

# **Kielen norminmukaisuus alakoululaisten kirjoitelmissa**

Riikka Lehtinen

Tampereen yliopisto

Kieli-, käännös- ja kirjallisuustieteiden yksikkö

Suomen kieli

Pro gradu -tutkielma

Elokuu 2013

Tampereen yliopisto  
Suomen kieli  
Kieli-, käännös- ja kirjallisuustieteiden yksikkö

LEHTINEN, RIIKKA: Kielen norminmukaisuus alakoululaisten kirjoitelmissa

Pro gradu -tutkielma, 92 sivua  
Elokuu 2013

---

Tutkielmassa tarkastellaan alakoulun toisen, neljännen ja kuudennen luokan oppilaiden kielen norminmukaisuutta Unelmien(i) päivä -kirjoitelmissa. Tutkielma kuuluu projektiin *Suomen kielen myöhempi hallinta: kouluikäiset ja nuoret aikuiset* (Pajunen), ja sen aineistona toimii osa projektin yhteydessä kerätystä Koulukorpuksesta. Tutkielman aineiston 544 kirjoitelmaa on kerätty vuonna 2006 Etelä- ja Länsi-Suomen alakoulun 2.–6. luokalta.

Kielen norminmukaisuutta mitataan fenomenologis-deskriptiivisen virheanalyysin keinoin. Tarkastelun kohteeksi on valittu 13 oikeinkirjoitusvirhettä, joista suurin osa on fonologisia, yksittäisiin kirjaimiin ja merkkeihin kohdistuvia virheitä, mutta mukana on myös morfologisia ja syntaktisia piirteitä. Piirteet on valittu tutkimukseen perinteisen sanelukirjoituksessa käytettävän virheanalyysin virhekatgorioita mukaillen, mutta perinteisen virheanalyysin mallia on tässä tutkimuksessa muokattu koulukirjoitelma-aineistoon paremmin sopivaksi.

Aineistosta tarkastellaan oikeinkirjoitusvirheiden esiintymistä ja esiintymäfrekvenssien kehityskulkuja luokka-asteilla kaksi, neljä ja kuusi sekä pohditaan syitä erilaisten virheiden esiintymiin. Pyrkimyksenä on tunnistaa todelliset virheet satunnaisista lipsahduksista. Hypoteesina on, että virheitä esiintyy kirjoitelmissa sitä vähemmän ja hallitummin, mitä ylemmältä luokka-asteesta on kyse. Lisäksi tutkimuksessa vertaillaan tyttöjen ja poikien eroja oikeinkirjoitustaidoissa. Oletuksena on, että poikien tutkitusti tyttöjä kielteisempi asennoituminen äidinkielen opetukseen sekä sanamääriltään lyhyemmät kirjoitelmat voisivat viitata siihen, että tytöt hallitsevat kirjoitetun kielen normit poikia paremmin.

Tutkimusanalyysi osoittaa, että tutkitut virheet ovat luonteeltaan hyvin erityyppisiä ja selittyvät eri luokka-asteilla eri tekijöillä. Eniten virheitä on kaikilla luokka-asteilla sanarajojen merkitsemisessä ja yhdyssanoissa. Varsinaisia virheitä on aineiston nuorimmilla oppilailla melko runsaasti, kun taas vanhempien oppilaiden virheet ovat todennäköisemmin satunnaisia lipsahduksia. Virheiden määrä vähenee radikaalisti toisen ja neljännen luokan välissä, mutta erot tasoittuvat neljännen ja kuudennen luokan aineistoja vertailtaessa. Tyttöjen ja poikien tekemien virheiden määrässä on melko selvä ero. Kirjoitelmien pituuksiin suhteutettu poikien oikeinkirjoitusvirheiden määrä on lähes kaksinkertainen tyttöihin verrattuna kaikilla luokka-asteilla.

**Avainsanat:** koulukorpus, alakoulu, kielen norminmukaisuus, kirjoittaminen, virheanalyysi

# Sisältö

<b>1. Johdanto</b> .....	1
1.1 Tutkimuksen tavoitteet.....	1
1.2 Tutkimusaineisto.....	3
<b>2. Luku- ja kirjoitustaito</b> .....	7
2.1 Lapsen myöhempi kielenkehitys.....	7
2.1.1 Kielellinen tietoisuus.....	8
2.1.2 Lapsen äänteellinen kehitys.....	9
2.2 Kirjoittamaan oppiminen.....	11
2.2.1 Oikeinkirjoitustaidon kehittyminen.....	14
2.2.2 Kirjoittamisen haasteet.....	15
2.3 Virheanalyysistä.....	18
2.3.1 Virheen käsite.....	21
2.3.2 Puhkekielisyyden ja virheen suhde.....	22
<b>3. Virheanalyysi Unelmieni päivä -kirjoitelmista</b> .....	25
3.1 Virheiden ryhmittely.....	25
3.2 Virheet sanarajoissa.....	27
3.2.1 Erilleen kirjoitettu yhdyssana.....	28
3.2.2 Puuttuva sanaväli.....	32
3.3 Puuttuva kirjain olettamuksia sanahahmon vaikutuksesta oppilaiden tekemiin virheisiin.....	35
3.3.1 Puuttuva konsonantti geminaatasta.....	36
3.3.2 Puuttuva vokaali pitkästä vokaalista.....	40
3.3.3 Muu puuttuva kirjain.....	42
3.4 Väärä kirjain.....	51
3.4.1 Väärä konsonantti.....	51
3.4.2 Väärä vokaali.....	55
3.5 Ylimääräinen kirjain.....	59
3.5.1 Konsonantin geminoituminen.....	60
3.5.2 Vokaalin pidentyminen.....	62
3.5.3 Muu ylimääräinen kirjain.....	63
3.6 Taivuttamiseen liittyvä virhe.....	65
3.6.1 Virheellinen taivutus.....	66
3.6.2 Virhe symbolissa.....	70
3.6.3 Monitulkintaiset virheet.....	71
<b>4. Lopuksi</b> .....	76
<b>Lähteet</b> .....	90

## 1. Johdanto

Kertovan tekstin kirjoittaminen on maailman ja oman ajattelun kielentämisen perustaito, jota harjoitellaan läpi peruskoulun. Perusopetuksen opetussuunnitelman perusteissa (2004: 57) todetaan, että yltääkseen hyvään äidinkielen ja kirjallisuuden arvosanaan peruskoulun päättyessä oppilaan on kyettävä muun muassa laatimaan suullisia ja kirjoitettuja tekstejä eri tarkoituksiin. Koulukirjoitelmien onnistuneisuutta sekä niissä saavutettua kielen ja kerronnan tasoa voidaan tarkastella monenlaisiin muuttujiin verraten. Tämän pro gradu -tutkielman aiheena on sanatason virheanalyysi alakoululaisten kirjoitelmista luokka-asteilla kaksi, neljä, ja kuusi. Virheanalyysia tekemällä halutaan tuottaa lisää tietoa siitä, miten hyvin alakoululaisten koulukirjoitelmissaan käyttämä kieli noudattaa normatiivisen kieliopin ohjeistusta ja kuinka se kehittyy vuosiluokalta toiselle siirryttäessä. Kouluikäisten lasten kielitaidon tarkastelu on erityisen kiinnostavaa, sillä tämä kehitysvaihe on jäänyt tutkimuksissa usein lapsen varhaisen kielenkehityksen varjoon. Tutkielman aineistona on käytetty lapsen myöhempää kielenkehitystä koskevan tutkimusprojektin yhteydessä kerättyä ja yhä kasvavaa koulukirjoitelmakorpusta (Pajunen, Puranen & Yli-Paavola 2007–2012), joka koostuu alakoululaisten kirjoitelmista aiheesta Unelmieni/Unelmien päivä. Tarkastelun kohteeksi on valittu 13 sanatason oikeinkirjoitusvirhettä, jotka esiintyvät äänne- ja muotoopin sekä syntaksin tasoilla.

Koulukorpus on kerätty lapsen myöhempää kielenkehitystä koskevan tutkimusprojektin yhteydessä. Kyseessä on laaja tutkimus ja laaja aineisto, joten siitä on jatkuvasti tekeillä useita opinnäytetöitä ja Tampereen yliopistosta on valmistunut useita pro gradu -tutkielmia, joiden aineisto on peräisin Koulukorpuksesta. Erilaisia oppilaiden kirjoitustaitoa mittaavia testejä on kehitelty runsaasti kautta tutkimushistorian, ja tämän tutkimuksen kannalta keskeisimmät ovat Ruoppilan, Römanin ja Västin virheanalyysi (1969) sekä Kuljun ja Mäkisen sanarakenneanalyysi (2009). Tutkimuksen aineiston käsittely tulee pohjautumaan ainakin osittain molempiin edellä mainittuihin menetelmiin.

### 1.1 Tutkimuksen tavoitteet

Tämä tutkimus tarkastelee alakoululaisten kirjoitustaitoa sanatason virheanalyysin kautta. Tutkimuksessa ryhmitellään erilaisia alakoululaisten teksteissä esiintyviä oikeinkirjoitusvirheitä eri ryhmiin ensinnäkin niiden yleisen luonteen mukaan; virhe voi olla esimerkiksi sanarajassa, äännepitäudessa tai erilaisissa kompleksissa rakenteissa. Lisäksi

virheet jaotellaan tarkempiin alaluokkiin niiden esiintymisympäristöjen mukaan ja nimetään näin syntyneet virheryhmät. Toiseksi vertaillaan eri virhetyyppien yleisyyttä eri luokka-asteilla ja tutkitaan, millaisia kehityslinjoja eri luokka-asteiden välillä voidaan havaita. Lisäksi suoritetaan myös tyttöjen ja poikien tuottamien tekstien vertailua toisiinsa juuri virhetyyppien yleisyyden näkökulmasta. Kolmanneksi tutkimuksessa pohditaan virheiden syitä ja sitä, mitkä suomen kielen rakenteet tuottavat erityisesti ongelmia oikeinkirjoituksessa alakoululaisilla. Oppilaan mahdollista lukivaikeutta voivat implikoida tietynlaiset virherakenteet hänen tuottamassaan tekstissä. Tällaista pohdintaa varten tutkimuksessa on myös perehdytty tyypillisiin lukihäiriöisillä kirjoituksessa esiintyviin virhetyyppisiin sekä lapsen äänteellisen kehityksen ongelmiin ja pyritty tunnistamaan niitä aineistosta. Tutkimuskysymykset ovat tiivistettynä seuraavat:

- Millaista alakouluikäisten kirjoitelmien kieli on normatiivisen kieliopin näkökulmasta? Millaisia oikeinkirjoitusvirheitä kirjoitelmissa esiintyy ja kuinka paljon?
- Miten paljon erilaisten virheiden esiintyminen vaihtelee kirjoitelmissa eri vuosiluokkien välillä, ja mitä syitä erojen taustalla on? Voidaanko eri piirteiden esiintymien välillä hahmottaa selkeitä muutoksia tai kehityskulkuja eri vuosiluokkien välillä?
- Miten paljon erilaisten virheiden esiintyminen vaihtelee kirjoitelmissa sukupuolten välillä? Minkälaisia kehityskulkuja hahmottuu eri sukupuolten välillä?

