääkirjoittaja ja jatko-opiskelija	17	Yhteneväisyys äänteissä ja vokaaleissa sekä oikein tuotetut konsonantit	91.6	Vokaaliset ilmaukset: 94.1	84.9

Tekijät, julkaisuvuosi	Kaikki litteroijat	Arviotujen näytteiden määrä kaikista näytteistä (%)*	Reliabiliteetin arviointitapa	Reliabiliteettiarvo, yhteneväisyys (%)		
				Koko foneeminen ja foneettinen litterointi	Vokaali- ja konsonanttitaso	Numeerinen menetelmä
Smith ym., 2006	2 puheterapeuttiopiskelijaa	16	Oikein tuotetut konsonantit eri kohdassa sanaa	-	-	94.2
Van Noort- Van Der Spek ym., 2010	2 puheterapeuttia	25	Yhteneväisyys sanoissa ja äänteissä	Sanoissa 87 Äänteissä 83	-	-
Watson & Terrell, 2012	Pääkirjoittaja ja jatko- opiskelija	17	Yhteneväisyys äänteissä	91	-	-

* Tarkasteltu kuinka suuri osuus prosentuaalisesti alkuperäisistä näytteistä on litteroitu uudelleen osana reliabiliteetin arviointia. Kaikissa tutkimuksissa prosentuaalista osuutta ei suoraan mainittu, joten näiden kohdalla prosenttiosuus laskettiin itse.

1 Tutkimus sisälsi kaksi erillistä koeasetelmaa, jossa vain toisesta arvioitiin litteroinnin reliabiliteetti. Tästä 50 % näytteistä arviointiin litteroinnin reliabiliteettiä.

2 Kuuden koehenkilön spontaanipuheen otoksista arvioitu 30 %, joista tehty vertailu.

4 Pohdinta

Tämän systemaattisen katsauksen tavoitteena oli tarkastella, millaisia numeerisia menetelmiä hyödynnetään 1–5-vuotiaiden lasten spontaanipuheen fonologian arvioinnissa ja kuinka luotettaviksi nämä menetelmät on arvioitu. Spontaanipuheen fonologian tutkimus numeerisin menetelmin on tärkeää, koska se antaa lisää keinoja kliiniseen fonologian arvioon. Nämä menetelmät tarjoavat objektiivista ja määrällistä tietoa lapsen puheen kehityksestä tavalla, joka mahdollistaa vertailun tyypilliseen kehitykseen sekä tarjoaa keinon seurata lapsen äänteellistä kehitystä.

Katsauksen tulosten perusteella kaksi menetelmää nousi selvästi esiin: PCC (*Percentage of Consonants Correct*) ja pMLU (*Phonological Mean Length of Utterance*). Molempien menetelmien reliabiliteettiarvot olivat pääosin korkeita. Tutkimusten välillä ilmeni kuitenkin vaihtelua reliabiliteettiarvioiden tehneiden arvioijien taustassa, analysoitujen otosten määrässä ja litteroinnin tasossa. Nämä ovat tekijöitä, jotka voivat vaikuttaa menetelmien tulkittavuuteen ja toistettavuuteen.

4.1 Tulosten tulkinta

4.1.1 Arvioijat

Artikkelien koetilanteiden litteroinnit tehneiden henkilöiden roolit sekä taustat ovat saattaneet vaikuttaa osaltaan reliabiliteettiarvoihin. Jokaisessa tarkastellussa tutkimuksessa oli raportoitu arviointiin osallistuneiden henkilöiden taustatiedot, kuten koulutustausta tai ammatillinen asema. On perusteltua olettaa, että arvioijan tausta, oli tämä puheterapeutti, opiskelija, tutkija, litteroija tai muu alan ammattilainen, on vaikuttanut havaintojen tulkintaan ja näin ollen myös saatuihin reliabiliteettiarvoihin ja arviointien yhteneväisyyksiin. Kokemus ja asiantuntemuksen taso voivat vaikuttaa esimerkiksi siihen, kuinka tarkasti ja johdonmukaisesti havaintoja tehdään vaikuttaen näin esitettyihin arvoihin ja niiden laatuun.

