

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Pajunen, Anneli; Honko, Mari

Title: Myöhempi kielenkehitys : sanatiedosta rakenne- ja tekstitietoon

Year: 2021

Version: Published version

Copyright: © 2021 Anneli Pajunen, Mari Honko ja SKS

Rights: CC BY-NC-ND 4.0

Rights url: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Please cite the original version:

Pajunen, A., & Honko, M. (2021). Myöhempi kielenkehitys : sanatiedosta rakenne- ja tekstitietoon. In A. Pajunen, & M. Honko (Eds.), Suomen kielen hallinta ja sen kehitys : peruskoululaiset ja nuoret aikuiset (pp. 8-57). Suomalaisen Kirjallisuuden Seura. Suomalaisen Kirjallisuuden Seuran toimituksia, 1472. <https://oa.finlit.fi/site/books/e/10.21435/skst.1472/>

Suomen kielen hallinta ja sen kehitys

Peruskoululaiset ja nuoret aikuiset

Toimittaneet

ANNELI PAJUNEN JA MARI HONKO



SUOMALAISEN KIRJALLISUUDEN SEURAN TOIMITUKSIA 1472

Teos on Suomalaisen Kirjallisuuden Seuran nimeämien asiantuntijoiden tarkastama.



VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

© 2021 Anneli Pajunen, Mari Honko ja SKS

Lisenssi CC BY-NC-ND 4.0 International

Kannen suunnittelu: Timo Numminen

Taitto: Maija Räisänen

EPUB: Tero Salmén

ISBN 978-951-858-407-3 (nid.)

ISBN 978-951-858-408-0 (EPUB)

ISBN 978-951-858-409-7 (PDF)

ISSN 0355-1768 (nid.)

ISSN 2670-2401 (verkkojulkaisut)

DOI <https://doi.org/10.21435/skst.1472>

Teos on lisensoitu Creative Commons CC BY-NC-ND 4.0 International lisenssillä. Tutustu lisenssiin englanniksi osoitteessa <https://creativecommons.org/licenses/by-nc-nd/4.0/> tai suomeksi osoitteessa <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.fi>.



Teos on avoimesti saatavissa osoitteessa <https://doi.org/10.21435/skst.1472> tai lukemalla tämä QR-koodi mobiililaitteella.



Hansaprint Oy, Turenki 2021

Myöhempi kielenkehitys: sanatiedosta rakenne- ja tekstitietoon

Anneli Pajunen

 <https://orcid.org/0000-0001-8597-7491>

Mari Honko

 <https://orcid.org/0000-0002-3240-622X>

Johdanto

Tämän teoksen aiheena on kielenhallinnan kehittyminen eli kieliyhteisölle yhteisen kielellisen tiedon vähittäinen lisääntyminen varhaislapsuuden jälkeen niin sanotun myöhemmän kielenkehityksen aikana (*later language development*, ks. esimerkiksi Berman 2004; Nippold 2006). Lukujen päähuomio on suomea ensikielenä oppineissa, äidin-kielisissä suomenpuhujissa.

Kielenhallinta edellyttää tiedostamatonta tai tietoista tietoa kielestä: intuitiivista tietoa kielen fonologisesta, morfologisesta ja syntaktisesta rakenteesta, näiden tasojen käyttökonteksteista ja käytön pragmaattisista ehdoista sekä lisäksi sanastosta. Riittävä kielenhallinta on edellytys kielenkäyttötaidoille, kommunikaatiolle sekä lukemaan ja kirjoittamaan oppimiselle. Kielenhallinnan kehitys alkaa puhutusta kielestä, mutta se ei ole riippuvainen puhutun ja kirjoitetun kielen modaaliteettieroista eikä varieteettieroista, vaan lapsi ja nuori oppii niitä äidinkieltensä merkityksiä,

muotoja ja käyttötapoja, joita hänellä on edellytys oppia ja joille hän ympäristössään altistuu. Kielenhallinta kattaa siis yksilön kaiken kielitiedon, se ei rajaudu kirjoitettuun yleiskieleen, vaikka hallinnan kehitystä on ehkä helpoin tutkia kirjoitetun kielen kontekstissa ja vaikka juuri myöhemmän kielenkehityksen vaiheissa hallinta laajenee puhutusta kielen kirjalliseen käyttöön. Kielenhallinta ja sen kehitys ei edellytä preskriptiivisen kieliopin normiston tai termien tuntemusta vaan kielitieto on puheyhteisölle yhteistä tietoa. Tähän kielen ydinalueeseen soveltuu yksimieliyssperiaate (*consensus principle*): yleisimmät sanat ja rakenteet hallitaan varmasti mutta muutoksenalaisten ja harvinaisten ilmiöiden osalta epävarmuus kasvaa (ks. Labov 1975: 8–9, 30–32; Itkonen 1978: 151–154; Hulstijn 2015; 2019). Kielenhallinta kehittyy vielä aikuisiässä mutta lapsilla ja nuorilla aikuistasoon pääsy kestää pitkään. Oman tutkimuksemme tavoitteena on ollut tehdä näkyväksi se, mikä on keskeistä yhteistä kielellistä tietoa peruskoulun eri vaiheissa oleville ja miten se iän mukaan kasvaa ja kehittyy erityisesti kielessä harvinaisempien ilmiöiden osalta. Oletus on, että harvinaisten ilmiöiden hallinnassa hajonta on suurta.

Pieni lapsi oppii kielenkehityksen alkuvaiheissa paitsi äidinkiensä (ensikielensä tai -kieltensä) äännerakenteen myös arkipäivän perussanastoa, sanontoja ja fraaseja sekä morfologian ja syntaksin yksinkertaisia rakenteita. Kehitys toisin sanoen alkaa varmasta tiedosta (tai siitä tiedon osasta, joka ensimmäisenä osoittautuu varmaksi). Puhe kohdistuu aluksi pääasiassa konkreettisiin, läsnä oleviin asioihin, ja yksittäisten ilmausten viittausala voi olla huomattavasti suppeampi tai laveampi kuin aikuisten kielessä (Hoff 2005). Kolmen neljän ikävuoden vaiheilla lapsi oppii systemaattisemmin irrottamaan puheensa sisällön puhehetkestä eli kertomaan myös asioista, jotka eivät ole läsnä tai totta (Doherty 2009; Suomessa esimerkiksi Heimo 2003). Myöhemmän kielenkehityksen alkuvaiheissa eli noin viisivuotiaana lapsen elämässä alkaa voimakas kielellisen kehityksen vaihe. Hän oppii kertomaan suullisia tarinoita ja huomioimaan kuulijoita, ja kaikkiaan kielelliset tiedot ja taidot monipuolistuvat ja rikastuvat monella tavalla. Lapsi oppii harvinaisempaa sanastoa, useampisanaisia ilmauksia sekä monimutkaisia rakenteita ja oppii hallitsemaan kielenkäyttöä erilaisissa tilanteissa. Tarkastelemme tässä teoksessa sekä sanaston että kompleksisten semanttisten ja raken-

teellisten ilmiöiden kielentämisen kehitystä eri-ikäisillä peruskoululaisilla ja nuorilla aikuisilla. Luku- ja kirjoitustaidon saavuttaminen vauhdittaa lasten kielenkehitystä mutta ne eivät ole kehityksen edellytyksiä.

Myöhemmän kielenkehityksen vaihe on todellisuudessa jatkumo (*life span development*) mutta sitä tarkastellaan yleensä vaiheittain siten, että erotellaan esikoulun ja ensiluokkien vaihe, alakoulun keski- ja loppuvaihe, yläkouluvaihe ja sen päätyminen sekä varhaisaikuisuus, josta edetään elinkaaren mukaisesti vanhuuteen asti (ks. esimerkiksi Nippold 2006). Aiemmin ajateltiin, että kyse on ikään kuin kehityskaaresta, jonka alku ja loppu ovat heikomman osaamisen vaihetta (ks. esimerkiksi Jakobson 1968). Ajateltiin toisin sanoen, että vanheneminen vaikuttaisi kielitietoa ja kielenkäyttötaitoja heikentävästi. Nykyisin ajatellaan ennemmin, että suotuisasti vanhentuneen eli esimerkiksi kielelliseen toimintaan vaikuttavilta sairauksilta säästyneen vanhan ihmisen kielenhallinta on nuoremman veroinen tai häntä parempikin, vaikka sanat eivät aina löydy yhtä nopeasti kuin ennen ja vaikka tarinasta tulee välillä syrjähdelyä. Samoin jos aiemmin ajateltiin, että nuori aikuinen on kielellisesti kuin valmis paketti, nykyisin tiedetään, että kielenkehitys ja -hallinta varioi huomattavasti jokaisessa ikävaiheessa. Esimerkiksi nuoret aikuiset ovat vielä vaiheessa, jossa kompleksisten rakenteiden hallinnassa on puutteita, rekisterit ovat melko kapeat ja sanavarasto parhaimmillaankin 20–30 prosenttiyksikköä pienempi kuin vanhimmilla ihmisillä (ks. esimerkiksi Kemper & Sumner 2001; Kavé & Halamish 2015; Keulers, Stevens, Mandela & Brysbaert 2015).

Tässä kirjassa keskitymme kielellisen kehityskaaren keskivaiheille, erityisesti 12-vuotiaista nuoriin aikuisiin, ja pyrimme esittelemään peruskouluikäisten Suomessa syntyneiden ja suomea äidinkielenään puhuvien lasten ja nuorten keskimääräistä kielenhallintaa ja toisaalta näyttämään ikäkausittaisen osaamisen vaihtelun. Nuoret aikuiset ovat sekä vertailukohtana että erillistutkimuksen kohteena. Oletus on, että maahanmuuttajataustaiset peruskoulua Suomessa käyneet lapset ja nuoret tavoittelevat samankaltaista osaamista. Osaamisen erot suhteutuvat Suomessa asumisen pituuteen sekä kielellisten kontaktien määrään ja laatuun mutta myös lukuisiin yksilöllisiin ja sosiaalisiin tekijöihin (ks. tarkemmin Honko 2013; Ruuska 2020).

Lähtökohtanamme on, että myöhempi kielenkehitys alkaa puhutun kielen pohjalta ja että puhutun ja kirjoitetun kielen ydin on sama mutta suhde muuttuu vähitellen vuorovaikutukselliseksi. Oletamme myös, että kielenhallinnassa ei ole suuria sukupuolieroja, vaikka niitä voi olla kielenkäyttötaidoissa, kuten lukemisessa ja kirjoittamisessa. Tutkimus perustuu tutkimushankkeeseen nimeltä ”Suomen kielen myöhempi kehitys: lapset, nuoret ja nuoret aikuiset”, jonka tuloksia on kirjallisuudessa esitelty jo jonkin aikaa. Hankkeen julkaisut etenevät ikään kuin elinkaaren mukaisesti: Alkuvaiheessa esillä oli alakouluvaihe (ks. Pajunen 2012; Honko 2013). Peruskoululaisista ja nuorista aikuisista on myöhemmältä 2010-luvulta useita julkaisuja (Pajunen, Itkonen & Vainio 2015; 2016; Vainio, Pajunen & Häikiö 2019; Pajunen & Itkonen 2019; Laasanen, Pajunen & Häikiö, arvioitavana) ja käsillä oleva kokoomateos keskittyy kokonaan tähän vaiheeseen. Vireillä hankkeessa on vanhempien ihmisten kielenhallinnan tutkimus.

Tässä yhteydessä keskitymme toisaalta semanttisen tason hallintaan (I osio), toisaalta muotojen hallintaan (II osio) peruskoululaisten ja nuorten aikuisten näkökulmasta. Sanatieto muodostaa kielenhallinnan perustan, josta kehitys alkaa ja josta se myöhemmän kielenkehityksen aikana huomattavasti täsmentyy ja tarkentuu määrällisesti ja laadullisesti myös vaikeiden semanttisten ilmiöiden kielentämiseen. Muodot tarjoavat ne ainekset, joilla merkityksiä kielennetään muiden ymmärrettäviksi ilmauksiksi ja teksteiksi. Näkökulmat luonnollisesti limittyvät, vaikka lähtökohta on – merkityksestä muotoon ja muodosta merkitykseen – vastakkainen. Metodisesti osiot I ja II eroavat selvästi: I osiossa analyysi alkaa semanttisten ilmiöiden kartoittamisesta, II osiossa lähtökohtana ovat muodot.

Sekä sanatieto ja kyky kielentää merkityksiä että tieto muodoista kehittyvät vähittäin. Tässä teoksessa pyrimme isoilla koehenkilömäärillä osoittamaan erilaisia kielellisen osaamisen ikäkausimediaaneja ja osaamisen vaihtelua. Tutkimus on toteutettu monimetodisesti käyttäen sekä korpustutkimuksen että kokeellisen tutkimuksen metodeja. Tutkimusaineisto on kerätty kynä-paperitestein, kokeellisesti sekä kirjoittamistehävillä. Perustelemme seuraavaksi tutkimuksen perusoletuksia, nimitäin sanatiedon keskeisyyttä, puhutun kielen mallia sekä sukupuolten

kielenhallinnan samanlaisuutta ja toisaalta ikäkausivaihtelun osuutta. Sen jälkeen esittelemme tutkimuksen aineistot ja koehenkilöt sekä hankkeen tutkijoiden työnjakoa ja kirjan lukujen näkökulmia. Osiot I ja II koostuvat kumpikin neljästä eri luvusta.

Tutkimuksemme kohdistuu suomen kieleen. Teoksessa käytetty käsitteistö on kuitenkin enemmän lingvististä kuin fennististä, koska vertailututkimus on pitkälti kansainvälistä. Myös tavoiteltu lukijakunta ulottuu fennistipiirin ulkopuolelle. Toivomme, että kirja tarjoaa uusia näkökulmia kenelle tahansa lingvistisestä tutkimuksesta sekä suomen kielen myöhemmän kehitysvaiheen tutkimisesta, oppimisesta ja opettamisesta kiinnostuneelle. Kirjan luvut sopivat hyvin yksittäinkin esimerkiksi yliopisto-opintojen kurssikirjallisuudeksi.

PERUSOLETUKSET KIELENHALLINNAN TUTKIMISEEN

Kielenhallinnan tutkimuksessamme perusoletuksia on ollut neljä: Kielenkehityksessä sanatiedon laatu ja määrä on keskeistä ja alussa kehityksen mallina on puhuttu kieli. Kielenhallinta ja sen kehitys ei eriydy sukupuolen mukaan, mutta ajallisesti se voi edetä eritahtisesti. Väestötasolla kielenhallinta on normaaliajakautunutta mutta varioi yksilötasolla ja kehityksellisesti. Varioinnin keskeinen taustatekijä on kielellinen älykyys, mutta siihen vaikuttaa myös muita kielenensisäisiä ja kielenulkoisia tekijöitä. Käsillä olevan teoksen luvut ja teemat ankkuroituvat näihin perusoletuksiin, ja niiden avulla voidaan hahmottaa lasten ja nuorten odotuksenmukaista kielenkehitystä ja poikkeamia siitä.

SANATIEDON KESKEISYYS

Kielen oppiminen on elinikäinen prosessi, joka alkaa sanojen oppimisesta ja jatkuu vanhuuteen (ks. esimerkiksi Ramscar, Hendrix, Love & Baayen 2013). Yksilön sanastollinen tieto ja taito onkin parhaimmillaan vasta iäkkäillä. Esimerkiksi laajoissa Alankomaissa toteutetuissa väestöpohjaisissa nettitesteissä yli 70-vuotiaat tunsivat vajaan 30 000 testisanasta yli 80 prosenttia mutta 13-vuotiaat vain vähän yli puolet ja nuoret aikuiset noin 60–65 prosenttia (Keulers ja muut 2015). Suomen osalta tilanne näyttää samansuuntaiselta (Pajunen, tekeillä). Kuudesluokkalaiset tunsivat testisanoina olleista melko yleisistä liikeverbeistä

hyvin noin puolet ja yhdeksäsluokkalaiset noin 60 prosenttia. Nuorten aikuisten ryhmässä liikeverbien tuttuus varioi erityisen paljon, jos sana oli harvinainen eli sana, jonka esiintymiä on miljoonassa tekstisanassa keskimäärin alle yksi (Pajunen & Itkonen 2019). Yksilön sanavaraston dynaamisuuden vuoksi voitaisiinkin oikeastaan puhua sanavarannosta, joka karttuu ja muovautuu jatkuvasti kielenkäyttökokemuksen myötä.

Ensimmäinen perusoletuksemme on, että kielen hallinnassa ja sen kehittämisessä juuri sanatiedon laatu ja määrä on keskeistä. Siksi sanatiedon arviointi on ollut käsillä olevassa varhaislapsuuden jälkeisessä myöhemmän kielenkehityksen tutkimushankkeessa keskiössä. Sanatiedon hallintaa on hankkeessa tutkittu monella tapaa, ja nämä tutkimustulokset ovat taustoittaneet hankkeen muuta tutkimusta. Sanatiedon hallintaa on arvioitu sekä morfologisten ja semanttisten testien avulla että kielenkäytöstä (ks. Honko 2013; Pajunen, Itkonen & Vainio 2015; 2016; Vainio, Pajunen & Häikiö 2019; tekeillä); toisin sanoen termiä *sanatieto* ei pidä ymmärtää liian kapeasti. Tämän teoksen ensimmäiset neljä lukua käsittelevät sanatiedon ja merkitysten hallintaa joko sana-analyysin tai semanttis-pragmaattisten ilmiöiden kielentämisen näkökulmasta. Kyse on siis itse asiassa kyvystä ilmaista olio- ja tapahtumamerkityksiä ja laajemmin tekstuaalisia merkityksiä.