Tutkimustiedoista voi olla käytännön apua opetuksen suunnittelussa, sillä tulokset havainnollistanevat ainakin joitakin vaikeuksia, joita alakouluikäiset kirjoittajat kohtaavat tavoitellessaan norminmukaista kieltä. Hypoteesina on, että oikeinkirjoitusvirheitä esiintyy kirjoitelmissa sitä vähemmän ja sitä hallitummin käytettynä, mitä ylemmästä vuosiluokasta on kysymys. Alempien vuosiluokkien teksteissä oikeinkirjoitusvirheitä esiintyy melko paljon ja tutkittavien piirteiden norminvastaiset muodot esiintyvät mielivaltaisesti sekaisin norminmukaisten kanssa samojen kirjoittajien teksteissä. Tutkimus on pääosin kvalitatiivinen, sillä koulukirjoitelmia analysoidaan niissä käytetyn kielen norminmukaisuuden ja virheiden laadun kautta. Kvantitatiivisen tutkimuksesta tekee se, että tarkastelussa on myös numeerisia tuloksia, joita havainnollistetaan kaavioiden ja taulukoiden avulla.

## 1.2 Tutkimusaineisto

Koulukirjoittamiskorpus on koottu myöhempää kielenkehitystä tutkivan projektin yhteydessä, ja tämä tutkimus kuuluukin projektiin *Suomen kielen myöhempi hallinta: kouluikäiset ja nuoret aikuiset* (Pajunen). Korpuksen keruu alkoi vuonna 2006, ja se koostuu tällä hetkellä yli 2000:sta alakoululaisen kirjoitelmasta. Koulukorpus on kerätty etelä-, itä- ja länsisuomalaisista kouluista, joista osa sijaitsee maaseudulla ja osa kaupungissa. Suurimman osan kirjoittajista äidinkieli on suomi; muuta kieltä kuin suomea äidinkielenään puhuvien tekstejä ei ole erotettu aineistosta omaksi ryhmäkseen, eikä esimerkiksi oppilaiden mahdollisista kehityshäiriöistä ole kerätty tietoja. Murrealueellisesti koulut sijoittuvat siten, että kirjoitelmissa esiintyy niin pääkaupunkiseudun kielen kuin hämäläismurteiden, eteläpohjalaisten murteiden ja ehkä savolaismurteidenkin piirteitä. Tässä tutkielmassa keskitytään tutkimaan kirjoitelmissa tehtyjä oikeinkirjoitusvirheitä, toisin sanoen sanamuotoja, jotka eivät noudata normatiivisen kieliopin mukaista kirjoitusasua. Varsinaiset virheet on siis pyritty huolellisesti erottamaan murteellisista ja puhekielisistä ilmaisuista, vaikka yksiselitteistä tällainen eronteko ei ole.

Tämän tutkimuksen aineisto koostuu osista Pajusen, Purasen ja Yli-Paavolan vuosina 2006–2012 keräämästä Koulukorpuksesta. Korpus koostuu alakoulun kaikkien luokka-asteiden oppilaiden (1.–6. luokka) tuottamista teksteistä. Kirjoitelmien pituus vaihtelee eri luokka-asteiden välillä. Keskimäärin kirjoitelmien pituus vaihtelee paristakymmenestä sanasta runsaaseen sataan. Luokka-asteelta toiselle siirryttäessä kirjoitelmien pituus kasvaa parillakymmenellä sanalla (Pajunen 2012: 7). Pisimmissä teksteissä sanoja on sadasta yli kolmeensataan. Etenkin ensimmäisen luokan oppilaiden kirjoitelmissa on huomattavan suuria pituus- ja tasoeroja, sillä lasten kirjoitustaito on vasta löytymässä.

Ensimmäisen luokan kirjoitelmia ei voi systemaattisesti verrata ylempien luokkien aineistoon ensinnäkin jo sen vuoksi, että korpus sisältää ensiluokkalaisten tekstejä jonkin verran muita luokka-asteita vähemmän (Pajunen 2012: 6). Lisäksi etenkin ensimmäisen luokan oppilaiden kirjoitelmissa on huomattavan suuria pituus- ja tasoeroja, sillä oppilaiden lukitaidoissa on jo kouluun tullessa ollut suuria eroja: osa ensiluokkalaisista on osannut lukea ja kirjoittaa jo ennen kouluun tulemistä ja osan luku- ja kirjoitustaidot ovat olleet vielä aineistonkeruun aikaan hyvin haparoivia. Ensimmäisen luokan kirjoitelmat onkin rajattu tämän tutkimuksen ulkopuolelle, sillä edellä mainitut seikat aiheuttaisivat aineiston epäsuhteellista painottumista ja vinoutuneita tutkimustuloksia. Lisäksi koulukorpusaineistossa on mukana kielikoulun oppilaita, jotka olivat korpuksen keräämisen aikana kolmannella

luokalla. Nämä oppilaat on valittu kielellisiin kriteereihin perustuvalla pääsykokeella toisen luokan keväällä. Aiemmat Koulukorpuksesta tehdyt tutkimukset osoittavat, että kolmannen luokan kielikouluoppilaiden kirjoitelmat ovat aineiston muihin kolmasluokkalaisten kirjoitelmiin nähden huomattavasti tasokkaampia, usein jopa neljäsluokkalaisten kirjoitelmia tasokkaampia, ja näin aiheuttavat pientä vinoumaa tutkimustuloksiin, kun halutaan tuottaa tietoa yleisestä alakoululaisten kirjoitustaidon kehityksestä alakoulun aikana (ks. mm. Salonen 2012). Tämä tutkimus on siis rajattu koskemaan alakoulun luokka-asteita kaksi, neljä, ja kuusi, jotta olisi helpompi tarkastella eri luokka-asteilla saavutettua kirjoitelman tasoa ja luokka-asteelta toiselle siirryttäessä tapahtuvaa kehitystä. Hypoteesina on, että luokka-asteelta toiselle edettäessä tekstien laatu paranee ja virheiden määrä pienenee. Toisin sanoen toisen luokan oppilaiden kirjoittamat kertomukset sisältänevät enemmän virheitä kuin kuudesluokkalaisten.

Kirjoitelmia on perusaineistossa noin 960, ja se perustuu noin 70 000 tekstisanan aineistoon. Tämän tutkimuksen aineisto sisältää osia kokonaisesta Koulukorpuksesta, ja tämän otannan kirjoittajien luokka-asteet ja sukupuolet jakautuvat taulukon 1 osoittamalla tavalla.

**Taulukko 1.** Aineiston kirjoittajat luokka-asteittain ja sukupuolittain.

Luokka-aste	Poikia (kpl)	Tyttöjä (kpl)	Yhteensä (kpl)
2.	102	113	215
4.	83	93	176
6.	77	76	153
<b>Yhteensä</b>	262	282	544

Tämän tutkimuksen aineistossa on siis tyttöjen kirjoitelmia 20 kappaletta enemmän kuin poikien. Ero on siinä mielessä huomattava, että tutkimustuloksina esitettyjä numeerisia arvoja tarkastellaan pelkkien virhemäärien lisäksi myös aineistojen kokoihin ja kirjoittajien määriin suhteutettuna.

Tähän tutkimukseen valikoitunut otanta Koulukorpuksesta sisältää 544 kirjoitelmaa, joiden yhteenlaskettu sanamäärä on 42 290. Kirjoitelmien sanamäärät on laskettu korpuksen kirjakielistetystä sanasarakkeesta, jotta saadut luvut olisivat vertailukelpoisia. Kirjakielistäminen koskee muun muassa useamman sanan yhteen kirjoittamista ja yhdyssanavirheitä. Tutkimuksessa analysoitavat virheet sen sijaan on poimittu sanojen

alkuperäisistä muodoista. Kirjoitelmien yhteenlasketut sanamäärät ja keskiarvopituudet on esitelty taulukossa 2.

**Taulukko 2.** Kirjoitelmien yhteenlasketut sanamäärät luokka-asteittain ja sukupuolittain.

<b>Luokka-aste</b>	<b>Pojat</b>	<b>Tytöt</b>	<b>Yhteensä</b>	<b>Kirjoitelmien keskiarvopituus sanoina</b>
<b>2.</b>	3926	6482	10408	48
<b>4.</b>	5522	8784	14306	81
<b>6.</b>	6802	10774	17576	115
<b>Yhteensä</b>	16250	26040	42290	78

Taulukko 2 osoittaa, että tutkimusaineistossa on suuria eroja myös kirjoitelmien sanamäärien suhteen: tyttöjen kokonaisaineisto käsittää lähes 10 000 sanaa enemmän kuin poikien, ja tätä eroa voidaan pitää todella merkittävänä. Taulukossa 2 on kuvattuna myös luokka-asteittainen kirjoitelmien keskipituus sanoina. Nämä luvut taas osoittavat, että kirjoitelmien sanamäärä lähes kaksinkertaistuu verrattaessa toisen ja neljännen luokan aineistoja, ja neljänneltä luokalta kuudennelle siirryttäessä kehitys hieman hidastuu, mutta kuudesluokkalaisten kirjoitelmat ovat edelleen selvästi pidempiä kuin neljäsluokkalaisten.