Osa arvioijista oli myös suorittanut International Phonetic Alphabet -koulutuksen (IPA), mikä voidaan nähdä merkityksellisenä tekijänä arvioinnin tarkkuuden sekä johdonmukaisuuden kannalta (Camarata & Schwartz, 1985; Nelson & Bauer, 1991; Watson & Terrell, 2012). IPA-koulutus antaa valmiudet täsmälliseen äänteiden tunnistamiseen ja kirjaamiseen. Koulutuksen tuoma osaaminen on todennäköisesti vaikuttanut suoraan litteroinnin laatuun sekä arvioijan kykyyn erotella puheen

piirteitä eri tavalla verrattuna sellaiseen arvioijaan, joka ei vastaavaa koulutusta ole käynyt. Tämä voi näkyä esimerkiksi Watson & Terrellin (2012) tutkimuksessa, jossa koko litteraatin reliabiliteetiksi saatiin korkea 91 %. Toisaalta koulutuksella on voinut olla myös päinvastainen vaikutus arvioijan suorittamaan litterointiin, kuten voidaan tulkita Nelsonin & Bauerin (1991) tutkimuksesta. Kyseisessä tutkimuksessa koko litteraatin reliabiliteetti jäi keskimäärin 72 prosenttiin, mikä on alle tähän katsaukseen sisältyneiden artikkelien reliabiliteettien keskiarvon. Osaavampi litteroija voi olla liian tarkka tai kriittinen analysoidessaan puheessa esiintyviä piirteitä. Toisaalta kokeneempi ja koulutetumpi litteroija voi luottaa liikaa omaan kliiniseen kokemukseensa eikä näin ollen kuuntele litteraatteja riittävän tarkasti. Nämä mainitut tekijät voivat vaikuttaa tehtyihin tulkintoihin ja siten koko litteraatin reliabiliteettiin.

4.1.2 Arvioitujen näytteiden määrä

Reliabiliteetin arvioinnissa on olennaista, kuinka suuri osuus puhenäytteistä valitaan uudelleenlitterointiin. Katsaukseen sisällytetyt tutkimukset vaihtelivat siinä, kuinka laajan osan näytteistä ne arvioivat uudelleen. Useissa tutkimuksissa tämä osuus oli noin 10–20 %, mutta joissain tutkimuksissa tarkkaa prosenttia ei voitu määrittellä. Parhaimmillaan osuus oli jopa 30 % koko tutkimuksen aineistosta.

Uudelleenlitteroidun osuuden suuruudella voi olla vaikutusta reliabiliteettiarvojen edustavuuteen. Pieni vertailumäärä näytteissä voi antaa keinotekoisesti korkean reliabiliteettiarvon, mikäli analysoidut näytteet ovat esimerkiksi poikkeuksellisen selkeitä tai edustavat helpommin analysoitavia tuotoksia. Vastaavasti suuremmat vertailumäärät vähentävät sattuman vaikutusta ja tarjoavat kattavamman kuvan arviointien johdonmukaisuudesta sekä otoksen edustavuudesta. Uudelleenlitteroitujen näytteiden valintakriteerit olivat useimmiten satunnaisia, mutta kaikissa tutkimuksissa ei selitetty valintaprosessia tarkasti, mikä voi heikentää reliabiliteettiarvioiden läpinäkyvyyttä ja toistettavuutta.

4.1.3 Litteroinnin taso - foneeminen vs. foneettinen

Katsauksen tutkimuksissa litterointia oli toteutettu joko foneettisella tai foneemisella tasolla. Yhdessä tutkimuksessa hyödynnettiin näitä molempia. Foneeminen litterointi kuvaa puheen äänteitä kielellisesti erottavien yksiköiden tasolla, kun taas foneettinen litterointi mahdollistaa tarkemman analyysin puheen äänteellisistä yksityiskohdista, kuten soinnillisuudesta ja artikulaatiotavasta.

