

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

Author(s): Honko, Mari

Title: Kirjoittamisen tarkkuuden kehittyminen ja yhteys kirjoittajan kokonaiskielitaitoon alakouluiässä

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:

Honko, M. (2021). Kirjoittamisen tarkkuuden kehittyminen ja yhteys kirjoittajan kokonaiskielitaitoon alakouluiässä. In A. Pajunen, & M. Honko (Eds.), Suomen kielen hallinta ja sen kehitys : peruskoululaiset ja nuoret aikuiset (pp. 302-366). 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

Kirjoittamisen tarkkuuden kehittyminen ja yhteys kirjoittajan kokonaiskielitaitoon alakouluiässä

Mari Honko

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

Johdanto

Kirjoittamisen taitodiskurssia painottava opetuksen, tutkimuksen ja arvioinnin perinne on Suomessa vahva. Tämä näkyy esimerkiksi menneiden vuosikymmenten opetus suunnitelmien perusteissa, oppimateriaaleissa ja arviointitottumuksissa sekä kielitieteen, erityispedagogiikan ja kasvatustieteen ja osin myös psykologian aloilla jo tehdyn tutkimuksen valinnoissa kuten aineistoissa, käytetyissä metodeissa ja teoreettisen viitekehysten hahmottelussa (ks. myös Kauppinen, Pentikäinen, Hankala, Kulju, Harjunen & Routarinne 2015). Taitodiskurssia painottaneessa opetuksessa korostuvat yksilön taidot. Kirjoittamista tarkastellaan ensisijaisesti kielellisten taitojen soveltamisena kirjoittamiseen, ja päähuomio kohdistuu kirjalliseen tuotokseen, kuten oikeinkirjoitukseen, tekstien yhtenäisyyteen, tarkkuuteen ja virheettömyyteen. Tätäkin diskurssia kiistatta tarvitaan, mutta sen asema ansaitsee uudelleen arviointia. (Ks. kirjoittamisen diskursseista Ivanič 2004.)

Taitodiskurssin puutteena voidaan pitää erityisesti siihen usein vähintään implisiittisesti liitettyä oletusta ikään kuin vaihtelemattomista, tilasteesta toiseen siirrettävissä olevista taidoista. Ongelmaksi taitodiskurssi muodostuu, mikäli sitä painotetaan opetuksessa ja arvioinnissa muiden näkökulmien kustannuksella ja mikäli unohdetaan yksilöiden erilaiset valmiudet, kirjoittamisen monet merkitykset ja päämäärät sekä kirjoittamistilanteiden ja -kontekstien vaihtelevuus. Kirjoittamisen opetuksen, arvioinnin ja tutkimuksen valinnat kytkeytyvät laajempaan kirjoittamiskäsitykseen, joka on muovannut niin kierrätettyjen käytänteiden, kirjoittamista koskevien uskomusten ja puhumisen tapojen kuin opetusta ohjaavien määräysasiakirjojen, arviointimallien ja oppikirjojenkin sisältöjä. (Ks. esimerkiksi Erra & Svinhuvud 2017). Käsite kirjoittamisesta ja kirjoitustaidosta on erityisesti tuoreimman opetussuunnitelmaudistuksen (OPH 2014) myötä laajentunut ja tutkimus moninaistunut jo tätä ennemmin. Yksilön taitoon ja erityisesti kirjoittamisen tarkkuuteen kiinnittyvä kirjoittamiskäsitys painottuu kuitenkin yhä vahvasti opetuksen ja arvioinnin käytänteissä sekä näistä käytävissä keskusteluissa, minkä olen havainnut myös työssäni kielen opettajana ja opettajankouluttajana. Näkökulma onkin tärkeä, sillä tekstittyneessä yhteiskunnassa toimimiseen tarvitaan vahvaa kirjoitustaitoa ja erilaisiin arviointitarpeisiin monenlaisia, myös yksilökeskeisiä, kirjoitustaidon arviointitapoja. Samalla tarvitaan ymmärrystä kirjoittamisen moninaisuudesta sekä kirjoitustaidon dynaamisesta ja lukuisiin vaikuttaviin tekijöihin kietoutuvasta luonteesta (ks. esimerkiksi Ivanič 2004; Kauppinen ja muut 2015).

Vaikka taitodiskurssilla on Suomessa ollut kaikessa kieltenopetuksessa vahva asema, puuttuu tutkimukseen perustuvaa tietoa – johon opetusta ja arviointia voitaisiin suhteuttaa – kirjoittamistaidon kehitymisestä erityisesti alkuopetusvaiheen jälkeen. Alakouluikäisten oppilaiden kirjoittamisen tarkkuutta on Suomessa tutkimuksen keinoin aikaisemmin tarkasteltu lähinnä suomea äidinkielenä puhuvien lasten näkökulmasta. Tyypillisesti on tutkittu kontrolloitua kirjoittamista, kuten toisto- tai sanelukirjoittamisessa sanatasolla ilmenevää teknisen kirjoittamisen taitoa, ja kehyksenä on usein ollut lukemisen ja kirjoittamisen erityisvaikeuksiin liittyvä tutkimus (ks. esimerkiksi Kulju & Mäkinen 2009; Aro, Eklund, Leppänen & Poikkeus 2011; Kauppinen ja muut

2015). Myös alakouluikäisten tuottavaa kirjoittamista on jo aikaisemmin tutkittu erilaisin painoituksin (esimerkiksi Saarela 1997; Mäki 2002; Mäkihonko 2006; Pajunen 2012). Suomi toisena kielenä -oppijoiden kehittyvää kirjoitustaitoa on samassa ikävaiheessa kuitenkin tarkasteltu vasta muutamissa tutkimuksissa (ks. Halonen 2009; Honko 2013; Sikiö, Siekkinen & Holopainen 2016).

Tässä luvussa tarkastelen alakouluikäisten lasten tuottavaa kirjoittamista, erityisesti sen tarkkuuden kehittymistä yhdessä kirjoitustaidon kiivaassa kehitysvaiheessa vuosiluokilta 2–3 vuosiluokille 5–6. Lisäksi tarkastelen kirjoittamisen tarkkuuden suhdetta tuottavan kirjoittamisen muihin osataitoihin. Tutkimusryhmässä on sekä suomea äidinkielenä että toisena kielenä puhuvia koululaisia eri puolilta Suomea. S2-ryhmässä vertaan kirjoittamisen tarkkuutta ja sen kehittymistä myös kirjoittajan suomen kielen sekä oman äidinkielen taidon tasoon.

Tutkimukseni kytkeytyy osaksi kielitaidon arviointitutkimusta, tarkemmin osaksi kirjoitustaidon arviointia (Huhta & Hildén 2016). Kontrolloituun kirjoittamiseen nähden tuottava kirjoittaminen antaa kattavamman kuvan oppijan toiminnallisesta kirjoitustaidosta, esimerkiksi kyvystä ilmaista itseään kirjallisesti. Koulunkäynnin alkuvaiheessa kokonaisen tekstin tuottaminen ja keskittyminen yhtä aikaa sekä tekstin sisältöön että kielelliseen muotoon on useimmille kirjoittajille toisaalta vielä vaikeaa (Mäki 2002). Kirjoitustaidon mittarina tuottavan kirjoittamisen tehtävät ovatkin kirjoittajia varsin hyvin erottelevia, minkä vuoksi ne sopivat kontrolloitua kirjoittamista paremmin etenkin alkuopetusvaihetta myöhemmän vaiheen ja tyypillisesti kehittyvän kirjoitustaidon arviointiin (ks. myös Aro ja muut 2011). Kielitaidon vertailumittarina käytän kirjoittajien sanastollisia taitoja mittaavia sanastotehtäviä (Honko 2013) sekä kirjoittajilta ja heidän opettajiltaan kerättyä kyselyaineistoa.

Etenkin S2-oppilaiden oman äidinkielen taidot huomioivaa kielitaidon arviointitutkimusta on Suomessa toistaiseksi tehty hyvin niukasti. Kansallisen koulutuksen arviointikeskuksen selvityksen mukaan perusopetuksen päättövaiheen oppilaat ovat oman arvionsa mukaan yleensä saavuttaneet vähintään hyvät vuorovaikutustaidot sekä suomen kielellä että omalla äidinkielellään (Kuukka & Metsämuuronen 2016). Kielten käyttöalassa, taidoissa ja merkityksissä on kartoituksen mukaan yksilöl-

listä ja tilanteistakin vaihtelua, mutta kirjallisen viestinnän taidot ovat tarkastellussa ikäluokassa keskimäärin vahvemmat suomen kielessä kuin omassa äidinkielessä (mts. 73–77). Tähän vaikuttanee ennen kaikkea vähäisemmät mahdollisuudet saada opetusta, ohjausta ja harjoitusta omalla äidinkielellä lukemiseen ja kirjoittamiseen. Kartoituksessa ei kuitenkaan ole tarkasteltu monikielisiä resursseja yksilöllisellä tasolla eli esimerkiksi tietyn kirjoittajan kirjoitustaitoa omalla kielellään suhteessa suomen kielen kirjoitustaitoon.

Kriittisen pedagogiikan mukaan kirjoittamisella voidaan viitata taitoon mutta myös luovuuteen tai prosessiin ja sitä voidaan tarkastella lukuisista eri näkökulmista, kuten tekstilajien tuottamisena, sosiaalisten käytänteiden noudattamisena tai sosiopoliittisena toimintana (Ivaničín 2004 pohjalta Erra & Svinhufvud 2017). On siis muistettava, että kokonaisvaltaisessa kirjoittamisen tutkimuksessa ja opetuksessa kirjoittamisen tarkkuus on vain yksi osa kirjoitustaitoa ja että taitodiskurssi puolestaan on vain yksi näkökulma kirjoittamiseen (Ivanič 2004, 225). Vaikka tämä luku kohdistuu kirjoittamisen tarkkuuteen, tarkoitus ei ole vahvistaa jo ennestään vahvan taitodiskurssin asemaa kirjoittamisen opetuksessa. Sen sijaan tavoitteena on auttaa hahmottamaan kirjoittamisen tarkkuuden tyypillistä kehittymistä ja siihen vaikuttavia syitä osana kirjoittamisen kokonaistaitoa, erityisesti toisella kielellä kirjoitettaessa.

Kun tarkkuuden kehitystä havainnoidaan diakronisesti ja samalla otetaan huomioon kielenoppimisprosessin eteneminen muilla kielitaidon eri osa-alueilla, on mahdollista saada tietoa sekä tyypillisestä kielenkehityksestä että kielenkehityksen tukemisen tarpeista (ks. esimerkiksi Ellis & Barkhuizen 2005; Martin 2010; Martin, Mustonen, Reiman & Seilonen 2010; Slabakova 2016). Kirjoittamistaidon tyypilliseen kehityskulkuun ja haasteisiin tutustuminen auttaa esimerkiksi ymmärtämään, millaiset seikat ovat kirjoittamisen kannalta ajankohtaisia tietyssä ikävaiheessa. Tämän tutkimuksen tieto voikin auttaa paitsi tunnistamaan kirjoittamisen tarkkuuteen liittyvän tuen tarpeita peruskoulun alaluokilla myös kohdentamaan tukea oikein ja vähentämään tarpeetonta huolta tilanteissa, joissa kehkeytyvään kielitaitoon liittyy tilapäisesti myös epä-tarkkuus.

TUTKIMUSKYSYMYKSET JA VIRHEEN KÄSITE

Tässä tutkimuksessa kirjoittamisen tarkkuutta tarkastellaan aluksi lingvistisen analyysin avulla. Sen jälkeen tuloksia peilataan kirjoittajakohtaisesti taustatietoihin sekä kirjoittamisen muihin osataitoihin. Tutkimuskysymyksiä on neljä:

- 1) Miten kirjoittamisen tarkkuus kehittyi seurantajaksolla alakoulun aikana?
- 2) Millaiset virheet ovat yleisimpiä ja mitkä tekijät niitä selittävät?
- 3) Onko kirjoittamisen tarkkuudessa eroja eri kirjoittajaryhmien välillä?
- 4) Onko kirjoittamisen tarkkuus yhteydessä muihin kirjoittamisen osataitoihin ja S2-ryhmässä tarkemmin kirjoittajan suomen kielen tai oman äidinkielen taitoon kielitaidon eri osa-alueilla?

Esittelen seuraavaksi tutkimuksen keskeiset käsitteet, perehdytän tarkkuuteen kehityksellisenä mittarina ja kokoan lyhyesti aiempaa tutkimusta. Aineisto ja tutkimusmetodi esitellään johdannon lopuksi. Tulokset esitetään analyysiluvuissa ja kootaan pohtien päätäntöluvussa.

Kun tutkitaan kielenkäytön tarkkuutta, otetaan samalla väistämättä kantaa siihen, että kieli ei aina täytä tarkan vaatimuksia. Kielellisiä epätarkkuuksia on nimetty ja määritelty lukemattomin eri tavoin, mutta eniten käytetty, joskin usein kovin ohuesti määritelty, lienee *virheen* käsite. Hieman yksinkertaistaen virheellä tarkoitetaan eroa odotuksenmukaisen ja toteutuneen muodon välillä eli esimerkiksi konventionaalista poikkeavaa kielipillistä muotoa, sanajärjestystä tai sanavalintaa. Englannin kieleen keskittyvässä toisen kielen oppimisen tutkimuksen perinteessä keskeistä on määrittelyssä ollut toistuvien ja satunnaisten poikkeamien erottaminen toisistaan (Corder 1967). Kiinnostavaa on, että suomen kielen taitoon kohdistuvissa tutkimuksissa on ollut varsin yleistä luokitella ja nimetä virheitä määrittelemättä kuitenkin itse virheen käsitettä. Ensimmäisessä suomi toisena kielenä -alan väitöskirjassa Martin (1995) kirjoittaa, että virheellä (*error*) on viitattu nimenomaan toisen kielen oppijan puheessa tai kirjoituksessa ilmenevään taidon puutteeseen, joka kontrastoituu äidinkielen puhujan tarkkaan tai korkeintaan lipsahduksia (*slips*) sisältävään kielenkäyttöön. Tämä näkemys ei nähdäkseni kuitenkaan ylety edellä mainittuun äidinkielen ke-

hityksen varhaisvaiheen tutkimukseen. Esimerkiksi Mäkihonko (2006) sekä Havukainen ja Salmi (2011) käyttävät virheen käsitettä synonyymina kirjoitusvirheelle tutkimuksissaan, jotka käsittelevät tuottavan kirjoitustaidon varhaisvaiheen kehittymistä.

On huomattava, että kielellisesti tarkkakaan muoto ei välttämättä ole sosiaalisesti hyväksyttävä, sillä kaikki virheettömät muodot eivät ole vakiintuneet kaikkiin kielimuotoihin ja kielenkäyttötilanteisiin; virheellinen muoto puolestaan voi olla tilanteisesti aivan toimiva. Neutraalissa puheessa sanommekin esimerkiksi *juuri tänään*, *puolukoiden kanssa* ja *järjetön mies* mutta luultavasti emme *aivan tänään*, *poulukkojen kansa* tai *hullu ukko*, vaikka kaikki muodot olisivatkin käyttöyhteydessään ymmärrettäviä.

Virheet kiinnostavat sekä maallikoita että kielentutkijoita, koska ne poikkeavat tutusta tai tavanomaisesta ja ovat usein helposti havaittavissa. Kielivirheet mutta myös niistä puhuminen nostattaa helposti voimakkaita tunteita, sillä virheiden tarkasteluun liittyy aina myös arvottamista ja ideologista painolastia. Virheiden ilmeisin merkitys arjessa perustuu siihen, että ne voivat aiheuttaa vuorovaikutuksessa kitkaa kuten tarkennuspyyntöjä, oikean muodon hakemista, epäröintejä ja keskustelun katkeilua. Vahingollisimpia virheitä eivät kuitenkaan yleensä ole huomion kohteeksi nousevat virheet vaan ne, jotka sivuutetaan ja joiden seurauksena vuorovaikutukseen osallistuneet voivat tahoillaan jäädä väärän ymmärryksen varaan.

Koska vuorovaikutus perustuu yksittäisten sanojen ja sananmuotojen sijaan huomattavasti laajempien kielellisten ja ei-kielellisten kokonaisuuksien tulkintaan, kielivirheet eivät aina ole vaaraksi merkityksen välittymiselle. Tarkkuus on silti tulkinnan sujuvuuden kannalta eduksi ja tietyissä konteksteissa sitä suorastaan vaaditaan. Esimerkiksi lääkäriltä voidaan valta-asemansa ja vastuullisen työnkuvansa vuoksi edellyttää tarkempaa kielenkäyttöä kuin monessa muussa ammatissa toimivalta. Ohjaavien ja velvoittavien tekstien itsenäinen laatiminen ja tulkinta taas edellyttävät lääkäriltäkin suurempaa tarkkuutta kuin vaikkapa epämuodollisempi vuorovaikutus tai työparin tuella toimiminen (Tervola 2019).

Kielenkäytön tarkkuus on hyväksi silloin, kun on tarve ilmaista itseä, tulla ymmärretyksi ja ymmärtää muita täsmällisesti. Lisäksi sillä on merkitystä sosiaalisen aseman ja hyväksytyyden näkökulmasta: virheitä tekevä voidaan tulkita esimerkiksi huolimattomaksi ja välinpitämättömäksi, kieli- ja kulttuuritaidoltaan muita heikommaksi, joskus myös vähemmän luotettavaksi tai älykkääksi (ks. esimerkiksi Jessmer & Anderson 2001; Kreiner, Schnakenberg, Green, Costello, & McClin 2002; Vignovic & Thompson 2010). Jos virheitä tarkastellaan suomen kielen oppimisen kannalta, tarkkuuden kehittymistä voidaankin pitää yhtenä tavoittelemisen arvoisena ja edistymisestä kertovana kielitaidon piirteenä (ks. myös Brunni, Jantunen & Skantsi 2019), johon sisältyvät niin oikean rekisterin eli tilanteeseen sopivan kielimuodon valinta kuin kielellisten ilmausten rekisterinmukainen käyttö.

OPPIJANKIELEN TARKKUUS TUTKIMUKSEN KOHTEENA

Oppijankielentutkimuksessa virheanalyttinen lähestymistapa saavutti nopeasti vankan aseman 1960- ja 1970-lukujen taitteessa Stephen Corderin artikkelin *The significance of learner's errors* (1967) innoittamana ja on sittemmin saanut osakseen paljon sekä ansaittua että ansaitsematonta kritiikkiä. Matka virheanalyttisen tutkimuksen alkua ajoista nykyhetkeen on kuitenkin ollut pitkä: aluksi virheitä tarkasteltiin lähinnä ”täydellisen kielisysteemin” hallinnan puutteina ja niiden syitä etsittiin 1960-luvun kontrastiivisen analyysin hengessä melko suorasukaisesti esimerkiksi puhujan äidinkielestä (Lado 1957; ks. myös esimerkiksi Gass & Selinker 2008, 89–90).⁷⁵ Sittemmin on kiinnostuttu enemmän oppijankielen kehityksellisestä tarkastelusta.⁷⁶ Esimerkiksi DIALANG-arviointijärjestelmään liittyvässä tutkimuksessa tarkkuuden

75 L2-tutkimuskentällä ei silti liene koskaan vallinnut täyttä yhteisymmärrystä siitä, mille kielitaidon osa-alueille äidinkielen vaikutukset voivat virheinä ilmentyä ja milloin on kyse kielten välisten suhteiden aiheuttamista virheistä (*interlingual errors*), milloin kohdekielen systeemin sekoittamisesta (*intra-lingual errors*) – ja kuinka mahdollinen tai tarpeellinen tämän eron tekeminen kulloisessakin tilanteessa on.

76 Varhaisen virheanalyttisen tutkimuksen päämääriä on kritisoitu osin aiheettomastikin, sillä jo Corder (1967, 167) itse korosti virheiden tarkastelussa kielen oppijan tukemisen näkökulmaa: Kielen oppijan tekemät virheet ovat merkityksellisiä sekä kielen opettajalle, tutkijalle että oppijalle itselleen. Ne kertovat oppimisen vaiheista ja siitä, miten kieltä opitaan eli millaisia strategioita oppija esimerkiksi käyttää kieleen tutustuessaan.

(kielioppi, sanasto, oikeinkirjoitus) on havaittu korreloivan voimakkaasti aikuisten kielenoppijoiden sanavaraston koon kanssa (Alderson 2005, 55, 89). Samansuuntaisia tuloksia on saatu muillakin arviointimetoodeilla sekä aikuisten englanninoppijoiden kielestä (ks. esimerkiksi Meara & Milton 2003; Milton 2009, 171) että kouluikäisten S2-oppijoiden kielestä (Honko 2013).

Dynaamisuu den myötä tarkasteltava ilmiö on muuttunut: kontekstistaan irrotettujen virheiden sijaan on ryhdytty tutkimaan muun muassa kielenkäytön epäidiomaattisuutta, virheiden esiintymisympäristöjä sekä esiintymistiheyttä, kielellisen tarkkuuden suhdetta tekstin tulkintaan ja tavoitemuodon neuvottelemista vuorovaikutuksessa muiden kielenkäyttäjien kanssa. Korpustutkimuksen keinot mahdollistavat nykyään paljon sellaista, mihin ei 1960- ja 1970-luvuilla ollut pääsyä, ja 2000-luvulla virheiden analyysi yhdistetään yhä useammin laajempaan kielitaidon profilointiin sekä ohjatun tai spontaanin kielenkäytön vaikuttavuuden arviointiin.

Myös oppimis- ja kielikäsitteet ovat muuttuneet. Modernin virheanalyttisen tutkimuksen lähtöoletuksena on, että kaikki kielenkäyttäjät äidinkielen ja kielenkehityksen vaiheeseen katsomatta tekevät virheitä ja että tarkkuus ei myöskään kehity suoraviivaisesti – vaan kielellisten valmiuksien uudistumiseen liittyy kokeiluja ja haparointia (ks. suomen oppijoiden kirjoittamistaidon kehitymisestä esimerkiksi Seilonen 2013; Mustonen 2015; Brunni ja muut 2019). Virheiden syynä voi olla huolimattomuus sekä se, ettei edes syntyperäisten aikuisten kielitaito automaattisesti noudata mitään standardiolemusta edes kieliopin osalta (ks. myös Chipere 2009, 190–191). Toiminnalliseen kielitaitokäsitykseen kuuluu ajatus siitä, että virheet itsessään myös edistävät oppimista, sillä ne mahdollistavat epämukavuusalueella toimimisen ja siten uuden oppimisen ja voivat saada oppijan kiinnittämään huomiota kehkeytymässä oleviin ilmaisuihin. Toisaalta aiemman tutkimuksen pohjalta tiedetään, että virheet muuttuvat oppimisen edetessä: alkuvaiheessa tehdään enemmän rakenteellisia ja sanastollisia virheitä, myöhemmin poikkeavuudet liittyvät lähinnä ilmiöiden yleisyyteen ja tyypillisiin esiintymisympäristöihin (Martin 2015).

TARKKUUDEN KEHITTYMINEN OSANA KARTTUVAA KIELITAITOA

Kirjoitustaidon ja kirjoittamalla viestimisen opettelua on verrattu uuden kielen opettelemiseen (esimerkiksi Halonen 2009, 330). Kirjoittamista opetteleva lapsi voi toisaalta hyödyntää kirjoittamisessa puhutun kielen pohjalta oppimaansa tietoa esimerkiksi suomen kielen sanastosta ja rakenteista, tekstilajeista sekä vuorovaikutuskulttuurista. Mitä monipuolisempi lapsen kielellinen ympäristö on varhaislapsuudessa ollut, sitä vahvempi perusta hänellä on kouluiässäkin toimia kielellä joustavasti erilaisissa tilanteissa, kartuttaa ilmaisuvarantoaan ja oppia uusia taitoja, kuten kirjoittamista. Esimerkiksi äänne–kirjain-vastaavuuden oppimisessa ja äännekestojen merkitsemisessä auttaa, jos jo osaa suomen kielen äänteet sekä niiden lyhyen ja pitkän keston merkitystä erottavan tehtävän. Kertomustekstin kirjoittamisessa puolestaan on hyötyä siitä, jos on tutustunut kertomuksiin kuuntelemalla ja puhumalla. Kakkoskielisellä puhujalla sekä suomen kielen systeeminen tieto että kielen käyttöön liittyvä kulttuurinen tieto on tyyppillisesti kapeampaa kuin äidinkielisellä suomenpuhujalla. Syynä tähän ei ole kakkoskielisyys tai monikielisyys itsessään vaan vähäisempi kokemus.