On mielenkiintoista tarkastella kirjoitelmia myös sukupuolten välisten erojen näkökulmasta. Tyttöjen tiedetään perinteisesti asennoituvan äidinkielen opiskeluun poikia suotuisammin, ja asenteiden vaikutus näkynee kirjoitelmien oikeakielisyydessä ja taidokkuudessa tyttöjen hyväksi. Tällaista tulosta ennustaa myös se, että sekä lyhyet että pitkät kirjoitelmat ovat aineiston tytöillä aina pidempiä kuin pojilla. Pajusen (2012: 8) mukaan alle 50 sanan kirjoitelmia on viidennellä ja kuudennella luokalla noin 15 %:lla pojista ja muutamalla prosentilla tytöistä. Erot sukupuolten välillä näkyvät myös esimerkiksi kertomusrakenteita vertailtaessa. Pajusen (2012: 13) mukaan noin kolmasosa pojista kirjoittaa vielä kuudennella luokka-asteellakin listamaisia kirjoitelmia ja vain noin 60 % kaikista kirjoittajista toteuttaa kertomusskeemaa. Kaikki edellä mainitut seikat viittaavat siihen, että poikien kirjoitelmissa saattaa esiintyä enemmän kirjoitusvirheitä kuin tyttöjen.

Kaikkien kirjoitelmien aiheena on Unelmieni/Unelmien päivä ja tekstilajitavoitteena on kertomus. Oppilaita on ohjeistettu kouluissa kirjoittamaan päivästä, jonka he ovat kokeneet tai haluaisivat kokea. Lisäksi on annettu mahdollisuus kirjoittaa unelmien päivästä maailmassa, jossa kaikki olisi mahdollista. Oppilaita ei ole ohjeistettu käyttämään mitään tiettyä kirjoitustyyliä, esimerkiksi tekstauskirjaimia tai kirjoituskirjaimia, ja etenkin nuorimmat ovatkin kirjoittaneet monenlaisilla tyyleillä. Tämä on luonnollista, sillä nuorimmat



oppilaat ovat vasta aloittelevia kirjoittajia, eikä heille ole muodostunut vielä samanlaista koulukirjoitelmissa käytettävää käsialanormia kuin vanhemmille oppilaille. Kaikissa kouluissa kirjoitelman laatimiseen on käytetty yksi oppitunti. Alakoululaisten kirjoitelmat eivät kauttaaltaan ole vielä erityisen kompleksisia. Lauserakenteet ovat niissä usein yksinkertaisia ja tekstissä on paljon toistoa. Vaikka kertomus tekstilajina on varmasti lapsille tuttu, ei suuri osa kirjoitelmista muistuta perinteistä kertomusta, vaan monesti teksteissä on vain luettelomaisesti kerrottu, mitä päivän aikana tapahtuu.

Tehtävänannon mukaan kirjoittaja on voinut kertoa jostakin jo tapahtuneesta päivästä, haaveiden unelmapäivästä tai hän on voinut keksiä jonkin täysin kuvitteellisen unelmapäivän mielikuvitusmaailmassa. Jo koettu unelmien päivä on aiheena helpoin ja täysin mielikuvituksellinen unelmapäivä haasteellisin. Vaikeimpaan vaihtoehtoon onkin tartuttu selvästi vähiten, sillä mielikuvituksellisia kirjoitelmia on aineistossa hyvin vähän. Jo koetuista unelmien päivistä kirjoittavat eniten alimpien luokka-asteiden oppilaat ja tulevista unelmapäivistä haaveilevat ylimpien luokka-asteiden oppilaat. Se, että kirjoitelmien tapahtumat sijoittuvat suurimmaksi osaksi arkisiin ympyröihin mielikuvitusmaailmojen sijasta, vaikuttaa varmasti osaltaan siihen, miten paljon arkikielisyyskirjoitelmiin on päätytty. Usein kirjoituksen aiheena on uusi lemmikki, kavereiden ja perheen kanssa puuhastelu, harrastaminen tai rahan kuluttaminen.

Tässä tutkimuksessa käytetyn Koulukorpuksen tekstiversio on siirretty ja koodattu Microsoft Excel -ohjelmaan, johon tehtiin myös tämän tutkimuksen aineiston koodaus. Kirjoitelmiin on merkitty tiedot esimerkiksi kirjoittajan koulusta, sukupuolesta ja luokka-asteesta sekä käytettyjen sanojen sanarakenteista ja sanaluokista, mikä mahdollistaa vertailun näiden muuttujien osalta. Tutkimuksen tavoitteena on kirjoitelmissa tehtyjen virheiden laskelmien lisäksi valottaa nuorten kirjoittajien käyttämää kieltä ja virheiden luonnetta myös esimerkein. Aineistosta poimituissa esimerkeissä ja taulukoissa on mukana identifointikoodi, josta selviää kirjoittajan luokka-aste ja sukupuoli (M = poika ja F = tyttö). Excel-aineistossa kirjoitelmista näkyvät vain lauseen rajat eikä esimerkiksi pilkun paikkoja tai yhdysmerkkejä ole merkitty, joten välimerkkien tutkiminen tästä aineistosta ei ole mahdollista. Näin ollen tämän tutkielman esimerkeihinkään ei ole merkkejä lisätty, eivätkä esimerkit siis erottele kirjoittajien osaamista tämän asian suhteen.

## 2. Luku- ja kirjoitustaito

Tämä tutkimus pyrkii tuottamaan tietoa alakoululaisten kirjoitustaidon kehittymisestä sanatason virheanalyysin kautta. Lapsen kirjoitustaidon kehittymiseen vaikuttavat monenlaiset seikat, ja pohja siihen luodaan jo ennen kouluikää varhaisen lapsuuden kielen oppimisen yhteydessä. Kirjoituksen perusta on puheessa, ja esimerkiksi sanelukirjoitusta voidaan pitää analyysiprosessina, jossa puhuttu kieli muutetaan kirjoitukseksi. Teknisesti kirjoittaminen on motorinen tapahtuma, jossa opitaan yhdistämään tiettyihin kuultuihin fonologisiin hahmoihin tietyt visuaaliset hahmot eli kirjaimet (Ahvenainen & Holopainen 2005: 33). Seuraavaksi luodaan tutkimukselle teoreettinen viitekehys ja tarkastellaan alakoululaisten luku- ja kirjoitustaitoon liittyviä tekijöitä tarkemmin.

### 2.1 Lapsen myöhempi kielenkehitys

Unelmieni päivä -tekstit on kerätty lapsen myöhempää kielenkehitystä tutkivaa projektia varten. Lapsenkielen tutkimuksen painopiste on perinteisesti ollut pienten, alle viisivuotiaiden lasten kielenoppimisessa. Yksilön kielenkehitys kuitenkin jatkuu läpi elämän, ja esimerkiksi kouluvuosina lapsen kielelliset taidot kehittyvät huomattavasti. Tolchinskyn (2004: 235) mukaan yli viisivuotiaan lapsen kielenkehitykseen vaikuttaa kaksi tavallaan toisilleen vastakkaista kehityssuuntaa: ensinnäkin kielenkäyttäjällä alkaa olla käsitys eri tilanteiden ja diskurssien vaatimuksista kieltä kohtaan ja hän harjaantuu käyttämään kieltä näiden vaatimusten mukaisesti, ja toisaalta kieli kehittyy entistä yksilöllisemmäksi.

Tolchinsky (2004: 236) esittää, että kumpaankin edellä mainituista kehityssuunnista – kielen sosiaaliseen ja yksilölliseen kehitykseen – liittyy sekä psykologisia että kulttuurisia tekijöitä. Kun lapsen kognitio kehittyy, hän omaksuu Tolchinskyn mukaan niin kutsutun mielen teorian eli oppii ymmärtämään ja tekemään päätelmiä toisten ihmisten mielen toimintaan liittyen. Kulttuuriset tekijät taas liittyvät ennen muuta kirjalliseen kulttuuriin: sen lisäksi, että lapsi koulutuksen piirissä oppii lukemaan ja kirjoittaman, hän tulee osalliseksi kirjallisen yhteisön eri tavoista kommunikoida. Kirjallinen kielitaito kehittyy suullisen kielitaidon rinnalle koulunkäynnin myötä.

Kirjalliseen kulttuuriimme on viime vuosikymmeninä syntynyt uusia tekstilajeja, joissa kieliopillisten sääntöjen noudattaminen ei ole yhtä säännönmukaista tai edes vaadittavaa kuin kirjoitetussa kielessä perinteisesti. Viimeaikainen tutkimus (ks. esim.

Kuorilehto 2003, Lumila 2000) on osoittanut, että puhutun kielen piirteet ovat siirtyneet yhä nuorempien oppilaiden kirjoittamiin teksteihin muun muassa IRC-keskusteluista, tekstiviesteistä ja nuortenpalstojen mielipidekirjoituksista. Entistä vapaampien kirjoitetun kielen ja puhutun kielen välimaastoon sijoittuvien tekstilajien on arveltu ja pelätty vaikuttavan negatiivisesti myös lasten oikeinkirjoitustaitoihin.

Pajusen (2012: 4) mukaan keskilapsuudessa (6–12 v.), myöhemmän kielenkehityksen aikana, lapsi oppii vähitellen käyttämään abstraktia sanastoa, kuvallisia ilmauksia, metaforia, idiomeja ja toisaalta vähäfrekvenssisia ja kompleksisia rakenteita. Lisäksi kyky hallita eri tyyllilajeja, tekstilajeja ja globaaleja rakenteita kehittyy. Keskilapsuudessa käsitteellinen ajattelu ja päättelykyky kehittyvät, ja lapsi oppii myös varioimaan ilmaisujaan. Unelmieni päivä -kirjoitelmien kirjoittajat ovat alakouluikäisiä, eli juuri Pajusen määrittelemän keskilapsuuden vaiheessa.

### 2.1.1 Kielellinen tietoisuus

Hyvä kielellinen tietoisuus tukee lukutaidon kehitystä. Kielellinen tietoisuus voidaan jakaa neljään alaryhmään: 1) kielen äännerakenteen oivaltaminen (fonologinen tietoisuus), 2) tietoisuus sanoista (morfologinen tietoisuus), 3) tietoisuus kielen säännöistä (syntaktinen tietoisuus) ja 4) tietoisuus kielen merkityksestä (pragmaattinen tietoisuus) (Takala 2006: 15).