Sanatason keskeisyyden kielenhallinnassa näkee hyvin esimerkiksi suomen johdoksista, joiden käytön määrä ja leksikaalinen monimuotoisuus lisääntyvät iän mukaan lineaarisesti. Tämä on sanaston kasvun kannalta merkitsevää, koska suomessa suuri osa sanastosta on joko johdoksia tai yhdyssanoja. Johtaminen on keskeisempi ja vaikeampi ilmiö kuin yhdistäminen, koska kaikkien avointen sanaluokkien sanoja johdetaan ja koska johtaminen on rekursiivista eli toistuvaa ja osin sisäkkäistä. Lisäksi sanavartalot varioivat johtamisen yhteydessä ja johtimia, derivaatiotapoja ja derivaatioketjuja on paljon. Alakoulun päättövaiheessa nuori osaa johdoksia sanakohtaisesti mutta ei yleensä sääntöperäisesti. Hän saattaa siis tuntea sanan *käveleskellä* mutta ei yleensä osaa yhdistää muita samalla tavalla johdettuja sanoja ryhmäksi tai ei tiedä, mistä kaikesta muista sanoista samalla johtimella voidaan tuottaa samantapaista lisämerkitystä ilmaisevia sanoja. Suomen tapaisessa synteettisessä kielessä vasta nuorilla aikuisilla johtaminen on sääntöperäistä siten, että

ymmärretään kielessä esiintymättömät sanat mahdollisiksi sanoiksi, jos ne on muodostettu sääntöjä (voimakkaasti) rikkomatta (Koivisto 2013; Vainio ja muut 2019; ks. myös Berman 2007). Sääntöperäisyyden ymmärtäminen on ikään kuin avain systemaattiseen analogiseen mallintamiseen (ks. Itkonen 2005), mutta peruskoululaisia näyttää vielä sitovan enemmän ”joko oikea sana tai sitten ei” -ajattelu. Pieni lapsi osaa toki hyödyntää analogioita mutta ei kovin systemaattisesti. Johdosten käytön lisääntyminen edellyttääkin suomen monimutkaisen sanajohdon sisäistämistä ja produktiivistumista (Bertram, Laine & Virkkala 2000; Kusnetsoff 2017; Vainio ja muut 2019). Sama pätee muista derivaattorikkaista kielistä (ks. Ravid & Avidor 1998; Carlisle 2000; Štekauer, Valeira & Körtvélyessy 2012; Clark 2014). Sanojen johtamisen hallinta eroaa vielä nuorten aikuisten ja peruskoulun päättövaiheessa olevien välillä, toisin sanoen nuoret eivät hallitse derivaatiota vielä 15-vuotiaanakaan täysin sääntötasoisesti tai laadullisesti nuorten aikuisten tapaan (Vainio ja muut 2019). Vanhempien kielenkäyttäjien osalta tutkimustieto puuttuu.

Sanatiedon hyvä hallinta ja laaja sanasto on vahvan luku- ja kirjoitustaidon edellytys (ks. esimerkiksi Carlisle 2000; Qian 2002; Nippold 2006, 25–29, 89; Milton 2009, 170–192; Schmitt 2010, 3–12; Lee 2011; suomen osalta Saarela 1997; Honko 2013; Kusnetsoff 2017). On luonnollista, että lähinnä ikäluokalle tuttua (tai kontekstista pääteltävää) sanastoa ymmärretään luettaessa ja käytetään kirjoituksessa. Luku- ja kirjoitustaito ovat myöhemmässä kielenkehityksessä keskeisiä saavutuksia, mutta ikäluokkatasoinen kielenhallinta on näiden taitojen hallintaan oton perusedellytys. Siten luku- tai kirjoitustaito eivät ole tässä tutkimuksessa varsinainen tutkimuskohde, mutta jos halutaan tukea kehittyvää luku- ja kirjoitustaitoa, on tarkoituksenmukaista seurata ja tukea myös kielenhallintaa.¹

1 Lukutaitoa on Suomessa tutkittu melko paljon – kirjoitustaitoa jonkin verran – mutta pääasiassa esikoulusta koulun alkuun, ei niinkään muiden peruskouluikäisten osalta (ks. Kulju, Kauppinen, Hankala, Harjunen, Pentikäinen & Routarinne 2017; ks. myös Tuovila, Kairaluoma & Mustonen, toim. 2020). Eniten hieman vanhempien peruskoululaisten luku- ja kirjoitustaitoa koskevaa tutkimusta on tehty Opetushallituksen toimeksiannoista. Nämä tutkimukset osoittavat samoja kehityslinjoja, joita tässäkin tutkimuksessa tulee esille. Lappalaisen monet raportit (ks. esimerkiksi 2003) olivat aikanaan yhtenä tämän tutkimuksen innoittajana.

Tässä yhteydessä katsomme eri näkökulmista nimenomaan sitä, mitä lapset ja nuoret keskimäärin osaavat tietyllä ikätasolla ja miten tämä osaaminen muuttuu ja oletettavasti kehittyi ikätasolta toiselle. Lasten ja nuorten osaamista verrataan nuoriin suomenkielisiin (ylioppilastutkinnon suorittaneihin) aikuisiin, joiden jo voi olettaa hallitsevan äidinkieltään riittävällä tasolla, vaikka niin kielenkäyttötottumuksissa kuin kieli- taidossakin tiedetään esiintyvän vaihtelua (ks. esimerkiksi Malin, Sulku- nen & Laine 2013). Tutkimuksemme nuorten aikuisten ryhmä ei edusta koko ikäluokkaansa, koska se ei ole tavoitettavissa yhtä helposti kuin peruskoululaiset. Lisäksi aikuisten osallistumiseen vaikuttaa tehtävän luonne ja sen vaatima aika ja vaivannäkö. Kirjoittaminen on tehtävä- tyyppinä vaativa ja karkotti tutkimukseemme jo muuten suostuneita. Aikuisryhmä edustaa tutkimuksessa lasten ja nuorten tavoitetasoa eli sitä, mihin he voivat pyrkiä ollakseen akateemisilta taidoiltaan itsenäisiä toimijoita.

Käsillä olevassa teoksessa sanatietoa ja merkitysten kiellentämistä kä- sittelee I osion neljä lukua. Niistä ensimmäinen tarkastelee sanaston diversiteettiä, yleisyystasoa ja semanttisia luokkia samasta aiheesta kirjoitetuissa kirjoitelmissa, toinen ja kolmas analysoivat semanttisesti sanatasoa kompleksisempien ilmiöiden kiellentämistä ja neljäs moniää- nisyysden ilmaisun kehittymistä tekstitasolla. Lisäksi sanatietoa käsitel- lään sanahahmon ja sanarakenteen näkökulmista II osiossa. Hankkeen sanatietoa koskevia aiempia julkaisuja ovat muun muassa Honko 2013; Pajunen ja muut 2015; 2016; Honko, Jarvis & Vainio 2019; Vainio ja muut 2019; Laasanen, Pajunen & Häikiö, arvioitavana; sekä Vainio, Pajunen & Häikiö, tekeillä.

PUHUTUN KIELEN MALLI KIRJOITETULLE

Toinen perusoletuksemme on, että lapsen kielenkehitys perustuu alku- vaiheessaan puhuttuun kieleen ja suulliseen syötökseen. Puhuttu kieli on lauserakenteeltaan yleensä melko yksinkertaista ja virkkeet koostuvat useammin ketjulauseista kuin alisteisista ja hierarkkisista rakenteista (ks. Pajunen & Palomäki 1984). Lapsi käyttää vielä myöhemmän kielen- kehityksen alkuvaiheissa puhutun kielen mallin mukaisia perusraken-

teita – teksti etenee peruslauseilla² – ja yhdistää niitä konnektiiveilla (ks. esimerkiksi Tolchinsky 2004, 236; Berman 2007, 357; Hulstijn 2015). Lapsi osaa jo kertoa tarinoita (ks. Suvanto 2012), mutta vain harva kirjoittaa niitä. Lapsen sanasto on yleistä ja konkreettista, ja lapsi saattaa jo ymmärtää konventionaalisia metaforia mutta alkaa tuottaa niitä itse vasta teini-iässä (vrt. Tolchinsky 2004, 238). Koulua aloittava lapsi on aika tyypillisesti kirjaimellisen merkityksen vaiheessa ja kaikki kuvainnollinen kielenkäyttö on vähissä. Tässä vaiheessa lasten väliset kielelliset kehityserot ovat hyvin suuret.

Myöhemmän kielenkehityksen aikana sekä puhutun että kirjoitetun kielen hallinta kehittyy. Lapsi oppii ymmärtämään puhutun ja kirjoitetun kielen modaliteettieroja ja oppii ottamaan erot huomioon tekstejä tuottaessaan. Lisäksi opitaan muut tekstityyppi- ja rekisterierot. Sekä sanaston että syntaksin kehitys on huomattavaa (ks. esimerkiksi Berman 2004; 2007; Nippold 2006). Syntaktisesti kehitys näkyy harvinaisempien rakenteiden haltuunotosta sekä lauseiden, lauserakenteiden ja tekstijaksojen kompleksistumisesta. Rakenteiden muutosta voi mitata erilaisilla tiheysmittareilla, koska muutokset ovat usein myös määrällisiä. Tekstitasolla kehitys näkyy esimerkiksi suunnittelun lisääntymisenä, lukijan huomioimisena, ajan ja paikan aiempaa parempana ankkuroimisena sekä viittausketjujen pitenemisenä. Myöhemmän kielenkehityksen aikana opitaan tekstityypit ja opitaan myös yhdistelemään niitä. Lapsi ainoastaan linkittää tapahtumia, mutta nuorella kertovaan tekstiin alkaa jo ilmestyä kuvailevia ja arvioivia jaksoja (ks. Tolchinsky 2004, 240).

Tässä tutkimuksessa kielellisen rakenteen kehittymistä yksinkertaisesta monimutkaisempaan tarkastellaan kirjoitelma-aineistolla. Aineistonkeruun yhteydessä kirjoitelmatehtävä ohjeistettiin tekstityyppejä tai -lajia määrittelemättä ”kirjoita päivästä – ” (eikä esimerkiksi ”kerro päivästä”). Tuloksena oli iän mukaan pitenevä kirjoitelma yhdestä sanasta parhaillaan kertomusrakenteen (Labov & Waletzky 1967) mukaiseen kirjoitelmaan ja edelleen kertoviin moniäänisiin kirjoitelmiin, joissa

2 Kielitieteessä peruslause (*simple basic sentence*) on yhden predikaation kielennys ja sekä rakenteeltaan että semantiikaltaan yksinkertainen. Peruslause on kaikissa kieliopeissa syntaksin kuvauksen lähtökohhta ja muodoltaan joko intransitiivi-, transitiivi- tai kopulalause (ks. esimerkiksi Itkonen 2008; 2009; Pajunen, Vainio, Hyönä & Bertram, tulossa).

oli kuvailevia ja arvioivia jaksoja. Tarinoiden suullinen kertominen voi olla hyvää jo kielenkehityksen varhaisvaiheessa, mutta niiden kirjallinen tuottaminen yleensä noudattaa kertomusrakenteen skeemaa – eli sisältö koostuu koherentisti – vasta noin 10-vuotiaana (ks. Berman 2007, 348). Omassa hankkeessamme ei ole erityisesti keskitytty kertomuksen kehittymiseen, mutta aineistossamme on havaittavissa selviä viivästyksiä tästä oletusiästä.

Tutkimuksemme varsinainen tavoite on tutkia merkityksien kielentämisen ja muotojen hallinnan kehitystä, ei kertovan tekstin tai tekstilajien hallinnan kehittymistä. Aineistoa analysoitiin kyllä myös tekstitasolla, vaikka se ei ollut ongelmatonta. Ensinnäkin aineiston laadullinen holistinen arvio (onnistuneisuus, tekstityypin tai -lajin toteutumisen aste) osoittautui alakouluikäisten tekstejä lukuun ottamatta vaikeaksi toteuttaa, ja arvioijareliabiliteetti vaihteli liikaa (50–80 %). Sisällön laadullinen arviointi onkin aina jossain määrin subjektiivista (ks. Kuuskorpi 2012). Omat arvioijamme tuntuivat painottavan eri tekijöitä, jotka voi niputtaa joko rakenne- tai sisältökriteereihin. Osa luokitteli tekstejä rakenteen ja koherenssin mukaan (perusasiat kunnossa tai sitten ei), ja osa taas painotti lisäksi tekstin synnyttämää vaikutelmaa (persoonallinen ja omaperäinen tai arkipäiväinen ja lattea). Holistisempi analyysi olisi edellyttänyt kriteerityypistä arviointia, joka olisi perustunut kirjoitelmien kokonaisrakenteen tai jaksoanalyysiin, mutta jo sadankin tekstin, sadoista puhumattakaan, kertomusskeema- ja jaksoanalyysi on valtava urakka (ks. tarkemmin esimerkiksi Juvonen 2004; 2010; Kalliokoski 2006; 2008; Karjalainen 2008; Juvonen, Virtanen & Voutilainen 2012). Toiseksi, vaikka hankkeen kirjoitelma-aineisto on luokiteltu jokseenkin systemaattisesti, kaikki tekstitason muuttujat eivät ole luokituksessa vielä mukana. Sisällöllisesti tekstitason hallinnan kehitys luonnollisesti näkyy koherenssin kautta. Koherenssin lisääntyminen on esillä I osion luvuissa, jotka osoittavat kertomuksellisuuden ja koherenssin yhteyden.

Valinnat ovat perusteltavissa myös siten, että kirjoittaminen on hyvin laaja-alainen ilmiö. Se edellyttää kognitiivista ja neuropsykologista kypsyystta, hienomotoriikan riittävää kehittyneisyyttä, kykyä suunnata tarkkaavaisuutta moniin asioihin yhdellä kertaa, toimivaa episodi- ja työmuistia sekä muun muassa kykyä suunnitella (ks. esimerkiksi Hooper

ja muut 2011). Kirjoittamiseen vaikuttaa voimakkaasti myös kirjoittajan motivaatio, käsitys itsestään kirjoittajana ja usko omiin saavutuksiin (ks. esimerkiksi Pajares & Valiante 2001; Routarinne & Absetz 2013). Lisäksi tiedetään, että taustatekijöinä on myös yhteiskunnallisia ja ympäristötekijöitä. Kehityksellisessä kontekstissa lukemisen ja kirjoittamisen tutkiminen edellyttää kognitiivisen profiilin (kuten muistin, motoriikan) kontrollointia. Suomessa tällaista tutkimusta ovat tehneet erityisesti psykologit ja logopedit (ks. esimerkiksi Niemi 2007; Torppa, Niemi, Vasalampi, Lerkkanen, Tolvanen & Poikkeus 2019). Omassa hankkeessamme emme lingvisteinä voineet kontrolloida neuropsykologisia ominaisuuksia ja kirjoittajien pystyvyydestä emme keränneet tietoa. Motivaation rajaamme tässä yhteydessä niin sanottuun ulkoiseen motivaatioon, esimerkiksi tilannekohtaiseen haluun yrittää tai tehdä jotain.

Puhutun kielen yksinkertaisempi rakenne kompleksistuu myöhemmän kielenkehityksen vaiheessa, kun sekä kielentämisen spesifisyys kasvaa että tekstin rakenne tiivistyy. Toisaalta kielennettävät ilmiöt ja niitä ilmaisevat rakenteet erilaistuvat (Berman 2007). Lapsen ja nuoren pitää kouluikäisenä saada hallintaansa kielellisiä ilmiöitä, jotka eivät ole puhutussa kielessä (yhtä) tärkeitä tai jotka eivät vaikuta puheessa. Usein kyseessä on heijastuma siitä, että puhetilanteessa osallistujat ovat läsnä ja tilanne on yhteinen, joten lisäinformaatiota ei tarvita. Kirjoituksessa lukija ei ole tilanteessa periaatteessa läsnä (vrt. kuitenkin Kallionpää 2014), joten hänen on saatava tilanteinen informaatio haltuunsa muulla tavoin kuin havainnoimalla. Siten puhutun kielen malli ei enää yksin riitä myöhemmän kielenkehityksen aikana.

Myöhemmän kielenkehityksen tutkimushankkeessa on osoitettu muun muassa, että kompleksisista kielenilmiöistä ei voi rakentaa muutoksen ennusteita puhutun kielen muutoksien varaan ja että hyvin kompleksisten ilmiöiden kielentäminen ei onnistu puhutun kielen hallinnan varassa. Pikkulapsilla puhutun kielen malli on ensikielen osalta yleensä melko samantapainen olosuhteista riippumatta – tavalliset rutiinit ja lähiympäristö – vaikka ei vaihtelematon. Lasten kasvaessa mallina on enemmän kirjoitettu kieli, ja tämä malli erilaistuu puhuttua kieltä paljon enemmän yksilöllisesti koulutuksen ja harrastusten mukaan eli yhä enemmän kulttuuristen, sosiaalisten ja poliittistenkin syiden varas-

sa (Tolchinsky 2004, 235–236). Kompleksisten ilmiöiden kielentämistä kuten kehystämistä, referentiaalisen koherenssin ja moniäänisyyden ilmaisua, tutkitaan erikseen kolmessa I osion luvussa. Ne osoittavat, että semanttis-pragmaattisten ilmiöiden kielentämisen hallintaan saanti on jokseenkin hidasta.

Kielen rakenteen hallinnan osalta keskityimme analyyttisiin, kehityksellisiin muuttujiin kuten muotojen kompleksistumiseen. Sanaston ja morfologian kompleksistumisen lisäksi myös syntaktisen kompleksisuuden eli esityksen tiivyyden kasvu osoittautuu kehitykselliseksi. Sanaston tai morfologisen ja syntaktisen tiivyyden lisääntymisessä ei ole mainittavia sukupuolieroja, joskin miespuoliset kirjoittajat suosivat hieman kompleksisempää ilmaisua. Muotomuuttujia käsitellään erityisesti teoksen II osiossa, jossa esillä on puhutun ja kirjoitetun kielen vuorovaikutuksellisuus, jota arvioidaan muun muassa possessiivisuffiksien käyttöön liittyen. Lisäksi osiossa analysoidaan muun muassa äännekirjain-vastaavuuden kehittymistä sekä morfologisten ja syntaktisten muotojen kompleksistumista.

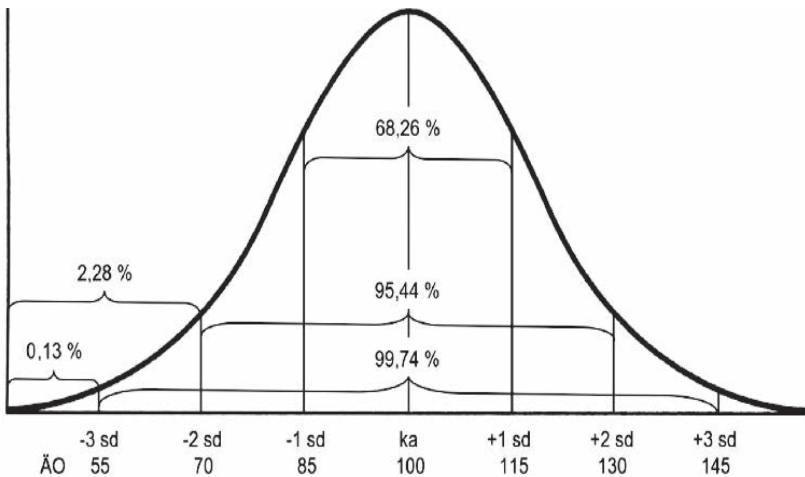
SUKUPUOLTEN SAMANLAISUUS

Kolmas keskeinen perusoletus hankkeessa on sukupuolten kognitiivisten kykyjen keskimääräinen samanlaisuus (*gender similarity hypothesis*, Hyde & Linn 1988; Hyde 2005). Yleisessä älykkyydessä ei katsota olevan sukupuolieroja, vaikka eroja voi olla joissain kognitiivisissa kyvyissä, jotka osaltaan vaikuttavat yleiseen älykkyyteen (Jensen 1998; Hyde 2005; Camarata & Woodcock 2006). Esimerkiksi Keith, Reynolds, Patel ja Ridley (2008) testasivat yleistä älykkyyttä ja kahdeksaa laaja-alaista kognitiivista kykyä ikäluokkiin 6–59 kuuluneella 7 000 koehenkilöllä käyttäen Woodcock-Johnson III -älykkyytestejä. Tuloksena löytyi vain kaksi johdonmukaisesti iästä riippumatta ilmenevää sukupuolieroja. Toinen koski prosessointinopeutta (*processing speed*), joka oli naisilla selvästi parempi, ja toinen verbaalista kykyä (*verbal ability*), joka oli miehillä hieman naisia parempi. Samaa älykkyydestä käyttäneet Camarata & Woodcock (2006) osoittivat, että lasten ja nuorten osalta on joitain sukupuolittuneita eroja mutta että ne ovat jo aikuisilla tasaantuneet. Kun omissa tutkimuksissamme voimme osoittaa, että derivaatiomorfologian tai semanttisen

tiedon keskimääräisessä hallinnassa ei ole mainittavia sukupuolittuneita eroja, tilanne on odotuksenmukainen. Pojilla oli reaktioaikoja ja virheitä mittaavassa derivaatiomorfologian testissä hieman tyttöjä enemmän virhevalintoja, mutta vastaavassa semanttisen tiedon testissä pojat olivat tyttöjä selvästi nopeampia. Reynoldsin, Schreiberin, Hojovskyn, Schwartzin ja Kaufmanin (2015) mukaan oikeinkirjoituksessa ja kirjoittamistaidoissa olisi suurimmat sukupuolierot tyttöjen hyväksi. Samat erot näkyvät aineistostamme ja saavat paljon huomiota samanlaisuusoletuksen vastaisina (ks. Pajunen 2016a).