Kirjoitustaidon oppiminen edellyttää puhutun kielen ja kirjoitusjärjestelmän yhteyden ymmärtämistä. Suomen kielen oppijalle tämä tarkoittaa muun muassa latinalaisaakkoston ja vastaavien äänteiden yhteyden sekä kirjoitussuunnan hahmottamista ja oikeinkirjoitusnormeihin tutustumista sekä kykyä pukea omat ilmaisutarpeet kirjalliseen muotoon. Taitodiskurssin näkökulmasta kirjoitustaitoon kuuluukin Mäen (2002, 17) mukaan kolme keskeistä osa-aluetta: kirjoituksen koostaminen kirjoitusmerkkien eli grafeemien avulla, oikeinkirjoitus sekä sisältöjen tuottaminen eli varsinainen kirjoittaminen. Käsin kirjoittaessaan kirjoittamistaitoa harjoitteleva lapsi siis ikään kuin piirtää näkyviin ajatuksiaan, samalla kun hän vielä harjoittelee kirjainten ja sanahahmojen muodostamista, tekstin asettelua sekä tuotettavan tekstilajin erityispiirteitä.

Luku- ja kirjoitustaidon tyyppillisen kehityksen tunnistaminen on avain lukemisen ja kirjoittamisen erityisvaikeuksien tunnistamiseen (Matilainen 1993, 5). Kakkoskielisyys tuo kirjoittamiseen oman erityisvärinsä, johon vaikuttaa paitsi oppijan kielellinen tausta myös muun

muassa kulttuuriset ja sosiaaliset tekijät. Myös suomea äidinkielenään puhuva on koulun alkuvuosina vasta ottamassa kirjoittamista haltuun, ja kirjoituksen tuottaminen vie huomiota suunnitelmalliselta sisältöjen tuottamiselta (ks. esimerkiksi Matilainen 1993). Kirjoittamisen alkuvaiheessa on tyypillistä, etteivät sekä tarkka oikeinkirjoitus että eheän ajatuskokonaisuuden luominen onnistu yhtä aikaa ongelmitta (Mäki 2002, 40). Vaikka esimerkiksi kirjoittamisen sujuvuus kasvaa nopeasti, virheiden määrä ei alkuopetuksen aikana juuri vähene (Mäkihonko 2006). Toisluokkalaaisella voi vielä olla vaikeuksia esimerkiksi äännekestojen kirjoittamisessa ja aloitetun virkkeen jatkamisessa eheänä loppuun saakka ilman, että kyseessä olisi kielellinen erityisvaikeus (ks. myös Kulju & Mäkinen 2009). Tekstistä voi myös puuttua kirjaimia, ja sanavälien merkitsemisessä voi olla horjuntaa. Havukaisen ja Salmen (2011) aineistossa lähes puolet (45 %) toisluokkalaisten tuottavan kirjoittamisen epätarkkuuksista liittyi oikeinkirjoituksen konventioihin, 40 prosenttia äännekirjain-vastaavuuteen (mukaan lukien pitkien äännekestojen merkintä, lisämerkit, puuttuvat kirjaimet, rauniosanat, puuttuvat tavut, rotaatiot ja reversaalit) ja vajaa viidennes (17 %) lausetason ilmiöihin. Vähitellen, Matilaisen (1993, 70–72) mukaan viimeistään neljänneltä luokalta alkaen, kirjoittamisen suunnitelmallisuus lisääntyy ja oppilaalle alkaa muodostua syvempi käsitys sekä käsiteltävistä asiasisällöistä että tekstistä kokonaisuutena.

Tiedetään, että erityisesti geminaattojen kirjoittaminen on äidinkiellekin suomenpuhujalle alkuvaiheessa vaikeaa (Kulju & Mäkinen 2009). Tämä voi liittyä auditiiviseen hahmottamiseen. Pitkä vokaali toteutuu puheessa ympäristönsä lyhyeen vokaaliin nähden pitkänä kestona, mutta geminaatassa pitkää kestoa havainnoidaan esimerkiksi kluusiilien tapauksessa lähinnä sulkeumavaiheen eli hiljaisen jakson kautta. Suomi toisena kielenä -oppijan aikaisemmin oppimisessa kielissä kestoerolla ei välttämättä ole lainkaan merkitystä erottavaa tehtävää tai kestoero voi olla distinktiivinen vain joissakin äänneissä. On siis huomattava, että esimerkiksi juuri äännekestojen kirjoittamisen vaikeudet eivät liity ainoastaan kirjoitetun kielen normien hallinnan puutteisiin vaan myös kielenhallintaan. Vasta, kun oppija oppii kuulemaan esimerkiksi suomen kielen merkitysten rakentamisen kannalta keskeiset kestoerot,

hän voi oppia myös tuottamaan näitä eroja puheessa ja kirjoituksessa tarkasti.

Piper (2001, 23) kritisoi lapsuusiän toisen kielen oppimista käsittelevässä teoksessaan opettajia virhepoliiseina toimimisesta eli siitä, että virheet nähdään koulussa epäonnistumisen merkkinä. Nykyisen oppimiskäsityksen mukaan virheet ovat kuitenkin oppimisen edellytys ja osoitus käynnissä olevasta oppimisprosessista eli ikään kuin osittaisesta menestyksestä uuden asian käyttöön siirryttäessä (esimerkiksi Martin 2009, 324). Kun aktiivinen kielenoppija testaa osaamistaan ja kielestä tekemiään päätelmiä, virheilta ei voi eikä tarvitse välttyä. Virheet voivat olla myös osoitus jo jollain tavalla omaksutun asian uudelleen prosessoinnista, sillä ”takapakki ja uudelleen analysoiminen” ovat oppijankielelle tyypillisiä piirteitä, kuten Martin (2003, 83) asian ilmaisee.

Kielitaidon eri osa-alueisiin pureutuvan tutkimuksen perusteella rakennetun CAF-mallin mukaan kielitaito kehittyi koko ajan rinnakkain kolmella osa-alueella: samalla, kun opitaan tuottamaan uusia, kompleksisempia muotoja, jo haltuun otettujen muotojen tuottaminen käy sujuvammaksi ja tarkemmaksi (ks. esimerkiksi Housen & Kuiken 2009; Housen, Kuiken & Vedder 2012). Keskittyminen yhteen osa-alueeseen tyypillisesti kuitenkin heikentää muita ainakin väliaikaisesti, ja esimerkiksi kirjoitusnopeuden kasvattamisesta seuraa helposti sekä käytetyn sanaston ja rakenteiden yksinkertaistuminen että pulmia ilmaisun tarkkuudessa. Kouluiässä kirjoittaminen kehittyy monialaisesti, mikä omassa tutkimusaineistossani näkyy muun muassa tuottamisen runsautumisena sekä harvinaisten ja rakenteeltaan kompleksisten sanojen käytön tihentymisenä vuosiluokalta toiselle siirryttäessä samaan aikaan, kun ilmaisu tarkentuu ja tekstien kokonaisrakenne kehittyi (Honko 2013). Kielen sujuvuus, tarkkuus ja kompleksisuus eivät kuitenkaan etene samaan tahtiin tai suoraviivaisesti, ja osa kielelliseen ilmaisuun yhtä aikaa vaikuttavista tekijöistä voi olla vaikutuksiltaan vastakkaisia.⁷⁷ Yksi esimerkki on sanapituuden ja tarkkuuden suhde. Yleinen kielen

⁷⁷ Suomi toisena kielenä -oppijoiden kirjoittamistaitojen kehittymistä taitotasoinnalla on tutkittu mm. Cefling- ja Topling-hankkeissa Jyväskylän yliopistossa, ks. esimerkiksi Seilonen 2013, Mustonen 2015 ja Lesonen 2020, ensikielisten peruskouluikäisten kirjoitustaitoja puolestaan Myöhemmän kielenkehityksen hankkeessa (ks. Pajunen 2012).

kehittymisen tendenssi on, että oppimisen myötä ilmaisupituus kasvaa ja tarkkuus lisääntyy. Tämän tutkimuksen kertomusaineistossa tarkimpia ovat keskimäärin kuitenkin tekstit, joissa suhteellinen sanapituus on pieni.⁷⁸ Ristiriitaa selittäviä tekijöitä on monia: 1) Ensinnäkin lyhyet sanat ovat tyypillisesti kielen yleisimpiä ja tutuimpia sanoja, joiden kirjoitusasu hallitaan muita sanoja paremmin. 2) Tekstit, joissa keskimääräinen sanapituus on pieni, yleensä myös rakentuvat suhteellisen pienen sanamäärän varaan, jolloin hallittavien sanojen ja sananmuotojen määrä pysyy kohtuullisena. 3) Sanastoltaan monimuotoisissa teksteissä taas tarvitaan yleisten lisäksi harvinaisia sanoja. Harvinaisten sanojen käytön myötä sanapituus väistämättä kasvaa, mutta myös virheiden todennäköisyys lisääntyy.⁷⁹ (Ks. tarkemmin Malvern ja muut 2004, 173.) 4) Oppimisen alkuvaiheessa kirjoittamisen lyhytsanaisuus yhdistyy usein suffiksaation eli esimerkiksi sanojen taivuttamisen vähäisyyteen. Oppimisen edetessä kielen yleisimmät sanat ja tutut sanamuodot osataan jo hyvin, mutta etenkin harvinaiset sanat sekä sanojen yhdistyminen ja kontekstissaan oikean muodon valinta voivat aiheuttaa haasteita pitkään.

Tarkkuuden mittaaminen

Tarkkuuden mittaaminen edellyttää mittayksikköä, jonka kautta osamista ja sen kehitystä tarkastellaan, ja valintaa yksikön esiintymien arvioinnissa. Määrittelen tässä luvussa oman lähestymistapani ja esittelen samalla tutkimuksen aineiston.

MIKÄ ON VIRHE?

Se, mitä kielessä pidetään virheellisenä, on aina sopimuksenvaraista ja ajassa muuttuvaa. Sekä maallikot että kielen ammattilaiset voivat arvioi-

78 Lisäksi aikuismaisesta kirjoittamisesta poiketen tekstien rakenteellinen kompleksisuus yhdistyy alaluokilla keskimääräistä pienempään sanapituuteen, mitä mahdollisesti selittää tekstin tuottamisen kuormittavuus (Honko 2013, 339).

79 Pitkät sanat tarjoavat suuremman grafeemimäärän vuoksi suhteessa enemmän mahdollisuuksia virheisiin kuin lyhyet. Toisaalta sanan pituus ei automaattisesti tee siitä vaikeaa, mihin vaikuttaa suomen kielessä etenkin pitkien sanojen monimorfeemisuus.

da kieltä eri tavoin: toinen pitää virheellisenä (*inaccurate*) ilmausta, joka toisen mukaan on aivan kelvollinen (*acceptable*) tai korkeintaan kontekstissaan sopimaton (*inacceptable*). Virheet, jotka viestin yhdessä vastaanottajassa saavat aikaan negatiivisia tunnereaktioita, eivät vaikuta samalla tavalla kaikkiin saman viestin tulkitsijoihin. (Ks. esimerkiksi Pääkkönen 1985; James 1998; Boland & Queen 2015, 2016; Al-Mahrooqi & Al-Aghbari 2016). Äidinkieleltään muunkielinen esimerkiksi saattaa arvioida oppijankieltä sekä tarkkuuden että pragmaattisen sopivuuden suhteen jopa korkeasti koulutautunutta äidinkielistä puhujaa kriittisemmin (Al-Mahrooqi & Al-Aghbari 2016, 7).

Kirjoitetun suomen kielen tarkkuuden tarkastelun tekee erityisen vaikeaksi virheiden ongelmallinen luokittelu. Puutteet oikeinkirjoituksessa, morfologian hallinnassa ja sanojen sekä sanontojen merkitysten osaamisessa limittyvät toisinaan niin, että virheiden analysointi tarkkarajaisesti voi olla mahdotonta (ks. myös Tarnanen 2002, 127–128). Vaikka toisen ja vieraan kielen tutkimuksessa systemaattiset virheet (*errors*) sekä satunnaiset lipsut (*mistakes*) on usein pyritty erottamaan toisistaan,⁸⁰ ei tällainen eronteko kuitenkaan ole mahdollista analysoitaessa lyhyitä tekstinäytteitä, esimerkiksi tämän tutkimuksen aineistoa. En myöskään erottele kontekstiriippuvaisia ja kontekstista riippumattomia virheitä – eli esimerkiksi rektiosijojen väärin valittuja sijamuotoja (*me emme pidän*) väärin muodostetuista (*me emme pidän sinästä*).

James (1998, 64–78) erittelee kielellisten ilmausten tarkkuutta paitsi ilmausten laadun myös niiden hallinnan asteittaisen kehittymisen kautta. Toisessa ääripäässä ilmausta ei osata vielä lainkaan, toisessa ääripäässä sitä käytetään aina oikein. Tyypillisen ja epätyypillisen kielenkäytön erottavan linjan sen enempiä kuin säännönmukaisuuden rajapyykkien asettelu ei kuitenkaan ole yksiselitteistä. Tyylipoikkeamia, esimerkiksi puhekielen muotojen soveltamista kirjoitukseen, en ole omassa tutkimuksessani katsonut perustelluksi tulkita virheiksi lainkaan, sillä kirjoittajat vasta opettelevat kielen eri rekisterien konventioita. Lisäksi kertomusgenre jo itsessään mahdollistaa puhekielen käytön esimerkiksi

⁸⁰ Perusteena on pidetty mm. sitä, että lipsahdukset eivät noudata systematiikkaa, joka voitaisiin järjkeväällä tavalla kuvata (ks. Corder 1967, 166–167).

henkilöhahmojen dialogeissa. Näyttää tosin siltä, että alakouluikäisten kertomustekstien dialogit ovat enemmän hypernormatiivisia ja noudattavat liian orjallisesti kirjoitetun kielen konventioita kuin ottavat vapauksia puhekielisyyden suuntaan, mikä muistuttaa Tannerin (2012) havainnotta oppikirjatekstien dialogeista.

Kirjoittamisen tarkkuutta mitattaessa on päädytty kolmenlaisiin menettelytapoihin (ks. esimerkiksi Ellis & Barkhuizen 2005). Holistinen arvio tekstin virheettömydestä on taloudellinen tapa saada tarkkuudesta yleiskäsitys. Se ei kuitenkaan anna lainkaan konkreettista tietoa kielellisen tuen tarpeesta eikä kerro sitä, kuinka laaja-alaisia pulmia kielen tarkkuudessa on. Jotkin virhetyypit, esimerkiksi sanojen sekoittuminen, saattavat myös painottua holistisessa arvioinnissa muita enemmän. Tämä ei kuitenkaan näy arviointituloksesta, joka on yleensä karkean mitta-asteikon mukainen numeerinen tai sanallinen arvio tarkkuuden tasosta. Arvioinnissa voidaan myös keskittyä tietyn piirteen tai tiettyjen piirteiden tarkkuuteen, jolloin saadaan tietoa juuri haluttujen kielenpiirteiden hallinnasta. Toisaalta valitut piirteet voivat edustaa hyvin suppeaa osaa oppilaan kieliopin tai sanaston hallinnasta puhumattakaan kokonaiskielitaidosta. Kolmas vaihtoehto tarkkuuden mittaamiseen on laskea virhetiheys eksplikoidulle virhejoukolle esittämällä näiden suhteen virheettömien sanojen tai lauseiden tai virkkeenkaltaisten T-yksiköiden prosenttiosuus. Tällöin ei yleensä oteta kantaa virheiden kokonaismäärään eikä virheiden laatua välttämättä eritellä tulosten esittelyn yhteydessä.

Yhdistämällä eri arviointitapoja, kuten tässä tutkimuksessa teen, on mahdollista verrata tarkkuutta erilaajuisissa aineistoissa ja pureutua tarkemmin muun muassa siihen, mihin muihin kielen osatekijöihin tarkkuus on yksilötasolla yhteydessä. Eri virhetyyppien jakauma voi toimia lähtökohtana toistuvien virhetyyppien tarkemmalle laadulliselle analyysille. Vastaavasti yksittäisten kielenpiirteiden tarkkuutta seuraamalla voidaan saada tietoa paitsi sen tyypillisestä vakiintumisajankohdasta oppijankieleen myös mahdollisista välivaiheista ja myöhemmistä uudelleenjärjestelyistä sekä tuen tarpeista silloin, kun tarkentumista ei tapahdu odotetulla tavalla.

Suomessa oppijankielen tarkkuutta on alimmilta vuosiluokilta lähtien perinteisesti havainnointi sanelukirjoittamisen yhteydessä joko laskemalla oppijan tuottama virhemäärä tai virheettömien tekstisanojen määrä. Tilalle Kulju ja Mäkinen (2009, 511) ovat ehdottaneet sanan fonologisen kokonaisrakenteen hallinnan arvioimista. Diagnostiikan välineenä on käytetty myös erilaisia sanelukirjoituksen sovelluksia sekä muita strukturoituja testejä, kuten virheiden paikantamiseen ja epäsanoihin liittyviä tehtäviä. Näiden suositeltavuus on sidoksissa paitsi itse testien ominaislaatuun ja käyttötarkoitukseen myös koulukäynnin vaiheeseen ja testatavan ryhmän erityispiirteisiin (ks. esimerkiksi Santala, Saaristo-Rinta, Metsola & Klippi 2013). Tämän tutkimuksen verrokkikorpuksen perustuvassa tutkimuksessa Lehtinen (2013) erittelee S1-kirjoittajien tekstien tarkkuutta toisella, neljännellä ja kuudennella vuosiluokalla kirjoitetuissa Unelmakirjoitelmissa. Lehtinen ei kuitenkaan käsittele tekstien virheteriheyttä suhteessa tekstien muihin ominaisuuksiin tai kirjoittamisen tarkkuutta yksilötason kielitaidon osa-alueena. Alakouluikäisten S2-oppiloiden teksteistä kattavaa tarkkuuden analyysiä ei tietääkseni ole tehty. Tässä luvussa havainnoidaan kirjoittamisen tarkkuutta monikielisessä lapsijoukossa sekä määrällisen että laadullisen analyysin avulla.

AINEISTO JA ANALYYSIMETODI

Tutkimusaineistoni koostuu yhteensä noin 600 kertomustekstistä,⁸¹ joista osa on suomea äidinkielenä puhuvien koululaisten ja osa edistyneiden,⁸² lukuisia eri kieli- ja kulttuuriryhmiä edustavien suomi toisena kielenä -oppiloiden (S2) kirjoittamia (ks. myös Honko 2013). Kaikki kirjoittajat ovat yleisopetuksen oppilaita. Kertomuskeruun yhteydessä oppijoille painotettiin, että tekstin tuottamisessa sisältö on tärkein, mutta kehoitettiin silti pyrkimään kirjoitetun kielen sääntöjen noudattamiseen ja tarkistamaan teksti kirjoittamisen lopuksi. Tekstit kirjoitettiin käsin ja syötettiin myöhemmin Excel-tietokannaksi.

81 Tekstit on kirjoitettu yhtenäisellä teemalla *Unelmien päivä*.

82 Kaikki tutkimukseen osallistuneet suomi toisena kielenä -oppilaat ovat Suomessa syntyneitä tai ennen esikouluikää Suomeen muuttaneiden maahanmuuttajavanhempien lapsia, jotka ovat tutkimuksen alkuvaiheessa asuneet Suomessa vähintään neljä vuotta ja sen loppuvaiheessa vähintään seitsemän vuotta (Honko 2013).

Tarkkuuden analysoinnissa keskityn kuuteen kirjoittamisen ongelma-kohtaan (ks. myös taulukko 44):

- 1) pulmat äänneiden merkitsemisessä
- 2) sanojen yhteen ja erilleen kirjoittamisen virheet
- 3) taivutusvirheet (myös puuttuva taivutus ja virheellinen sanavartalo)
- 4) sanajärjestyksen virheet
- 5) virheellinen sanavalinta, taivuttamaton tai puuttuva sana
- 6) muu lause- tai virkerakenteen ongelma

Tarkkuuden analysoinnin taustaksi olen seurannut sekä suomea ensimmäisenä (S₁) että toisena kielenä (S₂) puhuvien lasten kielitaidon muutoksia myös muilla kirjoittamisen osa-alueilla, kuten kirjoittamisen sujuvuuspiirteissä ja käytetyssä sanastossa. Lisäksi olen kerännyt opettajilta tietoa lasten kokonaiskielitaidosta ja kielellisistä erityisvaikeuksista (ks. Honko 2013). Laadullisen virheanalyysin jälkeen vertaan kirjoittamisen tarkkuutta muihin kirjoittamisen osataitoihin tilastollisten analyysien avulla. S₂-ryhmässä vertaan lasten yksilöllisiä kirjoittamistaitoja myös heidän kokonaiskielitaitoonsa käytettävissä olevan kyselytiedon avulla. Kyselytieto koostuu erillisestä suomen kielen sanaston hallintaan painotuvasta kielitestistä, S₂-arvosanasta sekä lomakevastauksista, joissa suomen kielen ja oman äidinkielen opettajat arvioivat lapsen kielitaitoa kielitaidon eri osa-alueilla sekä mahdollisia kielellisiä erityisvaikeuksia. Olen käyttänyt määrällisissä analyyseissa lähinnä ei-parametrisia tilastollisia analyysejä, sillä ne soveltuvat aineistoni laatuun parametrisia testejä paremmin. Lisäksi olen tutkinut aineistoa sirontakuvioiden avulla mahdollisten epälineaaristen yhteyksien paljastamiseksi.

On syytä huomata, että vaikka kirjoittamisen tarkkuutta tarkastellaan käyttöpohjaisesti, virheiden syntymisen täsmällisiin syihin ei tässä tutkimuksessa päästä käsiksi. Esimerkiksi äännevirheet voivat olla seurausta foneettisen tietoisuuden puutteista (väärin kuultu sanahahmo), sanahahmon väärästä muistiedustuksesta (oikein kuultu sana muistetaan väärin) tai kirjoitusjärjestelmän hallintaan liittyvän dekodauksen ongelmista (oikein muistettua sanaa ei vielä osata kirjoittaa oikein). Myös virheiden luokittelussa on toisinaan monitulkintaisuutta. Epävarmoissa tapauksissa olen katsonut ilmauksen mieluummin mahtuvan suomen

kielen variaation piiriin kuin poikkeavan siitä. Kun kirjoitettu muoto esimerkiksi reaalistuu partitiivia vaativassa lauseasemassa sanakirjamuodon kaltaisena taivuttamattomana sanana (*jalkapallo*), ei voida varmuudella sanoa, että kirjoittaja olisi pyrkinyt taivutusmorphologian käyttöön. Tapaukset on silti koodattu ainoastaan äännekestovirheenä, koska vokaaliyhtymän assimiloituminen (kuten *jalkapalloa* → *jalkapalloo*) on puhekielille tyypillinen piirre ja koska puhekielisyys muutoinkin voimakkaasti näkyy alakoululaisten kirjoittamisessa muun muassa juuri assimilaatioilmiöinä sekä loppukonsonanttien katona. Kielenhallinnan ja kirjoittamisen puutteet voivat toisin sanoen kirjoituksessa langeta yhteen, jolloin niitä ei voida luotettavasti erottaa yksittäisistä teksteistä.

Olen analysoinut tekstit yhdistetystä tekstikorpuksesta, jossa eri viiteryhmiä oppijoiden tekstit esiintyvät sekoitettuna. Analyysiä on ohjannut käyttöpohjaisuus, eli olen systemaattisesti ottanut huomioon myös poikkeavien muotojen esiintymiskontekstin. Tällainen lähestymistapa on oppijankielen tarkastelussa välttämätön, sillä vaikkapa urheiluselostusta jäljittelevä kielenkäyttö huudahduksissa *siisääään!* 'sisään' ja *maali!* 'maali' ei tietenkään ole osoitus kielen hallinnan puutteesta eivätkä ylipitkät vokaalit perusteltavissa kirjoituksen virheiksi.