Ponsilan (2011) mukaan kielellinen tietoisuus merkitsee kykyä ottaa kieli pohdinnan ja tarkastelun kohteeksi. Huomio kyetään suuntaamaan sanojen merkityksen sijasta niiden muotoihin ja funktioihin tai myös yksittäisten sanojen moniin merkityksiin. Kielellistä tietoisuutta voidaan tarkastella kielen kaikilla osa-alueilla: pragmatiikan, syntaksin, morfologian, semantiikan tai fonologian näkökulmasta. Tässä tutkimuksessa tutkitaan alakoululaisten kirjoitustaitoa kiinnittämällä huomiota lähinnä sanatason oikeinkirjoitusvirheisiin, ja näin ollen pohditaan kielellisen tietoisuuden ilmenemistä enimmäkseen morfologian ja fonologian näkökulmasta. Jotkin virheellisen taivutuksen virheet koskevat myös syntaktista kehittymistä. Koulukirjoitelma-aineistossa tosin myös pragmatiikka on väistämättä tutkinnanalaisena.

Fonologinen tietoisuus koostuu vaiheittain kehittyvistä ja monipuolistuvista kyvyistä kohdistaa huomio sanan semanttisesta sisällöstä sen äänteelliseen rakenteeseen, tehdä havaintoja sanan äänteellisestä rakenteesta sekä toteuttaa äännerakenteeseen kohdistuvia kielellisiä operaatioita (Ponsila 2011: 79). Kyseessä ei siis ole yksittäinen taito vaan joukko

taitoja, joiden kehittyminen alkaa jo varsin varhain ja on yhteydessä lapsen yleiseen kielelliseen kehitykseen ja sen eri osa-alueisiin. Fonologisen tietoisuuden kehittyminen liittyy läheisesti luku- ja kirjoitustaidon oppimiseen ja selittää osaltaan, miksi lukeminen ja kirjoittaminen on opittu ymmärtämään kehittyviksi kielellisiksi taidoiksi, joiden pohja luodaan jo varhaisen lapsuuden kielen oppimisen yhteydessä.

Fonologisella tietoisuudella on hyvin tärkeä rooli luku- ja kirjoitustaidon oppimisessa. Fonologiseen tietoisuuteen liittyy läheisesti myös sanatietoisuus, kyky eritellä lauseita tai useamman sanan kokonaisuuksia yksittäisiin sanoihin tai muodostaa sanoista suurempia kokonaisuuksia. Juuri sanatietoisuutta eli morfologista tietoutta tarkastellaan tässä tutkimuksessa erityisesti virhekategorioiden *puuttuva sanaväli* ja *erilleen kirjoitettu yhdysana* yhteydessä.

Ketosen (2010: 19) mukaan tietoisuus kirjoitetuista sanoista kehittyy lapsille varsin helposti, kun he tutustuvat kirjoitettuun tekstiin. Toisaalta lapset tulevat tietoisiksi puhevirrassa olevista yksittäisistä sanoista yleensä vasta silloin, kun heidän lukutaitonsa alkaa kehittyä. Käsite *sana* on monelle lapselle epäselvä, sillä määritelmät ovat heidän näkökulmastaan vaikeaselkoisia. Suomen kielen erityispiirteenä on, että sanat ovat monitavuisia ja pitkiä, koska kielessämme on runsaasti sanojen johdannaisia ja taivutuspäätteitä voi olla useita peräkkäin. Lisäksi yhdistämällä kaksi merkitykseltään erilaista sanaa yhdyssanaksi saadaan aikaan täysin uusia merkityksiä.

### 2.1.2 Lapsen äänteellinen kehitys

Niin eri äänteiden yhdistelyyn kuin äänteiden tuottamiseen eri sana-asemissa liittyy lapsilla voimakkaita rajoituksia. Puheessa ja ääntämyksessä esiintyvät ongelmat heijastuvat myös kirjoitettuun kieleen. Suomen kielessä on aivan erityisen rikas morfofonologia, ja se muodostaa selvän oppimishaasteen pienelle lapselle. Suomalaislapsen täytyy kieltä omaksuessaan oppia suuri määrä äännevaihteluita eli morfofonologisia vaihteluita (Savinainen-Makkonen & Kunnari 2004: 111). Tällaisia ongelmia suomen kielessä aiheuttavat erityisesti astevaihtelu, assibilaatio, tavuaseman muuttuminen, vokaalivaihtelut, vokaaliharmonia ja rajageminaatio<sup>1</sup>.

---

<sup>1</sup> Savinainen-Makkonen ja Kunnari käyttävät termiä *loppukahdennus*, mutta tässä tutkimuksessa puhutaan yhdenmukaisuuden vuoksi rajageminaatiosta Ison suomen kieliopin (ISK § 34) määritelmän mukaisesti.

Astevaihtelun oppiminen tapahtuu vähitellen, ja suurin osa lapsen käyttämistä sanoista on astevaihtelussa alusta saakka. Ongelmia tuottaa kuitenkin poikkeussääntöjen lisäksi muun muassa se, että useat lapsenkielen sanat ovat kokonaan astevaihtelun ulkopuolella (esim. [äitin] *äidin*), mikä hankaloittaa lapsen astevaihtelusääntöjen muodostamista. Assibilaatio on sananvartalossa tapahtuvaa sellaista *t:n* ja *s:n* vaihtelua, jossa sanavartalon *t:tä* vastaa *i:n* edellä *s*, esimerkiksi *vesi* : *vetenä* : *vesiä*. Assibilaation hallitseminen edellyttää kielioppisääntöjen soveltamista muun muassa imperfektimuotojen ja monikkomuotojen muodostamisessa, ja on siksi pienelle lapselle hyvin haasteellista. Tavuaseman muuttumisesta johtuvia vaihteluita ovat sellaiset vaihtelut, joissa sanavartalossa vaihtelevat erilaiset konsonantit ja konsonanttiyhtymät sen mukaan, ovatko ne tavun alussa vai tavun rajalla (esim. *lapsi* : *lapsen* : *lasta*, *varis* : *variksen* : *varista*). Ennen kuin tällaiset vaihtelut vakiintuvat kieleen, lapsi joutuu etsimään ja kokeilemaan monenlaisia vaihtoehtoja. Vokaaleista *i* aiheuttaa monenlaisia vaihteluita sanojen taivutuksessa. *i*-suffiksin edellä tietyt sanavartalon loppuvokaalit katoavat tai muuttuvat (esim. monikossa *kana* : *kanoissa*, imperfektissä *antaa* : *antoi*). Rungas vaihtelu voi johtaa analogian kautta virheellisiin muotoihin. Vokaaliharmonia ei juurikaan tuota ongelmia lapsen puheessa, mutta liitteiden *ko/kö*, *kaan/kään*, *han/hän* ja *pa/pä* käytössä esiintyy jonkin verran horjuntaa. Tämä johtuu ilmeisesti siitä, että liitteet saatetaan mieltää sanoista erillisiksi aineksiksi, mikä puolestaan kertoo sanatietoisuuden kehittymättömyydestä. Rajageminaatio on siitä mielenkiintoinen kielen piirre, että puheessa se on norminmukainen mutta kirjoituksessa sen merkitseminen on norminvastaista. Näin ollen lapsen puheessa rajageminaatio esiintyy usein samoin kuin aikuisen puheessa, mutta kirjoituksesta sen kitkeminen vaatii ilmiön tiedostamista, ja rajageminaation piiriin kuuluvien muotoryhmien merkitsemisessä on aikuisillakin runsaasti virheitä.

Lapsenkielelle ovat ominaisia erilaiset fonologiset prosessit. Fonologisella prosessilla tarkoitetaan lapsen luontaista tapaa muokata aikuiskielen sanoja systemaattisesti yksinkertaisemmiksi (Savinainen-Makkonen & Kunnari 2004: 102). Prosessit aiheutuvat erilaisista rajoituksista (kuten rajoituksista tuottaa vain yksikonsonanttisia ja/tai avotavuisia sanoja), joiden taustalla voi olla kypsymättömyys esimerkiksi puheen tuottoon (sekä yksittäisiin äänteisiin että erityisesti niiden yhdistelyyn) tai kuulohahmottamiseen. Erilaisia fonologisia prosesseja ovat *reduplikaatio*, *assimilaatio*, *sijaispidennys*, *omissio*, *temporaalinen jaottelu*, *substituutio*, *additio* ja *metateesi* (Savinainen-Makkonen & Kunnari 2004: 103).

Reduplikaatio määritellään usein tavun kahdennukseksi, jolloin yhtä tavua toistetaan yhden tai useamman kerran. Suomalaislapset yksinkertaistavat tyypillisesti pitkiä sanoja reduplikaation avulla (esim. [kottiti] *traktori*). Assimilaatiossa äänne samankaltaistuu osin tai kokonaan samassa sanassa esiintyvän toisen äänteen vaikutuksesta. Assimilaation alatyyppejä ovat kaukoassimilaatio (konsonanttiharmonia, esim. [papa] *napa*), lähiassimilaatio (konsonanttiyhtymän yksinkertaistuminen, esim. [tokku] *sotku*), ja vokaalien assimilaatiot (vokaaliharmonia, diftongin oikeneminen, esim. [veelä] *vielä*). Sijaispidennystä esiintyy tyypillisesti likvidayhtymissä (esim. [kuukku] *kurkku*, [kuuettaa] *kuljettaa*), joissa paikataan puuttuvaa konsonanttia tai koko konsonanttiyhtymää venyttämällä vokaalia. Omissiolla tarkoitetaan joko äänteen tai tavun katoamista lapsen tuotoksesta. Omission alatyypit ovat yksinäiskonsonanttien omissiot (esim. [una] *juna*), konsonanttiyhtymien osien omissiot (esim. [pokkana] *porkkana*) ja pitkien sanojen tavuomissiot (esim. [api] *apina*). Temporaalinen jaottelu ilmenee ilmausten epätavallisena jaksotteluna (esim. [kuo (.) mauto] *kuorma-auto*). Substituutiolla tarkoitetaan sellaista äänteen toisella äänneellä korvautumista, jota ei voida selittää assimilaatiolla. Yleensä substituution syynä on se, ettei lapsi osaa tuottaa tavoittelemaansa äännettä (esim. [hiili] *hiiri*). Additiolla tarkoitetaan äänteen lisäämistä. Äänteen lisäämistä esiintyy enimmäkseen vokaalialkuisten sanojen (esim. [poppu] *omppu*) tai hankalien vokaaliyhtymien välttämiseksi (esim. [lapilo] *lapio*). Metateesillä tarkoitetaan sitä, että äänneet vaihtavat paikkaa (esim. [havara] *harava*). Myös aivan vierekkäiset äänneet saattavat vaihtaa paikkaa erityisesti /ks/-yhtymissä (esim. [susket] *sukset*).