Sukupuolierot voidaan jakaa dimorfisiin eli kaksijakoisiin ja allomorfiisiin eli tilastollisiin (ks. Heilig 2019). Dimorfiset erot liittyvät yksinomaan lisääntymiseen eivät ominaisuuksien jakautumiseen. Allomorfiset sukupuolierot jakautuvat Gaussin käyrän mukaisesti siten, että enemmistö ihmisistä sijoittuu jakauman keskialueelle, jossa on paljon monien ominaisuuksien päällekkäisyyttä (ks. esimerkiksi kaavio 1). Pienillä jakaumeroilla ei ole päällekkäisyysalueella merkitystä, mutta ne muuttuvat sitä tärkeämmiksi, mikä kauemmas keskiarvosta edetään. On kuitenkin muistettava, että mikään yksittäinen testi ei määritä yksilön asemaa asteikolla; se voi poiketa suuntaan tai toiseen muun muassa vireystilan tai tehtävyyppien tuttuuden vaihdellessa. Mainittakoon vielä, että sukupuolten välisiä kielelliseen älykkyyteenkin kytkeytyviä eroja on pyritty tutkimaan myös aivokuvantamisen avulla. Niin sanotussa Philadelphia-tutkimuksessa noin tuhannen koehenkilön aivokuvantamistulokset osoittivat, että nais- ja miesaivoissa hermosoluyhteydet olisivat erilaisia siten, että miesaivoissa olisi vahvemmat yhteydet aivopuoliskon sisällä ja naisaivoissa aivopuoliskojen välillä (ks. Ingallhalikar ja muut 2014). Tuloksia ei ole yleisesti hyväksytty, koska alkuperäistutkimuksen efekti-koot olivat pienet ja koska toisaalta tuloksia on ylitulkittu jopa niin, että on aihetta puhua neuroseksismistä (ks. esimerkiksi Ribbon 2016).

Pitäydymme siis samanlaisuusoletuksessa. Tässä yhteydessä allomorfisten tilastoerojen voi ymmärtää tarkoittavan, että kielenhallinta on väestötasolla jatkuva ja normaali-jakautunut muuttuja ja että väestötasolla ei ole systemaattisia tai merkityksellisiä kielenhallinnan sukupuolieroja. Voimme olettaa, että vähintään 70 prosenttia suomenkielisestä väestöstä hallitsee äidinkielenä oppimansa kielen ikätasoonsa nähden vähintään



Kaavio 1. Keskihajonnat ja älykkyydosamäärät, alueiden prosentiosuudet (Psyko-logiilitto).

keskimääräisesti ja että pienellä osalla väestöstä (2–3 % asteikon kummassakin päässä) on kielen hallinnassa joko erityisiä vahvuuksia tai suuria ongelmia. Hallinta kehittyi iän mukaan tai voi muuttua patologisesti. Hallinnan kehitys voi sukupuolittain olla eriaikaista, mutta syyt eriaikaisuuteen voivat vaihdella. Oletuksemme on, että hallinnan erot näkyvät parhaimmin harvinaisten, vaikeiden tai kompleksisten ilmiöiden kielentämisestä ja että jakaumaeroja selittää kielellinen älykkyys, vaikka ei yksin. Yksilötasolla osaamista kuitenkin täytyy arvioida monipuolisemmin kuin väestötasolla, jotta satunnaiset kielenulkoiset tekijät tai mittauksen kapeus eivät vaikuta arviointiin.

Hankkeessa on muun muassa sanaston ja semanttisen tiedon osalta saatu melko odotuksenmukaisia tuloksia, kuten jo todettiin, mutta tulokset myös osoittavat, että peruskoululaisilla osaaminen kehittyi hie-man eriaikaisesti. Vaikka keskimääräinen kehitys on monen mitatun kielellisen ominaisuuden suhteen molemmilla lineaarista, tyttöjen kehitys alkaa aikaisemmin kuin poikien. Toisaalta pojat saavuttavat tyttöjä koko peruskoulun ajan ja sama linja jatkuu lukiossa (ks. Harjunen, Marjanen & Karlsson 2019, 37–38). Ajallinen kehitysero voi olla luonteeltaan biologinen, sosiaalinen tai kulttuurisiin syihin kietoutuva, ja syihin ei

sinänsä ole tässä yhteydessä tarve ottaa kantaa. Hankkeessa ei myöskään lähtökohtaisesti ole ollut tarvetta eritellä koehenkilöitä sukupuolen mukaan, eikä koehenkilöiltä ole suoraan kysytty heidän sukupuoltaan. Suomessa viimeisen kahdenkymmenen vuoden aikana käyty laaja keskustelu poikien heikosta äidinkielen taidosta ja erityisesti kirjoittamistaitojen heikkenemisestä (ks. esimerkiksi Lappalainen 2003; 2011) on kuitenkin nostanut sukupuolieron yhdeksi tutkimuksen muuttujaksi.

Uusimman PISA-tutkimuksen mukaan erot näkyvät nykyisin myös lukutaidossa ja muissakin oppiaineissa ja lisäksi tyttöjen ja poikien erot ovat OECD-maiden suurimmat (ks. Leino ja muut 2019; Saarinen 2020). Taitoeroja on havaittu tyttöjen ja poikien välillä niin suomen- ja ruotsinkielisillä kuin suomea toisena kielenä oppivillakin (Silverström & Rautopuro 2015; Kuukka & Metsämuuronen 2016; Kauppinen & Marjanen 2020). Tässä yhteydessä viittaamme äidinkielen hallinnan käsitteellä kuitenkin aineistomme laadun vuoksi ensisijaisesti suomea ensikielellään oppineisiin ja koulussa S1-oppimäärää (suomen kieli ja kirjallisuus) opiskelleisiin lapsiin, nuoriin ja aikuisiin. Tarpeen vaatiessa tarkennamme määrittelyä lukukohtaisesti. On huomattava, että PISA-raportin mukaan (mts. 124) myös oppilaiden välinen sosioekonomisesti selittyvä vaihtelu on kasvanut ja vastaavasti koulujen sijainnin ja oppimistulosten väliset yhteydet ovat voimistuneet (samoin Harjunen ja muut 2019; Saarinen 2020). Sama sosioekonominen vaihtelu näkyy myös omassa aineistossamme, vaikka se ei ole muuttujana mukana. PISA-raportissa kiinnitetään erityisesti huomiota heikkojen (suomenkielisten) osaajien määrän kasvuun, eli toisin sanoen keskimääräisesti osaavien joukko on pienentynyt ja heikkojen kasvanut yli ääriarvojen (vrt. myös Kusnetsoff 2017). Tilanne ei siten enää olisi normaalijakauman mukainen. Tällainen poikkeava muutos pitäisi kyetä selittämään.

Ensimmäiseksi on kuitenkin tärkeintä arvioida, minkälaisia kielellisiä sukupuolieroja on ja miten ne voisi selittää. Erot esimerkiksi morfologian tai semanttisen tiedon hallinnassa eivät selittäneet saman koehenkilöryhmän kirjoittamistaitojen eroja.³ Siten sukupuolierojen tausta-

3 Kuuselan (2011) mukaan heikoimmiksi arvioitujen 9.-luokan kirjoittajien lukutaito ja kielentuntemus on myös heikkoa, ja taso on näkyvissä samoilla oppilailla jo 7.-luokalla. Kielentuntemus määritellään tutkimuksessa melko yleisesti: kieli, sanasto ja peruskäsitteet.

tekijä ei ilmeisesti ole (pelkästään) kielitieto eikä kielellinen älykkyys vaan kyse on todennäköisesti enemmän jostain yhteiskunnallisesta tai opetuksellisesta tekijästä. Ulkoisen motivaation vähäisyys, kielteinen asenne ja pystyvyysuskon puutteet (*self-efficacy*, ks. Bandura 1997; Kuusela 2011; Harjunen ja muut 2019) ovat yhteiskunnallisesti latautuneita tekijöitä, joten vastuu ei voi olla yksin yksilötasolla. Testaamme tässä yhteydessä lähes kaikissa osatutkimuksissa eroja sukupuolen mukaan ja tuloksena on, että kirjoituksen kautta mitattavissa muuttujissa jakaumat eivät todellakaan aina ole odotuksenmukaisia vaan painottuvat liikaa pojilla alakvartiliin. On monia muuttujia, joissa suunta on alaspäin kaikilla peruskouluoppilailla. Kielitiedossa ero ei ole yhtäläinen, mikä vaikuttaa merkitykselliseltä, koska kielenhallinta on luku- ja kirjoitustaidon kehittymisen lähtökohta ja taustatekijä.

KEHITYKSELLISET MUUTTUJAT: HAJONTA VAI VARIAATIO

Neljäs taustaoletus on ollut, että vaikka väestötasolla kielenhallinta on normaalijakautunutta, yksilötasolla, kehityksellisesti ja suhteessa kielen resursseihin vaihtelu on suurta. Hankkeessa on tutkittu vaihtelun määrää ja laatua monella tapaa ja pyritty arvioimaan, liittyykö vaihtelu lähinnä tiettyyn ikätasoon vai liittyykö se myös muihin tekijöihin kuten yksilöihin tai kielennettävään ilmiöön. Edellisen tapauksen vaihtelua nimitämme (ryhmätason) hajonnaksi, jälkimmäistä (yksilötason) variaatioksi. Edellinen oletuksen mukaan vähenee iän mukaan, joten kyse on kehityksellisistä hajontamuuttujista. Jälkimmäinen voi oletuksen mukaan myös lisääntyä iän mukaan, joten kyse on myös ikäluokkatasolla monesta eri syystä varioivista muuttujista.

Pyrimme löytämään ne muuttujat, joissa vaihtelu liittyy ensi sijassa kielenhallinnan suuriin eroihin ikätasolla, ja ne muuttujat, joissa variaatio syntyy myös kielennyksen vaihtoehtojen – resurssien – suuresta määrästä. Edellisessä tapauksessa kyse on siis hajonnasta ja oletus on, että se tasaantuu ikäryhmittäin. Nuoremmilla ikäryhmäkohtainen hajonta voi olla suurempaa kuin normaalijakauma antaisi odottaa, mutta tyypillisessä kielenkehityksessä sen pitäisi vähetä selvästi iän mukaan (ks. esimerkiksi Metsämuuronen 2006; Harjunen, Marjanen & Karlsson 2019). Selvitämme, mikä on peruskoululaisilla se keskimääräisen hal-

linnan alue, jonka alittaminen on hyvä huomata ja johon opetuksessa voi reagoida. Vastaavasti pyrimme osoittamaan, missä määrin on kyse myös resurssihin liittyvästä variaatiosta. Luonnollisesti sekä hajonta että variaatio linkittyvät samoihinkin muuttujiin, mikä tekee arvioinnista hankalaa. Esimerkkinä voi mainita sanaluokkaindeksit, joissa sama arvo voi viitata kielitaidon kehittyneisyyteen tai kehittymättömyyteen. Esimerkiksi korkea substantiivitiheys voi toisaalta tarkoittaa tiivistä ja argumendoivaa tyyliä, toisaalta sanaluettelomaisuutta. Päätelmiä ei siis pidä tehdä yhden muuttujan jakaumista.

Oletus on, että hajonnan vähenemistä ilmaisevia kehityksellisiä muuttujia löytyy enemmän teksti- ja muototasolla, kehityksellisiä ja varioivia enemmän merkityksen kielentämisessä. Tämä selittyy yksinkertaisimmillaan siten, että yksikköjä on muototasolla suhteellisen vähän ja merkitystasolla paljon. Kehityksellisiä muuttujia tutkimme koululaisryhmässä pääasiassa näennäis- ja reaaliaikametodilla vertailukohtanamme nuoret aikuiset. Kielentämiserojen kautta syntyvä variaatio näkyy luotettavimmin aikuisikäluokkatasolla, jossa kielenhallinta on jo vähintään perustasolla eli jossa kielenhallinnassa ei ole oleellisia puutteita. Peruskouluikäisillä hajonta- ja varioivia muuttujia on paikoin hyvin vaikea pitää erillään ilman yksityiskohtaista laadullista analyysiä.

Varioivia muuttujia eli kielentämisen erojen lähteitä on viime aikoina tutkittu erityisesti tapahtumien kielentämisen yhteydessä ja on voitu osoittaa, että lähteitä on yksilöstä, stimuluksesta ja kielellisestä toimintaympäristöstä, havainnosta ja ympäristötekijöistä koulutuksen määrään (ks. Bohnemeyer ja muut 2014; Palmer, Lum, Schlossberg & Gaby 2017). Yksilöllistä variaatiota voi syntyä esimerkiksi sen mukaan, että perkeptiivinen ja kognitiivinen prosessi, jossa huomio (*attention*) kiinnittyy yksittäisiin ärsykkeisiin, on hyvin valikoiva ja yksilöittäin vaihteleva (Baker & Levin 2015). Koehenkilö voi tilanteittain yksinkertaisesti olla näkemättä jotain tai tulkita näkemänsä väärin (*change blindness*, ks. Gibbs, Davies & Chou 2016). Lisäksi siinä missä yksi näkee yksityiskohtia ja kielentää ne, toinen kielentää yleisemmällä tasolla. Muisti, näkökyky ja stressi vaikuttavat huomiokykyyn.

Kielentämisstrategiat voivat vaihdella yksilöllisesti mutta myös kehityksellisesti. Pieni lapsi voi kertoa lukijaa huomioimatta pitäen itseään

keskipisteenä, nuori voi kehystää tarinaansa enenevästi, mutta yleensä vasta aikuinen tiivistää ajallisesti ja vaihtelee tarinan tapahtumajärjestystä ja näkökulmia. Iän myötä kyky suunnitella kasvaa. Variaatiota syntyy myös kielennykseen käytettävissä olevista kielen resurssien määrästä ja niiden vaihtelevasta hallinnasta. Jos käytettävänä on vain yleisiä sanoja tai muotoja, kerronta on monen mitattavissa olevan muuttujan kannalta erilaista kuin jos tuntee ja käyttää yleisten sanojen tai muotojen spesifimpiä vastineita. Myös tekstityyppien iän mukana kasvava monipuolisuus vaikuttaa variaatioon ja toisaalta monet mittarit reagoivat tekstin ominaisuuksiin. Jos kirjoitetussa kielessä on paljon kehityksellistä variaatiota, variaatio voi perustua myös kulttuuriseen tai muuhun yksilölliseen erilaistumiseen.

Kehityksellinen hajontavaihtelu on esillä pääosassa teoksen luvuista, mutta myös kielentämisen (kehitykselliseen) yksilövariaatioon kiinnitetään kaikissa teoksen luvuissa huomiota. Erikseen sitä tarkastellaan spatiaalisen kehystämisen yhteydessä, muun muassa koska siihen suomen kielessä on tarjolla paljon kielen resursseja ja toisaalta yksilöllisiä kielentämisstrategioita. Tekstitasoista variaatiota tarkastellaan muun muassa possessiivisuffiksien käyttöä tutkimalla. Molemmissa luvuissa koehenkilöt ovat pääasiassa nuoria aikuisia. Kehityksellinen vaihtelu tulee esiin myös tekstin koherenssin – referentiaalisen, spatiaalisen, temaattisen ja temporaalisen – kehittymisen yhteydessä (ks. esimerkiksi Givón 1993; 2017).

Painopiste on kehityksellisissä varioivissa muuttujissa I osion luvuissa ja kehityksellisissä hajonnan pienenemistä ilmaisevissa muuttujissa II osion luvuissa. Edellisessä lähtökohtana on kielennettävän ilmiön semanttinen ja laadullinen analyysi, jälkimmäisessä muotojen käyttö. Edellisessä variaatio voi lisääntyä iän myötä, jälkimmäisessä vaihtelu voi vähetä, kun saavutetaan ikäluokkatasoinen keskimääräinen taso. Molemmissa tapauksissa taustalla on luonnollisesti myös kielelliset älykkyyserot.

TUTKIMUSKYSYMYKSET JA -METODIT

Edellä käsiteltyjen taustaoletusten kautta muotoutuvat myös tämän teoksen keskeisimmät tutkimuskysymykset:

- 1) Miten sanatiedon ja semanttisen tiedon hallinta kehittyi?
- 2) Miten kielen hallinta muotojen tasolla kehittyi?
- 3) Miten sukupuolierot näkyvät kielenhallinnassa?
- 4) Milloin vaihtelu tasaantuu ja milloin ei, eli milloin vaihtelun väheneminen on kehityksellistä ja milloin kyse on myös kielellisestä ja yksilöllisestä variaatiosta, jonka määrä ei ole sidoksissa pelkästään kehitykseen?

Kielenhallintaa on hankkeessa tutkittu monimetodisesti ajatuksena, että käyttämällä sekä korpus- että kokeellisia metodeja voidaan tutkittavasta ilmiöstä saada kokonaisvaltaisempi kuva. Korpusaineistoa on usein vaikea arvioida ilman tietoa koehenkilöreaktioista tai ilman laadullista etukäteistietoa, joten tietoa on kerätty myös kyselyin ja elisitoimalla (ks. Itkonen & Pajunen 2010). Aineiston ensivalinta ja analyysi on kuitenkin aina ollut luonteeltaan laadullista käsiteanalyysiä.