Foneettinen litterointi tarjoaa yksityiskohtaisemman kuvan lapsen tuottamasta puheesta ja voi siten lisätä analyysin herkkyyttä, erityisesti fonologisten poikkeavuuksien havaitsemisessa. Se edellyttää kuitenkin syvällistä fonetiikan osaamista ja yhtenäisiä käytäntöjä arvioijien välillä, jotta reliabiliteetti säilyy korkeana. Foneeminen litterointi puolestaan on usein suoraviivaisempaa ja soveltuu paremmin laajamittaiseen analyysiin tai tilanteisiin, joissa tavoitteena on tarkastella lapsen fonologisen järjestelmän rakennetta kokonaisuutena.

Reliabiliteetti-arvot olivat yleisesti litteroinnin tasosta huolimatta jokaisessa artikkelissa korkeat. Kuitenkin on huomioitavaa, että pääsääntöisesti tutkimuksissa käytettiin foneettista litterointia ja vain kolmessa tutkimuksessa käytössä oli foneeminen litterointi. Osassa tutkimuksissa litteroinnin tarkkuustasoa ei määritelty syvällisemmin lainkaan. Nämä litteroinnin tasoerot vaikeuttavat tulosten vertailua ja reliabiliteetti-arvojen luotettavaa arviointia. Tässä katsauksessa litteroinnin tasolla ei vaikuttanut olevan merkittävää vaikutusta saatuihin reliabiliteetti-arvoihin.

4.2 Vahvuudet ja rajoitukset

Tämän katsauksen vahvuus on sen systemaattinen toteutustapa: aineistonhaku ja tutkimusten valinta perustui tarkasti määriteltyihin sisään- ja ulosottokriteereihin. Tämä lisää työn toistettavuutta ja uskottavuutta. Katsauksella saattaa olla myös kliinistä merkitystä pienten lasten fonologisen kehityksen arvioinnissa spontaanipuheen perusteella. Aiheesta on julkaistu vain vähän aikaisempia katsauksia, joten työ mahdollisesti täydentää olemassa olevaa tutkimustietoa.

Katsauksen ensisijaisina rajoituksina ovat siihen valikoitunut pieni otoskoko sekä tutkimusten vähäiset osallistujamäärät. Rajoitteita ovat myös valikoitujen tutkimusten vaihtelevat koeasetelmat, esimerkiksi näytteiden määrä ja painotuserot reliabiliteetin arvioinnissa. Näiden lisäksi tutkimusten vertailtavuutta keskenään vaikeuttavat erot siinä, miten tarkasti arviointimenetelmät ja niiden sovellukset on kuvattu. Tutkimuksissa raportoidut korkeat ja yhteneväiset reliabiliteetti-arvot voivat viitata julkaisuharhaan. On mahdollista, että tutkimukset, joissa on saatu heikompia arvoja, eivät ole päätyneet julkaisuun tai niissä ei mainita reliabiliteetti-arvoja ollenkaan. Katsauksessa karsittiin pois useita tutkimuksia juuri siksi, ettei niissä ollut raportoitu reliabiliteettia. Nämä kaikki voi osaltaan vaikuttaa katsauksen yleistettävyyteen.

4.3. Johtopäätökset ja jatkotutkimusehdotukset

Tämän katsauksen perusteella PCC ja pMLU ovat tällä hetkellä tutkimuksissa yleisimmin käytettyjä numeerisia menetelmiä spontaanipuheen fonologian arvioinnissa 1–5-vuotiailla lapsilla. Molemmat menetelmät vaikuttavat luotettavilta, sillä niille on toistuvasti raportoitu korkeita reliabiliteettiarvoja tutkimuksissa, joissa reliabiliteetin arviointimenetelmä ja arvioijan tausta on selkeästi kuvattu.

Jatkossa olisi tärkeää tarkastella systemaattisemmin, kuinka arvioijan koulutus, litteroinnintaso ja tarkasteltavan näytteen koko vaikuttavat käytetyn numeerisen menetelmän reliabiliteettiin. Tieteen avoimuuden vuoksi olisi myös mielenkiintoista perehtyä tutkimuksiin, jotka käsittelevät heikompia yhteneväisyysarvoja tai jopa nolla-arvoja. Tämä tarkastelu voisi myös mahdollisesti tukea menetelmien kehitystä eteenpäin. Numeeristen menetelmien automatisointi on myös potentiaalinen kehityssuunta, joka voisi sekä parantaa arvioinnin tehokkuutta että lisätä toistettavuutta kliinisessä työssä ja tutkimuksessa.