Vaikka fonologiset prosessit ovatkin ensisijaisesti lapsen puheessa esiintyviä ääntämistä helpottavia tekijöitä, ovat ne selvästi fonologisen prosessin myötä siirtyneet myös kirjoitettuun kieleen. Varsinkin aloitteleva kirjoittaja pyrkii tunnistamaan omasta ääntämyksestään sanassa esiintyvät äänneet, jotka hän sitten äänne äänneeltä (kirjain kirjaimelta) koodaa kirjoitukseksi. Niinpä näitä fonologisia prosesseja on mielenkiintoista peilata aineiston oikeinkirjoitusvirheisiin ja pohtia niiden yhteyttä.

## 2.2 Kirjoittamaan oppiminen

Luku- ja kirjoitustaito ovat pitkälle kehittyneitä kognitiivisia taitoja, joiden normaalille kehittymiselle lasketaan perusta varhaisissa kielenkehityksen vaiheissa (Ahvenainen ja Holopainen 2005: 32). Ahvenaisen ja Holopaisen mukaan kirjoittamista voidaan tarkastella eri näkökulmista. Psykologiselta ja neurologiselta kannalta kirjoittamista voidaan analysoida

monimutkaisena prosessina ja lingvistiikasta käsin kielen muotona. Pedagogiikan näkökulmasta taas kirjoitus on äidinkielen osa-alue, oppilaan tuotos ja opittavissa oleva taito.

Tuotoksena kirjoitettu kieli voidaan jakaa muotoon ja sisältöön, jotka eivät käytännössä ole toisiaan poissulkevia. Muoto sisältää kielen ulkoasun, joka rakentuu kielelle ominaisista graafisista kirjoitusmerkeistä. Kirjoituksen sisältö tarkoittaa kirjoittamalla tallennettua sanomaa, jonka kirjoittaja haluaa välittää lukijalle. Kirjoitustyyliä voidaan pitää sisällön osana, kirjoittajan työvälineenä, jolla tekstin sanomaa voidaan korostaa. Lukemista ja kirjoittamista aletaan Suomessa opiskella 6–7 vuoden iässä. Alussa kirjoituksen opettelussa joudutaan harjoittelemaan paljon muotoon liittyviä asioita. Sitä mukaa kun kirjoittamisen perustekniikka opitaan, kohdistuu opetus kirjoituksen sisältöön (Ahvenainen ja Holopainen 2005: 33).

Kirjoituksen perusta on puheessa. Esimerkiksi sanelukirjoitusta (jonka tuotosta erilaiset virheanalyysimallit perinteisesti arvioivat) voidaan pitää analyysiprosessina, jossa puhuttu teksti muutetaan kirjoitukseksi. Teknisesti se on motorinen tapahtuma, jossa opitaan yhdistämään tiettyihin kuultuihin fonologisiin hahmoihin tietyt visuaaliset hahmot eli kirjaimet (Ahvenainen ja Holopainen 2005: 33). Kehittyessään kirjoitus eriytyy puhekielestä. Takalan ja Vähäpassin (1983) mukaan kirjoittaminen on erillinen lingvistinen toiminta ja eroaa puhutusta kielestä sekä rakenteeltaan että toiminnaltaan. Kirjoittaminen vaatii täsmällisempää ja paremmin muotoiltua kieltä kuin keskusteleminen. Kirjoitettu kieli vapauttaa ihmisen tilannesidonnaisuudesta, koska se ei ainoastaan edusta todellisuutta, vaan auttaa myös muovaamaan sitä.

Kirjoittaminen on Ahvenaisen ja Holopaisen (2005: 67) mukaan tuottavaa ja luovaa toimintaa. Kirjoittaminen on sidoksissa kirjoittajan ominaisuuksiin, kirjoitustilanteeseen, kirjoitustehtävän luonteeseen ja ympäröivään kulttuuriin. Lukemisella ja kirjoittamisella on paljon yhteisiä piirteitä, mutta prosesseina ne ovat erilaisia. Prosessien olennainen ero on niiden nopeudessa. Sujuva lukeminen on erittäin nopeaa, sillä äännettömässä lukemisessa nopeus on noin 300 sanaa minuutissa. Tähän verrattuna kirjoittaminen on noin kaksitoista kertaa hitaampaa (Takala 1990).

Kirjoitusprosessin hitaus antaa mahdollisuuden puuttua prosessin kulkuun aivan eri tavalla kuin lukemisen kohdalla on mahdollista. Ahvenainen ja Holopainen ovat kehittäneet sanatasoisen kirjoittamisen prosessimallin (2005: 68), jossa selvennetään sanan kirjoittamisen eri vaiheita. Eri vaiheiden hahmottaminen on tämän tutkimuksen kannalta tärkeää, sillä sanatason virheanalyysissä tulee ottaa huomioon monenlaisia oppilaiden kirjoitustaitoon vaikuttavia tekijöitä. Kirjoittaminen lähtee liikkeelle joko ulkoisen ärsykkeen tai oman

ajattelun pohjalta. Tämän tutkimuksen aineiston kaltainen, luovan ainekirjoitustyyppisen kirjoittamisen tuloksena syntynyt teksti perustuu siis omaan ajatteluun. Kuten lukemisessa myös kirjoittamisessa voidaan käyttää joko suoraa ortografisiin identiteetteihin perustuvaa tai fonologista strategiaa (Ahvenainen & Holopainen 2005: 69). Suoran väylän kirjoittamisessa ortografisen tunnistamisen jälkeen voidaan kirjoitusmerkit tuottaa ilman fonologista koodausta. Tätä strategiaa käytetään tuttujen sanojen kirjoittamisessa, jolloin kirjoitusprosessi on hyvin pitkälle automatisoitunut, esimerkiksi oman nimen kirjoittamisessa.

Aloitteleva kirjoittaja, tai taitavakin kirjoittaja kirjoittaessaan outoja, pitkiä ja vaikeita sanoja, käyttää fonologista strategiaa. Siinä fonologisen ja semanttisen tunnistamisen jälkeen kirjoitettava sana on jaettava pienempiin osiin. Segmentointi voidaan tehdä sanatasollakin erilaisten kielenyksiköiden pohjalta. Mitä kehittyneempi kirjoittaja, sen suurempia yksiköitä hän pystyy kerrallaan prosessoimaan. Suomen kielessä segmentointi on mielekästä tehdä tavuittain, koska puheessa painotus noudattelee tavujakoa ja tavukokonaisuus on aloittelevankin kirjoittajan hallittavissa.

Fonologisen väylän avainprosessi on Ahvenaisen ja Holopaisen (2005: 69) mukaan äänne-erottelu. Käsiteltävänä olevasta kielellisestä kokonaisuudesta on pystyttävä erottelemaan jokainen yksittäinen äänne ja pidettävä äänteiden muodostama rakenne hallinnassa. Tämän jälkeen kukin äänne on muutettava kirjoitetuksi kieleksi, jolloin on hallittava fonologinen koodaaminen ja kirjoitusmerkkien tuottaminen. Kirjoitettavan kokonaisuuden kokoamisessa tarvitaan fonologisen synteessin taitoa, joka on prosessina samanlainen vastaavan lukemisen osaprosessin kanssa. Myös oman kirjoituksen tarkistamiseen tarvitaan fonologisen synteessin taitoa.

Kirjoittamisprosessin eri vaiheissa tarvitaan työmuistia. Muistissa on kerrallaan pystyttävä käsittelemään useita kielellisiä yksiköitä, ja yksiköt on pystyttävä pitämään muistissa melko kauan. Jos esimerkiksi kirjoitetaan suomen kielen tavuja, työmuistin kapasiteetin on oltava vähintään neljä yksikköä, sillä merkkien lukumäärä tavussa vaihtelee yhdestä neljään. Unelmieni päivä -kirjoitelmien sanojen pitkistä tavuista puuttuvat kirjaimet saattavat siis hyvinkin liittyä myös työmuistin ongelmiin; varsinkin nuorimmille oppilaille on haastavaa käsitellä useita kielellisiä yksiköitä samanaikaisesti. Koska aloittelevalle kirjoittajalle kirjoitusprosessi on erittäin aikaa vievää, voi pitkää tavua kirjoitettaessa osa kirjaimista unohtua.

Kirjoitelman laatiminen edellyttää Ahvenaisen ja Holopaisen (2005: 34) mukaan monia erilaisia kognitiivisia prosesseja kuten tilanteeseen sopivia ajattelustrategioita, tietoja, kokemuksia ja motivaatiota. Strategia tarkoittaa intuitiivista tai tietoisesti harkittua



toimintatapaa, jonka välityksellä kirjoittaja käyttää hankkimiaan tietoja ja kokemuksia kirjoitustilanteessa. Valitut strategiat aktivoivat säilöntämuistista kirjoitelman aiheeseen liittyvää ainesta. Valitut ainekset siirtyvät työmuistiin käsiteltäväksi, ja niitä muokataan ja yhdistellään ennen tekstin tuottamista. Tämänäyttötyypiset vaiheet toistuvat kirjoitustapahtumassa useita kertoja, vaihtelevassa järjestyksessä, kirjoittajasta ja kirjoitelmasta riippuen.