Sekä korpus- että testiaineisto analysoidaan laadullisesti ja tilastollisesti. Laadullinen analyysi nojaa aineistojen sana-, lause- tai tekstityyppikoodaukseen ja tilastollinen analyysi on tässä yhteydessä ennen kaikkea kuvailevaa, koska tavoitteena on osoittaa osaamisen vaihtelu. Teoksen tehtävä ei ole osoittaa erojen tilastollista merkitsevyyttä, mutta isot koehenkilömäärät takaavat sen, että eri mittarien kautta arvioituina erot ikäryhmittäin ovat tilastollisesti pääasiassa merkitseviä. Muissa yhteyksissä julkaistuissa artikkeleissa tuloksia on analysoitu myös monimuuttujamenetelmin.

Tutkimushankkeen ja tämän teoksen keskeisin ja tärkein tutkimustulos on, että peruskoululaisilla ikäkausiosaamisen hajonta on hyvin suurta. Lisäksi tendenssinä on, että koko populaatioissa alakvartiilitason osajien määrä kasvaa. Myös muissa yhteyksissä on todettu, että äidinkielen hallinnan erot alkavat olla tasavertaisten opiskelumahdollisuuksien takaamisen kannalta liian suuria (vrt. myös Pöysä & Kupiainen 2018). Äidinkielen hallinta peruskoulun loppuvaiheessa näyttää ennakoivan melko voimakkaasti esimerkiksi menestymistä myöhemmin ylioppilaskokeissa (Harjunen ja muut 2019). Tämä on nähtävissä sekä äidinkielen ja kirjallisuuden kouluarvosanan ja tutkintomenestyksen välisenä korrelaationa että oppiainearvioinnista riippumattomalla tekstitaitomittarilla

arvioituna (mts. 11–12). Viimeisimmän lukutaitoa mittaavan PISA-kartoituksen perusteella heikkojen lukijoiden osuus on kasvanut Suomessa selvästi eli 8,1 prosentista 13,5 prosenttiin (OKM 2019). Tämä merkitsee, että Suomessa on yhä enemmän nuoria ja nuoria aikuisia, joiden kielilliset taidot voivat rajoittaa heidän itsenäistä toimintaansa tekstittyneessä yhteiskunnassa. Tilannetta ei voi selittää kielellisen älykkyyden väestötason jakaumalla, ja suurimmat ongelmat koskevat poikia ja heidän jatko-opiskelumahdollisuuksiaan.

AINEISTO

Hankkeen tutkimukset perustuvat empiiriseen aineistoon. Aineisto on kerätty useammassa vaiheessa, ja sisällöllisesti se on jaoteltavissa sähköiseen korpusaineistoon, kyselyin ja testein tuotettuun aineistoon sekä kirjoitustehtävillä koottuun aineistoon. Useimmiten yksi ja sama koehenkilö on tuottanut useammanlaista aineistoa samalla kertaa tai eriaikaisesti, mikä mahdollistaa vertailun kielen eri osa-alueilla tai kehityksellisesti. Pääosin kehitystä kuitenkin tutkitaan näennäisaikametodin avulla.

SÄHKÖISET KORPUKSET

Korpuukset koostuvat sanomalehtiaineuksesta, fiktiosta, alakoulun oppikirjoista ja nuorten aikuisten ylioppilaskirjoitelmista, ja ne ovat kaikki sähköisessä muodossa ja automaattisesti koodattu. Sanaesiintymämäärä on yhteensä noin 45 miljoonaa. Lehtiaines on *Helsingin Sanomista* (Pajunen & Virtanen 2002a), *Savon Sanomista* (Pajunen & Virtanen 2002b) ja *Turun Sanomista* (Laine & Virtanen 1999b). Fiktioaines on 2000-luvulla ilmestyneistä, lähinnä jännitys- tai salapoliisiromaaneiksi luokiteltavista teoksista ja oppikirjat edustavat 2000-luvun alakoulun oppikirjoja (ks. Jaakola 2004). Kirjoitelmakorpus on koostettu ylioppilaskirjoitelmia esittelevistä kirjoista ja sen sanamäärä on yhteensä miljoona.⁴ Kun muut korpusaineistot edustavat 2000-luvun kieltä, kirjoitelmakorpu-

4 Kirjoitelmakorpus on koostettu Suomalaisen Kirjallisuuden Seuran ja Äidinkielen opettajain liiton julkaisemista ainekokoelmista, jotka aluksi keskittyivät niin sanottuihin valioaineisiin eli erityisen hyviksi arvioituihin kirjoitelmiin. Myöhemmin valikoima on monipuolistunut eri arvosanoja saaneisiin kirjoitelmiin. Laadulliset sisällölliset erot eivät ole tutkimuksessamme mukana. Korpus ei ole julkinen.

sessä on tekstejä vuosilta 1948–2003. Lisäksi aineistossa on pieni määrä 1970-luvun Lauseopin arkiston materiaalia ja siihen pohjavia kirjoitettua ja puhutun kielen tietoja (ks. Pajunen & Palomäki 1984). Lauseopin arkiston (= LA) aineisto on 1970-luvun kirjoitettua yleiskieltä Oulun korpuksen jakauman mukaan koostettuna (sanomalehdet, aikakauslehdet sekä ei-fiktiivinen kirjallisuus, ks. Saukkonen 1982) tai se edustaa aluemurteita.

Korpuksset – LA:ta lukuun ottamatta – on koodattu morfosyntaktisesti FDG-jäsentimellä (*functional dependency grammar*, ks. Voutilainen 2002) ja ne on koottu konkordanssityyppisiksi tietokannoiksi (Laine & Virtanen 1999a; Virtanen & Pajunen 2001). Korpuksista voi tehdä hakuja lekseemin tai sanamuodon, sanaluokan, funktion, muodon ja frekvenssin avulla. Tarpeen vaatiessa hakusanan esiintymäympäristöä voi rajata paitsi edellä mainituin muuttujin myös etäisyyden tai pääsanan suhteen. Sähköisiä korpuksia on käytetty frekvenssien ja vertailulukujen laskemiseen ja testisanojen vertaistamiseen. Possessiivisuffiksien käytön tutkimus perustuu sähköiseen korpusaineistoon.

TESTIT JA KIELENTÄMINEN

Testien avulla on elisitoitu eli kirjoiteltu suunnitelluilla asetelmilla monenlaista kielellistä ainesta. Testien yhteydessä on tuotettu myös kielentämis- ja kirjoittamistehtäviä. Kirjoitustehtävänä on ollut joko kirjoitelma unelmapäivästä, kuvan tai tapahtuman kielentäminen tai kuvasarjatarinan kertominen. Aineistot on kerätty vuosina 2007–2020. Testauskokoontenpanoja on useammanlaisia:

- a) Määritelmätesti, sana-assosiaatiotesti, derivaatiotesti (kaikki paperi- ja kynätestejä), kirjoitelma unelmapäivästä, tietoa morfologian hallinnasta ja muun muassa ylioppilastutkintotodistuksesta. Koehenkilöinä on nuoria aikuisia yhteensä 240 (ks. myös Pajunen ja muut 2015; 2016; Laasanen ja muut, arvioitavana).
- b) Semanttisen ja morfologisen leksikaalisen päätöksen kaksi testiä, molemmat elektronisia online-testejä reaktioaikoinen, kirjoitelma unelmapäivästä. Koehenkilöinä on kuudes- ja yhdeksäsluokkalaisia sekä nuoria aikuisia yhteensä 300 (myös Vainio ja muut 2019; Vainio, Pajunen & Häikiö tekeillä).

- c) Määritelmätesti, sana-assosiaatiotesti, vapaa assosiaatiotesti, kirjoitustehtävänä kuvan kielentäminen, tuttuustesti, reitin kielentämistesti. Koehenkilöinä on 157 kuudes- ja yhdeksäsluokkalaista (Pajunen, tekeillä). Testit antavat tietoa peruskoululaisten sanastonhallinnasta ja erityisesti spatiaalisten ilmiöiden kielentämisestä.
- d) Sanatiedon hallintaa mittaavat testit ja kirjoitelma, taustatietolomake, tieto arvosanoista ja opettaja-arvio. Pitkittäistutkimus 3.–4.- ja 5.–6.-luokkalaista, kaikkiaan noin 600 S1- ja S2-koehenkilöä. (Tuloksista ks. Honko 2013). Luku norminmukaisuuden hallinnan kehityksestä perustuu tähän aineistoon, mutta Hongon laaja aineisto (2013) taustoittaa koko hanketta monella tavalla. Tämä on hankkeen ainoa osio, jossa S2-oppijoita on edustavasti mukana.
- e) Suullinen videoitujen liike- ja tilatapahtumien kielennys *Trajectoire*-videosarjan (ks. Ishibashi, Kopecka & Vuillermet 2006) ja fiktiivisen liikkeen (ks. Blomberg 2014) kuvasarjan avulla. Koehenkilöinä on 50 nuorta aikuista. Testit antavat tietoa nuorten aikuisten spatiaalisten ilmiöiden kielentämisestä. *Trajectoire* koostuu 76 videoklipistä, joista 55 esittää varsinaisen liiketilanteen, kaksi on tehtävänannon yhteydessä katsottuja harjoitusvideoita ja loput ovat ”fillereitä” eli mukana tutkimuksen tarkoituksen hämärtämiseksi. Videot ovat kymmenisen sekuntia pitkiä, ja niissä on näyteltyjä liiketilanteita vaihtelevissa luonnonympäristöissä. Kuvasarjassa on 36 tilaa ilmaisevaa piirrettyä kuvaa.
- f) Suullisia nimeämis- ja kielennystehtäviä; Olion- ja tapahtuman nimeäminen 800 kuvan avulla (vrt. Szekely, D’Amico, Devescovi, Federmeier, Herron, Iyer, Jacobsen, Arévalo, Vargha & Bates 2005), koehenkilöinä 60 aikuista. Nimeäminen 100 kuvan avulla, koehenkilöinä 50 nuorta aikuista. Kuvakirjakerronta Mayerin (1969) *Frog, where are you?* -kuvakirjasta (ks. Strömqvist ja Verhoeven 2004), koehenkilöinä 35 nuorta aikuista. Testeistä osa on toteutettu elektronisina. Testit antavat tietoa nuorten aikuisten sanastonhallinnasta.
- g) Joukko semanttisia subjektiivisia testejä, joissa on koehenkilöarvioita tiettyjen sanojen tuttuudesta, konkreettisuudesta, kuvitteellisuudesta, abstraktiudesta tai sanan herättämästä tunteesta tai sen

merkityksestä asteikolla positiivinen–negatiivinen (ks. esimerkiksi Söderholm, Häyry, Laine & Karrasch 2013; Citron, Cacciari, Kucharski, Back, Conrad & Jacobs 2016). Koehenkilöinä on yhteensä noin 200 peruskoululaista ja nuorta aikuista. Subjektiiivisten testien avulla on vertaistettu varsinaisia kokeita, mutta niiden avulla voi selittää myös lasten ja nuorten tekemiä valintoja ja arvioida sanaston kehitystendenssejä.

Testit ovat antaneet runsaasti tietoa sanatiedon laadusta ja määrästä kouluikäisillä ja nuorilla aikuisilla. Testeillä tuotettu aineisto on analysoitu soveltuvin kielensisäisin ja -ulkoisin muuttujin Excel-tietokantoihin siten, että samojen koehenkilöiden eri testitulokset voi yhdistää. Lisäksi on laskettu erilaisia frekvenssivertailulukuja. Elektroniset testit on toteutettu ohjelmalla Pohjankukka & Vainio (2013) tai ohjelmalla E-Prime-3 joko ryhmittäin samassa tilassa tai yksilötestauksin. Kaikki käytetyt testityypit, myös kynä-paperitestit, noudattavat yleisesti kielitietoa testaamaan käytettyjä testiformaatteja, jotka tutkimusryhmä on soveltanut suomen kieleen. Koska niitä on käytetty onnistuneesti samantapaisissa konteksteissa aiemmin, niitä voi pitää valideina kielenhallinnan osa-alueiden testaukseen.

UNELMAKIRJOITELMAT

Kirjoitelma-aineisto on kerätty vuosina 2006–2010 ja 2012–2014 ja se koostuu alakoululaisten, yläkoululaisten sekä nuorten aikuisten kirjoitelmista (ks. taulukko 1). Vanhimman aineiston ytimenä on Maija Purasen ja Annukka Yli-Paavolan keräämä alakoululaisaineisto, jota on muuttaman koulun erillisaineistolla täydennetty parina alkuperäiskeruuta seuranneena vuonna (= niin kutsuttu Koulukorpus, ks. Pajunen, Puranen ja Yli-Paavola 2010). Alakouluaineistossa on kaupunki-, taajama- ja haja-asutusalueen kouluja sekä kielikoulu- ja kielikylypyluokka. Koko alakoulun aineisto on kerätty maaliskuussa, jotta mahdollisimman moni ensimmäisen luokan oppilaista jo osaisi lukea ja kirjoittaa. Maaliskuu on myös alakoulun lukutestin vertailukuukausi (Lindeman 2000, 3). Vuosina 2013 ja 2014 kerättyyn kouluaineistoon kuuluu alakoulun kuudes luokka ja yläkoulun aineistoa. Keruu on toteutettu vastaavilta alueilta ja

myös siihen on valikoitunut yksi kieliluokka. Vanhempien koululais-ten osalta keruuajalla ei katsottu olevan enää merkitystä, koska kielelli-set valmiudet ovat paremmat. Tähän aineistoon eli niin kutsuttuun Unelmakorpuksen (ks. Pajunen 2016b) on yhdistetty myös aiempi kuudensien luokkien aineisto. Näin mahdollistuu eri aikoina kerätyn kuudensien kirjoitelmien vertailu. Unelmakorpuksen kuuluu paitsi pe-ruskoululaisten kirjoitelmia myös nuorten aikuisten kirjoitelmat, jotka on kerätty suomen kielen pääsykokeen yhteydessä tai ammattikorkea-koululaisilta vuosina 2012–2014. Koulukorpuksen rinnakkaisaineisto (= ns. S1–S2-korpus), jossa on yhteensä 550 S1- ja S2-kirjoittajien kirjoi-telmaa, on kerätty vuonna 2007 (toinen ja kolmas luokka) ja 2010 (vii-des ja kuudes luokka). Unelma-aineistoihin viitataan joko korpuksen nimellä tai lyhyemmin esimerkiksi sanalla Unelmakirjoitelmat.

Taulukko 1. Koulu- ja Unelmakorpuksen kirjoitelmien määrä.⁵

Ikäluokka	F	M	Kaikki
Alakoulu 1.–5.	481	466	947
Kuudennet	162	158	320
Yläkoulu	137	133	270
Aikuiset	244	98	342
Yhteensä	1024	855	1879

Unelmakirjoitelmat on kirjoitettu samasta aiheesta eli unelmien päivästä samalla tehtävänannolla: ”kirjoita unelmiesi päivästä, jonka olet kokenut, haluaisit kokea tai voisit kokea maailmassa, missä mikä tahansa olisi mahdollista” (ks. Pajunen 2012). Aikaa kirjoittamiseen on ollut yksi oppi-tunti ja etukäteissuunnittelua ei ole ollut.⁶ Kirjoitelmien kokonaissana-määrä kasvaa lineaarisesti peruskoulun ikäluokasta toiseen yhdestä sa-nasta muutamaan sataan, aikuisia puolestaan on pyydetty kirjoittamaan noin 300 sanan kirjoitelma (ks. taulukko 2). Sanaesiintymien määrä koko aineistossa on noin 234 000. S1–S2-korpus on kirjoitettu aiheesta Unelmien päivä (ks. Honko 2013), ja sen sanamäärä on noin 60 000.

5 F = tytöt ja naiset, M = pojat ja miehet.

6 Sintonen (2011) vertaili suunnittelemattomia ja suunniteltuja Unelmakirjoituksia pro gradu -tutkiel-massaan. Tuloksena oli, että suunnittelu toisaalta nostaa tekstin tasoa, toisaalta yhdenmukaistaa kirjoitelmia.

Tehtävänanto on ollut Unelmakirjoitelmia täsmällisempi ja kirjoittajia on opastettu kirjoittamaan satu tai tarina. Alussa on keskusteltu unelman käsitteestä. Aiheen ja ohjeistuksen ero näkyy aineistossa jonkin verran.

Taulukko 2. Sanaesiintymien määrä Unelmakirjoitelmissa.

Ikäluokka / Sanaesiintymä	n	Keskimäärin kirjoittajittain
Alakoulu 1.–5.	63 641	67,2
Kuudennet	42 605	133,1
Yläkoulu	38 365	142,1
Aikuiset	89 192	260,8
Yhteensä	233 803	

Kirjoittajien motivaatio (halu suostua kirjoittamaan tai viitseliäisyys) näyttää vaihdelleen yksilöittäin ja myös ryhmittäin.⁷ Uusimmissa keruissa motivaatiota on pyritty tasaamaan arpomalla kirjoittajien kesken pieniä palkintoja. Motivaation nostatus näyttää kasvattaneen sanamäärää mutta ei ole niinkään lisännyt rakenteen suunnittelua – ja kirjoitelman tasoa –, mikä näkyy, jos verrataan vuosien 2006 ja 2014 kuudensien aineistoja keskenään (ohjatun suunnittelun vaikutuksesta ks. Sintonen 2011, motivoinnista ks. myös Kyösti 2018; motivaation osatekijöistä ja vaikutuksesta ks. Troia, Harbaugh, Shankland, Wolbers & Lawrence 2013). Aikuisista ammattikorkeakoululaisille on maksettu rahapalkkio; pääsykokeeseen osallistuville itse koetilanne on ollut motivoiva.