Lähteet

- Ball, M. J., Müller, N., & Nelson, R. L. (with Ball, M. J., Nelson, R. L., & Müller, N.). (2014). *Handbook of Qualitative Research in Communication Disorders* (1. p.). Psychology Press. <https://doi.org/10.4324/9780203798874>
- Beers, M., Rodenburg-Van Wee, M., & Gerrits, E. (2019). Whole-word measures and the speech production of typically developing Dutch children. *Clinical Linguistics & Phonetics*, 33(12), 1149–1164. <https://doi.org/10.1080/02699206.2019.1596163>
- Berenthal, J. E., Bankson, N. W., & Flipsen, P. (2022). Speech sound disorders in children: Articulation & phonological disorders. Teoksessa *Speech sound disorders in children: Articulation & phonological disorders* (Ninth edition.). Paul H. Brookes Publishing Co.
- Bishop, D. V. M. (Dorothy V. M.), & Leonard, L. B. (2014). *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (1st ed.). Psychology Press. <https://doi.org/10.4324/9781315784878>
- Camarata, S. M., & Schwartz, R. G. (1985). Production of object words and action words: Evidence for a relationship between phonology and semantics. *Journal of Speech & Hearing Research*, 28(3), 323–330. APA PsycInfo. <https://doi.org/10.1044/jshr.2803.323>
- Crary, M. A. (1983). Phonological process analysis from spontaneous speech: The influence of sample size. *Journal of Communication Disorders*, 16(2), 133–141. [https://doi.org/10.1016/0021-9924\(83\)90044-8](https://doi.org/10.1016/0021-9924(83)90044-8)
- Edwards, J., Beckman, M. E., & Munson, B. (2004). The Interaction Between Vocabulary Size and Phonotactic Probability Effects on Children’s Production Accuracy and Fluency in Nonword Repetition. *Journal of Speech, Language, and Hearing Research*, 47(2), 421–436. APA PsycInfo. [https://doi.org/10.1044/1092-4388\(2004/034\)](https://doi.org/10.1044/1092-4388(2004/034))

- Heikkilä, T. (2008). *Tilastollinen tutkimus* ([7. uudistettu painos]). Edita.
- Ikonen, A., & Ström, U. (2009). Lasten kehitykselliset puhehäiriöt. Teoksessa O. Aaltonen, R. Aulanko, A. Iivonen, A. Klippi, & M. Vainio, *Puhuva ihminen: Puhetieteiden perusteet* (ss. 122–128). Otava, kustannusosakeyhtiö.
- Ingram, D. (1976). *Phonological disability in children*. Edward Arnold.
- Ingram, D. (2002). The measurement of whole-word productions. *Journal of child language*, 29(4), 713–733. <https://doi.org/10.1017/S0305000902005275>
- Ingram, D., & Ingram, K. D. (2001). A Whole-Word Approach to Phonological Analysis and Intervention. *Language, speech & hearing services in schools*, 32(4), 271–283. [https://doi.org/10.1044/0161-1461\(2001/024\)](https://doi.org/10.1044/0161-1461(2001/024))
- Kunnari, S., & Savinainen-Makkonen, T. (2009). Puheen, kielen ja kommunikoinnin kehitys lapsuudessa. Teoksessa O. Aaltonen, R. Aulanko, A. Iivonen, & M. Vainio, *Puhuva ihminen: Puhetieteiden perusteet* (1. p., ss. 114–121). Otava.
- Kunnari, S., & Savinainen-Makkonen, T. (2012). Systemaattisen fonologisen kehityksen kauden ajoitukset ja niitä palvelevat fonologiset prosessit. Teoksessa S. Kunnari, T. Savinainen-Makkonen, & kustantaja. PS-kustannus, *Pienten sanat: Lasten äänteellinen kehitys* (ss. 121–127). PS-kustannus.
- Kunnari, S., Savinainen-Makkonen, T., & Paavola, L. (2006). Kaksivuotiaiden suomalaisten konsonantti-inventaarit. *Puhe ja kieli*, 26(2), 71–79.
- Kunnari, S., Savinainen-Makkonen, T., & Saaristo-Helin, K. (2012). *Fonologiatesti: Lasten äänteellisen kehityksen arviointimenetelmä*. Niilo Mäki Instituutti.
- Martikainen, A.-L., Savinainen-Makkonen, T., Laukkanen-Nevala, P., & Kunnari, S. (2020). Intra-word accuracy and consistency in Finnish-speaking children with speech sound disorder