### 2.2.1 Oikeinkirjoitustaidon kehittyminen

Suomessa lasten lukemaan oppiminen tapahtuu verrattain nopeasti. Kaikille lukemaan oppiminen ei ole kuitenkaan helppoa. Lapsia, joilla on vaikea lukemisvaikeus (dysleksia), arvioidaan olevan Suomessa 6–10 prosenttia (Ketonen 2010: 9). Määrä vaihtelee sen mukaan, miten lukemisvaikeus määritellään. Nykykäsityksen mukaan dysleksian taustalla on usein niin sanottu fonologinen perusvaikeus, jolloin lapsella on puutteita fonologisessa tietoisuudessa, etenkin kirjain-äännevastaavuuksien automatisoitumisessa, mutta myös nopeassa sarjallisessa nimeämisessä ja fonologisessa työmuistissa. Tiedetään myös, että kirjainten nimeämisen taito ennustaa vahvasti varhaista lukutaitoa.

Ketosen (2010: 14) mukaan kirjoitustaitoa kuvaavat mallit muistuttavat lukutaitoa kuvaavia malleja, koska oikeinkirjoituksen ajatellaan myös kehittyvän vaihe vaiheelta ja liittyvän tiiviisti lukutaidon oppimiseen. Ehrin (1989) mallin ensimmäisessä vaiheessa (*esikommunikatiivinen*) lapsi harjoittelee kirjainten näköisten merkkien kirjoittamista (ns. pseudo-kirjaimet), ja vähitellen tuotokset alkavat muistuttaa oikeita kirjaimia. Seuraavassa, *osittaisfoneettisessa* vaiheessa lapsi oppii kirjainten nimiä ja äännteitä ja osaa ajoittain käyttää oikeita kirjain-äännevastaavuuksia kirjoittamissaan sanoissa. *Foneettisessa* vaiheessa lapsi hallitsee kirjain-äännevastaavuuden, ja hänen oikeinkirjoituksensa sisältää kaikki sanan äännteitä vastaavat kirjaimet. Alkava kirjoitustaito painottuu siten kirjain kirjaimelta kirjoittamiseen. Viimeisessä, *siirtymävaiheessa*, lapsi tulee tietoiseksi sanan visuaalisista piirteistä ja oivaltaa laajemmin sanojen kirjoitus- ja äänneasun yhteyksiä, ja luku- ja kirjoitustaito tulevat sujuvammiksi. Aloittelevalle kirjoittajalle asian ilmaiseminen kirjallisessa muodossa onkin aluksi haasteellista, koska häneltä vaaditaan kirjoitusmerkkien muistamista, selkeän kynäjäljen tuottamista ja oikeinkirjoitussääntöjen soveltamista.

Ehrin kirjoitustaidon kehittymistä kuvaava malli konkretisoituu Unelmieni päivä -kirjoitelmissa ensinnäkin siten, että kirjoitelmien sanamäärä kasvaa vertailtaessa eri luokka-

asteiden kirjoitelmia keskenään; kirjoitelmat pitenevät keskimäärin parillakymmenellä sanalla luokka-asteelta toiselle siirryttäessä. Kirjoitelmien sanamäärä ilmaisee paljon niiden sisällöstä ja rakenteesta, ja hyvässä kirjoitelmassa sanamäärä tuntuu olevan suurempi kuin huonossa. Sanamäärän kasvu kertonee myös kirjoitustaidon ja sujuvuuden kehittymisestä; alkava kirjoitustaito painottuu kirjain kirjaimelta kirjoittamiseen, ja näin ollen kirjoittaminen on siis kokeneemmalla kirjoittajalla sujuvampaa ja nopeampaa. Lisäksi etenkin nuorimpien oppilaiden kirjoitelmissa on huomattavan suuria pituus- ja tasoeroja, sillä lasten kirjoitustaito on vasta löytymässä.

Säännönmukaisissa kielissä kirjoitustapaa kutsutaan fonologiseksi, mikä tarkoittaa sitä, että jokaisella foneemilla on oma kirjainmerkkinsä tai merkintätapansa. Suomen kielessä jokainen kirjainmerkki äännetään aina samalla tavalla ja pitkä äänne merkitään kahdella ja lyhyt äänne yhdellä kirjainmerkillä. Oikeinkirjoituksessa käytetään siis yhtä montaa kirjainta kuin äännejärjestelmässämme on foneemeja ja lisäksi lainasanoissa käytettäviä kirjaimia. Poikkeuksena yhden kirjaimen sääntöön on *ŋ*-foneemi, joka merkitään *ng*-yhdistelmällä. Kirjoituksessa ei myöskään merkitä automaattisia sandhiääntämyksiä, kuten rajageminaatiota, esim. *tule tänne* [tulet tänne], *sinnekään* [sinnekkään], tai *n:n* assimiloitumista, esim. *kunpa* [kumpa] (ISK § 7). Koska suomen kielen oikeinkirjoituksessa voi lähes täysin hyödyntää kirjain-äännevastaavuutta, useimmat lapset oppivat oikeinkirjoituksen nopeasti.

Fonologisen kirjoitustavan ilmenemistä voidaan tarkastella Unelmieni päivä -kirjoitelmista siten, että kiinnitetään erityistä huomiota rakenteisiin, jotka eivät täysin noudata kirjain-äännevastaavuutta. Esimerkiksi juuri *ŋ*-äänteen merkitsemistä pidetään yhtenä suomen kielen ongelmarakenteena, jossa erityisesti nuorilla oppilailla ilmenee runsaasti hankaluuksia (Ahvenainen & Holopainen 2005: 107). Samoin äännepituuudet ja niiden merkitsemisessä ilmenevät ongelmat kertovat fonologisen tietoisuuden kehittymättömyydestä. Fonologinen tietoisuus on kehittyvä taito, joka etenee isompien yksiköiden, kuten sanojen ja tavujen, havaitsemisesta hienosäätöiseen äänneiden havaitsemiseen (Ketonen 2010: 19).

### 2.2.2 Kirjoittamisen haasteet

Kirjoitusvaikeuksisilla oppilailla on tavallisemmin ongelmia äänne-erottelussa. Lyhyiden tavujen äänne-erottelu ehkä hallitaan, mutta pitkiä ja vaikeita rakenteita sisältävien tavujen äänne-erottelu on vaikeaa. Tässä tutkimuksessa kiinnitetäänkin huomiota pelkkien virheiden määrien sijaan myös siihen, millaisissa sanarakenteissa ne ilmenevät. Yleisesti ongelmia

voidaan siis olettaa esiintyvän pitkissä tavuissa ja kompleksisissa sanarakenteissa, joissa äänne-erottelu on haastavaa. Usein näihin ongelmiin liittyy myös työmuistin ongelmia. Ahvenaisen ja Holopaisen (2005: 69) mukaan monien, ennen kaikkea aloittelevien kirjoittajien ongelmana on se, että he käyttävät ortografista strategiaa, vaikka eivät pysty hallitsemaan sanarakennetta yksittäisten merkkien tasolla. Usein syynä ortografisen väylän käyttämiseen ovat fonologisen väylän osaprosessin ongelmat. Sanan tunnistava lukeminen ei takaa sitä, että sana osataan kirjoittaa virheettömästi.

Kun oppilas harjoittelun jälkeen oppii äänne-kirjainyhdistämisen, kahden äänten yhteen liu'uttamisesta tulee usein todellinen haaste (Thuneberg 2006: 185). Liu'uttamisen vaikeudesta oppilaiden kirjoitelmissa kertovat erilaiset puuttuvien kirjainten virheet, erityisesti konsonanttiyhtymistä ja diftongeista. Näin ollen voidaan olettaa, että oppilaalla, jonka kirjoitelma sisältää paljon juuri virheitä konsonanttiyhtymissä ja diftongeissa, voi olla jonkinasteisia ongelmia äänteiden dekodauksissa. Alkuopetuksen aikana kehityserot ovat kovin suuret yleensäkin, kuten myös luku- ja kirjoitustaidon suhteen, joten on usein vaikea arvioida, onko hitaassa lukemaan oppimisessa kyse kehityksen varianssista, joka tasoittuu parin ensimmäisen vuoden aikana, vai ongelmasta, johon olisi mahdollisimman pian puututtava (Thuneberg, 2006: 178).

Alaluokkien suurimmat ongelmat ovat useimmiten teknisen lukemisen sujumisessa, mutta myöhemmin yksittäisiin sanoihin, erilaisiin teksteihin tai oppiaineisiin liittyvät ymmärtämisen ongelmat näyttäytyvät päällimmäisinä ongelmina. Useimmiten lisänä on tekstin tuottamisen pulmaa, uusien sanojen oppimisen vaikeutta sekä kuullun ymmärtämisen vaikeutta vieraisissa kielissä. Yleisesti tiedetään, että oikeinkirjoitusvirheet ovat todellinen ja näkyvä ongelma lukivaikeuksiselle henkilölle, sillä lukiessa on helpompi käyttää kompensatorisia strategioita, jolloin yksittäiset lukemisvirheet eivät välttämättä häiritse lukemista. (Holopainen & Savolainen, 2006: 206.)

Lukemisen ja kirjoittamisen vaikeuksien kenties merkittävimpinä taustatekijöinä ovat fonologiset ongelmat. Fonologisen, äänteellisen aineksen, prosessoinnin vaikeuksien tai fonologisen tietoisuuden ongelmat voivat olla lukemisen ja kirjoittamisen taitojen saavutuksen esteenä. Fonologinen tietoisuus sisältää herkkyyden havainnoida, käsitellä ja tietoisesti muokata puheäänteitä, ja sitä tarvitaan lukemisessa ja kirjoittamisessa yhdistämään puhuttu ja kirjoitettu kieli. (Holopainen & Savolainen 2006: 207.)

Lukivaikeuksia ryhmitellään tavallisimmin joko syytaustan tai oireiden perusteella. Lukivaikeudet voidaan syytaustan mukaan jakaa kehityksellisiin ja hankittuihin lukivaikeuksiin. Kehityksellisten lukivaikeuksien taustalla on jokin synnynnäinen

rakenteellinen poikkeavuus, jonka seurauksena kirjoitetun kielen taitojen oppimisessa on vaikeuksia. Vaikeuksilla on yleensä neurologinen syytausta, johon liittyy perinnöllisiä tekijöitä. Luku- ja kirjoitustaidon ja muun suoritustason välillä on selvä ero. Vaikeuksia on erityisesti fonologisten koodaustaitojen omaksumisessa.