Koko Unelma-aineisto (eli Koulu-, Unelma- ja S1–S2-korpus) on ensin kirjoitettu sähköiseen muotoon teksteiksi ja koodattu myöhemmin manuaalisesti sana kerrallaan Excel-tietokantaan. Jokainen sanaesiintymä on luokiteltu sanarakenteen ja sanaluokan mukaan sekä syntaktisesti että semanttisesti. Kaikki sanaesiintymät on myös kantasanaistettu eli muutettu lekseemeiksi. Aineisto on koodattu myös sanatasoa ja lausetta laajempiin yksiköihin, joista voidaan arvioida muun muassa kirjoitel-

7 Aineiston kerääjät välittivät usein koehenkilökommentteja ja kuvailivat keruutilanteen tunnelmaa tai häiriötekijöiden määrää; nämä antoivat tukea päätelmiin kirjoittajien (ulkoisesta) motivaatiosta. Koehenkilöt saattoivat myös kysyä, vaikuttiko suoritus oppiainearviointiin; vastaus luonnollisesti oli kielteinen. Sisäinen motivaatio ei ole tällaisilla havainnoilla arvioitavissa, joten viittaamme termillä *motivaatio* juuri ulkoiseen motivaatioon.

mien syntaktista hierarkkisuutta ja erilaisia leksikaalisia suhteita. Lisäksi esitetään tekstityyppiä ja koko kirjoitelmaa koskevia holistisia arvioita, mutta on syytä muistaa, että ei ole olemassa sellaisia holistisia hyvän kirjoitelman kriteerejä, joista vallitsisi yksimielisyys (ks. esimerkiksi Koutsoftas & Gray 2012, 407; Bouwer, Béguin, Sanders & van der Bergh 2015, 95). Tasomuuttujaa käytettiin vain alakouluaineistossa, koska siinä tasoerot olivat selvän kehitykselliset. Aineiston koodaajina on ollut useita henkilöitä, ja koodausta on testattu ja yhtenäistetty määrittelemällä tarkat sana-, sanaluokka- ja sanarakennekriteerit, semanttisen analyysin kriteerit sekä lause- ja tekstitason analysointikriteerit perus- ja rajatapauksineen. Muuttujien määrittelyä on usein jouduttu tekemään sekä sana- että tekstikontekstissa. Rajanveto on usein vaikea vaikkapa substantiivien ja adverbien välillä. Aineiston koodisuunnitelmassa on yli 60 sivua, ja juuri koodaus on ollut tutkimuksen työläin vaihe. Tulokset luonnollisesti perustuvat aineiston yksityiskohtaiseen koodaukseen, joka on sama kaikissa osatutkimuksissa. Korpuksia on käytetty myös kymmenissä opinnäytetöissä, ja opiskelijat ovat saaneet aineistonsa valmiiksi koodattuna. Siten nämä opinnäytetöiden laadulliset analyysit täydentävät hankkeen tutkimusta luotettavasti.

Koulu- ja Unelmakorpuksen tehtävänannossa ei viitata tekstityyppiin mutta otsikko on ohjannut kertovaan tekstiin, jossa tapahtumia kuvataan kronologisesti aamusta iltaan. S1–S2-kirjoitelmissa ohjeistettiin kirjoittamaan satu tai tarina, mutta lopputulos ei ratkaisevasti eroa muista. Aineistosta muodostuu kehityksellisesti kuvaava kaari kertomusrakenteen tavoittelusta aina enemmän kuvailevaa ja erittelevää ainesta sisältävään kerrontaan. Alakoulun yläluokilla kirjoitelmat alkavat tavoittaa kertomusrakenteen skeemaa (Pajunen 2012) ja kaikkein kehittyneimmät yksittäiset kirjoittajat arvioivat samalla unelmapäivän käsitettä jo kuudennella luokalla. Aikuisista suurin osa ainakin jossain määrin erittelee tai kuvailee. Vastaavasti irrottaudutaan tapahtumia pinnallisesti luettelevasta päiväskeemasta (*aamulla heräsin – illalla menin nukkumaan*). Petälän (2018) tutkimuksen mukaan keski-ikäiset kirjoittavat vastaavasta teemasta siten, että irtautuvat sekä tapahtumajärjestyksestä että kertomusrakenteesta. Teksti säilyy silti pääpiirteissään kertovana, vaikka evaluaation määrä nousee. Heinosen (2018) mukaan viisi yle-

sintä aihetta kuudes- ja yhdeksäsluokkalaisilla sekä nuorilla aikuisilla on harrasteet, ravinto, skeemaa jäsentävät triviaalirutiinit, ystävät sekä hyödykkeet. Nuorimmat kertovat juhlista, ystäväistä, lemmikeistä ja huippulomapäivistä; rahalle annetaan suuri merkitys. Vanhemmilla omaperäisyys lisääntyy, mutta toisaalta painotetaan aika samansuuntaisesti ystävyysuhteita ja arjen pienten tapahtumien onnea. Monet kirjoittelmissa ovat koskettavia, monet hauskoja, monet surullisia. Jo peruskoululaisilla mutta ennen kaikkea nuorilla aikuisilla kirjoitelma on elämäntarinatyyppinen (vrt. Habermas & de Silveira 2008; McLean & Breen 2009): se päivä elämässä, minä jotain käänteentekevää, merkityksellistä tapahtui tai odotetaan tapahtuvan. Kyse voi olla häistä, hautajaisista, opiskelemaan pääsystä, lapsen syntymästä tai vaikkapa unelmien matkasta. On tyypillistä, että tämäntyyppisissä kirjoitelmissa omaelämäkerrallisuus iän myötä lisääntyy (ks. Pasupathai & Mansour 2006).

Kirjoitelmien tekstityyppi on siis kertova ja kirjoitelmien tavoitemuoto on tyypillisesti kertomus tai omaelämäkerta. Kertomus on määritelmän mukaan tapahtumien representaatio ja tyypillisesti tapahtumat kerrotaan tapahtumajärjestyksessä. Kerrontaa sitoo sisäinen aikarakenne (ks. esimerkiksi Berman & Slobin 1994). Kirjoitelmissa suurimpia verbi-luokkia ovat odotuksenmukaisesti tapahtumia ja liikkeitä ilmaisevat tapahtuma- ja liikeverbit, ja suurin adverbiryhmä koostuu aikaa ilmaisevista. Käännekohdat kuljettavat kertomusta eteenpäin. Alakoulussa kertomustavoitemuodon saavuttamista mitattiin kertomusskeema-analyysillä (Pajunen 2012; Honko 2013) ja narratologisilla kriteereillä (Vesala 2012). Jos narratiivisuus määritellään hyvin tiukasti, suurin osa vielä 3.-luokkalaistakaan ei tuota kertomuksia. Noin 60 prosenttia kuudesluokkalaisten kirjoittelmissa on heikkoja tai vahvoja kertomuksia, mutta sana- tai lauseluettelaita tai muulla tavoin keskeneräisiä kirjoitelmia löytyy myös.

Kertomuksissa kertovan tekstin oheen syntyy arvioivia ja kuvailevia osuuksia; tämä on selvä kehityksellinen piirre. Örnmarkin (2011) mukaan argumentoivassa tekstissä monen tekstityypin yhdistelmät ovat ennemmin heikon kirjoittajan ominaisuus: erityisesti tukeutuminen (minäkeskeiseen) kertovaan kieleen argumentoivan sijasta on heikkouden signaali. Siten kehitykselliset tekstiominaisuudet näyttävät vaihte-

levan tekstityypeittäin. Esimerkiksi nuorilla aikuisilla yksinomaisesti kertova teksti vaikuttaa lapsekkaalta, koska lukija odottaa aikuiselta jo perusteltuja siitä, miksi jokin asia on tärkeä, tai edes tärkeyden kuvailemalla näyttämistä. Aineiston arvioivat kirjoitelmat, joissa ei ole lainkaan muiden tekstityyppien ominaisuuksia, rajoittuvat pariin tekstiin. Jos kirjoitelma keskittyy arviointiin, arvioinnin kohteeksi näyttää nousevan unelma, ei unelmien päivä. Kertova teksti on myös moniäänistä ja moniäänisyys lisääntyy kehityksellisesti, samoin lukijan huomioiminen ja vuorovaikutuksellisuus. Kertovalla tekstillä on toisin sanoen oikeastaan aika moninaisia ominaisuuksia, jotka saadaan hallintaan hitaan kehityksellisesti. Onnistunut kertova teksti on siten vaativa kirjoitustehtävä, vaikka se usein mielletään koulunkäynnin alkutaipaleille kuuluvaksi.

Alakoululaisia käsittelevissä luvuissa aineistona on joko Koulukorpus tai S1–S2-korpus. Yläkoululaisia ja nuoria aikuisia analysoidaan sekä Koulukorpuksen että Unelmakorpuksen aineistolla, mutta rakenteellinen kompleksisuus on tarkimmin koodattu Unelmakorpuksen. Osaksi käytössä on ainakin paikoin koko aineistoa, osaksi on tehty rajauksia. Norminmukaisuutta on arvioitu S1–S2-korpuksesta, jonka 650 kirjoittajalta on kerätty tietoa sekä eriaikaisesti että kyselytестein.

KOEHENKILÖT

Koehenkilöt ovat joko yleisopetuksen peruskoululaisia tai nuoria aikuisia. Peruskoululaiset edustavat tasaisesti koko ikäluokkaansa – lukuun ottamatta kahden kieliluokan oppilaita⁸ – mutta mukana ei ole erityisluokkien oppilaita. Aikuisilla on lähes kaikilla ylioppilastausta. S2-oppilaita on kerätty edustavasti vain osaan aineistoa. Kaikki suomi toisena kielenä -oppilaat ovat Suomessa syntyneitä tai ennen esikouluikää Suo-

8 Kolmasluokkalaiset kieliluokan oppilaat on valittu kielellistä ja matemaattista älykkyyttä mittaavan testin avulla peruskoulun 2. luokan keväällä. Nämä oppilaat kuuluvat monen muuttujan suhteen yläkvartiiliin ja ero valikoimattomaan kolmannen luokan aineistoon on keskimäärin 1,5–2 kehitysvuotta. Sekä ala- että yläkoulun aineistossa on yksi kieliluokka, alakoulun aineistossa on lisäksi kielikylpyluokkia. Alakoulun kieliluokasta *Helsingin Sanomien* antamien tietojen mukaan 60 % kirjoitti äidinkielestä laudaturin. Luokan kielellinen taso jo alakoulussa tietyksi osoittaa, että kyse ei ole yksin koulun vaikutuksesta (ks. myös Harjunen ja muut 2019). Hyvä osaamistaso ei kuitenkaan rajoitu kielikouluihin, vaan aineistossa on esimerkiksi eräs 6. luokka, joka suoriutui myöhemmin ylioppilaskirjoituksista lehtitietojen mukaan erinomaisesti. Vastaava osaaminen käy ilmi omasta tutkimuksestamme.

meen muuttaneiden maahanmuuttajataustaisten vanhempien lapsia, jotka ovat tutkimuksen alkuvaiheessa asuneet Suomessa vähintään noin 4 vuotta ja sen loppuvaiheessa vähintään noin 7 vuotta (Honko 2013, d)-aineisto). Koehenkilöiden kokonaismäärää on vaikea laskea tarkkaan, mutta heitä on vähintään 2 600. Luku nousee, jos mukaan lasketaan hankkeen opiskelijatöihin erikseen kerätty aineisto. Testeihin on osallistunut noin 1 400 koehenkilöä; ja Unelmakirjoitelmia on yhteensä runsas 2 400. Osa koehenkilöistä on osallistunut vain yhteen tehtävään, osa moneen. Sähköisistä korpuksista koehenkilöiden määrää ei voi arvioida, mutta koska mukana on useita sanomalehtiä, kirjailijoita ja ylioppilaskokeeseen osallistuneita, määrä on vähintään sadoissa. Sähköisten korpusaineistojen tuottajien ikä vaihtelee, mutta kyse on työikäisistä ihmisistä tai nuorista aikuisista.

Peruskoululaisten aineistot on kerätty eri puolilta Suomea, mutta keuruu painottuu eteläiseen ja läntiseen Suomeen. Kouluja on mukana keuruussa yli 20 ja luokkia kymmeniä, mikä nostaa aineiston edustavuutta. Peruskoulun aineistossa on mukana kaikenlaisia koehenkilöitä, ja heitä ei ole esikarsittu esimerkiksi (lievän) kielihäiriön perusteella.⁹ Otos edustaa siten koko yleisopetuksessa mukana olevaa suomenkielistä ikäluokkaa, samoin tulokset. Yleisopetuksen luokkien taso vaihtelee melko sattumanvaraisesti. Suorituksista ei voi aina päätellä, johtuvatko niiden puutteet ulkoisen motivaation, pystyvyyssuskon tai osaamisen puutteesta vai onko kyse kielihäiriöstä. Vaikka esimerkiksi suomalaiskoululaisten lukutaito on aikaisempien kansainvälisten PIRLS- ja PISA-kielitaloituskartoitusten perusteella huippuluokkaa, lukumotivaatio ja lukemiseen sitoutuminen ovat joillakin mittareilla mitattuna vertailumaiden heikoimpia (KTL 2017; Leino ja muut 2019). Poikien suhtautuminen sekä lukemiseen että kirjoittamiseen ja omiin tekstintaitoihin on tyttöjä kielteisempää (Harjunen & Rautopuro 2015; Leino ja muut 2019).

Asennetekijöiden merkitys oppimiseen voi olla suuri sekä suoraan että välillisesti: on vaikea syventyä asioihin, jotka eivät tunnu merkityk-

9 Karsinta ei olisi ollut mahdollistakaan, koska lieviä häiriöitä ei aina diagnosoida (vrt. Kuusela 2011). Lievät kielihäiriöt näkyvät tutkimuksen aineistossa muun muassa kirjoitusvirhetyyppien perusteella sekä toisaalta leksikaalisissa testeissä sattumanvaraisena vastaustapana (ks. Ahvenainen & Holopainen 2000; Kulju & Mäkinen 2009; Lehtinen 2014; vrt. Honko, tässä teoksessa).

sellisiltä, ja oppia sitä, millä ei ole yksilöllistä merkitystä tai minkä oppiminen ei tunnu mahdolliselta. Jos äidinkielen kouluopetus ja käytetyt materiaalit ovat erityisesti poikien näkökulmasta omaan elämänpiiriin ja tavoitteisiin nähden etäisiä (Sulkunen & Kauppinen 2018), koulu ei pysty myöskään parhaalla mahdollisella tavalla tukemaan positiivisten asenteiden tai henkilökohtaisen suhteen syntymistä opittaviin sisältöihin. Ulkoisen motivaation puute näkyy erityisesti haluttomuutena ylipäätään tarttua kirjoittamistehtävään mutta kietoutuu luonnollisesti myös osaamiseen, koska kirjoittamistaito ei kehity ilman harjoittelua. Motivaatiota pyrittiin nostamaan paitsi pienillä palkkioilla myös testien vaihtelevilla toteuttamistavoilla. Osaksi tässä onnistuttiin, ja erityisesti sähköiset testit kiinnostivat. Toisaalta kokeen toteuttajat raportoivat häiriökäyttäytymisestä myös sähköisissä testeissä.

Peruskoululaisia tarkastellaan osaksi luokkakohtaisesti jakaumalla 1.–9. luokkaan. Osaksi keskitytään peruskoulun päättövaiheisiin eli alakoulun kuudenteen ja yläkoulun yhdeksänteen luokkaan, ja tällöin ajatus on tutkia peruskoululaisia muutosvaiheessa. Lisäksi tutkimuksen ensimmäisiä tavoitteita oli tutkia juuri kuudesluokkalaisia, jotta saataisiin tietoa yläkoulun alkuun eli vaiheeseen, josta aineenopettajat aloittavat. Monissa osatutkimuksissa aineisto on valittu kahden luokkavuoden välein. Tämä siksi, että yhden ja saman ikäluokan sisällä kielellinen kehitysero voi olla jopa parisen vuotta, jolloin seurauksena on, että ikäluokkakehitys peittyy liian tiheällä koehenkilövalinnalla. Aineistossa on yleistä, että kolmannen ja neljännen, viidennen ja kuudennen sekä seitsemännen ja kahdeksannen luokan välillä muuttujissa on jonkinlaista häiriövaihtelua. Ero saattaa olla yhteydessä paitsi yksilölliseen kehitykseen myös murrosikään ja siihen liittyviin muutoksiin kirjoittamismotivaatiossa, keskittymiskyvyssä tai tehtäväorientaatioissa. Vuosien 2013–2014 keruusiin osallistuneiden motivaatiota pyrittiin nostamaan arpomalla luokittain pari elokuvalippua osallistumispalkkioksi. Tulosten perusteella näyttää siltä, että arviointitilanteessa motivaatioon voidaan vaikuttaa palkkioin vain hyvin rajallisesti. Palkkio nostaa vastaamishalukkuutta mutta ei niinkään vastaamisen laatua. Syvemmälle juurtuneet negatiiviset asenteet esimerkiksi oppiainetta, tekstitaitoja tai omaa osaamista kohtaan jäivät tämän vaikutuksen tavoittamattomiin, eikä palkkio

luonnollisestikaan voi testitilanteessa tasoittaa aikaisemmin syntyneitä taitoeroja.

Aikuisten testi- ja kirjoitelma-aineisto on kerätty ensin kolmessa Etelä-Suomen kaupungissa vuosina 2012, 2013 ja 2014. Nämä koehenkilöt eivät kaikki asu kyseisissä kaupungeissa vaan tulevat ympäri Suomea. Osa on ammattikorkeakouluopiskelijoita, osa on osallistunut suomen kielen pääsykokeeseen. Myöhempiin testeihin ja kielentämistehtäviin osallistuneet aikuiset on rekrytoitu pääosin korkeakouluopiskelijoista tai korkeakouluun pyrkijöistä, mutta jonkin verran mukana on myös muita vapaaehtoisia. Koehenkilöistä ammattikorkeakoululaisille maksettiin pieni osallistumispalkkio. Myös aikuisaineistossa taso vaihtelee, mutta otos ei edusta koko nuorten aikuisten ikäluokkaa, koska koehenkilöt ovat pääosin suorittaneet ylioppilastutkinnon ja he ovat vapaaehtoisia. Pääsykoeryhmä on nuorten aikuisten ryhmistä monimuotoisin, koska ryhmä ei keruun hetkellä koostunut korkeakouluopiskelijoista. Pääsykokeeseen osallistuneista kymmenisen prosenttia läpäisi pääsykokeen, muiden osalta ei ole tietoa myöhemmistä valinnoista. Ammattikorkeakoululaiset edustavat hyvin monia eri koulutusaloja humanistisista aloista terveydenhoitoon ja tekniikkaan, mikä lisää tulosten edustavuutta.

Yksittäisten testien koehenkilömäärät ovat vaihdelleet 20–600. Ryhmäkohtaisesti on yleensä pyritty 100 koehenkilöön. Peruskoulun kuudennen ja yhdeksännen luokan joukko, joka osallistui neljän testin sarjaan, käsitti yli 150 oppilasta. Yksi testeistä oli kirjoitettava kielentämistehtävä. S1- ja S2-koehenkilöiden kielitietotesteihin osallistui noin 650 alakouluikäistä, heistä 550 kirjoitti myös Unelmakirjoitelman. Lisäksi aineistossa on noin 50 kirjoitelmaa muilta samojen luokkien oppilailta. Sähköisiin semanttisiin ja morfologisiin testeihin osallistui yhteensä 300 koehenkilöä. He jakautuivat kuudes- ja yhdeksäsluokkalaisiin sekä nuoriin aikuisiin. Myös nämä koehenkilöt kirjoittivat Unelmakirjoitelman. Osa-aineistot pääsykoeaineistoa lukuun ottamatta sisältävät suunnilleen saman verran mies- ja naiskirjoittajia.