Koska on selvää, että pelkkä virheiden lukumäärä ei eripituisia tekstejä tarkasteltaessa kerro kirjoittamisen tarkkuudesta koko totuutta, olen laskenut virheiden tekstitiheyden suhteessa kunkin tekstin kokonaisosan määrään ja osin myös suhteessa kyseisen kielenpiirteen kokonaisuuteen aineistossa.⁸³ Teknisesti koodaus on toteutettu niin, että olen voinut tarkastella aineistosta sekä virheellisten tekstisanojen karkeaa määrää (ottamatta kantaa virhetyyppiin tai eri virheiden mahdolliseen liittymiseen saneessa) että virheiden tarkempaa laatua ja kasautumista yksittäisiin tekstisanoihin. Virheiden tekstitiheyksiä olen sitten analysoinut sekä kirjoittaja- että kirjoittajaryhmäkohtaisesti.

Usein toistuvista virheistä välimerkkivirheitä ei tutkimuksessa ole koodattu virheiksi, eikä välimerkkien norminmukaista käyttöä voi vielä alaluokilla pitää odotuksenmukaisenaan.⁸⁴ Yksilölliset erot välimerk-

83 Ks. virheanalyysistä oppijankielen tutkimuksen yhteydessä esimerkiksi Gass & Selinker 1994, 66–74.

84 Koska sidossanojen käyttö on alaluokilla vielä melko vähäistä ja yksipuolista, voitaisiin useat virkerajamerkinnot korjata myös sidossanojen avulla.

kien käytössä ovat valtavan suuria, sillä vaikka jo tois- ja kolmasluokalaisissa on kirjoittajia, joille välimerkkien käyttö on sekä onnistunut virheettömästi että ollut monipuolista, eivät kaikki kirjoittajat ole käyttäneet välimerkkejä lainkaan tai niiden käyttö on yhä ollut horjuvaa ja tietämys esimerkiksi välimerkin sijainnista puutteellinen (ks. esimerkki 1).

1) *Olipakeran toive päivä. Aamulla minulle tuotiin sänkyyn. Lahjan kun avasin sen sieälä oli. uusi pallo sen jälkeen me mentiin. uimaan serenaan sielä oli. Vesiliukumäkiä ja kun me. Tultiin sieältä me lähdettiin Lasketelemaan minä otin mukaan, Oman lumi laudan sieälä oli. Hiihto hissejä siten me lähdetiin. pois ja mentiin linanmäelle, me mentiin vuoristorataan, ja kahvikupeihin ja peisotiin. ja karuselliin siten. sen jälkeen me lähdettiin. Vuokraamaan lefaa ja pidettiin.lefa ilta ja mentiin vasta. Kahelta yöllä nukumaan. minä vielä valvoin vähän aikaa ja nukahdin. Loppu. (3. lk. PS1)*

Ylemmillä vuosiluokilla päättövälimerkkien käyttäminen sujuu jo melko hyvin, mutta sivulausepilkkutuksen ja erikoismerkkiä käytössä on yhä hyvin suuria yksilöiden välisiä eroja – kuten on välimerkit mahdollistavien rakenteiden käytössäkin. Runsaat välimerkkivirheet voivat siis olla merkki paitsi tarkkuuden ongelmista myös edistyneestä lause- ja virkerakenteiden käytöstä.

Tarkkuus alakoululaisten kertomusteksteissä

Alakouluikäisten kirjoittaminen on kehittyvää ja tuottaa monenlaista tarkasteltavaa. Sanastossa ja rakenteissa on puhekielisyyttä ja ilmaukset voivat olla monitulkintaisia ja hankalasti ymmärrettäviä. Keskityn tässä yhteydessä tarkemmin sanatason ongelmiin, lähinnä normipoikkeamiin, ja tyypittelen niitä tiheyden, kirjoittajaryhmien ja toistuvuuden kannalta. Muita, esimerkiksi tyyliin liittyviä ongelmia enemmän esittelen taustaksi kuin analysoin virheiksi. Kertomusrakenne ei ole suurella osalla nuorimmista alakoululaisista vielä hallinnassa, samoin lause- ja virke-

rakenne voi ikätasolla tuottaa ongelmia, joten tarkastelen rakenneseikkoja tarkemmin vain vanhempien alakoululaisten kirjoitelmista.

PUHEKIELI OSANA KERTOMUSTEKSTEJÄ

Kertomustekstien poikkeamat kirjoitetusta yleiskielestä voidaan jakaa karkeasti kahteen pääluokkaan: normi- ja tyylipoikkeamiin. Normipoikkeamilla tarkoitetaan muotoja, jotka eivät kuulu pysyvänä osana mihinkään kohdekielen tunnettuun varianttiin vaan esimerkiksi syntyvät tilanteisina lipsahduksina tai ovat osoitus keskeneräisestä oppimisprosessista (esimerkit 2 ja 4). Tyylipoikkeamat puolestaan syntyvät kielen varianttien sekoittumisesta. Kirjoitetussa kielessä normipoikkeamia edustavat selkeimmin oikeinkirjoitusvirheet sekä virheet sanojen sijaituvuudessa ja vartalomuutoksissa, tyylipoikkeamia esimerkiksi varieteettien sekoittumisesta johtuvat leksikaaliset erikoisuudet, kuten leimallisesti puhekieliset ja lapsenkieleen kuuluvat sanat tai sananmuodot.⁸⁵ Etenkin nuorten koululaisten tekstejä arvioitaessa on usein olennaista erottaa puhekielisyydet (ks. esimerkit 2–3) muista poikkeamista, sillä vaikka lapset jo alaluokilla oppivat, että kielellä on useita ilmenemismuotoja, he eivät välttämättä pysty pitämään variantteja johdonmukaisesti erillään. Sen sijaan esimerkiksi sanan grafeemirakenteen virheet johtuvat yleensä sanahahmon vieraudesta tai foneemisen rakenteen dekodeausongelmista.

2) *Me juhliitaan* [vrt. *me juhlivat] *ihanassa ravintolassa koko perhe.*

3) *Monet ihmetteli, että miks mulla on kepit* 'kainalosauvat'.

vrt. virhemuodot *ihmetelli, *etta, *mulle on, *tikkut

Alakoululaisten teksteissä puhekielisyyksien värittäminen kielen rakennetta ja näkyy myös sanavalinnoissa, vaikka kieli saattaa muutoin olla lähes virheetöntä. Puhekielisyyksien esiintyminen kertomusteksteissä on runsaimmillaan aivan kirjoitustaipaleen alussa mutta vähenee selvästi jo kolmannelle vuosiluokalle tultaessa. Salosen (2012) tutkimus tosin osoittaa, että kehitys ei ole lineaarista vaan tarkentuvan, kirjoitetun yleis-

85 Puhekielen muotoja en tässä yhteydessä ole sisällyttänyt virheanalyysiin.

kielen piirteitä tunnollisesti noudattavan vaiheen jälkeen puhekielisyys uudelleen lisääntyy ja myös muuttaa muotoaan. Pojilla puhekielisyudet ovat alakoululaisten aineistossa hiukan yleisempiä kuin tytöillä ja S2-kirjoittajien teksteissä selvästi yleisempiä kuin S1-kirjoittajien teksteissä. Tässä suhteessa myös toisen ja vieraan kielen opiskelijoiden tekstit poikkeavat toisistaan, sillä vieraskielisessä ympäristössä kieltä opitaan tiiviimmin kirjoitetun kielenmuodon välittämänä ja oppijan arkiseksi pulmaksi voi muodostua ennemmin puhekielen varianttien puute kuin niiden ylikäyttö. Kielen eri rekisterien haltuun ottamisen myötä puhekielisyys muuttuu vähitellen yhä selvemmin tietoisesti käytetyksi työvälineeksi, jolla muun muassa elävöitetään dialogeja, puhutellaan lukijaa ja joskus jopa protestoidaan kirjoittamisen konventioita vastaan (ks. myös Halonen 2009 ja Niinimäki tässä teoksessa). Silti kielen varianttien vaihtelu on alakouluikäisten teksteissä usein vielä hallitsematonta.

KERTOMUSTEKSTIEN YMMÄRRETTÄVYYS

Kertova teksti on todettu kouluikäisille suhteellisen helpoksi tekstilajiksi; vähäisilläkin taidoilla onnistuminen on mahdollista, vaikka kirjoittaminen iän ja karttuvan kokemuksen myötä kehittyä ja esimerkiksi arvioiva, analysoiva ote yleistyy (ks. esimerkiksi Pajunen 2012). Tekstien päätehtävä eli kommunikatiivisuus täyttyy unelmapäiväteksteissäkin jo alaluokilla hyvin, sillä koko aineistossa ainoastaan muutaman yksittäisen tekstikohdan tulkinta jää epävarmaksi (esimerkiksi *tardo* 'automerkki').⁸⁶ Tekstien tulkitseminen edellyttää silti lukijalta maailmantietoa ja toisinaan myös kielellistä soveltamiskykyä. Tätä havainnollistaa seuraava kuudesluokkalaisten tytön kertomuskatkelma (ks. esimerkki 4), jossa limittyvät puhekielen redusoituneet muodot sekä taivutusmuotojen valinnan ja sanahahmon tarkkuuden pulmat:

4) *Ja unelmoin että kukaa meistä perheestä tai joku muu_ ei tarvii olla jokin asma tai syöpä tai kaikki nämä jutut.* (6. lk. TS2, albania, 12 vuotta Suomessa)

86 Koko 60 000 saneen aineistossa epäselviä lekseemejä on vain kymmenkunta.

Toisinaan tulkinnanvaraisuuksissa on kyse laaja-alaisesta rakenteen tai tekstin jäsentelyn epäselvyydestä, mutta joskus myös yksittäisen leksikaalisen valinnan tulkinta aiheuttaa lukijalle päänvaivaa (esimerkiksi *Huoisin leikkiä 'haluaisin'/'voisin'*). Koheesion ja virketasoa laajemman kokonaisrakenteen pulmat on jätetty tarkkuusanalyysin ulkopuolelle. Tämäntyyppiset rakenteen pulmat vaikuttavat kuitenkin välillisesti kertomusrakenteen holistiseen arvioon, jota olen käyttänyt virheteiheyden vertailumuuttujana.

Joissakin virheellisen muodon sisältävissä tekstikohdissa tavoitemuodon tulkinta jää epävarmaksi. Esimerkissä 5 rakennevirheen taustalla saattaa olla taivutusmorfologian puutteellinen hallinta mutta myös tilanteeseen epäsoivat leksikaaliset valinnat:

5) *Ja kato Big ben, juuri avatessa ikkunan verhot ja kivat naapurit.*
(6. lk. PS2, vietnam)

Esimerkin 5 tulkinnaksi sopisi 'ja näin Big Benin heti, kun avasin ikkunanverhot' (Big Ben uutena puheenaiheena), mutta myös tulkinnat 'katsoin Big Beniä' tai 'ja katso' tai 'ja kuvittele' olisivat mahdollisia.⁸⁷ Esimerkissä 6 puolestaan subjektin tai subjektina toimivan nominilausekkeen merkitsemättä jättäminen on koheesion kannalta ongelmallista: Mikä on se, joka jatkuu?

6) *Olin voittanut lotossa illalla urheiluauton ja veneen. Ja ei kuin vain jatkunut kun minä sain viestin, jossa luki ettei minun tarvitse enää koskaan käydä koulussa.* (6. lk. PS1)

Lauseiden ketjuttaminen on jo alaluokilla tyylillisesti ominaista muutamille taitavillekin kirjoittajille. Rakenteellisen ongelman ketjuttaminen muodostaa esimerkiksi silloin, kun lukijan odottama päälause jää virkkeestä puuttumaan ja semanttinen sisältö karkaa (ks. esimerkki 7):

⁸⁷ Merkityslajantumista lapsenkielestä ks. esimerkiksi Tomasello & Akhtar 2000, 184.

7) *Kun meillä oli niin paljon rahaa ettemme voineet enää kantaa sitä mukana koska emme halunneet, että sitä valtavaa rahamäärää jonka voitimme, emme halunneet, että niitä varastetaan.* (6. lk. TS1)

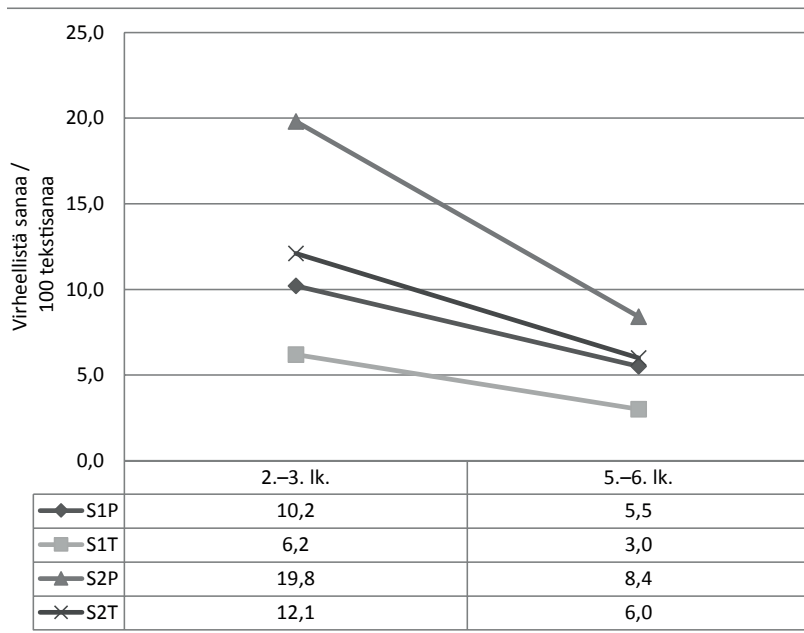
Alaluokilla monet tekstit ovat kuitenkin vielä sana- tai lausekelistoja, eikä lausetason rakenteellisille pulmille ole ollut luontevaa osoittaa määrää tai aina edes hahmoa (poikkeuksena jo mainitut erikseen koodatut virhetyypit). Siksi rakennevirheet on tässä tutkimusaineistossa koodattu ainoastaan viides- ja kuudesluokkalaisten kertomuksiin. Koska virkerakenteen puutteet yhdistyvät alaluokilla voimakkaasti muihin kielenhallinnan ja kirjoittamisen ongelmiin, ne heijastuvat joka tapauksessa tarkkuusanalyysin tuloksiin ja kertomusrakenteen holistiseen arvioon.

VIRHETIHEYS SEKÄ SEN MUUTOKSET RYHMÄ- JA YKSILÖTASOLLA
Tarkkuuden yleisindikaattorina käytän virhetiheyttä, joka on laskettu virheiden määrän ja tekstisanojen kokonaissanemäärän suhteesta siten, että yksittäiseen saneeseen liittyvät eri virheet on huomioitu vain kerran.⁸⁸ Vaikka osa virheistä todennäköisesti on seurausta kielitaidon kompleksistumisesta (esimerkiksi rektiovirheet uusien verbien yhteydessä) ja kirjoitusjuvuuden kasvusta (esimerkiksi lipsut oikeinkirjoituksessa), ei silti ole tarkoituksenmukaista tai oletettavaa, että epätarkkuudet kielessä lisääntyisivät säännönmukaisesti tai hallitsisivat ilmaisua oppijoiden siirtyessä vuosiluokalta toiselle. Onkin mielenkiintoista seurata, miten oppijankielen tarkkuus muuttuu kolmen vuoden seuranta-jakson aikana, jolloin sekä sujuvuuspiirteissä (tuottamisrunsaus, lekseemi- ja sanepituus) että sanaston kompleksisuudessa (sanarakenne, harvinaisten sanojen käyttö ja käytetyn sanaston vaihtelevuus) tiedetään tapahtuvan kehitystä (ks. myös Honko 2013).

Suurin virhetiheys, keskimäärin noin 20 virheen sisältävää tekstisanaa sataa juoksevaa tekstisanaa kohden, on alaluokilla S2-poikien teks-

88 Tekstien tarkkuutta on L2-kontekstissa yleisesti tutkittu laskemalla virheettömien lauseiden, virkkeiden tai T-yksiköiden suhteellisia määriä (ks. esimerkiksi Unsworth 2008, 320). Nuorimpien kirjoittajien tekstiaineiston tarkkuuden tarkasteluun lause- ja virkeajattelu soveltuu huonosti ennen kaikkea siksi, että lause- ja virkerakenteiden laajuudessa ja kompleksisuudessa on suuria yksilöiden välisiä eroja.

teissä, pienin (noin kuusi sataa tekstisanaa kohden) puolestaan S1-tyttöjen teksteissä (kaavio 13). S1-poikien ja S2-tyttöjen ryhmien välillä ei lähtötasossa ole merkittävää eroa, vaikka eri virhetyypit painottuvatkin ryhmissä eri tavoin. Sen sijaan kieliryhmien sisällä tyttöjen ja poikien kielellinen tarkkuus poikkeaa toisistaan voimakkaasti, sillä poikien teksteissä virheellisiä sananmuotoja tulee tekstissä vastaan lähes puolet useammin kuin tyttöjen teksteissä.



Kaavio 13. Virhetiheyden muutos alakoulun alku- ja loppuvaiheessa kirjoittajaryhmittäin.⁸⁹

Sanemäärään suhteutettu keskimääräinen virhetiheys laskee koko koululaisaineistossa jo toiselta kolmannelle vuosiluokalle lähes kolmanneksen

⁸⁹ Virhetiheys sisältää äänneiden merkitsemiseen ja äännevaihteluihin liittyvät virheet, yhteen ja erilleen kirjoittamisen virheet, taivutusvirheet (myös puuttuva taivutus), selkeät sanajärjestysvirheet ja sellaiset selkeät lausetason rakenteelliset virheet, joita ei voisi korjata välimerkein. Virhetiheys ei sisällä puhekielisyyksiä tai alkukirjain- ja välimerkkivirheitä.

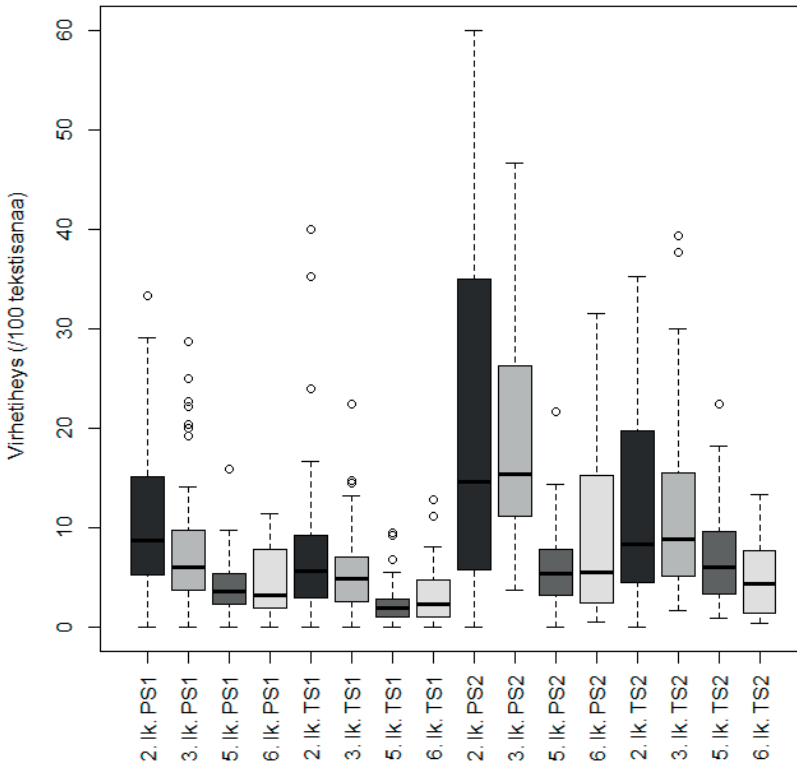
eli kahdestatoista noin kahdeksaan virheelliseen sanaan sataa juoksevaa tekstisanaa kohti. Ylemmillä vuosiluokilla virhetiheys on keskimäärin enää alle puolet alkuperäisestä eli noin viisi virheellistä tekstisanaa sataa juoksevaa sanaa kohti. Käytetyllä mittarilla arvioituna kirjoittaminen on viidennellä luokalla siis jo hyvinkin tarkkaa. Viidennen ja kuudennen vuosiluokan välillä virhetiheys ei kuitenkaan enää laske, eikä virhetyyppien keskinäisissä suhteissa tapahdu merkittävää muutosta.

Virhetiheyden muutos on selkeä kaikissa kirjoittajaryhmissä. Eniten kirjoittaminen on kolmen vuoden aikana tarkentunut S2-poikien ryhmässä, jossa lähtötaso oli heikoin. Kun vielä alaluokilla keskimäärin viidennes kaikista tekstisanoista sisälsi tässä ryhmässä yhden tai useamman virheen, on virhetiheys ylemmillä luokilla yli puolet pienempi. Absoluuttinen virheiden määrä on vähentynyt vielä enemmän, kun yksittäisiin saneisiin ei enää sisälly yhtä usein monia virheitä. Kaiken kaikkiaan virheellisten sanojen tekstitiheys on silti S2-oppijoilla vielä alakoulun loppuvaiheessa S1-oppijoita suurempi (Wilcoxonin testi $W = 3718.5$, $p < 0,001$) ja pojilla suurempi kuin tytöillä ($W = 8161$, $p < 0,01$).

Tuloksen tulkinnassa olennaista on muistaa, että lähes kaikki kirjoittajat ovat tutkimuksen ensimmäisessä vaiheessa vielä aloittelevia kirjoittajia, jolloin on luonnollista ja oletettavaa, että kirjoittamisen tarkkuudessa on paljon horjuntaa. Vaikka heikkojen kirjoittajien on perinteisesti havaittu poikkeavan vahvoista juuri suuremman virhemäärän perusteella, virheiden tiedetään myös kasautuvan tietyn tyyppisiin tekstiympäristöihin, esimerkiksi ylipitkiin tavuihin ja sanavartaloihin pääteaineksen sijaan (Kulju & Mäkinen 2009, 510–512, 522). Sekä aineis-toni laadullinen että määrällinen tarkastelu tukevat näitä havaintoja: edistyneiden alakouluikäisten teksteissä tavarakenteen karkeat virheet (kuten puuttuva, ylimääräinen tai tunnistamaton tavu) ovat vapaassa tuottamisessa kaikissa sana-asemissa ensinnäkin hyvin harvinaisia.⁹⁰ Lisäksi sana-asemalla ja sanan kokonaisrakenteella on merkitystä etenkin äänneiden kestoon liittyvien virheiden syntyymiseen: painottomassa asemassa epätarkkuudet ovat yleisempiä. Hajontakuvio (kaavio 14) puo-

90 Tutkimuksesta on rajattu pois dysfasia-diagnosin saaneet oppijat ja kokoaikaisessa erityisopetuksessa opiskelevat oppijat molemmista kieliryhmistä.

lestaan kertoo, että vaikka kirjoittajaryhmien väliset erot tarkkuudessa ovat suuret, jokaisessa kirjoittajaryhmässä on jo alaluokilla myös erittäin tarkkoja kirjoittajia.



Kaavio 14. Virhettiheyden kehitys eri oppijaryhmissä (julkaistu aiemmin Honko 2013, 153).