Virheanalyysissa voidaan eri virhetyyppien esiintymien perusteella pyrkiä määrittelemään dysleksian erilaisia tyyppisiä. Tämän ajattelun taustalla on oletus, että tietäntyyppiset dyslektikot tekevät tietäntyyppisiä virheitä lukiessaan ja kirjoittaessaan (Ahvenainen & Holopainen 2005: 88). Tällaisen virheluokituksen on laatinut Bakker (1978). Bakkerin mukaan tietyt virhetyypit heijastelevat lukemisvaikeuden taustalla olevia aivotoiminnallisia häiriöitä. Lukemisvirheet voidaan jakaa suoriin ja aikaavieviin virheisiin. Suorat virheet ovat todellisia virheitä, jolloin luetun lopputulos on väärä. Aikaavieviksi virheiksi luokitellaan lähinnä katkot ja toistot, jotka johtavat hitaaseen lukemiseen, mutta oikeaan lopputulokseen. Ahvenaisen ja Holopaisen esittämää lukemisvirheiden ryhmittelyä voidaan soveltaa myös oikeinkirjoitusvirheiden tarkastelussa, sillä kirjoitusvirheitä voidaan luokitella myös todellisiin virheisiin ja satunnaisiin lipsahduksiin.

Ahvenaisen ja Holopaisen (2005: 107) mukaan kieleemme kuuluvia sanoja samoin kuin lainasanoja voidaan ryhmitellä niiden rakenteen mukaisesti helppoihin ja vaikeisiin sanoihin. Lukivaikeuksisten oppilaiden ongelmat näkyvät juuri vaikeiden rakenteiden hallinnassa. Sanastotasolla voidaan suomen kielen ongelmarakenteet Ahvenaisen ja Holopaisen mallin mukaan ryhmitellä seuraavasti: pitkä vokaali ja diftongi (erityisesti reversoituvat yhdistelmät ei/ie, uo/ou), tavunloppuinen -h-, geminaatta (huomioitava erityisesti klusiiligeminaatat kk, pp, tt), konsonanttikasauamat (erityisesti reversoituvat yhdistelmät ts/st, sk/ks), η-äänne, vierasperäiset konsonantit ja rakenteet sekä kasautuvat rakenteet. Tässä tutkimuksessa pyritään vertaamaan Koulukorpuksessa esiintyviä oikeinkirjoitusvirheitä Ahvenaisen ja Holopaisen suomen kielen ongelmarakenteet -ryhmittelyyn. Heidän luokitteluunsa tukeutuen voidaankin olettaa, että alakoululaisilla esiintyy hankaluuksia erityisesti kasautuvissa rakenteissa ja pitkissä tavuissa. Näitä ongelmia selittänevät yhtäältä ongelmat fonologisessa koodaamisessa ja äänne-erottelussa ja toisaalta ongelmat työmuistin käytössä kirjoitusprosessin yhteydessä.

Johtamisen ohella sanojen yhdistäminen eli kompositio on toinen keskeinen sananmuodostuskeino suomen kielessä. Yhdyssana on kokonaisuus, joka koostuu kahdesta tai useammasta sanasta mutta on kielen yksikkönä yksi sana. Uusien merkitysten tuottaminen sanoja toisiinsa liittämällä ja prosessin ymmärtäminen vaativat sanatajun ja lausetajun kehittymistä sekä kielioppisääntöjen soveltamista. Sen lisäksi, että sanat kirjoitetaan yhteen,

jos ne muodostavat yhdessä tiiviin ja vakiintuneen kokonaisuuden, suomen kielessä on joukko varsinaisia sanojen yhteen kirjoittamista ohjaavia sääntöjä. Näitä sääntöjä opetellaan läpi peruskoulun, mutta alakoulun ensimmäisillä luokilla niihin ei juurikaan kiinnitetä huomiota, vaan yhdyssanoja pyritään opettamaan ensisijaisesti merkityksen kautta. Niinpä myös sanojen yhteen kirjoittaminen on eräänlainen suomen kielen erityisrakenne, jossa ilmenee runsaasti vaihtelua.

### 2.3 Virheanalyysistä

Virheanalyysillä tarkoitetaan kielellisen tietämättömyyden tutkimusta. Virheanalyysi on siis järjestelmällistä virheiden tarkastelua ja lähinnä virheitä aiheuttavien seikkojen tunnistamista. Virhetyyppejä voidaan tarkastella muun muassa kieliopillisesta tai hyväksyttävyyden näkökulmasta sekä puhuttuun kieleen verraten. Kielenoppiminen nähdään virheanalyysissä jatkumona, jossa virheet ovat oppimisen tietyssä vaiheessa luonnollinen osa. Vaikka virheiden määrä ja laatu eivät kuvaa suoraan oppilaan kielitaitoa, ne antavat tietoa oppilaan taidon luonteesta. Tutkimalla virheitä voidaan päätellä oppilaan senhetkinen tieto- ja taitotaso. Kuvaamalla ja luokittelemalla virheet on mahdollisuus tuoda esille ne kielen piirteet, joissa oppilaalla on oppimisongelmia. (Stenberg 2010: 11).

Virheanalyysin ensimmäisessä vaiheessa tunnistetaan mahdolliset virheet. Niistä pyritään erottamaan ensinnäkin kompetenssivirheet eli varsinaiset virheet, joissa ovat kyseessä sekä kielen järjestelmään liittyvät ongelmat ja virheet. Toiseksi kompetenssivirheistä pyritään erottamaan virheet, jotka ovat satunnaisia lipsahduksia kielenkäyttäjän tuotoksissa. Varsinaiset virheet ovat sellaisia kielen järjestelmään kuuluvia asioita ja käsityksiä, joita ei ole vielä opetettu, ja lipsahdukset virheitä, jotka oppija pystyy helposti korjaamaan hänellä jo olemassa olevan tiedon perusteella. Tunnistamisen jälkeen virheet luokitellaan, ja niitä pyritään selittämään.

Virheiden luokituksella eli virheanalyysillä tarkoitetaan myös oikeinkirjoitusvaiheessa tehtyjen virheiden jakamista kategorioihin. Virheiden luokituksessa on perinteisesti noudatettu kahta peruslinjaa: fenomenologis-deskriptiivistä ja funktionaalisi-etnologista virheanalyysia. Fenomenologis-deskriptiivisen virheanalyysin luokitus tapahtuu tehtyjen virheiden pohjalta. Tällainen luokittelu on selväpiirteinen eikä riipu psykologisista teorioista. Funktionaalisi-etnologisen virheanalyysin luokittelussa taas pyritään ottamaan huomioon se psyykinen funktio, jonka häiriöstä virheen katsotaan johtuvan. (Ruoppila ym. 1969: 7.)

Ruoppilan, Römanin ja Västin (1969) tarkoituksena oli luoda sellainen diagnostinen testi, joka sisältää virheiden määrän lisäksi tietoa virheiden luonteesta. Virheiden luokituksella eli virheanalyysillä Ruoppila ym. tarkoittivat oikeinkirjoituksessa tehtyjen virheiden jakamista kategorioihin eli virhetyyppeihin. He nimesivät luokituksen fenomenologis-deskriptiiviseksi virheanalyysiksi, jossa luokitus tapahtuu esiintyneiden virheiden pohjalta. Ruoppilan tutkimusryhmä piti luokittelua selväpiirteisenä ja psykologisista teorioista riippumattomana menetelmänä. Käytännön opetustyössä menetelmä onkin erotellut heikoimpia kirjoittajia ja analyysin tuloksia käytetään oppilaiden tuki- ja erityisopetuksen tarpeen määrittelyssä sekä korjaavan opetuksen suunnittelussa. Opettajille on tämän perinteisen arviointitavan ja kokemuksen perusteella muodostunut tuntuma niistä sanarakenteen piirteistä, jotka ovat haasteellisia alkavalle kirjoittajalle ja lukivaikeuksiselle opiskelijalle. Tällaisia piirteitä ovat esimerkiksi geminaatan, pitkän vokaalin ja η-äänteen kirjoittaminen.

Ruoppilan, Römanin ja Västin (1969) laatima diagnostinen testi on ollut Suomessa ensimmäinen opettajien käyttöön standardoitu testi, ja sitä on käytetty suomalaisissa peruskouluissa oppilaiden kirjoitustaitojen arvioinnissa jo pitkään. Virheanalyysia käytetään apuna opetustyössä, sillä testien avulla on mahdollista tarkastella oppilaiden kirjoittamistaitojen kehittymistä ja määrittellä tarvittavia tuki- ja erityisopetuksia (Kulju & Mäkinen 2009: 508–510). Virheanalyysia käytetään siis apuna seulontatyyppisessä analysoinnissa, jonka tarkoituksena on löytää ne oppilaat, joilla on lukemisen ja kirjoittamisen oppimisvaikeuksia. Tämän lisäksi arvioinnin tavoitteena on selvittää oppimisvaikeuksien taustalla olevia syitä ja antaa tietoa oppilaiden taitojen edistymisestä. (Ahvenainen & Holopainen 2005: 84)

Luku- ja kirjoitustaidoista tehty arviointi voi olla formaalista tai informaalista, ja virheanalyysiin perustuva kielitaidon arviointi on formaalista arviointia. Formaalista arviointia ovat systemaattiseen tiedon hankintaan perustuvat normeerattujen mittareiden ja testien käyttö ja informaalista arviointia esimerkiksi opettajan oma havainnointi ja sen kautta tehdyt arviot sekä erilaiset vapaamuotoiset mittarit (Ahvenainen & Holopainen 2005: 85). Oppilaan lukemis- ja kirjoittamistuotoksesta tarkasteltuihin virheisiin ja nopeuteen perustuva arviointi ymmärretään oiretason arvioinniksi. Tarkoituksena on päätellä lukemis- tai kirjoittamistuotoksessa esiintyvien virheiden perusteella, millainen ongelma lukemis- tai kirjoittamistapahtumassa on kyseessä (Ahvenainen & Holopainen 2005: 86). Oppilaiden kirjoittamistaitoja arvioitaessa perustana voidaan käyttää jäljentämiskirjoitusta, sanelukirjoitusta, toistamiskirjoitusta tai luovaa kirjoitusta. Kirjoittamisen perusprosessien

hallinnan selvittämisessä Ahvenainen ja Holopainen (2005: 91) pitävät käyttökelpoisimpana sanelukirjoitusta, jonka avulla voidaan systemaattisesti kontrolloida vaikeiden kielellisten rakenteiden hallintaa. Saneluun perustuvia testimalleja, sanalistoja tai kertomuksia, onkin laadittu eri vuosiluokille lukuisia.