Tutkimukseen on hankittu tarvittavat luvat, jotka pyydettiin yksittäin usealta eri taholta. Alle 15-vuotiaiden vanhemmilta on pyydetty lupa lasten osallistumiseen, ja kaikkien koululaisten osalta on pyydetty lupa alueviranomaisilta, koulujen rehtoreilta ja luokkien opettajilta. Jokainen

koululainen on osallistunut oppitunnilla tehtyyn testiin muiden tavoin, mutta hän on voinut kieltää aineistonsa käytön merkitsemällä kiellon koepaperiin. Kieltäytyjiä oli muutamia ja heidän aineistoaan ei ole sisällytetty tutkimukseen, vaan se on hävitetty. Aikuisten tuottamien aineistojen keruuseen on pyydetty suostumus paitsi koehenkilöiltä itseltään myös kouluviranomaisilta, alueviranomaisilta ja koulujen rehtoreilta, jos keruu on toteutettu koulun tiloissa.

Koehenkilöiltä on pyydetty paperitesteihin nimi. Osa testeistä on toteutettu sähköisesti, ja tällöin testiohjelma on antanut koehenkilölle yksilöllisen nimikoodin, jonka avaimella testauksen eri osat pystyttiin yhdistämään. Koehenkilöiden nimiä ei ole liitetty aineistotietokantoihin eikä niitä ole säilytetty. Osa oppilaista on alkuaankin identifioinut itsensä vain etunimellä. Koulujen nimiä tai koulupaikkakuntia ei mainita tietokannoissa. Aineisto on myös muuten anonymisoitu. Kaikki henkilönimet on kirjoitelmiin muutettu mutta yleisimmät paikannimet, huiipaikkojen nimet ja vastaavat ovat muuttamattomia, jos ne eivät viittaa koulujen sijaintikuntiin tai lähikuntiin.

Aineiston keruun aikana vuoden 2015 tasa-arvolain uudistus, joka laajensi sukupuoleen kohdistuvan syrjinnän estämisen myös identiteetin ja sukupuolen ilmaisuun, ei ollut vielä voimassa. Koehenkilöt eivät siten voineet tehdä kyselyn sukupuolta koskeviin taustatietoihin mahdollisesti haluamiaan valintoja. Tämä ei ole sikäli ongelma, että keneltäkään koehenkilöltä ei kysyty tietoa sukupuolesta, vaan aineisto on jaoteltu myöhemmin etunimien perusteella nais- ja mieskoehenkilöiden tuottamaksi, jotta väitetyjä kirjoittamistaitoeroja voidaan tutkia. Aineistoon ei osunut sukupuolineutraaleja nimiä, joten jaottelu ei tuonut ongelmia. S2-aineiston keruun yhteydessä oppilaan sukupuoli on koodattu oppilaan tapaamisen tai ilmoitetun tiedon perusteella. Monet aikuiskirjoittajat pohtivat kirjoitelmassaan identiteettiään, mutta näitä pohdintoja ei käsitellä erikseen kuten ei muutenkaan kirjoitelmien yksityisempiä sisältöjä. Tutkimuksen alkuperäistä paperiaineistoa tai sähköisiä testejä ei ole säilytetty. Aineistosta kootut tietokannat ovat sen tutkijan hallussa, jolla niihin on tekijänoikeus. Vanhimpaan aineistoon ei ole keruun yhteydessä pyydetty lupaa aineiston jatkokäyttöön, joten tämä osa aineistosta on jatkossakin yksityistä.

LUVUT

Teoksen luvut jakautuvat kahteen temaattiseen osioon, joista ensimmäisessä lukujen analyysit lähtevät merkityksestä ja toisessa muodosta. Perusteena on kielellisen eron lisäksi erilainen kielenhallinnan ja kielenkehityksen kulku. Esittelemme lyhyesti teoksen sisältöä kyseisen kahtiajaon ja itse lukujen argumentaation kannalta.

MERKITYKSESTÄ MUOTOON JA MUODOSTA MERKITYKSEEN

Teoksen luvut on jaettu kahteen temaattiseen osioon. Ensimmäisessä osiossa näkökulma ja analyysin lähtökohta on semanttinen. Kielenhallintaa käsitellään näissä luvuissa sen mukaan, miten peruskoululaiset ja nuoret aikuiset ilmaisevat sanamerkityksiä ja miten he kielentävät semanttis-pragmaattisia merkityksiä. Näitä merkityksiä luonnollisesti ilmaistaan muodoilla, mutta muodot saavat oikeutuksensa vasta merkityksien määrittelyn kautta. Luvut on järjestetty kehityksellisesti vaikeutuvaan järjestykseen siten, että I osio alkaa sanamerkityksistä (Pajunen & Honko) ja laajenee sitten spatiaalisen merkityksen ilmaisuun (Tuuri) ja edelleen referentiaalisten viittauksien (Kouvo, Honko & Pajunen) ja moniäänisyyden ilmaisuun (Niinimäki). Sanamerkityksissä kehityksellinen linja kulkee asteikoilla yleisyys, konkreettisuus ja spesifisyys. Toisin sanoen iän mukaan käytössä olevien sanojen määrä kasvaa ja niiden yleisyys laskee, merkitykset muuttuvat abstraktimmiksi (samasta teemasta kirjoitettaessa) ja sanamerkitykset spesifisoituvat, niiden granulariteetti eli merkityksen täsmällisyys kasvaa. Nämä yleistykset koskevat myös spatiaalisen merkityksen ilmaisun hallintaa, mutta spatiaalisilla merkityksillä luodaan myös koherenssia ja ilmaisu laajenee sosiokulttuuriseen ja yksilöpsykologiseen suuntaan. Referentiaaliset ilmaukset lisääntyvät kertomuksellisuuden kehittyessä; vasta samojen diskurssireferenttien riittävän moni uudelleenmaininta lukijan kannalta identifioitavasti on (yksi) merkki kertomusrakenteen hallinnasta. Jos yksiäänisen diskurssin viittaussuhteiden ilmaisu on hallinnassa vasta alakoulun ylemmillä luokilla, moniäänisen kerronnan kerrostuneiden viittaussuhteiden hallinta kehittyy vielä hitaammin.

Semanttisten ilmiöiden kielentäminen edellyttää niiden semanttisen luonteen ja erilaisten funktioiden ymmärtämistä ja monien hierarkkis-

ten tai vaihtoehtoisten kielentämisen keinojen hallintaa. Niiden kielentämisessä kuvausta ei voi perustaa yhden ryhmän analyysien varaan, koska polku aikuisilmaisuuksiin on niin mutkikas: usein ilmaisu, joka pienellä alakoululaisella on kehityksellinen saavutus, on yläkoululaisella heikko suoritus ja aikuiselta jo kadonnut.

Toisessa osiossa lukujen näkökulma ja analyysin lähtökohta on kirjoitelmien muodollisissa yksiköissä. Näillä muodoilla luonnollisesti ilmaistaan merkityksiä, mutta muotolähtökohta rajaa analyysin näkökulmaa esimerkiksi kompleksistumisasteikon mukaisesti. Kielenhallintaa käsitellään sen mukaan, miten lapset, nuoret ja nuoret aikuiset hallitsevat morfologisia ja syntaktisia muodollisia yksiköitä ja miten hallinta kehityksellisesti muuttuu. Osio on järjestetty muotojen yksiselitteisyyden mukaiseen järjestykseen siten, että ensimmäinen luku (Laasanen & Pajunen) käsittelee muodollisesti hyvin yksiselitteisiä possessiivisuffikseja ja pyrkii selvittämään paitsi käytön määrää, myös sen suhdetta puhutusta kielestä esitettyihin suffiksien katoamisennusteisiin. Toisessa luvussa (Honko) analysoidaan alakouluikäisten kirjoitelmien lähinnä sanatasolla esiintyviä normipoikkeamia, ja niiden iän mukaista vähenemistä sekä vähenemisen suhdetta kielenhallintaan suomea toisena kieleinä puhuvilla. Kolmas ja neljäs luku (Pajunen & Vainio) käsittelevät kompleksistumisasteikolla morfologisia ja syntaktisia yksiköitä – sanarakennetta, sanaluokkajakaumiin perustuvia indeksejä sekä lauserakennetta. Tavoitteena on selvittää, miten sanatasolla ja kertovan rakenteen tasolla mitattavat yksiköt hallitaan ja mitä muutosta niissä iän mukaan tapahtuu. Muotojen kompleksistuminen tarkoittaa käytännössä niillä ilmaistujen merkityksien tiiviimpää – ja abstraktimpaa – esittämistä. Tiivistyminen puolestaan nostaa kertovan rakenteen globaalikoherenssia.

Jos I osiossa kehitys on hyvin hidasta, II osiossa se on nopeampaa. Keskeinen ero on myös, että I osiossa analyysi on enemmän laadullista, II osion muotoja voidaan helpommin poimia ja mitata ja mittauksen avulla osoittaa ikäkausittaisia osaamisen mediaaneja, koska tiivistyvien muotojen hallinta näkyy kertovan tekstin samanlaistumisena toisin kuin merkitysten ilmaisujen hallinta, joka näkyy vaihtoehtojen monipuolistumisena. Ero johtuu jo yksinkertaisesti siitä, että muototasolla on yksiköitä suhteellisen vähän, merkitystasolla niitä on paljon. Ajatus merkityk-

sen ja muodon epäsymmetriasta ei ole tietenkään uusi, mutta tässä se on pyritty esittämään täsmällisesti.

LUKUJEN ABSTRAKTIT

Myöhempi kielenkehitys: sanatiedosta rakenne- ja tekstitietoon
Anneli Pajunen ja Mari Honko

Johdantona toimiva ja tutkimusta taustoittava luku esittelee Myöhemmän kielenkehityksen tutkimushankkeen, jonka yhteydessä koko teos – sekä lukuisat opinnäytteet – on kirjoitettu. Pajunen ja Honko määrittelevät teoksen keskeisen käsitteistön, yhteiset tutkimuskysymykset, aineiston ja koehenkilöt, teoksen rakenteen ja yksittäisten lukujen teeman. Teoksen keskeisenä teemana on suomen kielen hallinta lapsilla, nuorilla ja nuorilla aikuisilla ja sen kehityksellinen eteneminen peruskouluikäisillä verrattuna nuoriin aikuisiin. Kielenhallinnan kehittymistä arvioidaan merkityksien kielentämisestä ikäkausittain ja toisaalta muotojen kielentämisestä. Merkitystason hallintaa tarkastellaan sanastotiedon ja semanttis-pragmaattisten ilmiöiden kielentämisen kannalta ja muotojen hallintaa morfologian ja syntaksin kautta.

I Sanat ja merkitykset

Osion I analyysiluvut vastaavat etenkin tutkimuskysymyksiin 1 (sanasto ja semanttinen tieto) ja 4 (kehityksellinen, kielellinen ja yksilöllinen variaatio).

Unelmieni päivä -sanasto kehityksellisestä näkökulmasta
Anneli Pajunen ja Mari Honko

Pajunen ja Honko tarkastelevat Koulu- ja Unelmakorpuksen kirjoittajien käyttämää sanastoa laadullisesta ja määrällisestä näkökulmasta. Koehenkilöiden ja kirjoitelmien määrä vaihtelee tarkastelussa vajaasta tuhannes-

ta lähes pariin tuhanteen. Tarkimmin on analysoitu niiden koehenkilöiden kirjoitelmia, jotka osallistuivat myös semanttisen ja morfologisen prosessoinnin testaukseen. Sanatiedon osalta sukupuolierot eivät ole osoittautuneet kovin suuriksi, ja sukupuoli ei siksi ole kaikissa sanastoanalyysissä mukana. Frekvenssilaskelmat osoittavat, että koululaiset käyttävät paljon kielen yleisintä sanastoa ja että ero on suurempi peruskoululaisten ja aikuisten välillä kuin peruskouluryhmien. Tämä voi johtua vahvasta koulukirjoittamisen konventiosta mutta heijastaa tietysti myös aikuisten laajempaa elämänpiiriä ja kokemusta. Diversiteettierot kasvavat aineistossa kehityksellisesti, eli käytetty sanasto monimuotoistuu karttuvan iän ja kielenkäyttökokemuksen myötä. Sukupuolen mukainen ero on melko pieni – mikäli sitä ylipäättään voidaan testata – ja aikuisilla toisensuuntainen kuin peruskoululaisilla. Semanttinen analyysi osoittaa, että eri ikäluokat kirjoittavat samasta aiheesta ensisilmäyksellä samoilla aineksilla vaikka täsmällisemmin. Tarkempi analyysi osoittaa, että yleistasolla on jatkuva siirtymä konkreettisesta abstraktiin, joten elollisista ja konkreettisista elottomista kirjoittaminen vähenee iän mukaan, mutta että lisäksi abstraktistumiskehitys kulkee konkreettisiin tarkoitteisiin viittaavien ryhmässä siten, että peruskoululainen kirjoittaa tietyn luokan kuten syötävän aineen esiintymästä (*karkki, limsa, sipsi, pizza*), aikuinen enemmän itse luokan yleistasolla (*ruoka, juoma, vesi*). Semanttinen analyysi osoittaa kirjoitelmat myös selvän kertoviksi: tapahtumaverbit ja ajan adverbit vallitsevat.

Liiketilanteen väylän ja suunnan kielentäminen suomessa Emilia Tuuri

Tuuri keskittyy spatiaalisuuden alaan kuuluvien liikkeen väylän ja suunnan kielentämiseen nuorilla aikuisilla. Jos ryhmätason hajonta yleensä kehityksellisesti vähenee, yksilötason variaatio toisaalta kasvaa, ja se kasvaa erityisesti spatiaalisessa kontekstissa. Siinä vaikuttavia tekijöitä ovat tila ja liike, liikkeen rajattuus, kielentämisen resurssit, tapahtumien havaitsemiseen liittyvät tekijät ja lisäksi yksilölliset kielentämisstrategiat. Aineistona on videolle kuvattuja lyhyitä liiketapahtumia eli videostimulusia ja niiden kielentäjinä 50 nuorta aikuista. Väylä ja suunta eroavat

toisistaan juuri rajattuuden kannalta: väylän ilmauksissa liikeradalle ilmaistaan alkupiste (esimerkiksi *ulos talosta*), päätepiste (esimerkiksi *sisään taloon*) tai keskivaihe (esimerkiksi *pellon poikki*), joka rajaa muuttujan liikkeen suhteessa kiintopisteeseen. Suunnan ilmauksissa (esimerkiksi *taloa kohti*, *poispäin*) liikettä ei rajata kielellisesti. Luku keskittyy erityisesti väylän keskivaiheeseen ja suuntaan, jotka asettuvat monien videostimulusten kohdalla juuri rajattuuden kannalta vaihtoehtoisksi tavoiksi jäsentää liiketilannetta: käveleekö poika *sillan yli* vai *siltaa pitkin*, nouseeko nainen *portaat* vai *portaita ylös*? Koehenkilöt ilmaisevat videostimuluksissa näkyviä liiketilanteita joko melko yhdenmukaisesti tai hyvin varioiden, ja Tuuri analysoi myös tähän variaatioon liittyviä taustatekijöitä. Käy myös ilmi, että yhdeksäsluokkalaisilla on pääosin samat resurssit jo käytössään kuin nuorilla aikuisilla mutta kuudesluokkalaisilla ei vielä ole. Kielitypologiselta kannalta Tuuri osoittaa, että suomessa kielennetään väylää ja suuntaa rikkaasti moniin muihin kieliin verraten. On odotuksenmukaista, että tällainen moninaisuus kehittyi verrattain hitaasti.

Referentiaaliset määräiset ilmaukset ja niiden suhde kertomusrakenteeseen alakouluiässä

Anna-Kaisa Kouvo, Mari Honko & Anneli Pajunen

Kouvo, Honko ja Pajunen tarkastelevat kertovan tekstin koherenssin eli tekstuaalisen yhtenäisyyden syntyä. Luku keskittyy referentiaaliseen koherenssiin ja tutkii kertomuksellisuuden kehittymistä tarkastelemalla tarkoitteiden määräisyyden ilmaisemisen muutoksia alakouluikäisten teksteissä. Nuorimmat alakouluikäiset kirjoittavat atomistisia tekstejä, mikä näkyy muun muassa siten, että tekstiin tuodaan koko ajan uusia tarkoitteita, referenttejä, mutta niistä ei ole enempää sanottavaa. Iän myötä samoista tarkoitteista kerrotaan pidempään eli kerrotaan useammista toisiinsa linkittyvistä tapahtumista samojen henkilöiden tai asioiden näkökulmasta. Samalla kun samoihin tarkoitteisiin viittaaminen lisääntyy, lisääntyvät myös keinot osoittaa kielellisin keinoin, että kyse on jo tekstiin esitellyistä tarkoitteista eli tutusta tiedosta. Yksijäseniset ilmaukset korvautuvat toisin sanoen erityyppisillä määritteillä tarkenne-

tuilla monijäsenisillä ilmauksilla. Luku osoittaa myös, että puhutun kiel-
len keinot ilmaista spesifisyyttä artikkelinomaisella *se*-pronominilla eivät
kuulu alakouluikäisten kirjoitettuun kieleen. Jo nuorimmilla koululaisil-
la on siis tässä suhteessa taito erottaa puhuttu ja kirjoitettu kielimuoto
omassa kirjoittamisessaan, jos kirjoittaminen ylipäätään vielä onnistuu.

Moniäänisyyden ilmaisun kehittyminen peruskoulun päättövaiheiden
kirjoitelmissa

Anneli Niinimäki

Niinimäellä tarkastelun kohteena on moniäänisyyden tuottaminen kir-
joitettuun kieleen. Moniäänisyys on semanttis-pragmaattisesti komplek-
sinen ilmiö, jonka avulla ilmaistaan kerronnan kielellistä moniaineksi-
suutta. Moniäänisyyttä on tekstiin rakennettu moninäkökulmaisuus ja
tulkintojen vaihtoehtoisuus. Muutokset kertojan tai puhujan osallistuja-
roolissa rakentavat moniäänisyyttä, samoin etääntyminen omasta ker-
ronnasta ja viittaukset tekstin ulkopuolelle. Moniäänisyyden kielentä-
mistä opitaan puhutussa kielessä melko varhain, mutta sen hallinta
kirjoitetussa kielessä etenee hitaasti. Tämä prosessi voi olla kesken vielä
peruskoulun päättövaiheissa. Kehitys suhteutuu kielellisen koherenssin
luomiseen ja kertomusrakenteen tai yleensä tekstin rakentamisen oppi-
miseen. Moniäänisyyden tyypeiksi on alakoululaisaineiston analyysin
avulla määritelty intertekstuaalisuus, metakieliset viittaukset tapahtuma-
kontekstin ulkopuolelle, suora ja epäsuora esitys sekä vapaa epäsuora
esitys muunneltuihin. Varsinaisena aineistona on kuudes- ja yhdeksäs-
luokkalaisten Unelmakirjoitelmia. Kouluikäisillä tarkastelun keskiöön
nousee referointi, jonka kautta kirjoittajat tuovat kertovaan tekstiinsä
toisen äänen. Kuudesluokkalaisilla referoinnin funktio on tyypillisesti
kerrontaa eteenpäin vievä, ei pohdintaan liittyvä, kerronnan pysäyttävä.
Yhdeksäsluokkalaisilla taas funktiona on episteeminen modaalisuus,
minäkertojan ajatuksien ja tuntemuksien kuvaaminen. Samalla suoran
esityksen määrä vähenee ja epäsuoran kasvaa. Toinen tyypillinen moni-
äänisyyttä luova ilmiö on kertomuksen metatasoon kuuluva vuorovaiku-
tuksen luominen lukijaan. Referoinnin taidoissa ja vuorovaikutuksen
rakentamisessa on puutteita vielä peruskoulun päättövaiheessa. Osalla

kirjoittajista teksti kuudennella vuosiluokalla ja vielä peruskoulun päätösvaiheessa on täysin yksiäänistä. Referoitua puhetta ei eroteta muusta kerronnasta, ja vuorovaikutuksen taidot puuttuvat. Toisaalta on myös joukko kirjoittajia, jotka osaavat lähestyä tarkoituksenmukaisesti lukijaa ja luoda kokemuksiinsa perustuvan kuvan omasta sisäisestä ajatusmaailmastaan.