Jo varhaisessa kielenkehityksessä 2,5-vuotiailla S1-oppijoilla ilmenneet sanarakenteen prosessoinnin vaikeudet voivat näkyä kouluiässä lukutaidon tasossa (ks. tarkemmin Kulju & Mäkinen 2009, 512). Kirjoittamiseenkaan alkutaipaleella vaikeudet eivät ole satunnaisia, sillä aiemman tutkimuksen perusteella ne yhdistyvät muun muassa juuri hitaaseen ja takeltelevaan lukemiseen. Omassa tutkimuksessani kirjoittajaryhmien välinen järjestys kirjoittamisen tarkkuudessa säilyi samana alakoulun

edetessä. Näin on lähes kaikkien tarkasteltujen piirteiden hallinnassa, vaikka kussakin kirjoittajaryhmässä tapahtuu samaan aikaan huimaa kehitystä. Merkille pantavaa on, että myös yksilötasolla kirjoittamisen tarkkuus on 60 koululaisen seurantaryhmässä kohtalaisen stabiili ominaisuus, vaikka vertailtavana on kultakin kirjoittajalta vain yksi tekstipari. Spearmanin korrelaatioanalyysin tulos ($r_s = 0,50$, $p < 0,001$) kertoo, että virheettömmiin alaluokilla kirjoittaneet oppilaat ovat todennäköisimmin kirjoittaneet ikäluokkaan verrattuna tarkasti myös kolme vuotta myöhemmin.⁹¹ Havainnot vastaavat lukemisen vaikeudesta saatua tutkimustietoa: vaikka yksittäisen kielenpiirteen hallinnassa saattaa tapahtua nopeitakin muutoksia, tekstitaitojen kokonaisvaltainen kehittyminen on hidaskäyttöprosessi, jossa roolit suhteellisen taitavien ja heikkojen kielenkäyttäjien välillä ovat melko pysyviä (ks. esimerkiksi Eloranta, Närhi & Ahonen 2009; Undheim 2009; Stothers & Klein 2010).

On muistettava, että kirjoittamisen tarkkuus ei aina ole merkki erinomaisesta kirjoittamisesta vaan se saattaa toisen ja vieraan kielen oppimisen kontekstissa yhdistyä tekstin leksikaaliseen ja rakenteelliseen yksinkertaisuuteen (ks. esimerkiksi Skehan 1998). Mukavuusalueelta kurkistelua – kuten monipuolisten sanavalintojen yhteyttä kirjoittamisen tarkkuuteen – voidaan tarkastella suunta-antavasti yksittäisten tekstien eri piirteitä vertaamalla. Tähän palataan seuraavassa analyysiluvussa, kun tekstien yksilöllistä virhetiheyttä verrataan muun muassa sanaston monimuotoisuuteen, harvinaisten sanojen osuuteen ja keskimääräiseen sanepituuteen.

VIRHEETTÖMÄT TEKSTIT JA YKSILÖLLISET EROT

Kokonaan virheettömiä kirjoitelmia on 5.–6. luokan aineistossa vajaat 5 prosenttia kaikista kirjoitelmista ($n_{PS1} = 4$, $n_{TS1} = 9$) tai noin 6 prosenttia, mikäli alkukirjainvirheet jätetään huomioimatta. Ylemmillä vuosiluokilla kaikki virheettömät tekstit löytyvät S1-kirjoittajien aineistosta. Vaikka kirjoittamisen tarkkuus on kehityksellinen piirre, siihen liittyy

91 Esimerkiksi leksikaalisen diversiteetin eli sanaston monimuotoisuuden suoritustason pysyvyys on samassa aineistossa ja samoilla kirjoittajilla selvästi matalampi, vaikka sekin on tilastollisesti merkitsevä ($r_s = 0,33$, $p < 0,01$).

erottamattomasti sujuvuuden ja usein myös kompleksisuuden kärsiminen (Housen & Kuiken 2009, 465). Alimmilla vuosiluokilla virheetömät tekstit ovatkin molemmissa kieliryhmissä sekä sana- että lauserakenteiltaan yksinkertaisia ja tyyppillisesti hyvin lyhyitä, monet vain yhden kokonaisen virkkeen mittaisia (ks. esimerkit 8–9).

8) *Minä unelmoin että minä osaisin mennä kauppaan.* (2. lk. PS2, albania)

9) *Minä nukuin, nukuin ja nukuin* (2. lk. PS1).

Silloin, kun tekstissä esiintyy virheitä, niitä esiintyy yleensä useissa sanoissa. S1-pojilla virheen sisältävien sanojen tyyppiarvot ovat yläluokilla 1 ja 3 (maksimi 18), S2-pojilla 1, 2 ja 8 (maksimi 34), S1- ja S2-tyttöillä 4 (maksimi arvot 21 ja 36). Tyyppiarvot ja maksimiarvot eivät erilaajuisissa aineistoissa ole suoraan vertailukelpoisia, mutta ne kuvaavat silti yksilöiden välisen vaihtelun suuruutta. Hajonta eri tekstien tarkkuudessa on erittäin suurta, sillä kun koko 5.–6. luokkien aineistossa virheen sisältävien tekstisanojen esiintymistiheys on noin 5,4 sataa tekstisanaa kohden, on keskihajonta lähes 5,0. S2-oppijoille monivirheisyys on kuitenkin tyyppillisempää.⁹²

Poikien kertomuksista vain kolme on kokonaan kirjoitettu virheettömällä yleiskielellä, sillä muissa puhekielen rakenteet ja sanasto joko liittävät kirjoitetun kielen muotoihin tai puhekielisyys hallitsee muotoja kauttaaltaan (esimerkki 10). Puhekielen konventioiden käyttö teksteissä on kuitenkin jo alaluokilla toisinaan tarkasti säädeltyä esimerkiksi niin, että se kattaa vain henkilöhahmojen puheenvuorot tai säilyy tyyllisesti eheänä minäkertojan äänessä läpi tekstin.

10) *Mä haluaisin et mun elämäis ois ees yks koulupäivä jollo mua ei kiusattais. Ja mä toivoisin et saman päivän olis sisäliikkaa. Ja mä toivoisin et saman päivän ois ruokan perunoit ja nakkikastiket.* (3. lk. PS1, Uusimaa)

92 Kun vähintään viisi virheen sisältävää tekstisanaa esiintyy 5. ja 6. vuosiluokalla keskimäärin puolella kirjoittajista, S1-ryhmässä osuus on 40 prosenttia ja S2-ryhmässä 64 prosenttia.

Kirjoitetun kielen konventioiden haltuun ottamisen haastavuudesta kertoo toisaalta se, kuinka myös hyvin taitavasti rakennetuissa teksteissä esiintyy epäsystemaattisesti käytettyjä puhekielen muotoja. Puhekielisyudet ovat lisäksi huomattavasti yleisempiä kertomusten loppu- kuin alkupuoliskoilla, mikä voi kertoa kirjoittajan väsymisestä tekstin loppua kohden.

S1-tytöillä virheettömät kertomukset ovat jo alaluokilla selvästi pidempiä kuin S1-pojilla, ja myös niistä löytyy puhekielelle tyypillisiä piirteitä. Poikien tekstien tapaan puhekielisyys esiintyy enemmän kertomusten loppu- kuin alkupuolella. S1-tyttöjen alaryhmässä on silti sekä alimmilla vuosiluokilla että alakoulun päättövaiheessa suhteessa eniten sellaisia vahvoja kirjoittajia, joiden teksteissä sujuvuus, tarkkuus ja kompleksisuus yhdistyvät: monet tekstit ovat sekä pitkiä, sanastoltaan ja rakenteiltaan monipuolisia että virheettömiä, ja kielen eri rekisterien käyttö on hallittua.

Sekä poikien suurempi virhetiheys että virheettömien tekstien pituus ja laatuero vahvistavat aiempien tutkimusten käsitystä poikien suuremmista vaikeuksista äidinkielellään kirjoittamisessa. Kielenhallinta on verrokkitestinä toimineessa sanastotestissä pojilla samaa tasoa kuin tytöillä tai jopa korkeampi (Honko 2013), mutta poikien kirjoittaminen on oikeakielisuuden ja rekisterivalintojen suhteen epätarkempaa ja rakenteeltaan eheät tekstit ovat harvinaisempia. Tekstitaitotutkimuksessa rajaukset kuten valittu genren tai tekstin käsittelemisen kanava voivat vaikuttaa tuloksiin. Silti sekä kansallisten arviointien että PISA-kartoitusten kautta on nähtävissä, että sukupuolten väliset erot luku- ja kirjoitustaidossa eivät tasoitu pelkästään tarkastelukontekstia modifioimalla (ks. esimerkiksi Lappalainen 2011; Arffman & Nissinen 2015).

YLEISIMMÄT VIRHETYYPIT

Virheiden luokittelu (ks. taulukot 44–45) osoittaa, että eniten tarkkuuden pulmia alakoululaisten kirjoituksessa aiheuttavat sanojen yhteen ja erilleen kirjoittaminen, kaksoiskonsonantit sekä isot ja pienet alkukirjaimet.⁹³

93 Taulukoihin 44–45 on listattu vain kummankin keruuvaiheen 22 yleisintä virhetyppiä. Aineistoon koodatuista virhetyypeistä osa on jäänyt harvalukuisuuden vuoksi taulukon ulkopuolelle. Tällaisia piirteitä

Käsin kirjoitettaessa tyypillisiä ovat myös kirjainten piirtämiseen liittyvät haasteet. Lähes kaiken tyyppiset kirjoittamisen epätarkkuudet käyvät alakoulun loppuvaiheeseen tultaessa silti harvinaisemmiksi. Myös virheiden kokonaismäärä on pienempi, vaikka ylempien luokkien tekstiaineisto on sanamääräisesti noin kolmanneksen alempia vuosiluokkia suurempi. Muutaman yleisimmän virhetyypin (ylimääräinen ja puuttuva sanaraja, puuttuva geminaatta) osuus kaikista virheistä on kummassakin keruuvaiheessa suuri, mutta ylempillä luokilla virheiden jakauma on tässä suhteessa hiukan tasoittunut yhdysanavirheitä lukuun ottamatta. Taivutukseen ja sanavalintoihin liittyvät virheet ovat sikäli kiinnostavia, että vaikka nekin harvinaistuvat, niistä tulee muihin virheisiin nähden tyypillisempiä. Taivutuksen virheiden lisääntyminen voi liittyä siihen, että haltuun otetaan monimutkaisempia ilmaisukeinoja samalla, kun kertomuksen sisällöllinen ja rakenteellinen tasokkuus kasvaa. Havainto vastaa dynaamisten systeemien viitekehyksessä tehtyä suomen kielen oppimisen kulkua esittelevää tutkimusta (Lesonen 2020).

Taulukko 44. Yleisimmät virhetyypit kertomusteksteissä 2.–3. luokalla.⁹⁴

	Virhetyyppi	f	%	Kumulatiivinen tiheys / 100 sanaa
1	Ylimääräinen sanaraja	523	19,9	2,0
2	Puuttuva geminaatta	373	14,2	3,4
3	Puuttuva sanaraja	346	13,1	4,8
4	Alkukirjainvirhe ⁹⁵	236	9,0	5,7
5	Piirtovirhe	171	6,5	6,3
6	Puuttuva pitkä vokaali	116	4,4	6,8
7	Väärä taivutus	102	3,9	7,2
8	Puuttuva vokaali	77	2,9	7,5

ovat äänneiden erotteluun liittyvä vokaalin ja konsonantin sekoittuminen (esimerkiksi i/j) sekä sanan kirjoitusasuun liittyvät reversaali, rotaatio, reduplikaatio ja sanan osittainen toisto (kuten altaaseeneen 'altaaseen', voitthisin 'voittaisin').

94 Puhekielisyksistä virhetilastoissa ovat mukana ainoastaan virheellisenä lekseemivariantana tulkittavissa olevat tapaukset (*kun/kuin, vaan/vain, niin/tosi, joka/mikä/kuka*). "Muu lauserakenteen ongelma"-kohta kattaa sanatasoa laajemmat virhetyypit (esim. lausekkeen toisto tai puuttuva lauseke). Omina kohtinaan on koodattu yleisimmät lausetason ongelmat eli a) väärä sanajärjestys, b) puuttuva/ylimääräinen (yksittäinen) sana sekä c) virheellinen lauseenvastike. Kohta puuttuva/virheellinen taivutus sisältää myös kongruenssivirheet ja väärin valitut rektiosijat.

95 Alkukirjainvirheet eivät ole mukana virhetiheyttä koskeissa analyyseissa.

9	Puuttuva konsonantti	77	2,9	7,8
10	Väärä sana	77	2,9	8,1
11	Ylimääräinen geminaatta	73	2,8	8,3
12	Puuttuva sana (muu)	71	2,7	8,6
13	Puuttuva taivutus	62	2,4	8,9
14	Ylimääräinen pitkä vokaali	42	1,6	9,0
15	Väärä vokaali	41	1,6	9,2
16	Väärä konsonantti	37	1,4	9,3
17	Rauniosana	37	1,4	9,5
18	Irrallinen sananosa	68	2,6	9,7
19	Ylimääräinen vokaali	29	1,1	9,8
20	Ylimääräinen konsonantti	27	1,0	9,9
21	Soinnillisuus	25	0,9	10,0
22	Astevaihtelu	24	0,9	10,1
	Yhteensä	2634	100	

Taulukko 45. Yleisimmät virhetyypit kertomusteksteissä 5.–6. luokalla.

	Virhetyyppi	Vrt. 2.–3. luokka	f	%	Kumulatiivinen tiheys /100 sanaa
1	Ylimääräinen sanaraja	1	323	19,0	0,9
2	Väärä taivutus	7	183	10,8	1,4
3	Puuttuva sanaraja	3	161	9,5	1,9
4	Alkukirjainvirhe	4	120	7,1	2,2
5	Puuttuva geminaatta	2	103	6,1	2,5
6	Väärä sana	10	93	5,5	2,8
7	Puuttuva pitkä vokaali	6	92	5,4	3,1
8	Piirtovirhe	5	82	4,8	3,3
9	Puuttuva sana (muu)	12	81	4,8	3,5
10	Ylimääräinen geminaatta	11	59	3,5	3,7
11	Puuttuva taivutus	13	56	3,3	3,8
12	Puuttuva konsonantti	9	46	2,7	4,0
13	Puuttuva vokaali	8	46	2,7	4,1
14	Ylimääräinen konsonantti	20	40	2,4	4,2
15	Muu lauserakenteen ongelma		40	2,4	4,3
16	Sanajärjestys		33	1,9	4,4
17	Astevaihtelu	22	33	1,9	4,5
18	Väärä konsonantti	16	31	1,8	4,6
19	Ylimääräinen pitkä vokaali	14	25	1,5	4,7
20	Väärä vokaali	15	22	1,3	4,7
21	Irrallinen sananosa	18	16	0,9	4,8
22	Rauniosana	17	13	0,8	4,8
	Yhteensä		1698	100	

FREKVENTIT VIRHETYYPIT KIRJOITTAJARYHMITÄIN

Poikien kirjoittaminen on molemmissa kieliryhmissä ja useiden eri kielipiirteiden osalta keskimäärin epätarkempaa kuin tyttöjen: virheitä tehdään useammin (ks. kaavio 13). Tämä vastaa aikaisempien tutkimusten havaintoja kirjoitustaidosta eri ikävaiheissa (ks. esimerkiksi Mäki 2002; Pajunen 2012). Myös lukivaikeusriskiryhmässä oikeinkirjoitustaidoissa on havaittu alkuopetusvuosina eroja sukupuolten välillä tyttöjen eduksi, ja varhain lukemaan oppineilla pojilla kiinnostus kirjaimiin oli jo esiopetusvaiheessa tyttöjä vähäisempää ja tehtävää välttävä strategia tyttöjä yleisempi (Lerikkanen ja muut 2010). Koska poikien kirjoittaminen tarkastellussa aineistossa on myös sanamääräisesti niukempaa ja kertomusrakenteen saavuttaminen harvinaisempaa, on perusteltua kuvata poikien kirjoittamista tässäkin tutkimuksessa keskimäärin selvästi tyttöjä heikommaksi. Mikäli tarkastellaan yksinomaan kirjoittamisen virheitä, S1-pojilla vaikuttaa kuitenkin olevan ongelmia lähinnä saneen foneemisen esityksen koodaamisessa kirjoitukseksi,⁹⁶ kun S2-ryhmässä erot poikien ja tyttöjen välillä ovat vielä alakoulun alkuvaiheessa laaja-alaisempia. Tarkastelluissa S2-poikien teksteissä esimerkiksi sanojen taivuttamisen hallinta näyttää kehittyvän keskimäärin tyttöjä myöhemmin.

Maahanmuuttajataustaisia oppilaita opettaneet opettajat ovat listanneet tyypillisiksi S2-kirjoittamisen virhepiirteiksi pitkien vokaalien ja geminaattojen kestot, rektiovirheet, verbien ja adjektiiviattribuuttien kongruenssivirheet sekä korvakuulolta kirjoitettujen vasta opittujen sanojen virheellisen kirjoittamisen (Lyly 2008, 153).⁹⁷ Kirjoitetun kielen sijataivutuksessa lukioikäisillä S2-oppijoilla eniten ongelmia tuottaviksi piirteiksi on havaittu objektin sijavalinnat ja paikallissijat (Elo 1998, 36–38). Mannisen ja Martinin (2000) mukaan venäjänkielisten ylioppilaskokelaiden tyyppivirheitä ovat kvantiteettiin, viittaussuhteisiin, sanavallintoihin, vokaalisointuun, kongruenssiin, objektiin, verbimuotoihin sekä sijataivutukseen liittyvät ongelmat. Vaikka S2-kirjoittajilla tiedetään olevan kirjoittamisen haasteita myös virketasolla muun muassa lausera-

96 Väite voidaan perustella myös sadutusnauhoitusten eli toisen ja kolmannen vuosiluokan oppilailta kerätyn suullisen kertomusaineiston avulla (ks. myös Honko 2017a ja 2017b).

97 Mahdollisia oman äidinkielen vaikutuksia kirjoittamisen puutteisiin opettajat eivät kuitenkaan välttämättä tarkemmin pysty erittelemään, mikä ei ole ihme, sillä tehtävä on lingvistillekin useimmiten vaikea.

kenteissa ja virkerajojen hahmottamisessa, suomen opettajat eivät pidä näitä leimallisina piirteinä juuri maahanmuuttajataustaisille oppilaille (Lyly 2008, 153). Kaikille oppilaille yhteisinä kirjoittamisen pulmapiirteinä opettajat ovat Lylyn kyselyssä maininneet välimerkkivirheet, virheet yhteen ja erilleen kirjoittamisessa sekä puhekielisyydet ja ongelmat viittaussuhteissa (*joka* ja *mikä*). (Lyly 2008, 153–154). Yleisistä virheistä on koululaiskirjoittamisen yhteydessä tapana mainita erikseen myös omistusliitteen merkitsemättä jättäminen (ks. esimerkiksi Elo 1998, 36; Turunen 2012). Tässä tutkimuksessa olen katsonut possessiivisuffiksin merkitsemättä jättämisen puhekielen vaikutukseksi ja varsinaisen virheanalyysin ulkopuolelle.

Kirjoittajaryhmittäinen tarkastelu kertoo, että sekä alakoulun alkuaikana loppuvaiheessa keskeiset virhetyypit ja niiden väliset voimasuhteet ovat ehkä hieman yllättäenkin pitkälti samanlaiset molemmissa kieliryhmissä. Äännekesto- ja sanavalinnoissa ja taivutuksessa sekä erityisesti sanarajojen ja alkukirjainten merkinnöissä on pulmia paitsi S2-ryhmässä myös syntyperäisten kielenkäyttäjien ryhmässä. Unelmapäivätekstien virheitä ei toisin sanoen voida ainakaan ilman tarkentavaa laadullista analyysiä luontevasti luokitella siten, että jotkin virheistä olisivat vain S1- tai S2-kirjoittajien virheitä (ks. myös Hamp-Lyons 2003, 173), vaan kaikkia yleisimpiä virhetyyppejä esiintyy molemmissa kieliryhmissä. Näin on siitä huolimatta, että kyseiset virheet ovat S1-kirjoittajilla huomattavasti harvinaisempia ja liittyvät useammin puhekielen vaikutukseen eivätkä varsinaisesti kielenhallinnan puutteisiin eli esimerkiksi suomen foneemiston hallintaan. Kaikki S2-ryhmän kirjoittajat ovat asuneet Suomessa jo useita vuosia ja noin puolet heistä syntymästään saakka. Tämä tasoittaa mutta ei poista kieliryhmien välistä eroa.

Yleisimpien virhetyyppien kirjoittajaryhmäkohtaisiin taulukoihin (ks. taulukot 46 ja 47) on koottu koko aineiston tyypillisimmät virheet: toisen ja kolmannen luokan aineistossa 22 yleisimmän virhetyypin listalle mahtuvat kaikki koko korpuksessa yli 20 kertaa esiintyneet virheet, viidennen ja kuudennen luokan aineistossa kaikki yli 10 kertaa esiintyneet virheet. Luvuissa ei ole mukana puhekielisyyksistä johtuvia poikkeamia yleiskielen kirjoitusasuun eikä alkukirjainvirheitä, joiden määrä molem-

missa kieliryhmissä ja molemmissa keruuvaiheissa on hyvin suuri sekä virkkeenalkuisten että erisnimistä puuttuvien isojen kirjainten vuoksi. Juuri virkkeenalkuisten alkukirjainvirheiden koodaaminen olisi kuitenkin keinotekoista välimerkkivirheiden suuren määrän vuoksi.

Taulukko 46. Yleisimmät virhetyypit 2.–3. vuosiluokalla kirjoittajaryhmittäin.⁹⁸

	Virhetyypin osuus %	S1	PS1	TS1	S2	PS2	TS2	Yht.
1	Puuttuva yhdyssana	27,4	24,2	30,3	10,4	9,4	11,8	19,6
2	Puuttuva sanaraja	16,9	20,7	13,5	9,2	9,5	8,6	13,4
3	Puuttuva geminaatta	11,2	10,7	11,7	21,9	23,4	20,0	16,1
4	Piirtovirhe	7,2	6,5	7,7	7,6	7,2	8,2	7,4
5	Irrallinen sananosa	4,1	4,5	3,8	1,5	2,0	0,9	2,9
6	Puuttuva vokaali	4,0	3,5	4,4	2,5	2,6	2,4	3,3
7	Puuttuva konsonantti	3,1	3,0	3,2	3,6	4,4	2,4	3,3
8	Väärä taivutus	3,1	2,8	3,3	5,9	5,1	7,1	4,4
9	Ylimääräinen geminaatta	3,1	3,7	2,6	3,2	3,1	3,3	3,2
10	Puuttuva pitkä vokaali	2,9	3,3	2,6	7,5	7,9	6,9	5,0
11	Väärä konsonantti	2,4	2,2	2,6	0,7	0,5	0,9	1,6
12	Väärä sana	2,2	1,7	2,7	4,6	3,8	5,8	3,3
13	Puuttuva sana	1,9	1,5	2,3	4,4	3,3	6,0	3,1
14	Soinnillisuus	1,7	1,5	2,0	0,3	0,2	0,4	1,1
15	Puuttuva taivutus	1,6	1,5	1,7	4,0	3,6	4,4	2,7
16	Rauniosana	1,4	2,3	0,6	1,8	2,8	0,4	1,6
17	Ylimääräinen vokaali	1,4	1,5	1,4	1,0	1,2	0,9	1,3
18	Väärä vokaali	1,4	1,7	1,1	2,3	2,0	2,7	1,8
19	Ylimääräinen konsonantti	1,2	0,5	1,8	1,1	1,2	1,1	1,2
20	Ylimääräinen pitkä vokaali	0,8	1,5	0,2	3,0	3,8	2,0	1,8
21	Astevaihtelun virhe	0,5	0,5	0,5	1,7	2,0	1,3	1,0
22	Väärä sanavartalo	0,4	0,7	0,2	1,7	1,2	2,4	1,0
	% yhteensä	100	100	100	100	100	100	100
	kpl yhteensä	1 258	599	659	1 059	608	451	2 317

98 Ei sisällä puhekielisyyskiä eikä alkukirjainvirheitä. Taulukossa on esitetty kunkin virhetyypin esiintymien osuus kaikista taulukkoon sisällytävistä virhetyypeistä. Tarkemmat selitteet löytyvät käsittelyluvuista.