compared to their typically developing peers. *Clinical Linguistics & Phonetics*, 34(8), 718–733. <https://doi.org/10.1080/02699206.2019.1696610>

McLeod, S., van Doorn, J., & Reed, V. A. (2001). Consonant cluster development in two-year-olds: General trends and individual difference. *Journal of Speech, Language, and Hearing Research*, 44(5), 1144–1171. APA PsycInfo. [https://doi.org/10.1044/1092-4388\(2001/090\)](https://doi.org/10.1044/1092-4388(2001/090))

Morris, H., & Ozanne, A. (2003). Phonetic, phonological, and language skills of children with a cleft palate. *Cleft Palate-Craniofacial Journal*, 40(5), 460–470. APA PsycInfo. [https://doi.org/10.1597/1545-1569\(2003\)040](https://doi.org/10.1597/1545-1569(2003)040)

Nelson, L. K., & Bauer, H. R. (1991). Speech and language production at age 2: Evidence for tradeoffs between linguistic and phonetic processing. *Journal of Speech & Hearing Research*, 34(4), 879–892. APA PsycInfo. <https://doi.org/10.1044/jshr.3404.879>

Owens, R. E., Metz, D. E., & Farinella, K. A. (2015). *Introduction to Communication Disorders* (5. p.). Pearson Education UK.

Peña-Brooks, A., & Hegde, M. N. (Mahabalagiri N.). (2023). *Assessment and treatment of speech sound disorders in children: A dual-level text* (4th edition.). PRO-ED.

Ramig, L. A., & Fox, C. (2012). The integral role of speech production in the science and treatment of Parkinson's disease. *Movement Disorders*, 27(7), 811–813.

Roberts, J., Rescorla, L., Giroux, J., & Stevens, L. (1998). Phonological skills of children with specific expressive language impairment (SLI-E): Outcome at age 3. *Journal of Speech, Language, and Hearing Research*, 41(2), 374–384. APA PsycInfo. <https://doi.org/10.1044/jslhr.4102.374>

Rvachew, S., & Brosseau-Lapr e, F. (2018). *Developmental phonological disorders: Foundations of clinical practice* (Second edition.). Plural Publishing, Inc.

- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L. (1997). The Percentage of Consonants Correct (PCC) Metric: Extensions and Reliability Data. *Journal of speech, language, and hearing research, 40*(4), 708–722. <https://doi.org/10.1044/jslhr.4004.708>
- Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders: III. A procedure for assessing severity of involvement. *Journal of Speech & Hearing Disorders, 47*(3), 256–270. APA PsycInfo. <https://doi.org/10.1044/jshd.4703.256>
- Smith, B. L., McGregor, K. K., & Demille, D. (2006). Phonological development in lexically precocious 2-year-olds. *Applied Psycholinguistics, 27*(3), 355–375. <https://doi.org/10.1017/S0142716406060310>
- Stoel-Gammon, Carol., & Dunn, Carla. (1985). *Normal and disordered phonology in children*. PRO-ED.
- Stokes, S. F., Klee, T., Carson, C. P., & Carson, D. (2005). A Phonemic Implicational Feature Hierarchy of Phonological Contrasts for English-Speaking Children. *Journal of speech, language, and hearing research, 48*(4), 817–833. [https://doi.org/10.1044/1092-4388\(2005/057\)](https://doi.org/10.1044/1092-4388(2005/057))
- Storkel, H. L., & Morrisette, M. L. (2002). The Lexicon and Phonology: Interactions in Language Acquisition. *Language, Speech, and Hearing Services in Schools, 33*(1), 24–37. [https://doi.org/10.1044/0161-1461\(2002/003\)](https://doi.org/10.1044/0161-1461(2002/003))
- Tomas, E., & Dorofeeva, S. (2019). Mean Length of Utterance and Other Quantitative Measures of Spontaneous Speech in Russian-Speaking Children. *Journal of Speech, Language, and Hearing Research, 62*(12), 4483–4496. https://doi.org/10.1044/2019_JSLHR-L-18-0339
- Torvelainen, P. (2007). *Kaksivuotiaiden lasten fonologisen kehityksen variaatio* [Väitöskirja]. Jyväskylän yliopisto.