II Muodot ja rakenteet

Toisessa osiossa vastataan erityisesti tutkimuskysymyksiin 2 (muotojen hallinnan kehitys) ja 3 (sukupuolierot) ja osin 4 (kehityksellinen variaatio).

Meneillään oleva kielenmuutos puhutusta kielestä kirjoitettuun ja suomen possessiivisuffiksi
Mikko Laasanen & Anneli Pajunen

Laasanen ja Pajunen tarkastelevat puhutun ja kirjoitetun kielen modaaliteettieroja ja toisaalta niiden samanlaisuutta ja erilaisuutta ja toteavat, että vaikka kirjoitetun kielen kehitys on alisteinen puhutulle myöhemmän kielenkehityksen alkuvaiheissa, myöhemmin suhde kehittyy vuorovaikutukselliseksi. Argumentaatio rakentuu erityisesti puhutun kielen perusteella meneillään olevan muutoksen – possessiivisuffiksien käytön vähenemisen ja mahdollisen kadon – varaan ja keskeinen tarkasteltava luvussa on, siirtykö puhutussa kielessä meneillään oleva muutos suoraan tai viipeellä kirjoitettuun kieleen. Keskeistä on siis, miten pitkälle kompleksisten ilmiöiden kielentämisen yhteydessä oleva kirjoitetun kielen muotoyksikkö saa vaikutuksia puhutusta kielestä. Tarkastelu ei luvussa ole kehityksellistä, jotta kehityksellinen vaihtelu ja muutos eivät sekoittuisi, vaan perustuu nuorten aikuisten eriaikaisiin kirjoitelmiin. Joitain kehityksellisiäkin havaintoja kyllä esitetään. Ajallinen ulottuvuus kattaa noin 70 vuotta 1940-luvulta 2000-luvulle. Vertailuaineistona on lehtikieltä, fiktiota ja oppikirjatekstejä. Possessiivisuffiksien käyttöä arvioidaan määrällisesti ja sanaluokkakohtaisesti. Lisäksi tarkastellaan suffiksien käyttöä niin sanottujen indeksisanojen eli eroavaa tai eroamatonta

ta omistusta ilmaisevien sanojen yhteydessä. Koska kyse on inherentisti tai juridisesti omistettavia objekteja ilmaisevista sanoista, voi olettaa, että possessiivisuffiksien käyttö juuri niiden yhteydessä olisi välttämättömintä ja vähäinen tai entistä vähäisempi käyttö ilmaisisi helposti muutoksen. Tutkimus osoittaa, että kyseisenä ajanjaksona possessiivisuffiksien käytössä ei ole selvää ajallista muutosta, eikä käyttö myöskään aineiston eri osissa ole vähentynyt. Aineisto osoittaa sen sijaan kiistatta, että kirjoitetun kielen eri osa-aineistoissa possessiivisuffikseja käytetään eri määrin ja eri tavoin. Indeksisanoissa käyttö on yleisintä, muuten kyse on tekstilaji- ja tyyppieroista.

Kirjoittamisen tarkkuuden kehittyminen ja yhteys kirjoittajan kokonaiskielitaitoon alakouluiässä
Mari Honko

Honko keskittyy tarkastelemaan norminmukaisen kirjoitetun kielen halluunottoa eli kirjoittamisen tarkkuuden kehittymistä alakouluikäisillä koehenkilöillä ja vertaa suoriutumista kielenhallintaa mittaaviin testiartoihin, todistusnumeroihin ja opettaja-arviointeihin. Koehenkilöinä on yhteensä 600 suomea äidinkielenä tai toisena kielenä puhuvaa. Kaikki suomea toisena kielenä puhuvat ovat Suomessa syntyneitä tai pitkään ennen kouluikää Suomessa asuneita. Hongon tutkimus on pitkittäistutkimus, ja samoja koehenkilöitä on testattu kirjoituttamalla Unelmakirjoitus kahtena eri ajankohtana, ensin alakoulun alkuvaiheessa ja myöhemmin loppuvaiheessa. Muuttajat ovat äänne- ja muotorakenteesta, mutta tarkastelussa on mukana myös sana- ja lausetason piirteitä. Äännetasolla kirjoittamisen kehittyminen edellyttää äänne–kirjainvastaavuuden eli auditiivisen diskriminaation kehittymistä, toisin sanoen kykyä analysoida puhutun kielen yksiköitä. Alakouluikäisillä norminmukaisuus kehittyi selvästi sekä puhutun kielen että kirjoitetun kielen vaikutuksessa; kirjoittamisessa edellinen heikkenee iän myötä, jälkimmäinen voimistuu. Aineistosta kuitenkin nousee oletuksen vastaisesti pääosin pojista koostuva ryhmä, jonka tulokset viittaavat lieviin kielihäiriöihin, vaikka kielihäiriödiagnosi puuttuu. Honko osoittaa, että norminmukaisuuden kehittyminen on sidoksissa kielenhallintaan. Norminmukaisuus-

den astetta voisi siten pitää alustavana kielenhallinnan kriteerinäkin, jos kirjoittamiskonteksti, tekstin luonne ja muut tekijät otetaan huomioon. Tulosten pohjalta voi myös arvella, että aina heikot kirjoittamistaidot eivät riipu pelkästään motivaation puutteesta.

Kielen rakenteen hallinta peruskoululaisilla ja nuorilla aikuisilla
Anneli Pajunen ja Seppo Vainio

Pajunen ja Vainio keskittyvät teoksen kahdessa loppuluvussa kielenhallinnan kehitystrendeihin, joita voi kuvata asteikolla yksinkertaisesta – eli puhutun kielen kaltaisesta – rakenteesta kompleksiseen rakenteeseen. Ensimmäisessä kuvaus etenee sanarakenteesta lauserakenteeseen ja osoittaa, että sanojen pidentyessä ja lauserakenteiden kompleksistuessa tekstit tiivistyvät. Samalla niiden rakennetta kuvaavissa indekseissä tulee näkyviin ikäkausikohtaisia tavanomaisen suoriutumistason mediaaneja, joiden avulla voi arvioida ikäluokkatasoista osaamista ja kielenhallintaa. Nuorimmilla alakouluikäisillä osaamisen hajonta on hyvin suurta, mutta jo kuudesluokkalaisilla muuttujille on laskettavissa osaamisen tavanomaista tasoa ilmaisevia mediaaneja ja myös kehitys seuraavaan ikäluokkaan on ennustettavissa. Aineisto tosin osoittaa, että monissa muuttujissa alakvartiiliosuudet ovat normaalijakauman odotusarvoihin nähden liian suuret ja että monilla peruskoulun pojilla muuttujat sijoittuvat suoralle viivalle nousevan sijaan. Tämä on huolestuttava tulos, vaikka syitä ei saa aineistolla esiin. Kaikkiaan muotoon keskittyvät kirjan luvut osoittavat, että kielen hallinta etenee ikäluokasta toiseen melko tasaisesti paljastaen ikäkausiosaamisen keskitason. Puhutun kielen rakenne on pitkälti kehityksen lähtökohtana, mutta sen vaikutus ei ole kehitystä rajaava tekijä. Voi tosin arvella, että jos jo puhutun kielen hallinnassa on ongelmia, ne heijastuvat myös kirjoitetun kielen hallintaan. Pajunen ja Vainio arvioivat kirjoitelmaa muodollisten ja rakenteellisten mittarien avulla.

Syntaktista rakennetta tiivistävät ja kompleksistavat tekijät
peruskoululaisilla ja nuorilla aikuisilla
Anneli Pajunen ja Seppo Vainio

Pajunen ja Vainio osoittivat edellisessä luvussa, että kirjoitetun kielen syntaktinen rakenne tiivistyy iän myötä. Nyt he erittelevät tarkemmin, mikä ominaisuus milloinkin vaikuttaa syntaktisen rakenteen kompleksisuuteen, joka selvästi kasvaa ikäryhmittäin. Toisin sanottuna, mikä nostaa eri kompleksisuusmuuttujien arvoja ja mitä muodon hallinnan kannalta näistä ominaisuuksista voidaan päätellä. Sukupuolen mukaan arvioiden kompleksisuuserot ovat vähäisempiä kuin monet muut määrälliset erot. Keskeisin ero näyttää olevan se, että mieskirjoittajat tuottavat hieman kompleksisempää tekstiä jo kuudennelta luokalta lähtien. Aikuisryhmässä erityisesti miehille ominaista on tiivistää kerrontaa lauseenvastikkeita käyttäen. Kertova teksti on miehillä usein tiivistahtinen tapahtumasarja, jossa erityisesti temporaalirakenteilla taustoitetaan ja taikaistetaan tapahtumia. Myös naiskirjoittajilla on paljon tapahtumasarjakirjoitelmia, mutta tekstin kulku on toinen. Jos sukupuolieroja mitataan sanamäärinä, erot ovat naisten hyväksi, mutta jos muotoeroja katsotaan tarkemmin, tilanne ei olekaan aivan yksinkertainen. Aineistosta nousee näkyviin myös yhdeksännen luokan poikien kirjoitelmia, joissa muuten nousevana linjana näkyvä kehitys on monen mittarin kautta mitattuna suoralla viivalla: arvot ovat usein samat kuin kuudennen luokan pojilla. Äkkipitäen näyttää, että mitään kehitystä ei olisi kielenhallinnassa tapahtunut, mutta tämä ei voi olla koko totuus, koska elektronisissa testeissä tulokset ovat linjassa ikäluokkaan ja osin parempia kuin tytöillä. Ero voi liittyä myös ymmärtämisen ja tuottamisen väliseen eroon, sillä nämä kehittyvät eriaikaisesti. Motivaation puute on yksi taustatekijä, mutta se vaikuttaa usein enemmän tekstin koherenssiin kuin muotoyksiköihin. Väestötasolla täytyy tietysti huomata, että noin 2–3 prosentin alisuoriutuminen on odotuksenmukaista.

Lähteet

- Ahvenainen, O. & Holopainen, E. 2000. *Lukemis- ja kirjoittamisvaikeudet: Teoreettista taustaa ja opetuksen perusteita*. Jyväskylä: Special Data.
- Babayigit, S. 2015. The dimensions of written expression: Language groups and gender differences. *Learning and Instruction* 35, 33–41.
- Baker, L. & Levin D. 2015. The role of relational triggers in event perception. *Cognition* 136, 14–29.
- Bandura, A. 1997. *Self-efficacy: The exercise of control*. New York: Freeman.
- Berman, R. A. 2004. Between emergence and mastery: The long developmental route of language acquisition. Julkaisussa: R. A. Berman (toim.) *Language development across childhood and adolescence*. Amsterdam: John Benjamins, 9–34.
- Berman, R. A. 2007. Developing linguistic knowledge and language use across adolescence. Julkaisussa: E. Hoff & M. Shatz (toim.) *Blackwell handbook of language development*. Oxford: Blackwell Publishing, 347–367.
- Berman, R. A. & Slobin, D. 1994. Narrative structure. Julkaisussa: R. A. Berman & D. Slobin (toim.) *Relating events in narrative: A cross-linguistic developmental study*. Hillsdale, NJ.: Lawrence Erlbaum Ass., 39–84.
- Bertram, R., Laine, M. & Virkkala, M. 2000. The role of derivational morphology in vocabulary acquisition: Get by with a little help from my morpheme friends. *Scandinavian Journal of Psychology* 41, 287–296.
- Blomberg, J. 2014. *Motion in language and experience: Actual and non-actual motion in Swedish, French and Thai*. Lund: Lundin yliopisto.
- Bohnemeyer, J., Donelson, K. T., Tucker, R. E., Benedicto, E., Capistrán Garza, A., Eggleston, A., Hernández Green, N., de Jesús Selene Hernández Gómez, M., Herrera Castro, S., O’Meara, C. K., Palancar, E., Pérez Báez, G., Polian, G. & Romero Méndez, R. 2014. The cultural transmission of spatial cognition: Evidence from a large-scale study. *Proceedings of the 36th annual meeting of the cognitive science society*, 212–217.
- Bouwer, R., Béguin, A., Sanders, T. & van den Bergh, H. 2015. Effects of genre on the generalizability of writing scores. *Language Testing* 32: 1, 83–100.
- Camarata, S. & Woodcock, R. 2006. Sex differences in processing speed: Developmental effects in males and females. *Intelligence* 34, 231–252.
- Carlisle, J. F. 2000. Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing* 12, 169–190.
- Citron, F., Cacciari, C., Kucharski, M., Back, L., Conrad, M. & Jacobs, A. 2016. When emotions are expressed figuratively: Psycholinguistic and affective norms for 619 idioms for German. *Behavior Research Methods* 48, 91–111.
- Clark, E. 2014. Acquisition of derivational morphology. Julkaisussa: R. Lieber & P. Štekauer (toim.) *The Oxford handbook of derivational morphology*. Oxford: Oxford University Press, 424–439.
- De Zeum, E. L., van Beijsterveldt, C. E. M., Glasner, T. J., de Geus, E. J. C. & Boomsma, D. I. 2016. Arithmetic, reading and writing performance has a strong genetic component: A study in primary school children. *Learning and Individual Differences* 47, 156–166.

- E-Prime-3. Psychology Software Tools. <https://pstnet.com>. Viitattu 31.1.2020.
- Doherty, M. J. 2009. *Theory of mind: How children understand others' thoughts and feelings*. Hove: Psychology Press.
- Gibbs, R., Davies, G. M. & Chou, S. 2016. A systematic review of factors affecting the likelihood of change blindness. *Crime Psychology Review* 2: 1, 1–21.
- Givón, T. 1993. *Coherence in text, coherence in mind*. Amsterdam: John Benjamins.
- Givón, T. 2017. *The story of zero*. Amsterdam: John Benjamins.
- Habermas, T. & de Silveira, C. 2008. The development of global coherence in life narratives across adolescence: Temporal, causal and thematic aspects. *Developmental Psychology* 44: 3, 707–721.
- Harjunen, E. & Rautopuro, J. 2015. *Kielenkäytön ajattelua ja ajattelun kielentämistä. Äidinkielen ja kirjallisuuden oppimistulokset perusopetuksen päättövaiheessa 2014: keskiössä kielentuntemus ja kirjoittaminen*. Helsinki: Kansallinen koulutuksen arviointikeskus.
- Harjunen, E., Marjanen, J. & Karlsson, J. 2019. *Äidinkielen pieni pitkittäisarviointi 2014–2017: Perusopetuksen päätöstä lukion päättöön*. Helsinki: Kansallisen koulutuksen arviointikeskus.
- Heilig, M. 2019 [2018]. *Naisen aivot, miehen aivot: Onko aivoilla sukupuolta?* Kääntäjä: S.-A. Ahvonen [Alkuteos *Hon, han och hjärnan*]. Jyväskylä: Atena.
- Heimo, H. 2003. ”Kaikki on enemmän kuin joka toinen”: *Oppimisvaikeudet ja suhdekäsitteet esiopetuksessa*. Helsinki: Helsingin yliopisto.
- Heinonen, S. 2018. *Narratiivisuuden sisällöllinen kehittyminen aiheanalyysin kautta tarkasteltuna: Kvantitatiivinen tutkimus aiheiden esiintymisen vaihtelusta iän ja sukupuolen mukaan ”Unelmieni päivä” -kirjoitelmissa*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Hoff, E. 2005. *Language development*. Belmont: Thomson Wadsworth.
- Honko, M. 2013. *Alakouluikäisen leksikaalinen tieto ja taito: Toisen sukupolven suomi ja S1-verrokki*. Tampere: Tampereen yliopisto.
- Honko, M., Jarvis, S. & Vainio, S. 2019. Lukijat sanaston monimuotoisuutta määrittämässä: Leksikaalisen diversiteetin tarkastelua määrällisen ja laadullisen tutkimuksen rajapinnassa. *Virittäjä* 123: 1, 44–78.
- Hooper, S., R., Costa, L.-B., McBee, M., Anderson, K. L., Yerby, D. C., Knuth, S. B. & Childress A. 2011. Concurrent and longitudinal neuropsychological contributors to written language expression in first and second grade students. *Reading and Writing* 24, 221–242.
- Hulstijn, J. 2015. *Language proficiency in native and non-native speakers: theory and research*. Amsterdam: Benjamins.
- Hulstijn, J. 2019. An individual-differences framework for comparing nonnative with native speakers: perspectives from BLC theory. *Language Learning* 69: 1, 157–183.
- Hyde, J. S. 2005. The gender similarity hypothesis. *American Psychologist* 60, 581–592.
- Hyde, J. S. & Linn, M. C. 1988. Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin* 104: 1, 53–69.
- Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., Hakonarson, H., Gur, R. E. & Verma, R. 2014. Sex differences in the structural connectome of the human brain. *PNAS* 111, 823–828.
- Ishibashi, M., Kopecka, A. & Vuillermet, M. 2006. *Trajectoire: Matériel visuel pour élicitation*