Taulukko 47. Yleisimmät virhetyypit 5.–6. vuosiluokalla kirjoittajaryhmittäin.⁹⁹

	Virhetyypin osuus %	S1	PS1	TS1	S2	PS2	TS2	Yht.
1	Puuttuva yhdyssana	32,4	29,5	35,1	11,2	13,6	9,4	19,2
2	Puuttuva sanaraja	16,4	21,5	11,3	6,6	7,5	5,9	10,3
3	Puuttuva sanaraja (muu)	6,1	4,4	7,4	4,6	6,1	3,6	5,2
4	Piirtovirhe	5,2	8,8	2,7	5,2	5,8	4,8	5,2
5	Puuttuva geminaatta	4,6	4,0	5,1	7,8	10,0	6,2	6,6
6	Ylimääräinen geminaatta	4,4	4,4	4,5	3,4	2,2	4,3	3,8
7	Väärä taivutus	3,9	3,2	4,5	16,4	14,8	17,6	11,7
8	Ylimääräinen konsonantti	3,5	2,0	4,8	1,9	0,7	2,8	2,6
9	Puuttuva konsonantti	3,2	4,4	2,4	2,8	2,9	2,7	2,9
10	Väärä sana	3,2	1,6	4,5	7,6	5,1	9,4	5,9
11	Puuttuva vokaali	2,9	4,0	2,1	3,0	4,4	2,0	2,9
12	Sanajärjestys	1,9	1,6	2,1	2,3	1,2	3,0	2,1
13	Väärä konsonantti	2,0	2,0	2,1	1,9	1,7	2,1	2,0
14	Puuttuva pitkä vokaali	1,7	0,8	2,4	8,4	6,3	9,9	5,9
15	Rakenne	1,7	1,6	1,8	3,1	2,9	3,2	2,6
16	Puuttuva taivutus	1,5	1,2	1,8	4,8	5,1	4,6	3,6
17	Pitkä vokaali	1,5	1,2	1,8	1,3	1,7	1,1	1,4
18	Ylimääräinen pitkä vokaali	1,4	1,2	1,5	1,7	1,0	2,3	1,6
19	Astevaihtelun virhe	0,5	0,4	0,6	3,1	2,9	3,2	2,1
20	Irrallinen sananosa	1,0	1,6	0,6	1,0	1,2	0,9	1,0
21	Sanarakenne	0,2	0,0	0,3	1,2	1,7	0,9	0,8
22	Ylimääräinen vokaali	0,8	0,8	0,9	0,6	1,2	0,2	0,7
	% yhteensä	100	100	100	100	100	100	100
	kpl yhteensä	592	251	336	975	412	563	1 567

TOISTUVAT VIRHETYYPIT

Tässä alaluvussa esittelen tarkemmin aineistossa useimmin toistuvia virhetyyppejä: a) sanojen yhteen ja erikseen kirjoittamisen pulmia, b) ongelmia äännekestoissa, c) grafeemien valinnassa sekä d) grafeemien piirtämisessä ja siihen liittyen etu- ja takavokaalien erottamisessa, e) väärässä ja puuttuvassa taivutuksessa sekä f) sanajärjestyksessä ja g) sanavalinnoissa. Samalla erittelen kirjoittajaryhmien välisiä eroja kirjoittamisen tarkkuudessa ja sen kehitymisessä.

99 Ei sisällä puhekielisyyksiä eikä alkukirjainvirheitä. Taulukossa on esitetty kunkin virhetyypin esiintymien osuus kaikista taulukkoon sisällyvistä virhetypeistä.

SANARAJOJEN MERKITSEMINEN

Yhdyssanojen virheellinen merkintä eli osien kirjaaminen erillisinä on S1-oppijoilla tyypillisin virhe läpi alakoulun (*lajin vaihto koje, uima allas*), ja se on myös S2-ryhmässä kummassakin keruuvaiheessa toiseksi yleisin virhetyyppi. Jo alaluokilla yhdyssanoista on silti onnistuneesti kirjoitettu yhteen noin 60 prosenttia jokaisessa kirjoittajaryhmässä, ylemmillä vuosiluokilla poikien ryhmissä lähes 80 prosenttia ja tyttöjen ryhmissä yli 80 prosenttia. Tämä vastaa aikaisempia havaintoja yhdysanatajun kehittymisestä erityisesti kolmannen vuosiluokan jälkeen (Stenberg 2010).

Yhdyssanavirheet ovat lähes poikkeuksetta yhdyssubstantiivivirheitä, mikä johtuu muiden sanaluokkien yhdyssanojen vähäisestä esiintymismäärästä. Niihin yhdistyy molemmissa kieliryhmissä sekä tytöillä että pojilla muutakin kirjoittamisen epätarkkuutta – alemmilla luokilla noin 20 prosentissa tapauksista (*huipu auton, jatski kioski*), ylemmillä noin 10 prosentissa. Luvut peilaavat sekä pitkien sanojen kirjoittamisen vaikeutta, sillä virheitä on yhdyssanoissa enemmän kuin tekstisanoissa keskimäärin, että pitkienkin sanojen kirjoittamisen hallinnan kehittymistä alakoulun aikana.

Ylimääräisiä sanarajoja esiintyy alakoululaisten aineistossa myös yhdistämättömissä sanoissa, mutta nämä virheet eivät ole sattumanvaraisia. Ilmiö on yleisin koulunkäynnin alkuvaiheessa, jolloin niin liitepartikkeleita, johtimia kuin tavujakin kirjoitetaan välillä omiksi yksiköikseen. Molemmissa keruuvaiheissa useamman esiintymän tapauksia ovat additiivinen *-kin* (*ennen kin, kavereita kin*) sekä *-lAinen*-loppuiset adjektiivijohdokset (*kaiken laisia, saman laisia*). Toisella ja kolmannelle vuosiluokalle muutamit oppilaat myös tavuttavat vielä, tosin lähinnä paikallisesti, sanoja sanavälien avulla mahdollisesti oikeinkirjoituksen tai sanan taipumisen hahmottamisen helpottamiseksi (*per hee seen kuuluu, mun lem pi ruo ka on pa ler mon pasta, nukku ma an*). Vain hyvin pieni osa sanarajavirheistä viittaa silkkaan huolimattomuuteen, piittaamattomaan kirjoitustyyliin tai laaja-alaiseen sanarajojen hahmottamisen pulmaan (*siitäalkoisei kkailu*).

Sanojen ketjuttaminen on kertomusaineistossa tyypillisempää pojilla kuin tytöillä, jotka näyttävät oppivan sanarajojen tarkan merkitsemisen

keskimäärin varhemmin. Vaikka ketjuttaminen vähenee ylemmillä luokilla, puuttuu sanarajoja molemmissa kieliryhmissä enemmän poikien kuin tyttöjen teksteistä. Ketjutusvirheiden syntymiseen ja tulkintaan saattaa vaikuttaa tekstien kirjoittaminen käsin. Niissäkin on silti havaittavissa säännönmukaisuutta, kuten pikapuhemuototyyppisiä yhdistelmiä (*jolvoi*), määritteen ja edussanan yhdistelmiä (*kokopaikan, pahaolo, parituntia*) sekä taipumattomien pikkusanojen, esimerkiksi *kanssa*-postposition ja konjunktioiden sekä kieltoverbin yhdistämistä muihin sanoihin (*iskänkaa, juurikun, eituo*). Ketjuttamisvirheiden kasautuminen merkinnee, että sanojen ominaislaadulla ja toisaalta sanojen painottamisella on kirjoittamistavasta riippumattomia systemaattisia vaikutuksia siihen, mitkä sanat mielletään yhteen kuuluviksi.

Sanarajojen kirjoittamisen vaikeudet voivat olla satunnaisia lipsahduksia tai kertoa normien tuntemuksen puutteista tai piittaamattomuudesta normeja kohtaan. Myös sanan ominaispiirteet vaikuttavat sen kirjoittamisen vaikeuteen, sillä lyhyet ja nominatiivimuotoisen määritteen omaavat sanat kirjoitetaan yleensä muita tarkemmin (ks. myös Stenberg 2010). Sanojen yhteen ja erilleen kirjoittamisen taito kehittyä, kun kirjoitetun yleiskielen normit tulevat tutuiksi. Yhdyssanojen ja sanaliittojen kirjoittamisen tarkkuudella on kuitenkin tekstin pintarakennetta syvemmälle yltävä tehtävä: Sanarajojen tehtävä on auttaa kirjoituksen kuvaaman käsitteistön ja näistä syntyvien merkityskokonaisuuksien hahmottamista. Jos muodon ja merkityksen yhteys horjuu, lukija joutuu ponnistelemaan merkitysten tulkinnassa. Sanojen yhteen ja erilleen kirjoittaminen ei ole ”pelkkä oikeinkirjoitusasia”, vaan kyse on myös tavasta hahmottaa sanarakenteita ja jäsentää maailmaa kielen keinoin (Eronen, Maamies & Räikkälä 1996). Yhdyssanojen kirjoittaminen kertoo, miten hahmotamme maailmaa, ja tyypillisessä tilanteessa yhdyssana ilmaisee yhden käsitteen.

ÄÄNNEKESTOT

S2-ryhmässä eniten epätarkkuutta on alaluokilla geminaattojen kirjoittamisessa. S2-pojilla geminaatta on jäänyt yksinäiskonsonantiksi yli puolet useammin kuin yhdyssana kirjoittamatta yhteen. Ero on merkittävä, sillä S1-pojilla suhde on lähes yhtenevästä yhdyssanatiheydestä huoli-

matta päinvastainen. Myös S₂-tytöillä geminaattavirheet ovat virheistä yleisimpiä, vaikka niiden tekstitiheys on pienempi kuin pojilla.

Suhteellisesti eniten vaikeuksia on tuottanut geminaatta *kk* (*jääkieko, karkia, marokoon*). Yleisiä ovat myös virheet merkinnöissä *ll*, *mm*, *nn*, *pp*, *ss* ja *tt* (*sielä, myöhemmin, Särkänniemeen, apelsiini, kansa, kiltti*), harvinaisempia muutoinkin harvinaisemmissa yhdistelmissä *ff* ja *rr* (*lefaa, harastamaan*). Tuloksia voivat selittää oppijan äidinkielen ja mahdollisesti hallitsemien muiden kielten suomesta poikkeava äänteiden yhdistyminen: kaksoiskonsonantteja ei välttämättä ole tai niiden joukko on rajatumpi. Geminaattatyypeillä voi olla myös akustisia eroja, jotka vaikuttavat tuloksiin. Yksi vaikeutta selittävä tekijä saattaa lisäksi löytyä suomen kielen astevaihtelun yliyleistämistä.

Alaluokilla kesto-*virheitä* on toisinaan useita samassa sanassa (*tykän, pikusiskole, joukuessa*). Tähän voivat vaikuttaa niin pitkien sanojen kompleksisuus (prosessoinnin kuormittavuus), harvinaisuus (ei-tuttuus) kuin silkkä sanapituuskin (*virheitä* mahtuu sanaan enemmän). Toisinaan kirjoittaja on hahmottanut, että sanaan kuuluu jokin pitkäkestoinen äänne, mutta pitkän keston sijainti on paikallistunut väärin (*etää, leenän, mennisime, semoisselle*). Tällaiset kompensatiot näyttäivät kasautuvan samoille kirjoittajille ja olevan ominaisia alimmilla vuosiluokilla käsillä olevalle kehitysvaiheelle. Myös oikein merkittyyn pitkään vokaaliin yhdistyy usein geminaattavirhe, ja vastaavasti vokaalin kesto-*virhe* reaalistuu usein sananmuodossa, jossa on geminaatta (*kunnella, liikkelle*). Ilmiö kuvastaa kesto-*virheiden* merkintöjen tarkentumisen asteittaista luonnetta: silloinkin, kun kirjoittaja tiedostaa kahdella kirjainmerkillä kirjoittamisen konvention idean ja erottaa ainakin osan puhutun kielen kesto-*virheistä*, hän ei välttämättä erota pitkiä kesto-*virheitä* kaikista äänneistä tai kaikista sana-*virheistä* tai merkitse kesto-*virheitä* säännönmukaisesti.

Sekä sukupuoli että kielitausta vaikuttavat geminaattavirheiden suhteelliseen määrään, vaikka yksilölliset erot ovat hyvin suuria. Alaluokilla geminaattojen onnistumisprosentti on S₁-kirjoittajien teksteissä sekä tyttöjen että poikien ryhmissä keskimäärin yli 95 kaikissa muissa yhtymissä paitsi *mm:ssä*, joka on pojilla jäänyt yksinäiskonsonantiksi noin joka kymmenennessä tapauksessa. S₂-ryhmän pojilla geminaattavirheet ovat yleisimpiä: eniten virheitä ovat aiheuttaneet *kk* (onnistumisprosentti

noin 60 %, S2-tytöillä vastaavasti 85 %), *pp* (70 % / > 95 %), *ss* (70 %, > 95 %), *tt* (83 %, 93 %), *mm* (87 %, 71 %; mukana ei ole monikon 1. persoonaa korvaavia passiivimuotoja) ja *ll* (90 %, 93 %).

Vaikka geminaattavirheet ovat suhteellisen yleisiä, alaluokillakin vain viiden oppijan tekstissä niitä esiintyy runsaasti, yli 10 kappaletta. Kyseisistä oppijoista yksi on toisluokkalainen S1-tyttö, yksi kolmasluokkalainen S1-poika, kaksi arabiaa äidinkielenään puhuvaa toisluokkalaista poikaa ja yksi toisluokkalainen vironkielinen tyttö. Havainnot ovat linjassa aiemman tutkimustiedon kanssa: Geminaattojen tiedetään yleisesti tuottavan pulmia aloitteleville kirjoittajille (Kulju & Mäkinen 2009, 510–516, 524). Toisaalta geminaattavirheet saattavat olla merkki lukivaikkeudesta tai S2-ryhmässä suomen kielen kestoasteiden omaksumisen keskeneräisyydestä. Oli merkintöjen epätarkkuuden taustalla mitä syitä tahansa, niiden vaikutukset ovat alakoulun loppuun mennessä jo hyvin vähäiset: 5.–6. luokalla S2-poikienkin ryhmässä kaikkien geminaattojen merkintä on onnistunut yli 90-prosenttisesti. Vaikeimpia ovat alakoulun loppuvaiheessa tässä kirjoittajaryhmässä olleet *nn* (oikein 92 %) sekä *kk* (oikein 94 %).

Pitkien vokaalien merkinnät ovat kertomusteksteissä jo koulunkäynnin alkuvaiheessa huomattavasti tarkempia kuin pitkien konsonanttien merkinnät, ja jokaisen vokaalin merkinnät tarkentuvat edelleen alakoulun aikana. Vokaali itsessään vaikuttaa suuresti virheiden määrään: *yy:n* ja *öö:n* merkinnöissä ei esiinny virheitä, koska tavoitemuotoinkin nämä ovat harvinaisia. Määrällisesti eniten virheitä on molemmissa keruuvaiheissa aiheuttanut *aa:n* kirjoittaminen. Alaluokilla *ee:n* suhteellinen virheprosentti on kuitenkin suurin (noin 3,4 %). Myös sana-asemalla on merkitystä, sillä pitkät vokaalit redusoituvat kirjoituksessa lähinnä painottomista asemista sananlopuista ja keskeltä, eivät sanojen alusta. Alakoulun päättövaiheessa enää muutama tapauksista on sanassa (pää) painollisessa asemassa.

Pitkiin vokaaleihin liittyvät virheet ovat Suomessa pitkään asuneilla S2-oppijoilla huomattavasti yleisempiä kuin S1-ryhmässä. Vielä toisella ja kolmannella vuosiluokalla pitkän vokaalin redusoituminen lyhyeksi on lisäksi selvästi yleisempää S2-pojilla kuin S2-tytöillä ja hieman yleisempää pojilla myös S1-ryhmässä. Alakoulun loppuvaiheessa tyttöjen ja poikien välinen ero on kuitenkin kadonnut molemmissa kieliryhmissä.

Pitkien vokaalien ja konsonanttien kirjoittamisen haastavuus etenkin koulunkäynnin alussa on todettu myös lukuisissa sanelukirjoittamiseen perustuvissa aikaisemmissa tutkimuksissa. Näissä tutkimuksissa on havaittu, että sanan harvinaisuus ja taivutustyyppin tunnusmerkkisyys sekä tavun pituus kasvattavat virheen todennäköisyyttä kestojen merkinnässä (ks. Kulju & Mäkinen 2009, 510–511). Sanelukirjoittamisessa geminaattojen ja konsonanttiyhtymien kirjoittaminen on aiemminkin todettu äidinkielisille aloitteleville kirjoittajille pitkien vokaalien ja diftongien kirjoittamista haastavammaksi (mts. 520).

Virhetilastoissa huomiota kiinnittää se, että ylimääräisten geminaattojen (esimerkiksi *halussi*, *päästä*, *omaann*) tekstitiheys on S1-kirjoittajien teksteissä yhtä suuri kuin puuttuvien pitkien vokaalien, vaikka S2-ryhmässä pitkien vokaalien merkintöjen puutteet ovat huomattavasti yleisempiä. Molemmissa kieliryhmissä ylimääräisistä geminaatioista suuri osa liittyy alakoulun päättövaiheessa jo joko astevaihtelun puuttumiseen (*pikkuserkkun*), yliyleistämiseen (*kauppungille*) tai liitepartikkelin sisältävien sananmuotojen kirjoitusasuun rajageminaation yhteydessä (*kenellekkään*). Ylimääräiset pitkät vokaalit (*uneessa*) puolestaan ovat molemmissa kieliryhmissä harvinaisia, ja suhteessa eniten tällaisia virheitä ovat tehneet S2-pojat alakoulun alaluokilla. Erityisryhmän muodostavat vierassanat, joiden toisen tavun vokaali tai toisen tavun alkukonsonanti on kirjoitettu pitkänä (*Austraaliaan*, *robootti*; *vamppyyri*, *internetin*).

Kertomusaineiston pohjalta voidaan arvioida, että myös osalla Suomessa syntyneistä S2-oppijoista on alakouluvaiheessa jonkin verran vaikeuksia paitsi pitkien äännekestojen merkinnöissä myös niiden havaitsemisessa. Akustisesti selvimmät tapaukset (painolliset sananalkuiset asemat) ovat alakoulun päättövaiheessa kaikissa kirjoittajaryhmissä jo hyvin hallussa. Horjunta sekä pitkistä äännekestoista lyhyisiin että lyhyistä pitkiin saattaa myös tarkoittaa, että oppijat, jotka tiedostavat äännekestojen merkitystä erottavan tehtävän ja merkitsemisen tärkeyden, päätyvät toisinaan myös ylikorjaamaan kirjoitustaan: he hahmottavat kestojen merkitsemisen systeemitasolla mutta eivät vielä kuule tai osaa merkitä kestoja tarkasti.

GRAFEEMIN VALINTA

Puuttuva konsonantti (muu kuin geminaattavirhe) on koko aineistossa yleisempi konsonanttiyhitymissä kuin vokaalien välissä. Puolivokaalin *j* puuttuminen tekijänimissä selittää jälkimmäisistä tapauksista noin puolet (*palvelia, tutkiana*). Konsonanttiyhitymissä etenkin akustisesti heikko *h* on jäänyt toistuvasti merkitsemättä (esimerkiksi *pereeni* 'perheeni', *voikaisi* 'voihkaisi'), samoin *η*-äänteen *n* (esimerkiksi *kunigas, sägystä*). Yleisiä nämä virheet eivät tutkimuksessa aineistossa silti ole. Ylemmillä vuosiluokilla selkein yksittäinen konsonanttien poisjääntiä aiheuttava tekijä näyttääkin olevan konsonantin akustisen havaittavuuden heikkous.

Virheellisesti valittujen konsonanttien joukko on hyvin kirjava. Läpinäkyvimpiä kirjoittamisprosessin kannalta ovat vierasperäisten kirjainten merkintöjen pulmat (*asvalttia* 'asfalttia', *twstä* 'tv:stä', *Nikaracuassa* 'Nicaraguassa') sekä *n/m*-merkintöjen epätarkkuus (*esin* 'esim.', *kauenpana* 'kauempana'), joita esiintyy kaikilla luokka-asteilla. Lähes yhtä yleisiä ovat alakoululaisten teksteissä satunnaisen oloiset kirjainten sekoittumiset (*joulupuppi* 'joulupukki', *koka* 'joka', *kuninhas* 'kuningas', *savoim* 'sanoin').

Syyt vokaalien sekoittumiseen pirstaloituvat samaan tapaan kuin virheellisten konsonanttien tapauksessa. Kirjaimen muodon epätarkkuus selittää muutaman tapauksista (*lappuun* 'loppuun'), ja osa *S2*-kirjoittajien vääristä vokaaleista näyttäytyy tekstissä todennäköisesti kielitaustasta johtuviksi. Vokaalit sekoittuvat silti vain harvakseltaan molemmissa kieliryhmissä.

PIIRTOVIRHEET JA VOKAALISOINNUN HALLINTA

Molemmissa kieliryhmissä yli 80 prosenttia kaikista piirtovirheistä on puuttuvia *ä:n* pisteitä, muita esimerkiksi puuttuva *t:n* viiva. Alimmilla vuosiluokilla *ä:n* ja *ö:n* pisteiden merkitsemättä jättäminen lienee melko yleisesti oikeinkirjoituksen lipsahdus, vaikka puutteellinen suomen kielen vokaaliharmonian hallinta saattaa ilmentyä tekstissä yhtenevällä tavalla, eikä vokaaliharmonian hallinnassa ilmene edistyneillä *S2*-oppijoilla tai äidinkielisillä kirjoittajilla juuri edes satunnaisia puutteita.

Molemmissa aineistoissa sananalkuiset ja sanansisäiset *ä:t* ovat redusoituneet *a:ksi* loppuasemaista *ä:tä* useammin, mikä on loogista

kirjoitussuuntaa ja kirjoitusjärjestystä ajatellen: monille kirjoittajille on tyypillistä palata merkitemään pisteet sanahahmon valmistuttua. Molemmissa kieliryhmissä *a/ä*-vaihtelua esiintyy sekä sanavartaloissa että sijapäätteissä ja enimmäkseen aivan tutuissa mutta myös harvinaisemmissa sanoissa. Molempien kieliryhmien teksteissä esiintyy tapauksia, joissa virheeseen yhdistyy muita äännetason virheitä. S2-ryhmässä useamman erityyppisen virheen kasautuminen samaan tekstisanaan on tosin kaiken kaikkiaan yleisempää. Ylemmillä vuosiluokilla piirtovirheet ovat samantyyppisiä kuin alaluokilla, mutta niiden tekstitiheys on molemmissa kieliryhmissä selvästi vähentynyt. S2-teksteissä niiden määrä on yhä tekstimassaan suhteutettuna noin kaksinkertainen verrokkiaineistoon verrattuna.

TAIVUTTAMISEN ONGELMAT

Kontekstissaan virheellisiä taivutusmuotoja esiintyy kaikkien kirjoittajaryhmien teksteissä. Taivutuksen puuttuminen (nomini perusmuodossa tai verbi 1. infinitiivissä) on jo alaluokilla kuitenkin harvinaisempi virhe kuin muuhun virheelliseen taivutusmuotoon päätyminen. Yleisimpiä taivutusvirheet ovat S2-ryhmässä, jossa siinäkin taivutuksen virhetiheys jää noin yhteen sanaan sataa tekstisanaa kohden. S1-kirjoitelmissa taivutusvirheiden esiintymistiheys laskee alakoulun aikana edelleen jo lähtökohtaisesti hyvin matalasta tasosta (tekstitiheys 0,2 → 0,1). S2-oppijoilla taivutusvirheiden tekstitiheys hiukan kasvaa (0,9 → 1,0), mutta kun tarvittavan taivutusaineksen puuttuminen kokonaan on aiempaa harvinaisempaa, kokonaisuudessaan taivuttaminen on tarkentunut (1,4 → 1,3). Näin on siitä huolimatta, että teksteissä käytetty sanasto on monipuolistunut, lekseemi-, sane- ja tekstipituudet kasvaneet ja kompleksisten rakenteiden käyttö yleistynyt.

Nominien taivutusvirheistä eniten on objektivirheitä, joita esiintyy S2-kirjoittajilla paikallissijan valintavirheitä enemmän mutta S1-kirjoittajilla vain muutamia. Objektin sijavalinta ja muodostus tiedetäänkin yhdeksi suomen kielen oppijoita eniten haastavaksi rakennepiirteeksi (ks. esimerkiksi Martin 1995; Spoelman 2013). Nominivirheistä subjektin vääriä sijavalintoja sekä vääriä paikallissijamuotoja esiintyy niin ikään harvakseltaan molempien kirjoittajaryhmien teksteissä. Sekä nominien

että verbien taivutuksessa esiintyy myös jonkin verran yksikön ja monikon valintaan liittyviä luvun virheitä, vaikka niitäkin korpuksen kokoon nähden vähän. Lukuun ja subjektin sijavalintaan liittyviä virheitä S2-oppijoiden teksteissä ei esiinny juuri verrokkiryhmää useammin, paikallissijan valinnassa sen sijaan on vaikeuksia hiukan enemmän. Kongruenssivirheitä (esimerkit 11–12) esiintyy kaikkien kirjoittajaryhmien teksteissä niin alimmilla luokilla kuin alakoulun loppuvaiheessaakin. Kongruenssivirheet eivät kuitenkaan ole systemaattisia yhdelläkään kirjoittajalla, ja niiden kokonaismäärä on vähäinen.