- Turunen, P. (2003). *Production of word structures: A constraint-based study of 2;6 year old Finnish children at-risk for dyslexia and their controls* [Väitöskirja]. Jyväskylän yliopisto.
- Van Noort-Van Der Spek, I. L., Franken, M.-C. J. P., Wieringa, M. H., & Weisglas-Kuperus, N. (2010). Phonological development in very-low-birthweight children: An exploratory study. *Developmental Medicine & Child Neurology*, 52(6), 541–546. APA PsycInfo. <https://doi.org/10.1111/j.1469-8749.2009.03507.x>
- Vihman, M. May. (2014). *Phonological development: The first two years* (Second edition.). Wiley-Blackwell.
- Watson, M. M., & Terrell, P. (2012). Longitudinal changes in phonological whole-word measures in 2-year-olds. *International Journal of Speech-Language Pathology*, 14(4), 351–362. APA PsycInfo. <https://doi.org/10.3109/17549507.2012.663936>
- Yang, X. (2024). The Phonological Development Process in Early Childhood. *Journal of Phonetics & Audiology*, 10(4). <https://doi.org/10.35248/2471-9455.24.10.255>

Liite 1. pMLU:n laskeminen ja PCC:n laskukaava

pMLU:n laskemiseen liittyvät menetelmät tiivistettynä

- Analyysiin valitaan vähintään 25 sanaa.
- Jokainen segmentti eli konsonantti ja vokaali lapsen tuottamassa sanassa lasketaan, ja kustakin annetaan yksi piste.
- Lisäpiste annetaan jokaiselle oikealle konsonantille verrattuna aikuisen tuottamaan vastineeseen.
- pMLU lasketaan lisäämällä saadut pisteet valitulle sananäytteelle ja jakamalla kokonaispistemäärä sanojen lukumäärällä.
- Esimerkiksi jos lapsi saa 70 pistettä 25 sanan näytteestä, pMLU lasketaan seuraavasti:

$$pMLU = \frac{70}{25} = 2.8$$

PCC:n laskukaava

$$PCC = \frac{\textit{oikeiden konsonanttien määrä}}{\textit{oikeat konsonantit + väärät konsonantit}}$$

Liite 2. *Tarkat hakulausekkeet tietokantakohtaisesti*

Pubmed:

("phonological dev*" OR "phonology" OR "phonetic analysis" OR "phonological complexity" OR "speech analysis" OR "speech processing") AND ("child*" OR "preschool*" OR "toddler*" OR "infant*" OR "Child"[Mesh]) AND ("spontaneous speech" OR "spontaneous speaking" OR "spontaneous talk*" OR "spontaneous language" OR "spontaneous speech samples" OR "natural* speech" OR "unscripted speech")

PsychInfo:

("phonological dev*" OR "phonology" OR "phonetic analysis" OR "phonological complexity" OR "speech analysis" OR "speech processing" OR DE "Speech Development" OR DE "Phonology" OR DE "Phonetics") AND ("child*" OR "preschool*" OR "toddler*" OR "infant*") AND ("spontaneous speech" OR "spontaneous speaking" OR "spontaneous talk*" OR "spontaneous language" OR "spontaneous speech sample*" OR "natural* speech" OR "unscripted speech")