- des données linguistiques*. Laboratoire Dynamique du Langage (CRNS/Université Lyon 2) – Fédération de Recherche en Typologie et Universaux Linguistiques, CRNS, France.
- Itkonen, E. 1978. *Grammatical theory and metascience*. Amsterdam: John Benjamins.
- Itkonen, E. 2005. *Analogy as structure and process: Approaches in linguistics, cognitive psychology and philosophy as science*. Amsterdam: John Benjamins.
- Itkonen, E. 2008. *Maailman kielten erilaisuus ja samuus. Osa I*. Turku: Turun yliopisto.
- Itkonen, E. 2009. *Maailman kielten erilaisuus ja samuus. Osa II*. Turku: Turun yliopisto.
- Itkonen, E. & Pajunen, A. 2010. *Empiirisen kielitieteen metodologia*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Jaakola, M. 2004. *Peruskoulun 2. ja 3. luokan oppikirjojen sanasto – dynamiikasta luettavuuteen*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Jakobson, R. 1968 [1940–1942]. *Child language, aphasia and phonological universals*. The Hague: Mouton.
- Jensen, A. R. 1998. *The g factor: The science of mental ability*. Westport: Praeger Publisher.
- Juvonen, R. 2004. *Kertomus ylioppilaslainsäädännössä: Kevään 1999 tehtävistä 23 kirjoitettujen aineiden tekstilajianalyysiä*. Suomen kielen pro gradu -tutkielma. Helsingin yliopisto.
- Juvonen, R. 2010. *Evaluoiva että-yhdyslause ja retoriset rakenteet suomenkielisissä ylioppilaslainsäädännössä*. *Virtittäjä* 114, 39–70.
- Juvonen, R., Virtanen, M. & Voutilainen, E. 2012. *Fennistisen tekstilajitutkimuksen suuntia*. Julkaisussa: V. Heikkinen, E. Voutilainen, P. Lauerma, U. Tiililä & M. Virtanen (toim.) *Genre-analyysi: Tekstilajitutkimuksen käsikirja*. Helsinki: Gaudeamus, 453–469.
- Kalliokoski, J. 2006. *Tekstilajin taju ja toisella kielellä kirjoittaminen*. Julkaisussa: A. Mäntynen, S. Shore & A. Solin (toim.) *Genre – tekstilaji*. Helsinki: Suomalaisen Kirjallisuuden Seura, 240–265.
- Kalliokoski, J. 2008. *Suomi toisena kielenä ja koulussa kirjoittaminen*. Julkaisussa: S. Routarinne & T. Uusi-Hallila (toim.) *Nuoret kielikuvassa: Kouluikäisten kieli 2000-luvulla*. Helsinki: Suomalaisen Kirjallisuuden Seura, 349–370.
- Karjalainen, M. 2008. *Yläkoululaiset tekstin rakentajina*. Julkaisussa: S. Routarinne & T. Uusi-Hallila (toim.) *Nuoret kielikuvassa: Kouluikäisten kieli 2000-luvulla*. Helsinki: Suomalaisen Kirjallisuuden Seura, 290–306.
- Kauppinen, M. & Marjanen, J. 2020. *Millaista on yhdeksäsluokkalaisten kielellinen osaaminen? Suomen kielen ja kirjallisuuden oppimistulokset perusopetuksen päättövaiheessa 2019*. https://karvi.fi/app/uploads/2020/08/KARVI_1320.pdf. Viitattu 20.11.2020.
- Kavé, G. & Halamish, V. 2015. *Doubly blessed: Older adults know more vocabulary and know better what they know*. *Psychology and Aging* 30, 68–73.
- Keith, T. Z., Reynolds, M. R., Patel, P. G., Ridley, K. P. 2008. *Sex differences in latent cognitive abilities ages 6 to 59: Evidence from the Woodcock-Johnson III tests of cognitive abilities*. *Intelligence* 36, 502–525.
- Kemper, S. & Sumner, A. 2001. *The structure of verbal abilities in young and older adults*. *Psychology and Aging* 16, 312–322.
- Keuleers, E., Stevens, M., Mandela, P. & Brysbaert, M. 2015. *Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment*. *The Quarterly Journal of Experimental Psychology* 68: 8, 1465–1492.
- Koivisto, V. 2013. *Suomen sanojen rakenne*. Helsinki: Suomalaisen Kirjallisuuden Seura.

- Koutsoftas, A. & Gray, S. 2012. Comparison of narrative and expository writing in students with and without language-learning disabilities. *Language, Speech, and Hearing Services in Schools* 43, 395–409.
- KTL 2017. Suomalaislapset edelleen parhaiden lukijoiden joukossa. Koulutuksen tutkimuslaitoksen tiedote 5.12.2017. <https://www.jyu.fi/ajankohtaista/arkisto/2017/12/tiedote-2017-12-05-12-10-32-101275>. Viitattu 6.8.2020.
- Kulju, P. & Mäkinen, M. 2009. Sanelukirjoituksen arviointi – tyyppivirheistä sanarakenteiden fonologiseen analyysiin. *Virittäjä* 113: 4, 508–532.
- Kulju, P., Kauppinen, M., Hankala, M., Harjunen, E., Pentikäinen J. & Routarinne, S. 2017. Reviewing Finnish studies on writing in basic education. Towards a pedagogy for diversity. Julkaisussa: N. Pyry, L. Tainio, K. Juuti, R. Vasquez & M. Paananen (toim.) *Changing subjects, changing pedagogies: Diversities in school and education*. Helsinki: Helsingin yliopisto, 110–127.
- Kusnetsoff, T. 2017. *Suomenkielisten nuorten morfologinen tietoisuus ja sen yhteys luettun ymmärtämiseen*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Kuukka, K. & Metsämuuronen, J. 2016. *Perusopetuksen päättövaiheen suomi toisena kielenä (S2) -oppimäärän oppimistulosten arviointi 2015*. Helsinki: Kansallinen koulutuksen arviointikeskus.
- Kuusela, J. 2011. Kun kirjoittaminen ei suju. Julkaisussa: E. Harjunen, R. Juvonen, J. Kuusela, B. Silén, M. Sääkslahti & M. Örnmark (toim.) *Miten peruskoululaiset kirjoittavat? Näkökulmia ja kysymyksiä. Perusopetuksen 9. luokan äidinkielen ja kirjallisuuden oppimistulosten seuranta-arvioinnin aineistoa 2010*. Helsinki: Kansallinen koulutuksen arviointikeskus.
- Kuuskorpi, T. 2012. *Psykologisten testien käyttö Suomessa*. Turku: Turun yliopisto.
- Kyösti, J. 2018. *Motivoinnin vaikutuksia kuudesluokkalaisten kirjoittamiin kertomuksiin*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Laasanen, M., Pajunen, A. & Häikiö, T. Arvioitavana. Kielen natiivihallinnan variaatio ja sosiolingvistinen meneillään olevan kielenmuutoksen tutkimus.
- Labov, W. 1975. *What is a linguistic fact?* Lisse: The Peter de Ridder Press.
- Labov, W. & Waletzky, J. 1967. *Narrative analysis: Oral versions of personal experiences*. Julkaisussa: J. Helm (toim.) *Essays on the verbal and the visual arts*. Seattle: American Ethnological Society, 12–44.
- Laine, M. & Virtanen, P. 1999a. WordMill. Lexical Search Program. Turku: University of Turku.
- Laine, M. & Virtanen, P. 1999b. Turun Sanomat 1993–1994 [= TS-korpus]. Morfosyntaktisesti analysoitu tietokanta, 23 miljoonaa sanamuotoa.
- Lappalainen, H.-P. 2003. *Osaat lukea – osaatko kirjoittaa? Perusopetuksen 6. vuosiluokan suorittaneiden äidinkielen ja kirjallisuuden oppimistulosten arviointi 2002*. Helsinki: Opetushallitus.
- Lappalainen, H.-P. 2011. *Sen edestään löytää. Äidinkielen ja kirjallisuuden oppimistulokset perusopetuksen päättövaiheessa 2010*. Helsinki: Opetushallitus.
- Lee, J. 2011. Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics* 32: 1, 69–92.
- Lehtinen, R. 2013. *Kielen norminmukaisuus alakoululaisten kirjoitelmissa*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.

- Leino, K., Ahonen, A., Hienonen, N., Hiltunen, J., Lintuvuori, M., Lähteinen, S., Lämsä, J., Nissinen, K., Nissinen, V., Puhakka, E., Pulkkinen, J., Puro, J., Sirén, M., Vainikainen, M.-P., Vettenranta, J. 2019. *PISA18 Ensituloksia*. Opetus- ja kulttuuriministeriön julkaisuja 2019: 40.
- Lindeman, J. 2000. *Ala-asteen lukutesti: Käyttäjän käsikirja*. Jyväskylä: Gummerus.
- Malin, A., Sulkunen, S. & Laine, K. 2013. *PIAAC 2012 Kansainvälisen aikuistutkimuksen ensituloksia*. Helsinki: Opetus- ja kulttuuriministeriö.
- Mayer, M. 1969. *Frog, where are you?* New York: Dial Press.
- McLean, K. C. & Breen, A. V. 2009. Processes and content of narrative identity development in adolescence: Gender and well-being. *Developmental Psychology* 45: 3, 702–710.
- Metsämuuronen, J. 2006. Äidinkielen ja kirjallisuuden oppimistulosten ja asenteiden muuttuminen perusopetuksen ylempien luokkien aikana. Oppimistulosten arviointi 3/2006. Helsinki: Opetushallitus.
- Milton, J. 2009. *Measuring second language vocabulary acquisition*. Bristol: The Multilingual Matters.
- Nippold, M. A. 2006 [3. painos]. *Later language development: School-age children, adolescents and young adults*. Austin: Pro-Ed.
- Nippold, M. A., Hesketh, L. J., Duthie, J. K. & Mansfield, T. C. 2005. Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research* 48, 1048–1064.
- Niemi, P. 2007. Lukemistutkimuksen arvoitus: opetusta hylkivät oppilaat. *NMI-bulletin* 17: 3, 8–12.
- OKM 2019. PISA 2018: Suomi lukutaidossa parhaiden joukossa. Opetus- ja kulttuuriministeriön tiedote 3.12.2019. <https://minedu.fi/-/pisa-2018-suomi-lukutaidossa-parhaiden-joukossa>. Viitattu 17.11.2020.
- Pajares, F. & Valiante, G. 2001. Gender differences in writing motivation and achievement of middle school students: A function of gender orientation. *Contemporary Educational Psychology* 26, 366–381.
- Pajunen, A. 2012. Kirjoittamistaitojen kehitys 8–12-vuotiailla: Alakoululaisten unelmakirjoitelmat. *Virittäjä* 116: 1, 3–34.
- Pajunen, A. 2016a. Sukupuolierot kielen hallinnassa (etenkin kirjoittamisessa). Esitelmä. Språkets funktion 6.–7.6. 2016. Åbo Akademi.
- Pajunen, A. 2016b. Unelmakorpus. Peruskoululaisten ja nuorten aikuisten kirjoitelmia aiheesta Unelmieni päivä. Excel-tietokanta.
- Pajunen, A. 2019. Kielenhallinnan odotuksenmukaisuus peruskoululaisilla kielen eri osa-alueilla. Esitelmä. Språkets funktion 4.6.2019. Åbo Akademi.
- Pajunen, A. Tekeillä. 6.- ja 9. luokkalaisten sanaston hallinta sana- ja kielentämistehtävin arvioituna.
- Pajunen, A. & Palomäki, U. 1984. *Tilastotietoja suomen kielen rakenteesta 1*. Helsinki: Kotimaisten kielten tutkimuskeskus.
- Pajunen, A. & Virtanen, P. 2002a. Helsingin Sanomat 2000–2001 [= HS2000-korpus]. Morfosyntaktisesti analysoitu tietokanta, 24 miljoonaa sananmuotoa.
- Pajunen, A. & Virtanen, P. 2002b. Savon Sanomat 2001 [SS2000-korpus]. Morfosyntaktisesti analysoitu tietokanta, 8 miljoonaa sananmuotoa.

- Pajunen, A., Puranen, M. & Yli-Paavola, A. 2010. Koulukorpus. Alakoululaisten kirjoitelmia aiheesta Unelmieni päivä. Excel-tietokanta.
- Pajunen, A., Itkonen, E. & Vainio, S. 2015. Sanamerkityksen hallinta nuorilla aikuisilla. *Virittäjä* 119: 2, 160–187.
- Pajunen, A., Itkonen, E. & Vainio, S. 2016. Nuorten aikuisten kyky määrittellä sanojen merkityksiä. *Virittäjä* 120: 4, 477–515.
- Pajunen, A. & Itkonen, E. 2019. Intuition and beyond: A hierarchy of descriptive methods. Julkaisussa: A. Mäkilähde, V. Leppänen, E. Itkonen (toim.) *Norms and normativity in language and in linguistics*. Amsterdam: Benjamins, 213–234.
- Pajunen, A., Vainio, S., Hyönä, J. & Bertram, R. Tulossa. The role of derivative morphemes and lexical frequency in accessing transitivity information in reading: An eye tracking study in Finnish. Julkaisussa: A. Mäkilähde, T. Alho, V. Leppänen & A. Pajunen (toim.) *Unity and diversity of linguistics*. Routledge.
- Palmer, B., Lum, J., Schlossberg, J. & Gaby, A. 2017. How does the environment shape spatial language? Evidence for sociotopography. *Linguistic Typology* 21: 3, 457–491.
- Pasupathai, M. & Mansour, E. 2006. Adult age differences in autobiographical reasoning in narratives. *Developmental Psychology* 42: 5, 798–808.
- Petälä, I. 2018. *Kertomuksen rakenne, arviointi ja kompleksisuus aikuisten käännekohtakertomuksissa*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Pohjankukka, J. & Vainio, S. 2013. The internet based language testing program. Department of Psychology, University of Turku. <http://kielitestit.utu.fi/>. Viitattu 31.5.2020.
- Psykologiliitto. Kognitiivisten testien pistemäärien sanalliset kuvaukset. https://www.psyli.fi/files/836/sanalliset_kuvaukset.pdf. Viitattu 31.5.2020.
- Pöytä, S. & Kupiainen, S. (toim.) 2018. Tytöt ja pojat koulussa: Miten selättää poikien heikko suoriutuminen peruskoulussa? Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 36/2018.
- Qian, D. D. 2002. Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning* 52, 513–536.
- Ramscar, M., Hendrix, P., Love, B. & Baayen, R. H. 2013. Learning is not decline: The mental lexicon as a window into cognition across the lifespan. *The Mental Lexicon* 8: 3, 450–481.
- Ravid, D. & Avidor, A. 1998. Acquisition of derived nominals in Hebrew: Developmental and linguistic principles. *Journal of Child Language* 25: 229–266.
- Reynolds, M., Schreiber, C., Hajovsky, D. B., Schwartz, B. & Kaufman, A. S. 2015. Gender differences in academic achievements: Is writing an exception to the gender similarity hypothesis? *The Journal of Genetic Psychology* 176: 4, 211–234.
- Ribbon, G. 2016. The trouble with girls? *The Psychologist* 29: 12, 918–922.
- Routarinne, S. & Absetz, P. 2013. Writing performance and writing self-efficacy: The case of Finnish fifth graders. Julkaisussa: S. Tavares (toim.) *Why do we write as we write?* Oxford: Inter-Disciplinary Press, 51–66.
- Ruuska, K. 2020. *At the nexus of language, identity and ideology: Becoming and being a highly proficient second language speaker of Finnish*. Jyväskylä: Jyväskylän yliopisto.

- Saarela, L. 1997. *Peruskoululaisten kirjoitelmien kehittyminen sanastotutkimuksen valossa*. Oulu: Oulun yliopisto.
- Saarinen, A. 2020. *Equality in cognitive learning outcomes: The roles of educational practice*. Helsinki: Helsingin yliopisto.
- Saukkonen, P. 1982. *Oulun korpus. 1960-luvun suomen yleiskielen tutkimusmateriaali*. Oulu: Oulun yliopisto.
- Schmitt, N. 2010. *Researching vocabulary: A vocabulary research manual*. London: Palgrave Macmillan.
- Silverström, C. & Rautopuro, J. 2015. *Språk och skrivande i årskurs 9. En utvärdering av läsresultat i modersmål och litteratur våren 2014*. https://karvi.fi/app/uploads/2015/09/KAR-VI_1815.pdf. Viitattu 1.6.2020.
- Sintonen, M. 2011. *Suunnittelemisen vaikutus kertomuksen rakenteeseen viides- ja seitsemäsluokkalaissa kirjoittajilla*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arévalo, A., Vargha, A. & Bates, E. 2005. Timed action and object naming. *Cortex* 41, 7–25.
- Štekauer, P., Valera, S. & Körtvélyessy, L. (toim.) 2012. *Word-formation in the world's languages: A typological survey*. Cambridge: Cambridge University Press.
- Suvanto, A. 2012. *Lapsi tarinaa rakentamassa: Kielihäiriöisten lasten kerrontataidot ja niiden kuntouttaminen*. Oulu: Oulun yliopisto.
- Strömqvist, S. & Verhoeven, L. (toim.) 2004. *Relating events in narrative*. Volume 2. Typological and contextual perspectives. New Jersey: Lawrence Erlbaum Ass.
- Sulkunen, S. & Kauppinen, M. 2018. ”Ainakin muutaman pojan luokallani tiedän lukevan” – Poikien (osin piiloiset) luku- ja kirjoitustaidot. Julkaisussa: A. Kivijärvi, T. Hyyki, H. Lunabba (toim.) *Poikatutkimus*. Tampere: Vastapaino, 146–173.
- Söderholm, C., Häyry, E., Laine, M. & Karrasch, M. 2013. Valence and arousal ratings for 420 Finnish nouns by age and gender. *Plos One* 8: e72859.
- Tolchinsky, L. 2004. The nature and scope of later language development. Julkaisussa: R. A. Berman (toim.) *Language development across childhood and adolescence*. Amsterdam: John Benjamins, 233–247.
- Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M.-K., Tolvanen A. & Poikkeus, A.-M. 2019. Leisure reading (but not any kind) and reading comprehension support each other: A longitudinal study across grades 1 to 9. *Child Development* 91: 3, 876–900.
- Troia, G. A., Harbaugh, A. G., Shankland, R. K., Wolbers, K. A. & Lawrence, A. M. 2013. Relationships between writing motivation, writing activity, and writing performance: Effects of grade, sex, and ability. *Reading and Writing* 26, 17–44.
- Tuovila, S., Kairaluoma, L. & Mustonen V. (toim.) 2020. *Luku- ja kirjoitustaidon pedagogiikka yläkouluun*. Rovaniemi: Lapin yliopisto.
- Vainio, S., Pajunen, A. & Häikiö, T. 2019. Acquisition of derivational morphology: School-aged children and young adults. *First language* 39: 2, 139–157.
- Vainio, S., Pajunen, A. & Häikiö, T. Tekeillä. Acquisition of semantic knowledge: School-aged children and young adults.
- Vesala, S. 2012. *Kerronnan keinot alakoululaisten teksteissä*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.

- Virtanen, P. & Pajunen, A. 2002. ContextMill. The easy to use concordancer. Aurit Software, Turku.
- Voutilainen, A. 2002. Conexor FDG [= Functional dependency grammar] for Finnish. Linguistic description. Conexor – Natural knowledge. Moniste.
- Örnmark, M. 2011. Tyylitaju ja genretietoisuus 9. luokalla. Julkaisussa: E. Harjunen, R. Juvonen, J. Kuusela, B. Silén, M. Sääkslahti & M. Örnmark (toim.) *Miten peruskoululaiset kirjoittavat? Näkökulmia ja kysymyksiä. Perusopetuksen 9. luokan äidinkielen ja kirjallisuuden oppimistulosten seuranta-arvioinnin aineistoa 2010*. Raportit ja selvitykset 2011: 2, 40–51. www.oph.fi/julkaisut. Viitattu 22.4.2020.