11) *onhan meilläkin tänä viikollakin kivaa* (2. lk. PS1)

12) *Ei olisi koulua puoli vuoteen* (5. lk. TS2, vietnam, SS)

Verbintaivutuksessa pulmat kietoutuvat lähes aina kompleksisten rakenteiden, kuten kieltomuotojen ja verbiketjujen, yhteyteen. Satunnaisia puutteita on silti molemmissa kieliryhmissä myös predikaattiverbin persoonamuotojen valinnassa ja satunnaisesti moduksen ja tempuksen valinnassa.

Sanojen taivuttaminen ei ole kertaluonteisesti opittava ilmiö. Yksi ja sama kielenoppija saattaa ymmärtää jonkin taivutuspuutteen funktion, vaikka ei itse vielä käytä päätettä lainkaan, tai hän saattaa käyttää sitä puutteellisessa muodossa tai vain tutun konstruktion osana. Senkään jälkeen, kun tietty taivutusmuoto on ilmaantunut oppijan kirjoitukseen, käyttö ei siis välttämättä ole kattavaa. Äidinkielenään arabiaa puhuvan mutta Suomessa syntyneen toisluokkalaisen tytön kertomustekstin esimerkit 13–15 havainnollistavat ilmiötä: monisanaisissa nominilausekkeissa edussana on toistuvasti päätteetön. Toisaalta kongruoiva määrite taipuu näissäkin tapauksissa usein oikein ja myös edussanasta esiintyy virheettömästi taivutettuja objektimuotoja.

13) *Eka on se että haluaisi oman kannettava tietokone.* – – *toivoisi vielä oman leikki huone.*

14) *Menin kauppaan ja ostin vaatteita ja ostin kenkä ja sain jäätelöä.*

15) *Neljäs on se että saisi ison trampoliinin. Ja kultainen talo.*

Rektiovirheitä esiintyy niin ala- kuin yläluokilla ja molemmissa kieliryhmissä tytöillä ja pojilla, vaikkakin jälleen tiheämmin S2-teksteissä. Lähes kaikki rektiovirheet liittyvät verbitäydennyksien taivutusmuodon valintaan, vaikka muitakin esimerkkejä löytyy (*Talo oli tehty kultalla, se unelmoi kissaa; kolmekymmentä euron maksava, ennen lähdömmä, kaikki ikäiset* 'kaiken ikäiset').

Kaikki S2-oppijoiden teksteissä useita kertoja esiintyvät taivutusvirhetypit ovat sellaisia, joiden oppiminen aiemman tutkimuksen perusteella tiedetään haasteelliseksi (esimerkiksi Martin 1995; Ivaska 2009; Niiranen 2010; Reiman 2011; Spoelman 2013; Mustonen 2015). Subjektin sijavalintavirheissä huomiota kiinnittää partitiivisubjekti. Kieliryhmille yhteinen piirre on, että partitiivista esiintyy muutaman kerran sekä yliettä alikäyttöä. Mitenkään yleisiä subjektin sijavalintaan liittyvät pulmat eivät tässä edistyneiden kielinoppijoiden ryhmässä kuitenkaan ole. Paikallissijoista ovat tuottaneet vaikeuksia lähinnä olo- ja tulosijat (inessiivi, adessiivi, illatiivi, allatiivi), mutta kaiken kaikkiaan tämäntyyppiset valintavirheet ovat jo alaluokilla harvinaisia molemmissa kieliryhmissä, ja hyvin harvoin paikallissija on kokonaan merkitsemättä (ks. esimerkit 16–18). Lause- ja virkerakenteiden kompleksistuessa paikallissijojen valintaan liittyvien virheiden määrä on säilynyt ylemmillä vuosiluokilla ennallaan. Taivutusvirheet eivät toisin sanoen katoa ylemmillä luokilla, mutta yhä useammin ne ilmenevät aiempaa kompleksisempien lauserakenteiden yhteydessä, mikä vastaa myös Lesosen (2020) tuloksia.

16) *Oli kaunis ilma harjoitella golfia yksin kauniissa kentässä.*

(5. lk. TS1)

17) *Ja niinpä IsoBritannia siihen aion lähteä* (6. lk. PS2, vietnam, SS)

18) *kun Emilianäi jäätelöäsil tuli pääsä unelmin.* (2. lk. TS2, arabia, 5 vuotta Suomessa)

SANAJÄRJESTYS JA SANASTON TARKKUUS

Sanajärjestykseen liittyvät epätarkkuudet lisääntyvät ylemmillä luokilla. Se on hyvä esimerkki uuden asian haltuun ottamisen vaiheesta, jossa tarkkuuden kehittyminen seuraa kompleksistumiskehitystä. Sanajärjestyksen epätarkkuudet yhdistyvät lähes poikkeuksetta kasvavaan virke-

pituuteen useampilauseisissa virkkeissä, joissa lisäksi on saatettu käyttää kompleksisia rakenteita kuten lauseenvastikkeita ja verbiketjuja (ks. esimerkit 19 ja 20).

19) *Noustuani päätän mennä rinneravintolaan syömään, joka on Ursuksen yläpuolella.* (6.lk. S1P)

20) *Päätimme vielä yrittää liikunnanopettajalle tehdä metkun ja se onnistui.* (6.lk. S1T [A])

Ongelmavirkeissä on usein käytetty haastavaa kieltomuotoa ja sen yhteydessä indefiniittipronominia sekä erityistapauksissa vielä objektia. S2-kirjoittajille pulmia ovat verrokkiryhmää enemmän tuottaneet myös lausetta täydentävien ajanilmausten sijainti ja sisäinen rakenne. Koko aineiston tasolla sanajärjestys hallitaan kuitenkin molemmissa kieliryhmissä erittäin hyvin (vrt. muun muassa Elo 1998, 36).

Leksikaalisiin valintoihin liittyviä virheitä on totuttu pitämään oppijankielessä yleisinä ja ongelmallisina virheinä, koska niiden on ajateltu johtavan helposti väärin tulkintoihin (Read 2000, 200–201). Etenkin leksikaaliset virheet yhdessä rajoittuneen kokonaissanaston kanssa voivat myös ohjata lukijan tulkintaa voimakkaan negatiiviseksi tekstin mahdollisista muista ansioista huolimatta (Meara, Rodgers & Jakobs 2000, 346). Agustín Llachin (2011, 165, 172) tutkimuksessa kouluikäisten englanninoppijoiden kirjoitelmien laatu (holistinen arvio) korreloi voimakkaasti juuri leksikaalisten virheiden esiintymisen kanssa.¹⁰⁰ Alakoululaisten kirjoitelmissa leksikaaliset virheet eivät lähellekään aina silti aiheuta väärinymmärtämisen vaaraa, ja vain harvoin valittu lekseemi on tarkastelluissa kertomuksissa täysin sopimaton. Selkeästi virheellisiä sanavalintoja kertomusteksteissä esiintyy hyvin vähän, sekä alemmilla että ylemmillä vuosiluokilla alle sata tapausta. Syntyperäisillä kirjoittajilla virheellisiksi merkittyjä sanavalintoja on molemmissa keruuvaiheissa saman verran kuin S2-kirjoittajilla. Tämä havainto poikkeaa osin toisen kielen oppimisen aiemmista tutkimuksista, joissa leksikaalisten vir-

¹⁰⁰ Yhteys ei välttämättä säily samankaltaisena oppimisen eri vaiheissa. Agustín Llachin (2011, 178) tutkimuksessa korrelaatio kääntyy osittain toisensuuntaiseksi ylemmillä vuosiluokilla, mitä voi selittää leksikaalisen riskinoton positiivinen vaikutus tekstin sisällölliseen antiin.

heiden on havaittu olevan virheistä tyypillisimpiä (ks. Gass & Selinker 1994, 270–272, tuoreemmin esimerkiksi Pietilä 2012). Eroon vaikuttavat suomen kielen ominaispiirteet (taivutusjärjestelmän vaativuus), mutta myös poikkeavuudet tutkimustehtävissä ja tutkittujen oppijoiden iässä ja kielitaitotasossa sekä mahdollisesti itse virheiden luokittelutapa (Neff ja muut 2007, 208).

Kertomustekstit osoittavat, että vaikka oppijankieli ei leksikaalisten merkitysten suhteen aina ole kovin täsmällistä, sanavalintojen tarkkuus on pitkään Suomessa asuneiden S2-kirjoittajien teksteissä kuitenkin kehittyneempi kuin oikeinkirjoituksen tarkkuus. Tekstiyhteydessään suoranaisesti epäsopivia sanoja käytetään harvoin.

Rauniosanalla tarkoitetaan sellaista sananmuotoa, jonka oletettu tavoitemuoto on kaukana toteutuneesta sananmuodosta eli jonka kirjoitusasu poikkeaa tavoitemuodosta esimerkiksi tietyn grafeemimäärän verran (Matilainen 1985) tai jonka virhetyyppiä on virheiden lukuisuuden tai limitteisyyden vuoksi mahdotonta määrittellä tarkasti (Kulju ja Mäkinen 2009, 315). Omassa tutkimuksessani olen käyttänyt rauniosanan määrittelyyn kolmen grafeemin sääntöä: mikäli sanavartalossa on vähintään kolmen grafeemin poikkeama tavoitemuotoon tai virheet sanamuodon kirjoitusasussa ovat sellaisia, että lekseemi muutoin vaatii ympäröivän tekstin avulla päättelyä, se on määritetty rauniosanaksi. Vaikka käytetty rauniosanamääritelmä on tiukka, rauniosanat osoittautuvat kaikkien kirjoittajaryhmien teksteissä hyvin harvinaisiksi. Alemmilla vuosiluokilla rauniosanojen esiintyvyys on tekstimäärään suhteutettuna noin kolme kertaa suurempi kuin ylemmillä. Alimmilla luokilla rauniosanoja on enemmän poikien kuin tyttöjen teksteissä, mutta sukupuolen välinen ero tasoittuu alakoulun aikana: S1-kirjoittajilla rauniosanoja ei enää viidennellä ja kuudennella luokalla esiinny ja S2-ryhmässäkin pojilla huomattavasti aiempaa vähemmän.

Puuttuviksi sanoiksi on katsottu tapaukset, joissa lauserakenteen ongelmat voitaisiin korjata lisäämällä tekstiin puuttuva sana, yleensä subjekti, predikaatti tai objekti. S1-kirjoittajilla tällaiset tapaukset ovat alimmilla vuosiluokilla huomattavasti harvinaisempia kuin S2-kirjoittajilla, ja lauseesta puuttuu pääjäsenäni harvoin. Virkkeiden pidentyessä ja virkerakenteiden kompleksistuessa erityisesti verbiketjujen virheet

lisääntyvät (ks. esimerkit 21–23). S₂-pojilla puuttuvien sanojen teksti-tiheys on alakoulun päättövaiheessa muita suurempi, mihin vaikuttaa sekä virheiden esiintyminen useammalla oppijalla että puuttuvan sanan tyyppivirhe muutamalla kirjoittajalla.

21) *Se on minun lempi pelaaja ja hän Ø minun mielestäni maailman paras.* (5. lk. PS₂, albania, SS)

22) *Olen unelmoinut Ø delfiinin kouluttajaksi.* (5. lk. TS₁)

23) *Minun unelmien päivä on, että juuri perjantaina meillä oppilailla on vapaapäivä ja Ø nukkua paljon* (6. lk. PS₂, albania, SS)

Tarkkuuden yhteys kirjoittamisen muihin osataitoihin ja kirjoittajan kokonaiskielitaitoon

Tarkastelen lopuksi kirjoittamisen tarkkuuden yhteyttä sanastonhallintaan, suomen kielen taitoon sekä oman äidinkielen taitoon. S₁-oppijat ovat mukana vertailussa soveltuvin osin.

YHTEYS KIRJOITTAMISEN OSATAITOIHIN

Tarkkuuden vertailumuuttujina toimii kolme kertomustekstien kielellistä kehittyneisyyttä kuvaavaa muuttujaa: a) sanaston monimuotoisuus¹⁰¹, b) harvinaisten sanojen osuus¹⁰² sekä c) tavoitemuodon mukaan korjattu sanepituus, jotka kaikki tiedetään kielellisen kehityksen mukaan karttuviksi ominaisuuksiksi (ks. tarkemmin esimerkiksi Malvern & Richards 2004; Crossley ja muut 2010). Tiiveimmin näistä tekstin ominaisuuksista kirjoittamisen tarkkuus on tutkitussa koululaisaineistossa yhteydessä sanaston monimuotoisuuteen: Mitä monipuolisempaa sanastoa kirjoitelmassa on käytetty, sitä pienempi virhetiheys siinä on todennäköisesti ollut. Toisen vuosiluokan jälkeen yhteys on kaikilla vuosiluokilla tilastollisesti merkitsevä tai erittäin merkitsevä ($p < 0,01$ tai $p < 0,001$). Yhteys selittyy etenkin S₁- ja S₂-kirjoitelmien voimakkailla keskimää-

101 D-indeksiin perustuva *sums of probabilities* -diversiteettiarvo.

102 Yleisyystasoa 1000 harvinaisemmat sanat, vertailukorpuksena CSC:n lekseemilista.

räisillä eroilla molempien piirteiden suhteen: sanaston monimuotoisuus on S1-kirjoitelmissa selvästi korkeampi ja virhetiheys pienempi (mediaanien erot erittäin merkitsevät). Alaryhmätasolla (PS1, TS1, PS2, TS2) yhteys on kuitenkin kohtalaisen voimakas ainoastaan viidennen vuosiluokan S1-tyttöjen ja kuudennen vuosiluokan S2-tyttöjen ryhmissä ($r_s \leq -0,350$, $p < 0,05$, $n_{TS1} = 46$, $n_{TS2} = 27$). Tulosta voi tulkita niin, että sekä kirjoittamisen tarkentuminen että sanaston käytön monimuotoistuminen ovat kehityksellisiä piirteitä, jotka tuovat osaamistason erot esille sitä selvemmin, mitä suuremmat ryhmän sisäiset taitoerot ovat ja mitä paremmin kirjoittajat ovat kirjoittaneet omalla taitotasollaan. Esimerkiksi heikko kirjoitusmotivaatio tai tarkkaavuus tilanteessa kuitenkin estää kirjoittajaa osoittamasta parhaita taitojaan; tosin motivaatiotekijät näyttävät vaikuttavan voimakkaammin tekstin tarkkuusominaisuuksiin kuin sanaston monimuotoisuuteen.

Saneen tavoitepituuden ja virhetiheyden välillä ei ole alimmilla vuosiluokilla lineaarista yhteyttä missään yksittäisessä kirjoittajaryhmässä, mutta koko kertomuskorpuksessa tarkkuus korreloi löyhästi sekä sanapituuden että runsaampien taivutuselementtien käytön kanssa toisella ja viidennellä vuosiluokalla ($r_s \approx 0,4$, $p < 0,05$). Kompleksisuus ja tarkkuus näyttävät silti kehittyvän eri tahdissa: toisella vuosiluokalla keskimääräisen sanepituuden ja virhetiheyden välinen yhteys on positiivinen, kolmannella vuosiluokalla enää hyvin heikosti positiivinen ja kääntyy viidennellä negatiiviseksi. Havaintoa on tulkittava siten, että koulunkäynnin alkuvaiheessa pisimpiä sananmuotoja tavoittelevissa kirjoitelmissa esiintyy keskimäärin myös eniten virheitä, ja vasta ylemmillä vuosiluokilla mekaanisten kirjoitustaitojen alkaessa vakiintua sanepituus lakkaa laajamittaisesti kasvattamasta virheiden todennäköisyyttä. Erityisesti koulun aloitusvaiheessa suomen kielen pitkät sanat kuormittavat sana- ja tavurakenteiden prosessointia, ja samoilla lapsilla on usein vaikeuksia sekä näiden tarkassa kirjoittamisessa että sujuvassa lukemisessa (Mäkinen 2007). Tulos on linjassa myös Malvernin kumppaneineen esittämän havainnon kanssa: oppimisen alkuvaiheessa oikeakielisyysvirheiden määrä on pienin teksteissä, joissa keskimääräinen sanapituus on pienin (Malvern ja muut 2004). Tämän tutkimuksen aineistossa korkea sanepituus alkaa (vasta) viidennellä vuosiluokalla yhdistyä alhaiseen

virheteriheyteen, ja silloinkin alaryhmistä ainoastaan S1-oppijoiden ryhmässä ($r_s = -0,322$, $p < 0,01$, $n = 77$).

Kiintoisaksi osoittautuu myös harvinaisten sanojen tekstitiheyden ja kirjoittamisen tarkkuuden välinen yhteys. Koko aineiston tasolla mittarit asettavat tekstit toisella vuosiluokalla samansuuntaiseen paremmuusjärjestykseen: mitä enemmän harvinaisia sanoja tekstissä on käytetty, sitä suurempi tekstin virheteriheys todennäköisesti on ollut. Aikaisemmin samansuuntaisia havaintoja suomen sanojen yleisyyden ja niiden kirjoittamisen tarkkuuden suhteesta on tehnyt Pekonen (2007). Korrelaatio on kuitenkin hyvin matala eli harvinaisten sanojen käytön selitysvoima virheiden määrään on alhainen, vaikka suurehkoissa ryhmässä tulos on tilastollisesti merkitsevä ($r_s = 0,212$, $n = 198$). Yhteys muuttujien väliltä katoaa jo kolmannella vuosiluokalla ja korrelaatio pysyy lähellä nollaa alakoulun loppuun saakka. Ainoastaan S1-poikien ryhmässä harvinaisten sanojen runsas käyttö lisää virheiden todennäköisyyttä vielä ylemmillä vuosiluokilla.¹⁰³

YHTEYS KIRJOITTAJAN SUOMEN KIELEN TAITOON

On mahdollista, että virheettömästi kirjoittavat koululaiset ovat kirjoittajina pedantteja mutta eivät erityisen taitavia kokonaiskielitaidoltaan. Aiempien tutkimusten perusteella kielitaidon tekijät kuitenkin kumuloi-tuvat niin, että kielitaidoltaan vahvat oppijat ovat tyypillisesti vahvoja hyvin monilla eri osa-alueilla (esimerkiksi Alderson 2005). Kielellisesti lahjakkaat eivät tosin aina motivoidu irrallisista kirjoittamistehtävistä, mikä aiemman tutkimuksen valossa pätee erityisesti poikiin. Tutkin omassa aineistossani alakoululaisten kielitaitoprofilin tasaisuutta asian valaisemiseksi kahden erityyppisen vertailumuuttujan avulla. Vertasin kirjoittamisen tarkkuutta ensinnäkin erillisellä testillä mitattuun suomen kielen sanaston hallintaan sekä toisekseen kyselytiedon¹⁰⁴ avulla arvioituun suomen kielen yleistason: lukemis-, puhumis- ja kirjoitus-taitoon, S2-arvosanaan sekä erityisvaikeuksiin.

103 Epälineaarinen yhteys muissa, suuremmissa kirjoittajaryhmissä on sinänsä looginen, sillä harvinaisten sanojen käyttöön ja niiden käytön tarkkuuteen vaikuttaa lukuisia ristikkäisiä tekijöitä.

104 S2-opettajan/luokanopettajan arvio S2-oppilaiden kielitaidosta.

Kirjoittajakohtainen kertomustekstin virhetiheys korreloi jokaisella vuosiluokalla kohtalaisen voimakkaasti kirjoittajan leksikaalisten taitojen kanssa, mikäli tarkastellaan koko kirjoittajaryhmää ($r_s \approx -0,5$, $p < 0,001$; ks. taulukko 48). Yhteys on kaikilla vuosiluokilla negatiivinen eli tarkimmin kirjoittavat sanastollisilta taidoiltaan vahvimmat lapset.¹⁰⁵ Tulos on linjassa toisen kielen oppimisesta aiemmin havaitun kanssa (ks. esimerkiksi Llach 2011, 190, 192). Korrelaatio on kuitenkin kaukana maksimiarvostaan, mihin vaikuttaa riippuvuuden ehdollisuus. Vaikka vahva leksikon hallinta – sekä mahdollisesti tarkkaavuus ja tehtäviin suuntautuminen testitilanteessa – yhdistyy virheettömään tai melko virheettömään kirjoittamiseen, kaikki tarkat kirjoittajat eivät ole sanastollisilta taidoiltaan vahvoja.

Pienemmissä tyttöjen ja poikien alaryhmissä yhteys on aina samansuuntainen. Sanastotehtävistä parhaiten suoriutuneet lapset ovat todennäköisimmin myös tarkkoja kirjoittajia. Koska sanastotehtävät ovat edellyttäneet hyvin niukasti kirjoittamista, kirjoitustaito itsessään ei riitä yhteyden selittäjäksi, vaan syytä on haettava kielenhallinnan eroista. Erot tehtävöorientoituneisuudessa saattavat tosin osittain selittää suoriutumista molemmissa tehtävissä ja samalla myös alaryhmien välisiä korrelaatioeroja: etenkin sanastollisesti taitavat S1-pojat eivät näytä motivoituneen ylemmillä vuosiluokilla kirjoittamaan parhaan osaamisensa mukaisesti, ja juuri tässä ryhmässä taitojen välinen yhteyskin on heikoin.

Taulukko 48. Virhetiheyden ja sanastotehtävuoriutumisen välinen yhteys (r_s = Spearmanin korrelaatio).

r_s	Toinen	Kolmas	Viides	Kuudes
Kaikki	-0,450***	-0,450***	-0,568**	-0,565***
S1	-0,400***	-0,197*	-0,286*	-0,462**
S2	-0,486***	-0,417**	-0,430***	-0,582***
Pojat (S1)	-0,261*	-0,144	-0,155	-0,165
Tytöt (S1)	-0,454***	-0,193	-0,418**	-0,686***
Pojat (S2)	-0,579**	-0,548	-0,269	-0,719***
Tytöt (S2)	-0,280	-0,181**	-0,471**	-0,297

105 On syytä huomauttaa, että sanastollisten taitojen arvioinnissa fonologisella tai morfologisella tarkkuudella ei ole ollut merkitystä.

Sekä S₁- että S₂-kirjoittajien opettajille lähetetyssä taustatietokyselyssä on kartoitettu oppijoiden kielellisiä erityisvaikeuksia. Näistä kyselymuuttujista muodostetun summamuuttujan oppijakohtaisen arvon vertaaminen saman oppijan tekstin virhetiheyyteen antaa lisätietoa eri oppijaryhmien kielellisten erityisvaikeuksien ja kirjoittamisen tarkkuuden välisestä suhteesta. Niissä kirjoittajaryhmissä, joissa sanastollisten taitojen ja kielen virheettömyyden yhteys osoittautui selvimmäksi, kielellisten erityisvaikeuksien ja tarkkuuden välillä ei ilmene yhteyttä. Tulosta selittää etenkin arvioitujen erityisvaikeuksien vähäinen määrä aineistossa.

S₂-oppijoiden ryhmässä on ollut mahdollista tutkia myös kirjoittamisen tarkkuuden ja suomen kielen edellisen todistusarvosanan välistä yhteyttä. Niin S₂-tytöillä kuin pojillakin edellinen suomen kielen todistusarvosana on yhteydessä kirjoittamisen tarkkuuteen (S₂: $r_s = -0,37$, $p < 0,001$, $n = 91$), mutta arvosanaan luonnollisesti vaikuttaa moni muukin asia, ja yhteys on heikompi kuin tarkkuuden tilastollisesti erittäin merkitsevä yhteys sanastollisiin taitoihin. Erityisen mielenkiintoinen on opettajan arvioiman kirjoitustaidon yhteys kirjoittamisen tarkkuuteen: tytöillä lineaarinen yhteys on kohtalaisen voimakas ja tilastollisesti merkitsevä ($r_s = -0,438$, $p < 0,001$, $n = 63$) mutta poikien aineistossa sisäinen hajonta on suurempi eikä riippuvuutta ilmene ($r_s = -0,136$ $p > 0,05$, $n = 42$). Tulos tukenee havaintoa siitä, että pojat eivät ole osanneet tai halunneet osoittaa parasta osaamistaan unelmakirjoitelmaa kirjoittaessaan tai ainakaan tehneet sitä yhtä systemaattisesti kuin tytöt. Kirjoittamistaidon kartoituksissa olisikin jatkossa tärkeää kiinnittää erityistä huomiota kirjoittamistehtävään ja sen merkityksellisyyteen erilaisten kirjoittajien näkökulmasta. Toisaalta voi olla, että koska tyttöjen kirjoittaminen on keskimäärin poikien kirjoittamista kehittyneempää, juuri tyttöjen kirjoitustaidon arvioinnissa opettajat kiinnittävät enemmän huomiota myös tarkkuuspiirteisiin.

Opettajien arvioimista osataidoista suomen kielen puhetaito tai eurooppalaista viitekehystä mukaileva yleisempi taitotasoarvio eivät tarkastellussa aineistossa korreloi kirjoittamisen tarkkuuden kanssa. Pääsyynä ensimmäiseen lienee vertailuaineiston luonne: monet kirjoittamisen virheistä liittyvät nimenomaan kirjoitetun kielen normien hallintaan.

Lähes kaikki pikkulapsesta asti Suomessa asuneet koululaiset ovat lisäksi jo hyvin sujuvia puhujia, minkä vuoksi puhetaidon karkea arvio ei myöskään erottele oppilaiden kielenhallintaa S2-ryhmän sisällä (ks. mittarista tarkemmin Honko 2013). Eurooppalaista viitekehystä mukaillevan taitotasoarvion ja tarkkuuden heikkoon yhteyteen voidaan nähdä kaksi hyvin erityyppistä syytä, joista ensimmäinen liittyy toisen kielen oppijan kielenoppimisprosessin vaiheisiin: on mahdollista, että virheet lisääntyvät aaltomaisesti eri kielitaitotasolla, jolloin lineaarista tai edes monotonista yhteyttä ei korrelaatioanalyysillä voida havaita (ks. myös esimerkiksi Brunni ja muut 2019). Toinen mahdollinen selitys liittyy itse mittarin validiteettiin: alakouluikäisen suomenoppijan kielitaidon arvioinnissa viitekehysten mukaiset taitotasot eivät ole olleet kovin helppoja soveltaa.

Kielitaidoltaan vahva ei välttämättä hallitse kirjoitetun kielen normeja tai muusta syystä kirjoita tarkasti. Esimerkiksi kuudesluokkalaisten Suomessa syntyneen ja ensikieleltään vietnaminkielisen pojan kertomuksessa virheettömät fraasit ja idiomaattiset rakenteet (*pienestä asti, jonakin päivänä, mitä vielä, ne vasta ovatkin herkkuja*) sekä sofistikoituneet sanat (erityisesti harvinaisten sanojen käyttö) ja rakenteet (kuten *elämä ilman autoja*), konditionaalin ja perfektin käyttö, erilaiset sivulausetyypit ja lukuisat onnistuneet objektitavutukset vartalomuunnoksineen osoittavat oppijan paitsi käyttävän suomea monipuolisesti myös taitavan kompleksisia rakenteita. Toisaalta kirjoittamisessa on vielä puutteita muun muassa välimerkityksessä, alkukirjaimissa sekä yhteen ja erilleen kirjoittamisessa. Vaikka kielen rekisteritkin välillä sekoittuvat (*naarmuks*), teksti osoittaa, että kirjoitetun kielen konventioiden oppiminen on käynnissä. Tämä näkyy esimerkiksi verbien infiniittimuotojen standardinmukaisena viimeistelynä (*haaveillut, lähteä* vrt. *haaveillu, lähtee*) sekä joidenkin lähinnä kirjoitetussa kielessä esiintyvien konstruktioiden käyttönä. Juuri näiden konstruktioiden hiomattomuus (*eihän voi olla paha mennä, mitä mahtavempia*) toisaalta kuvastaa oppimisprosessin dynaamisuutta. Kun oppija käyttää ilmaisuväriä monipuolisesti, rohkeasti ja luovasti ja samalla vasta harjoittelee kirjoitetun kielen normeja, on luonnollista, että ilmaisuun sisältyy myös epäidiomaattisuutta ja epätarkkuutta.

YHTEYS KIRJOITTAJAN OMAN ÄIDINKIELEN TAITOON

Koulunkäynnin alussa lasten luku- ja kirjoitustaidossa on suuria eroja, mutta alkuopetus tasaa eroja nopeasti, kuten tämänkin tutkimuksen aineisto osoittaa. Vaikka suomenkielisessä ympäristössä vietetty aika tukee kielitaidon kehittymistä erityisesti kouluikässä, myös koko peruskoulunsa Suomessa opiskelleilla S2-oppilailla voi olla puutteita suomen kielen taidoissa (myös Kuukka & Metsämuuronen 2016, 85). Kirjoitustaidon osalta yksilölliset erot ovat kuitenkin suuremmat oman äidinkielen taidossa. Koulun alkuvaiheessa vain harvat toisen kielen oppijat ovat ehtineet oppia kirjoittamaan ensikielellään. Noin 17 prosenttia monikielisistä oppilaista arvioi oman äidinkielen kirjoitustaitonsa vielä peruskoulun päättövaiheessa joko huonoksi tai olemattomaksi (Kuukka & Metsämuuronen 2016, 73). On huomattava, että myös yksilöllisissä mahdollisuuksissa toiminnallisen kielitaidon kehittämiseen on suuria eroja.

Tähän tutkimukseen osallistuneista vain noin joka kymmenes koululainen piti ensikielen kirjoitustaitoaan toisen tai kolmannen vuosiluokan kevätlukukaudella suomen kielen kirjoitustaitoaan parempana, ja kolme vuotta myöhemmin näin ajattelevien osuus oli edelleen hieman pienentynyt (Honko 2013). Siksi on ollut erityisen kiinnostavaa tarkastella sitä, miten varhaislapsuudesta asti Suomessa asuneiden S2-oppiloiden suomen kielen kirjoitustaito kehittyi suhteessa suomea ainoana kotikielenä puhuneisiin ikätovereihin ja toisaalta suhteessa S2-oppiloiden oman äidinkielen taitoon.

Noin puolet tutkimuksessa mukana olleista S2-oppilaista osallistui aineiston keruuhetkellä oman äidinkielen opetukseen. Näiden oppilaiden oman äidinkielen opettajista 21 vastasi kyselyyn, joka oli samansisältöinen kuin suomen kielen opettajille lähetetty kysely. Vähäiseen vastausmäärään vaikuttaa se, että kaikilla oppilailla ei ole mahdollisuutta tai halua osallistua järjestettyyn oman äidinkielen opetukseen sekä se, ettei opetusta etenkin vähän puhutuissa kielissä ja pienemmällä paikkakunnalla ollut tarjolla kaikille. Kysely ei myöskään tavoittanut kaikkia opettajia, eivätkä kaikki tavoitetut opettajat vastanneet kyselyyn. Vastausmäärä on tilastollisten analyysien kannalta hyvin pieni. Nostan tässä yhteydessä kuitenkin esille joitakin kiinnostavia havaintoja, jotka toimivat avauksi-

na keskusteluun ja mahdollisesti tukena laajempiin pitkittäistutkimuksiin tulevaisuudessa.

Opettajien arvioima oman äidinkielen taito viidennen ja kuudennen luokan keväällä ei ainakaan tässä pienessä tutkimusryhmässä vaikuttanut millään käytetyistä mittareista (tarkemmin Honko 2013) olevan yhteydessä saman ikävaiheen suomen kielen kirjoittamisen tarkkuuteen. Hieman yllättäen oman äidinkielen taidon ja aikaisemman eli toisen ja kolmannen luokan kirjoittamisen tarkkuuden väliltä paljastui kuitenkin melko voimakas ja tilastollisesti merkitsevä korrelatiivinen yhteys. Tämä yhteys ilmeni erikseen sekä opettajan arvioiman oman äidinkielen puhetaidon ($r_s = -0,51$, $p < 0,05$, $n = 21$), kirjoitustaidon ($r_s = -0,63$, $p < 0,01$, $n = 21$) että sanastollisten taitojen ($r_s = -0,62$, $p < 0,01$, $n = 21$), edellisen oman äidinkielen todistusarvosanan ($r_s = -0,56$, $p < 0,01$, $n = 20$) ja suomen kielen kirjoittamisen tarkkuuden välillä. Yhteys oli aina samansuuntainen: varhaisten kouluvuosien suomen kielen kirjoittamisen tarkkuus yhdistyi myöhemmin vahvaksi arvioituun oman äidinkielen taitoon.

Positiivinen yhteys saattaa tukea ajatusta toiminnallisen monikielisyiden eduista eli siitä, että yksilön vahva kielitaito yhdessä kielessä voi tukea muidenkin kielten taitojen kehittymistä. Kyse voi olla myös yleisemmin kielellisestä älykkyydestä tai taipumuksista, jotka syntyvät paljolti ympäristötekijöiden vaikutuksesta mutta osin myös geneettisesti (ks. esimerkiksi Kaprio 2006). Yhteyden katoaminen myöhemmässä vaiheessa voi merkitä ainakin kahta asiaa. Ensinnäkin kirjoittamisen tarkentuessa yläluokilla tarkkuusmittarin erottelevuus luonnostaan vähenee. Toiseksi varhaisen vaiheen vahva suomen kielen taito sekä ehkä myös käytettyjä arviointitapoja suosivat yksilölliset ominaisuudet, kuten tarkkaavuus, ovat voineet olla eduksi oman äidinkielen taidon kehittymiselle. Lisäksi on mahdollista, että opettajan kyselyyn vastaamisen yhteydessä soveltama arviointitapa on suosinut tarkkuutta tai sellaista tehtäväorientoituneisuutta, jota tämän tutkimuksen kieliaineiston keruussakin on vaadittu ja joka on erotellut oppilaita erityisesti alakoulun alkuvuosina.

Lopuksi

Kirjoittaminen tarkentuu alakoulun aikana selvästi kaikissa kirjoittajaryhmissä. Voimakkainta kehitys on S2-pojilla, joilla kirjoittamisen epätarkkuus alakoulun alkuvaiheessa oli suurinta. Yleisimpiä ovat sekä alakoulun alku- että loppuvaiheessa sanojen yhteen ja erilleen kirjoittamisen haasteet. Etenkin alimmilla vuosiluokilla yleisiä ovat myös puuttuvat geminaatat ja pitkät vokaalit sekä kirjainten piirtämiseen liittyvät epätarkkuudet. Alakoulun loppuvaiheessa sanojen taivuttamiseen liittyvät virheet ovat jo hieman näitä yleisempiä. Puhekieli lomittuu kirjoitettuun yleiskieleen kaikissa kirjoittajaryhmissä – mutta osin hyvinkin hallitusti ja kerrontaa tukien.

Koko kertomusaineistossa virhetiheydellä ilmenee toisen vuosiluokan jälkeen negatiivinen yhteys kielenhallintaan, jonka mittarina tässä tutkimuksessa toimi sanaston käytön monimuotoisuus: virhetiheys on suhteessa suurin niillä kirjoittajilla, jotka käyttävät yksinkertaisinta sanastoa. Tätä selittää ennen kaikkea kieliryhmien välinen kielitaitoero: sanasto monimuotoistuu ja kirjoittaminen tarkentuu alakoulun aikana mutta S2-ryhmässä keskimäärin S1-ryhmää hitaammin. Tarkkuus ja sanarakenteen kompleksisuus puolestaan näyttävät kehittyvän eri tahdisa; pitkien sananmuotojen tavoittelu lisää virheiden todennäköisyyttä vain hieman ja vain alimmilla vuosiluokilla. Lineaarisen yhteyden puuttumista selittää muun muassa suomen kielen pitkien sanojen systemaattinen rakenne: monimorfeemisuus ei automaattisesti tee sanasta vaikeaa. Myös harvinaisten sanojen käyttö lisää virheiden todennäköisyyttä tutkimuksessa aineistossa vain aivan alakoulun alkuvaiheessa.

Tarkimmin kirjoittavat sanastollisilta taidoiltaan vahvat lapset. Tämän yhteyden voimakkuuteen voi kielenhallinnan lisäksi kuitenkin vaikuttaa myös yleinen tarkkaavuus ja tehtäväorientaatio. Opettajan arvioimien kielellisten erityisvaikeuksien määrään ei ilmene yhteyttä kummassakaan kieliryhmässä tai kummankaan kielen osalta (suomi, monikielisten oppilaiden oma äidinkieli), mitä osaksi selittää arvioitujen erityisvaikeuksien vähäinen määrä. Muista opettajan arviointiin nojautuvista suomen kielen taitomittareista ainoastaan edellinen todistusarvosana ja holistinen arvio kirjoitustaidosta ovat yhteydessä kirjoittamisen tarkkuu-

teen tutkimusaineistossa, näistäkin jälkimmäinen ainoastaan tyttöjen ryhmässä. On mahdollista, että tytöillä kirjoittamisen tarkkuus vastaa paremmin yleistä kirjoitustaidon tasoa ja että poikien ryhmässä tekstitaidoista irrotettavat tekijät, kuten tarkkaavuuden tai motivaation puute kirjoittamistilanteessa, heikentävät tätä yhteyttä. Kenties yllättävin, melko voimakas ja lisätutkimukseenkin rohkaiseva korrelatiivinen yhteys S2-aineistossa ilmeni koulunkäynnin alkuvaiheen suomen kielen kirjoitustaidon (virhetiheys) sekä myöhemmän oman äidinkielen taidon (puhuminen, kirjoittaminen, sanasto) välillä. Oman äidinkielen taitojen arviointi perustui useiden eri opettajien tekemiin ja useiden eri kielten taitoja koskeviin holistisiin arviointeihin.

Yleisen virhetiheyden laskeminen ja kaikentyyppisten oikeakielisyysvirheiden väheneminen tarkoittavat, että alakouluiässä kirjoittamisen tarkkuus kasvaa laaja-alaisesti siitä huolimatta, että kirjoittaminen yhtä aikaa sujuvoituu ja kompleksistuu. Vaikka pinnistely yhdellä kielitaidon osa-alueella saattaa hetkellisesti laskea suorituskyykyä toisella,¹⁰⁶ pidemmällä aikavälillä kehitys on kaikkiaan nousujohteista. Monet virheet selittyvät alaluokilla sillä, että kirjoitetun kielen konventioiden tuntemus on vielä puutteellista ja kirjoittamisen rutiinit puuttuvat. Esimerkiksi geminaattojen kirjoittamisessa äidinkielisilläkin on horjuntaa aluksi mutta myöhemmin vain satunnaisesti, kun taas yhdyssanojen kirjoittamisessa norminmukaisten konventioiden oppiminen ja käyttö on kompleksimpi ilmiö, joka haastaa myös aikuisia kirjoittajia.

Aineisto osoittaa myös selvästi, että poikien kirjoitustaito kehittyy keskimäärin tyttöjä hitaammin kaikilla tarkkuuden osa-alueilla ja että pitkäkään maassaoloaika ei ole aina auttanut kaikkia Suomessa koulupolkunsa aloittaneita S2-oppilaita saavuttamaan monipuolista ja tarkkaa kielellistä ilmaisu- ja viestintäkykyä. Virheiden olemus ja esiintymisympäristötkin muuttuvat alakoulun aikana, mutta muutokset tapahtuvat eri kirjoittajilla ja kirjoittajaryhmittäin eri rytmissä. Tarkkuustarkastelun tulokset vastaavat nykykäsitystä kielenomaksumisesta: eri kielenoppimisen tasoilla on

106 Esimerkiksi kirjoitustehtävän sisältöön suunnattu ennakkoharjoittelu saattaa viedä huomiota kielen tarkkuudesta ja yllättäen jopa kasvattaa virheiden määrää sen sijaan että vähentäisi niitä, ks. myös Skehan 1998, 69, 70–71.

omat haasteensa ja omat tyypilliset virhekeskittymänsä (ks. esimerkiksi Ellis & Barkhuizen 2005, 78). Voimakkainta kirjoittamisen tarkentuminen on juuri pojilla sekä toisen kielen oppijoiden ryhmissä, joissa lähtötaso koulunkäynnin alkuvaiheessa on matalin. Näin on siitä huolimatta, että juuri pojilla vaikuttaa olleen vaikeuksia osoittaa parasta osaamistaan.

Kielenoppijoille on luontevaa käyttää teksteissään ilmauksia, jotka he jo hallitsevat hyvin (ks. esimerkiksi Folse 2004, 130). Vaikuttaa siltä, että tämä pätee kirjoitetussa suomen kielessä erityisesti sanojen merkityksiin, sillä sanojen käyttö selvästi virheellisessä merkityksessä on tutkitussa kirjoitetun kielen aineistossa harvinaista. Kielen yleisimmät sanat kuitenkin kattavat hyvin suuren osan teksteistä, mikä väistämättä heikentää tekstien ilmaisuvoimaa. Tulosta voi pitää positiivisenakin, sillä sanaston käytön virheet ja epätäsmällisyydet on katsottu keskeisiksi, usein vuorovaikutuksen kannalta merkittävimmiksi kielitaidon puutteiksi (Ellis 1994, 65; Agustín Llach 2011, 91).¹⁰⁷ Myös kielenkäyttäjät itse merkitsevät puheessaan juuri sanastollisia ongelmia epäröintien ja puheen uudelleenmuotoilujen kautta (Halonen & Kokkonen 2008, 132) sekä mieltävät leksikaaliset virheet vakavimmiksi (Gass & Selinker 1994, 270).¹⁰⁸

Oppimisen kannalta pelkästään tutun aineksen käyttäminen ei kuitenkaan ole etu. Virheitä voidaan nimittäin aina tarkastella paitsi osaamisen puutteina myös kielellisten resurssien karttumisesta kertovana ilmiönä (esimerkiksi Seilonen 2013; Mustonen 2015; Lesonen 2020). Esimerkiksi verbien finiittimuotojen virheitä ei synny, ellei oppija ylipäätään käytä infiniittisiä muotoja vaativia rakenteita. Tarkkuusorientoitunut tai ylenpalttisesti korjaavaa palautetta saava oppija voi oppia välttelemään virheitä ja epävarmoiksi kokemiaan ilmaisutapoja. Lopputuloksena voi syntyä virheettömiä tekstejä, mutta samalla tuottamisen sujuvuus ja

107 Huomattakoon, että suomen kielessä myös virheelliset rakenteet toki saattavat haitata viestin välitymistä esimerkiksi englannin kieltä useammin jo runsaamman taiputusmorphologian vuoksi. Lisäksi englanninkielisessä kirjallisuudessa leksikaalisista virheistä puhuttaessa sisällytetään virhetyyppiin usein myös mm. sanan oikeinkirjoituksen virheet ja toisinaan myös taiputuksen pulmat (ks. Agustín Llach 2011).

108 Leksikaalisia virheitä ja niihin suhtautumista on tosin tutkittu lähinnä aikuisten kielenoppijoiden kautta, ja kouluikäisen kielestä tietoa on huomattavasti vähemmän (ks. kuitenkin Agustín Llach 2011).

monimuotoisuus vähenevät eikä oppija pääse hyödyntämään oppimisen koko potentiaaliaan (ks. esimerkiksi Martin ja muut 2010).

Monipuolisen funktionaalisen kielitaidon saavuttamisen tavoitteen sisältyy Suomessa kirjallisten konventioiden ja muodollisen kielen hallinta, sillä sitä edellyttävät niin jatko-opinnot, työelämä kuin toimivan arjen viestintäkin (ks. myös Halonen & Kokkonen 2008, 123–124, 137–139; Kalliokoski 2006; 2008). Esimerkiksi suomi toisena kielenä -ylioppilaskokeesta tiedetään, ettei sisällöllisesti ansiokkaalla ja sekä rakenteellisesti että sanastollisesti monipuolisella kirjoittamisella yletä korkeisiin pistemääriin, mikäli kielivirheitä on tekstissä runsaasti (Elo 1998, 36). Toisaalta norminvastaisten muotojen käyttäminen voi toisinaan olla tietoista ja tarkoituksenmukaista (ks. esimerkiksi Halonen 2009; Lehtonen 2014).

Kielelliselle tarkkuudelle voidaan funktionaalisen kielikäsitteksen mukaisesti asettaa ainakin yksi laaja oppimistavoite: rakenteita ja sanastoa on opittava ymmärtämään ja tuottamaan niin hyvin, että merkityksen kannalta olennaiset erot tulevat selviksi (Martin 2003, 84; Halonen & Kokkonen 2008, 131). Koska kieltä ei kuitenkaan voi opetella tällaisten merkityksiä erottavien esimerkkien tai sääntöjen listoina, vaatimukseen sisältyy loppujen lopuksi valtaisa määrä opittavia asioita: sanoja, sanontoja, rakenteita, prosodiikkaa, sävyjä, tyylejä, käytänteitä, vaihtelua ja vaihtoehtoisuutta – mutta myös sääntöjä, vaihteluttomuutta ja vaihtoehdottomuutta. Vasta, kun yksilö hallitsee kieltä todella hyvin, hänen on mahdollista tietää, miten yksityiskohdat sitoutuvat kokonaisuuksiksi, kuinka paljon kokonaisuuksissa on jouston varaa ja mikä lopulta milloinkin on merkityksen kannalta olennaista. Tutkimuksen keinoin on esimerkiksi leksikaaliseen välttelyyn vaikea saada otetta.¹⁰⁹ Tarkkuus voi olla sidoksissa kielellisten taitojen ja tilannetekijöiden lisäksi myös yksilöllisiin kielenkäyttötapoihin. Erityisesti vieraalla kielellä kirjoitettaessa kohdekielen äänne-, muoto- ja lauserakenteen sekä kirjoitusjärjestelmän etäisyys jo osatuista kielistä mutta myös monet paljon laajemmat tekijät kuten kirjoittamisen konventiot ja vuorovaikutuskulttuuri vaikuttavat

109 Apuna voidaan käyttää esimerkiksi ääneenajattelumenetelmää tai tietoa ilmausten esiintymistodennäköisyyksistä. Rakennetuissa testiympäristöissä suomen kielen oppijoiden turvallisuushakuisuudesta on näyttöä (ks. esimerkiksi Nissilä 2011, 150).

toisaalta siihen, kuinka paljon opittavaa on. Mitä erilaisempaan järjestelmään ja kulttuuriin siirrytään, sitä enemmän kohdekielen mukaisten muotojen omaksuminen vaatii aikaa ja harjoittelua.

Kirjoitustaito ei kehity lineaarisesti tai toteudu samassakaan tekstissä aina samalla tavalla, ja etenkin rajatusta aineistosta tehdyt havainnot antavat kirjoittamisen tarkkuudestakin toisinaan ristiriitaisen kuvan. Tämän tutkimuksen tuloksista ei voi tehdä kauaskantoisia päätelmiä yksilön taidoista – tavoite on ollut ryhmätasolla –, sillä kultakin kirjoittajalta on tutkittu molemmissa vaiheissa vain yksi teksti. Lekin ja muiden (2008) mukaan juuri emotionaalinen teksti, kuten kirjoittaminen unelmien päivästä, kuitenkin vaatii tarkkojen ilmausten miettimistä ja vetää kirjoittajan huomion tyypillisesti tekstin paikallisiin seikkoihin globaalien sijaan. Siksi aineistoa voi pitää tutkimustehtävän kannalta perusteltuna.

Mäkisen (2007) tutkimustulokset rohkaisevat hyödyntämään virheanalyysien tuloksia erityisesti oppilaan puutteellisten kirjoittamisstrategioiden tunnistamiseen. Virheanalyysiin kytketyn sanarakenneanalyysin avulla on Kuljun ja Mäkisen (2008, 524–246) mukaan mahdollista saada tietoa erityisesti siitä, minkälaisia ongelmia heikoilla kirjoittajilla on sanarakenteen hahmottamisessa. Sanatasolla alakoululaisilla on vaikeuksia erityisesti suomen kielen pitkien äännekestojen kirjoittamisessa, mutta on muistettava, että oppimisen kannalta keskittyminen pelkästään näiden kirjoittamisen harjoitteluun ei ole tehokasta. Tarvitaan myös kielellisen tietoisuuden vahvistamista, esimerkiksi sanarakenteen tajun harjoittamista eri tavoin myös puhutun kielen kautta, sekä lukemisen ja kirjoittamisen taitojen tarkastelemista ja tukemista vuorovaikutteisesti.

On mahdollista, että tässä tutkimuksessa havaitun kaltaisia taitoeroja eri kirjoittajaryhmien välillä voitaisiin vähentää tunnistamalla kielellisen prosessoinnin ja opiskelutaitojen puutteita yli kielirajojen ja tukemalla tehokkaita lukemisen, kirjoittamisen ja opiskelun strategioita yhtä aikaa oppilaan kaikissa kielissä. Tiedetään, että sanarakenteen hahmottaminen vaikuttaa niin lukemisen kuin kirjoittamisen oppimiseen (ks. esimerkiksi Kulju & Mäkinen 2009). Tiedossani ei kuitenkaan ole tutkimuksia, joissa olisi tarkasteltu aloittelevien monikielisten lukijoiden ja kirjoittajien taitoja tai kielenhallintaa ja niiden tukemista yhtä aikaa suo-

men kielessä, omassa äidinkielessä ja mahdollisesti jo opiskeltavissa muissa kielissä. Tarvitaan lisätietoa myös siitä, suosiiko kirjoittamisen opetus ja arviointi tietäntyyppisiä opiskelu- tai vuorovaikutustapoja niin, että oppimisprofiilien tasoittamisen sijaan opetus voi jopa ruokkia kirjoittamisen taitoerojen syntymistä. Parhaimmillaan opetuksessa pystytään aidosti tukemaan erilaisia ja erilaisissa oppimisen vaiheissa olevia kirjoittajia, esimerkiksi tunnistamaan taitoerojen lisäksi yksilölle merkityksellisiä kirjoittamisen tapoja ja tavoitteita sekä muokkaamaan opetusta ja arviointia näiden havaintojen pohjalta.

Kirjoitetun kielen normien tuntemus karttuu kouluaikana nopeasti. Toiminnallisen kirjoitustaidon näkökulmasta sääntöpohjainen opetus ei kuitenkaan ole kirjoittamistaidon kehittymisen kannalta tehokkainta. Shoren ja Rapatin (2014) pohjalta voidaan esittää, että kirjoittamisen tarkkuutta lisää ja kirjoitustaitoa vahvistaa pintatason virheiden korjaamista tehokkaammin keskittyminen merkitysten selventämiseen. Tarkkuuden kannalta keskeistä on ymmärtää sääntöjen merkitys ja oppia ohjatusti harjoittelemalla, miten sääntöjä on tarkoitus soveltaa erilaisissa teksteissä ja vuorovaikutustilanteissa. Kirjoittamisen opettamisessa ja ohjauksessa on tärkeää tukea kirjoittamista tasapuolisesti sekä kognitiivisena, kulttuurisena että sosiaalisena pääomana ja toimintana (Luukka 2009). Samalla tulisi huomioida kirjoitustaidon kehityksen jatkumot ja kirjoitustaidon dynaaminen luonne eli esimerkiksi kirjoittamisen kytkökset kirjoittamismotivaatioon ja minäpystyvyyden kokemuksiin (Hankala, Harjunen, Kauppinen, Kulju, Pentikäinen & Routarinne 2015). Tähän ohjaavat myös voimassa olevat Perusopetuksen opetussuunnitelman perusteet (OPH 2014), joissa tuottamistaidot kytetään osaksi monilukutaitoa.

Kirjoitustaidon opettajilta ja arvioijilta vaaditaan herkkyyttä havainnoida kirjoittamisen yksilöllisiä tarpeita ja taitoja eli esimerkiksi sitä, että tietyt kirjoittamisen sisällöt ja taidot eivät ole yhtä merkityksellisiä, ajankohtaisia tai tarpeellisia kaikille kirjoittajille. Kaikissa tilanteissa ei myöskään tarvita samanlaista kirjoittamista. Kommunikatiivisen kirjoitustaidon näkökulmasta kaikki kirjoittamisen tekstilajit ja funktiot eivät esimerkiksi edellytä yhtä suurta tarkkuutta, eikä kaikkia tekstejä tuoteta yksin tai tulkita pelkästään kirjallisen tuotoksen perusteella.

LÄHTEET

- Agustín Llach, M. P. 2011. *Lexical errors and accuracy in foreign language writing*. Bristol: Multilingual Matters.
- Alderson, C. 2005. *Diagnosing foreign language proficiency. The interface between learning and assessment*. London: Continuum.
- Al-Mahrooqi, R. & Al-Aghbari, K. 2016. Refusal strategies among Omani EFL students. *SAGE Open* October-December 2016, 1–10. <http://journals.sagepub.com/doi/pdf/10.1177/2158244016672907>. Viitattu 23.5.2018.
- Afffman, I. & Nissinen, K. 2015. Lukutaidon kehitys PISA-tutkimuksissa. Julkaisussa: J. Välijärvi & P. Kupari (toim.) *Millä eväillä osaaminen uuteen nousuun? PISA 2012 tutkimustuloksia*. Opetus- ja kulttuuriministeriön julkaisuja 2015: 6, 28–49.
- Aro, T., Eklund, K., Leppänen, P. & Poikkeus, A.-M. 2011. Lukivaikeusriskin arviointi ja lukivaikeuden tunnistaminen suomen kielessä. *Psykologia* 46, 92–98.
- Boland, J. E. & Queen, R. 2015. I think your going to like me: Exploring the role of errors in email messages on assessments of potential housemates. *Linguistics Vanguard* 1: 1, 283–293.
- Boland, J. E. & Queen, R. 2016. If you're house is still available, send me an email: Personality influences reactions to written errors in email messages. *Plos One*, 11: 3, e0149885. <https://doi.org/10.1371/journal.pone.0149885>.
- Brunni, S., Jantunen, J. & Skantsi, V. 2019. Korpusavusteinen virheanalyysi tarkkuuden kehityksestä EVK:n taitotasolla A2-B2. *Puhe ja kieli* 39: 3, 275–304.
- Chipere, N. 2009. Individual differences in processing complex grammatical structures. Julkaisussa: G. Sampson, D. Gil & P. Trudgill (toim.) *Language complexity as an evolving variable*. Oxford: Oxford University Press, 178–191.
- Corder, S. P. 1967. The significance of learner's errors. *International Review of Applied Linguistics* 5, 161–170.
- Crossley, S. A., Salsbury, T., McNamara, D. S. & Jarvis, S. 2010. Predicting lexical proficiency in language learner texts using computational indices. *Language Testing* 28: 4, 561–580.
- Ellis, R. 1994. *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. & Barkhuizen, G. 2005. *Analysing learner language*. Oxford: Oxford University Press.
- Elo, E. 1998. Kvantiteetti vai kvaliteetti – vai molemmat? Hyvän kirjoittamisen syntaksia. Julkaisussa: N. Rekola & H. Korpela (toim.) *Toisella kielellä – näkökulmia lasten ja nuorten suomi toisena kielenä -opetukseen*. Helsinki: Opetushallitus, 35–40.
- Eloranta, A.-K., Närhi, V. & Ahonen, T. 2009. Oppimisvaikeuksien pysyvyys – seurantatutkimus yhdeksästä lukivaikeustaustaisesta nuoresta naisesta. *NMI-bulletin* 19: 1, 4–19.
- Eronen, R., Maamies, S. & Räikkälä, A. 1996. Yhdyssanat. *Kielikello* 1996: 4.
- Erra, S. & Svinhufvud K. 2017. Kirjoittamisen diskurssit lukion äidinkielen ja kirjallisuuden opetussuunnitelman perusteissa ja äidinkielen ylioppilaskokeen määräyksissä. *Virittäjä* 121: 3, 316–354.
- Folse, K. S. 2004. *Vocabulary myths. Applying second language research to classroom teaching*. Ann Arbor: The University of Michigan Press.

- Gass, S. M. & Selinker L. 1994. *Second language acquisition. An introductory course*. Hillsdale: Lawrence Erlbaum Ass.
- Gass, S. M. & Selinker, L. 2008. *Second language acquisition. An introductory course*. 3., uudistettu painos. Hoboken: Taylor & Francis.
- Halonen, M. 2009. Puhutun kielen variantit resurssina monikielisten koululaisten kirjoittelussa. *Virittäjä* 113: 3, 329–355.
- Halonen, M. & Kokkonen, M. 2008. ”Loppu-ännät puuttuu!” ”Toivotaan sujuvaa puhetta!” – Lapsi- ja aikuisoppijoiden kielenoppimis- ja arviointikontekstien erot. Julkaisussa: J. Tanner & M. Kokkonen (toim.) *Kakkoskieli 6. Suomenopetus, kielitaito ja tutkimus*. Helsinki: Helsingin yliopisto, 109–140.
- Hamp-Lyons, L. 2003. Writing teachers as assessors of writing. Julkaisussa: B. Kroll (toim.) *Exploring the dynamics of second language writing*. Cambridge: Cambridge University Press, 162–189.
- Hankala, M., Harjunen, E., Kauppinen, M., Kulju, P., Pentikäinen, J., & Routarinne, S. 2015. Tutkimuksesta tukea kirjoittamisen opettamiseen. Julkaisussa: E. Harjunen (toim.) *Tekstit puntarissa: Ajatuksia äidinkielen ja kirjallisuuden oppimistuloksista perusopetuksen päättöarvioinnissa 2014 ja 2010*. Helsinki: Kansallinen koulutuksen arviointikeskus, 73–84.
- Havukainen, M. & Salmi, L. 2011. *Oppilaiden tuottavan kirjoittamisen taso toisella luokalla*. Kasvatustieteen pro gradu -tutkielma. Itä-Suomen yliopisto.
- Honko, M. 2013. *Alakouluikäisten leksikaalinen tieto ja taito: toisen sukupolven suomi ja S1-verrokki*. Tampere: Tampere University Press.
- Honko, M. 2017a. Sadutettu sanasto: puhutun kielen leksikaalinen diversiteetti arviointikohteena. Julkaisussa: M. Kuronen, P. Lintunen & T. Nieminen (toim.) *AFinLA-e*. Soveltavan kielitieteen tutkimuksia 10, 163–192.
- Honko, M. 2017b. Kieli- ja kielitaitokäsitykset tutkivan opettajan kenttäpäiväkirjamerkinnoissä. *Puhe ja kieli* 37: 4, 215–238.
- Housen, A. & Kuiken, F. 2009. Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics* 30: 4, 461–473.
- Housen, A., Kuiken, F. & Vedder I. (toim.) 2012. *Performance and proficiency: Complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins.
- Huhta, A. & Hildén, R. 2016. Kielitutkinnot ja muu laajamittainen kielitaidon arviointi Suomessa. Julkaisussa: A. Huhta & R. Hildén (toim.) *AFinLA-e*. Soveltavan kielitieteen tutkimuksia 9, 3–26.
- Ivanič, R. 2004. Discourses of writing and learning to write. *Language and Education* 18: 3, 220–245.
- Ivaska, I. 2009. *Eksistentiaalilauseen ilmiäsuut edistyneiden suomenoppijoiden kirjoituksessa*. Suomen kielen pro gradu -tutkielma. Turun yliopisto.
- James, C. 1998. *Errors in language learning and use: Exploring error analysis*. Essex: Longman.
- Jessmer, S. L., & Anderson, D. 2001. The effect of politeness and grammar on user perceptions of electronic mail. *North American Journal of Psychology* 3: 2, 331–346.
- 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.
- Kaprio, J. 2006. Lahjakkuuden monet ulottuvuudet. *Duodecim* 122: 23, 2871–2877.
- Kauppinen, M., Pentikäinen, J., Hankala, M., Kulju, P., Harjunen, E. & Routarinne, S. 2015. Systemaattinen katsaus perusopetusikäisten kirjoittamisen opetusta ja osaamista koskevaan tutkimukseen. *Kasvatus* 46: 2, 160–175.
- Kreiner, D. S., Schnakenberg, S. D., Green, A. G., Costello, M. J., & McClin, A. F. 2002. Effects of spelling errors on the perception of writers. *Journal of General Psychology* 129: 1, 5–17.
- Kulju, P. & Mäkinen, M. 2009. Sanelukirjoituksen arviointi – tyyppivirheistä sanarakenteiden fonologiseen analyysiin. *Virittäjä* 113: 4, 508–532.
- Kuukka, K. & Metsämuuronen J. 2016. *Perusopetuksen päättövaiheen suomi toisena kielenä (S2) -oppimäärän oppimistulosten arviointi 2015*. Helsinki: Karvi.
- Lado, R. 1957. *Linguistics across cultures*. Ann Arbor: The University of Michigan.
- Lappalainen, H. 2011. *Sen edestään löytää. Äidinkielen ja kirjallisuuden oppimistulokset perusopetuksen päättövaiheessa 2010*. Helsinki: Opetushallitus. http://koulutustoimikunnat.fi/download/132347_Sen_edestaan_loytaa.pdf. Viitattu 31.5.2020.
- Lehtinen, R. 2013. *Kielen norminmukaisuus alakoululaisten kirjoitelmissä*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Lehtonen, H. 2015. *Tyylitellen: nuorten kielelliset resurssit ja kielen sosiaalinen indeksisyys monitieteisessä Helsingissä*. Helsinki: Helsingin yliopisto.
- Lerkkanen, M.-K., Poikkeus, A.-M., Ahonen, T., Siekkinen, M., Niemi, P. & Nurmi, J.-E. 2010. Luku- ja kirjoitustaidon kehitys sekä motivaatio esi- ja alkuopetusvuosina. *Kasvatus* 41: 2, 116–128.
- Lesonen S. 2020. *Valuing variability: Dynamic usage-based principles in the L2 development of four Finnish language learners*. Jyväskylä: Jyväskylän yliopisto.
- Lyly, T. 2008. Opettajien kokemuksia tekstien arvioinnista peruskoulussa. Julkaisussa: *Kakkoskieli* 6. Suomenopetus, kielitaito ja tutkimus. Helsinki: Helsingin yliopisto, 141–162.
- Malvern D., Richards, B. Chipere, N. & Durán, P. 2004. *Lexical diversity and language development: Quantification and assessment*. Hampshire: Palgrave Macmillan.
- Manninen, T. & Martin, S. 2000. Vieraskielisten ongelmia ylioppilasaineessa. Julkaisussa: S. Martin & H. Sulkala (toim.) *Tutkielmia oppijankielestä*, 73–90.
- Martin, M. 1995. *The map and the rope: Finnish nominal inflection as a learning target*. Jyväskylä: Jyväskylän yliopisto.
- Martin, M. 2003. Kieli on kuin lammikko – johdatusta toisen kielen oppimiseen. Julkaisussa: L. Nissilä, H. Vaarala & M. Martin (toim.) *Suolla suomea. Perustietoa maahanmuuttajien suomen kielen opettajille*. Helsinki: Äidinkielen opettajain liitto, 75–90.
- Martin, M. 2009. Yhteistä ja erityistä – johdantoa teemanumeron kirjoituksiin. *Virittäjä* 113: 3, 321–328.
- Martin, M. 2010. Sujuvuus, tarkkuus ja kompleksisuus kielitaidon kehityksessä. Julkaisussa: S. Stolt, M. Lehtihalmes, S. Tarvainen & K. Launonen (toim.) *Suomalainen monikielisyys ja sen haasteet*. Helsinki: Puheen ja kielen tutkimuksen yhdistys.
- Martin, M. 2015. VIRHE! Professori Maisa Martinin jäähyväisluento. Jyväskylän yliopisto 25.9.2015.

- Martin, M., Mustonen, S., Reiman, N. & Seilonen, M. 2010. *On becoming an independent user*. Julkaisussa: I. Bartning, M. Martin & I. Vedder (toim.) *Communicative proficiency and linguistic development, intersections between SLA and language testing research*. EUROSLA Monograph Series 1. European Second Language Association, 57–80.
- Matilainen, K. 1985. Lukemaanopettamismenetelmien yhteydet oikeinkirjoitustaidon oppimiseen ensimmäisen lukuvuoden aikana. Kasvatustieteiden tiedekunnan tutkimuksia 5. Joensuu: Joensuun yliopisto.
- Matilainen, K. 1993. *Peruskouluilaisten kirjoittamisen hahmottamisstrategioiden kehittyminen ja kuvaus oppimisvaikeuksisten oppilaiden kirjoittamisen ja luetun ymmärtämisen strategioista*. Joensuu: Joensuun yliopisto.
- Meara, P. & Milton, J. 2003. *X_Lex, The Swansea Levels Test*. Newbury: Express.
- Meara, P., Rodgers, C. & Jacobs, G. 2000. Vocabulary and neural networks in the computational assessment of texts written by second-language learners. *System* 28: 3, 345–354.
- Milton, J. 2009. *Measuring language vocabulary acquisition*. Bristol: Multilingual Matters.
- Mustonen, S. 2015. *Käytössä kehittyvä kieli. Paikat ja tilat suomi toisena kielenä -oppijoiden teksteissä*. Jyväskylä: Jyväskylän yliopisto.
- Mäki, H. 2002. *Elements of spelling and composition. Studies on predicting and supporting writing skills in primary grades*. Turku: Turun yliopisto.
- Mäkinen, M. 2007. Fonologinen analyysi toisluokkalaisten luku- ja kirjoitusprofileista. *Puhe ja kieli* 27: 4, 149–170.
- Mäkihönkö M. 2006. *Luetun ymmärtämisen ja tuottavan kirjoittamisen kehittyminen alkuopetuksen aikana*. Joensuu: Joensuun yliopisto.
- Neff, J., Ballesteros, F., Dafouz, E., Martinez, F., Rica, J.-P., Diez, M. & Prieto, R. 2007. A contrastive functional analysis of errors in Spanish EFL university writers' argumentative texts: A corpus-based study. Julkaisussa: E. Fitzpatrick (toim.) *Corpus linguistics beyond the word. Corpus research from phrase to discourse*. Amsterdam: Rodopi, 203–225.
- Niiranen, L. 2010. Tapaustutkimus kolmen suomenoppijan kompleksisista verbikonstruktiosta. Julkaisussa: A. Kaivapalu, P. Muikku-Werner & M.-M. Sepper (toim.) *Lähivördlusi. Lähivertailuja* 20, 155–190.
- Nissilä L. 2011. *Viron kielen vaikutus suomen kielen verbien ja niiden rektioiden oppimiseen*. Oulu: Oulun yliopisto.
- OPH 2014. Perusopetuksen opetussuunnitelman perusteet. Helsinki: Opetushallitus.
- Pajunen, A. 2012. Kirjoittamistaitojen kehitys 8–12-vuotiailla. Alakouluilaisten nelmakirjoittelmat. *Virittäjä* 116: 1, 4–32.
- Pekonen, V. 2007. *Suomen kvantiteetti ja kirjoitusprosessi: Mallinnusta ja empiriaa*. Jyväskylä: Jyväskylän yliopisto.
- Pietilä, P. 2012. Lexical vs. syntactic competence in L2 English. *Journal of Language Teaching and Research* 3: 2, 314–320.
- Piper, T. 2001. *And then there were two: Children and second-language learning*. Toronto: Pippin Publishing Corporation.
- Pääkkönen, I. 1985. Kielivirheiden asema ja merkitys ylioppilasaineiden arvostelussa. Ylioppilaskirjoituksen aineiden arvostelijan puheenvuoro. *Kielikello* 1, 19–21.
- Read, J. 2000. *Assessing vocabulary*. Cambridge: Cambridge University Press.

- Reiman, N. 2011. Transitiivikonstruktio ikkunana syntaksin kehitykseen: Infiniittiset rakenteet ja passiivi taidon indikaattoreina S2-oppijoiden teksteissä. Julkaisussa: E. Lehtinen, M. Aaltonen, M. Koskela, E. Nevasaari & M. Skog-Södersved (toim.) *AFinLA-e Soveltavan kielitieteen tutkimuksia* 3, 142–157.
- Saarela, L. 1997. *Peruskoululaisten kirjoitelmien kehittyminen sanastotutkimuksen valossa*. Oulu: Oulun yliopisto.
- Salonen, I. 2012. *Puhekielen piirteiden esiintyminen alakoululaisten kirjoitelmissa*. Suomen kielen pro gradu -tutkielma. Tampereen yliopisto.
- Santala, R., Saaristo-Rinta, S., Metsola, M. & Klippi, A. 2013. Suomenkielisten 13–16-vuotiaiden nuorten kielellisten taitojen arviointi. *Puhe ja kieli* 33: 1, 15–41.
- Seilonen, M. 2013. *Epäsuora henkilöön viittaaminen oppijansuomessa*. Jyväskylä: Jyväskylän yliopisto.
- Shore, S. & Rapatti, K. (toim.) 2014. *Tekstilajitaidot. Lukemisen ja kirjoittamisen opetus koulussa*. Helsinki: Äidinkielen opettajain liitto.
- Sikiö, S., Siekinen, M. & Holopainen, L. 2016. Literacy development among language minority background and dyslexic children in Finnish orthography context. *Reading Psychology* 37: 5, 706–727.
- Skehan, P. 1998. *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Slabakova, R. 2016. *Second language acquisition*. Oxford: Oxford University Press.
- Spoelman, M. 2013. Prior linguistic knowledge matters: The use of the partitive case in Finnish learner language. Oulu: Oulun yliopisto.
- Stenberg, P.-R. 2010. *Yhdysänatajain kehittyminen ja lingvistinen tiedonrakentelu peruskoulusta yliopistoon*. Suomen kielen pro gradu -tutkielma. Itä-Suomen yliopisto.
- Stothers, M. & Klein, P. 2010. Perceptual organization, phonological awareness, and reading comprehension in adults with and without learning disabilities. *Annals of Dyslexia* 60, 209–237.
- Tarnanen, M. 2002. *Arvioija valokeilassa. Suomi toisena kielenä -kirjoittamisen arviointia*. Jyväskylä: Jyväskylän yliopisto.
- Tanner, J. 2012. *Rakenne, tilanne ja kohteliaisuus. Pyyntö S2-oppikirjoissa ja autenttisissa keskusteluissa*. Helsinki: Helsingin yliopisto.
- Tarnanen, M. 2003. Kielitaidon arviointi: valintoja ja neuvottelua. Julkaisussa: L. Nissilä, H. Vaarala & M. Martin (toim.) *Suolla suomea. Perustietoa maahanmuuttajien suomen kielen opettajille*. Helsinki: Äidinkielen opettajain liitto, 143–160.
- Tervola, M. 2019. *Maahanmuuttajalääkärien suomen kielen taito ja kielitaitotarpeet: Erityisalan kielitaidon näkökulma lääkärin työhön*. Tampere: Tampereen yliopisto.
- Tomasello, M. & Akhtar, N. 2000. Five questions for any theory of word learning. Julkaisussa: R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. L. Woodward, N. Akhtar, M. Tomasello & G. Hollich (toim.) *Becoming a word learner: A debate on lexical acquisition*. Counterpoints: cognition, memory, and language. Oxford: Oxford University Press, 179–186.
- Turunen, U. 2012. *Suomi toisena kielenä -oppijoiden possessiivisuffiksien käyttö*. Suomen kielen pro gradu -tutkielma. Jyväskylän yliopisto.
- Undheim, A. 2009. A thirteen-year follow-up study of young Norwegian adults with dyslexia in childhood: Reading development and educational levels. *Dyslexia* 15: 4, 291–303.

- Unsworth, S. 2008. Comparing child L2 development with adult L2 development: How to measure L2 proficiency. Julkaisussa: B. Haznedar & E. Gavruseva (toim.) *Current trends in child second language acquisition: A generative perspective*. Amsterdam: John Benjamins, 301–333.
- Vignovic, J. & Thompson, L. F. 2010. Computer-mediated cross-cultural collaboration: Attributing communication errors to the person versus the situation. *Journal of Applied Psychology* 95: 2, 265–276.