Virheanalyysin etuna on se, että sen avulla on helppo verrata oppilaan suoriutumista samanikäisten suoriutumiseen ja tämän perusteella selvittää perustaitojen tasoa ja seuloa oppilaita lukiopetukseen. Toisaalta taas pelkkien virheiden perusteella voi olla toisinaan vaikea päätellä, miksi lukijan tai kirjoittajan tuotos ylipäänsä on vääränlainen, ja virheanalyysia onkin moitittu siitä, että virheiden lukumäärä ja tyypillisten virheiden kartoittaminen eivät kuvaa kirjoittamisvaikeuksien laatua riittävän tarkasti.

Vaikka Ruoppilan ym. virheanalyysiin pohjautuvilla testeillä on edelleen vankka asema kirjoitustaidon arvioinnissa, sen antamaa tietoa ovat kriittisesti arvioineet Kulju ja Mäkinen artikkelissaan (2009). Heidän mukaansa kirjoittamisprosessi nähdään usein turhankin yksinkertaisesti »lineaarisen» foneemisen esityksen koodaamisena kirjaimiksi johtuen ehkä suomen kielen selkeästä kirjain-äännevastaavuudesta. Perinteisen virheanalyysin rinnalle he ovatkin kehittäneet arviointimenetelmän, jossa sanarakenteiden hallintaa tarkastellaan sanan hierarkkisen kokonaisrakenteen näkökulmasta (jatkossa sanarakenneanalyysi). Sanarakenneanalyysi on luonteeltaan fonologinen ja perustuu tavoitesanan fonologiseen kompleksisuuteen siten, että sanarakenteita tarkastellaan tavumäärän, tavorakenteen, fonotaksin ja yksittäisten kirjainten tasolta. Tällä pyritään selvittämään sitä, mitkä fonologiset piirteet ovat aloittelevalla kirjoittajalle ongelmallisia. Analyysissä on siten keskeistä kirjoittamisprosessiin liittyvä sanan rakenteen hahmottaminen ja sen koodaaminen kirjoitukseksi. Koska sanarakenneanalyysissä on tietoisesti keskitytty ainoastaan sanan fonologiseen rakenteeseen, erikseen ei määritelty virhetyyppejä, jotka voivat olla esimerkiksi teknisiä kuten kirjaimen kääntyminen (rotaatio), sosiolingvistisiä (lapsi kirjoittaa tuntemansa variantin) tai foneettisia (esim. *vormula* »formula»). Tämä ratkaisu poikkeaa virheanalyysistä, jossa virheluokkien määrä on suuri. Suuri virheluokkien määrä voi toisaalta johtaa siihen, että kadotetaan kokonaiskuva lapsen taidosta hahmottaa sanarakenne. (Kulju & Mäkinen 2009: 508–509.)

Joistakin ongelmista ja heikkouksista huolimatta virheanalyysi on käyttökelpoinen ja vuosikymmeniä käytössä ollut analyysimalli. Se on pätevä oikeinkirjoituksen tasoa mittaava diagnostinen testi, joka sisältää tietoa paitsi virheiden määrästä myös niiden luonteesta. Koska tässä tutkimuksessa tarkoituksena on nimenomaan selvittää alakoululaisten tekemien kirjoitusvirheiden määrää ja virheiden luonnetta, virheanalyysi soveltuu hyvin sen tarpeisiin.

Analyysi pohjautuu ensisijaisesti Ruoppilan ym. (1969) virheanalyysiin mutta mukailee Kuljun ja Mäkisen (2009) sanarakenneanalyysia niin, että tässä tutkimuksessa pyritään myös osin selvittämään virheiden sana-asemia ja näin selittämään virheitä. Tässä tutkimuksessa käytetty virheanalyysi on siten fenomenologis-deskriptiivinen, ja keskittyy sekä virheiden määrään että niiden luonteeseen.

### 2.3.1 Virheen käsite

Kieltä opetteleva henkilö ei pysty noudattamaan sellaisia lingvistisiä sääntöjä, joita hänelle ei ole opetettu, eikä näin ollen pysty korjaamaan mahdollista virhettä. Tällöin oikean lingvistisen säännön noudattamisessa voi olla kyse sattumasta, arvauksesta tai päättelystä. Virheet voidaan jakaa ennustettaviin, järjestelmällisiin virheisiin ja ennustamattomiin, umpimähkäisiin virheisiin (Stenberg 2010: 12). Oppijalla on kielenoppimisessa erilaisia välivaiheita, välikieliä, jolloin hän tekee järjestelmällisiä virheitä. Sitten kun asia on hänelle opetettu ja hän oppii sen, virheen tulisi poistua. Harjoitteluvaiheessa oppija voi olla epäjohdonmukainen virheissään, ja toisinaan hän suoriutuu jostakin tehtävästä oikein ja toisinaan väärin. Sattumasta on kyse silloin, kun oppilas vastaa useimmiten väärin ja vain harvoin osuu oikeaan muotoon. Tällöin voitaisiin olettaa, että oppilas tietää vain epämääräisesti, mitä hänellä on opittavanaan. (Corder 1976: 264, 275.)

Koulukirjoittamisprojektissa sovelletaan käsitteitä normipoikkeama ja tyyli-poikkeama. Mari Hongon (tulossa) mukaan normipoikkeamia edustavat selkeimmin oikeinkirjoituspoikkeamat ja tyyli-poikkeamia leksikaaliset erikoisuudet, kuten vieraskieliset sanat ja näiden mukautukset sekä leimallisesti puhekieliset ja lapsenkieleen kuuluvat sanat. Normi- ja tyyli-poikkeamat saattavat kuitenkin limittyä sekä sanamuodon vuoksi että lähdetulkinnan kannalta.

Normipoikkeamien ja tyyli-poikkeamien kirjo on Hongon mukaan lasten kirjoittamissa teksteissä moninainen. Toisinaan poikkeamat ovat viite huolimattomuudesta tai puutteellisesta taidosta, toisinaan viite kehittyvästä kielitaidosta ja riskinotosta. Mukavuusalueelta kurkistelua, riskinottoa, ja sen yhteyttä normipoikkeamiin ei voida todentaa pelkkää kirjallista tuotosta tarkastelemalla aukottomasti, mutta kylläkin suuntaantavasti ilmausten sofistikoitumisen kautta. Normipoikkeamia, tyyli-poikkeamia ja sofistikoituneisuutta (tässä laajasti käsittäen niin pituussujuvuuden, kompleksisuuden kuin frekvenssin) vertailemalla voidaan luoda yleiskuva kirjoittamisen tasosta, ehkä jopa saada



viitteitä kirjoittamistaidosta. Honko tutkii myös koululaistekstejä ja hänen tavoitteenaan on saada laadullista tietoa koululaisten kirjoittamiskäyttäytymisestä puutteineen, ja yhdistettynä sanastonhallintaa mittaavien tehtävien tuloksiin osoittaa a) keinoja tunnistaa keskimääräiset resurssit ja b) alueet, joilla yksittäisillä oppilailla tai oppilasryhmillä on puutteita tai nollaedustusta sekä c) keinoja suunnata opetusta ja rakentaa harjoittelumateriaalia, ehkä - tilanteitakin.

Honko muistuttaa, että etenkin nuorten koululaisten kohdalla on olennaista erottaa puhemodaliteetin aiheuttamat normipoikkeamat muista poikkeamista. Toisinaan myös esimerkiksi äänteelliset normipoikkeamat voidaan tulkita joko puhekielisyydeksi (aluerekisteri) tai oikeinkirjoitusvirheeksi (väärät, ylimääräiset tai puuttuvat äänteet: *viälä, miäs*). Yleensä nämä voidaan Hongon mukaan kokonaistekstikontekstin avulla sekä kirjoittajan identiteettitietojen perusteella luokitella varsin pitkälle. Jos poikkeamat eivät selity asuinpaikkakunnan aluepuhekielen piirteillä tai poikkeamat eivät ole puhujalle luonteenomaisia (toistuvuus, josta voitaisiin arvella esim. paikkakunnalle muutto toisenlaisen aluevariaation vaikutuspiiristä), ne on perusteltua tulkita niin sanotuiksi normipoikkeamiksi, eli oikeinkirjoituksesta, puutteellisesta kielitaidosta tai rakenteen hallinnan keskeneräisyydestä johtuviksi virheiksi. Toisaalta poikkeamien analysointi ei usein onnistu myöskään ilman teksti(laji)kontekstin tarkistamista, sillä kertomusrakenne mahdollistaa monenlaisten varianttien käytön dialogisuuden kautta, ja myös esimerkiksi lyriikan keinojen käyttämisen kautta.

### 2.3.2 Puhekielisyyden ja virheen suhde

Tässä tutkimuksessa alakoululaisten tekemiä virheitä tarkastellaan pääosin Hongon luokittelun mukaisina normipoikkeamina, vaikka selvää rajaa tyyli- ja normipoikkeaman välille ei voidakaan vetää (ks. luku 2.3.1). Virheanalyysia tehtäessä tulee siis jokaisen virheen kohdalla ottaa huomioon virheen taustalla vaikuttavat tekijät. Pitää varoa tuijottamasta tiukasti vain virheen muotoa, sillä muodollisesti samalta näyttävät virheet voivat selittyä eri tekijöillä. Jotkin virheet ovat muodollisesti erityisen hankalia nimetä tyyli- tai normipoikkeamaksi, koska ne saattavat selittyä myös puhekielen ilmiöillä. Tästä syystä tämän tutkimuksen analyysissa pyritään varsinaisista kirjoitusvirheistä erottamaan ainakin yleisimmät puhekielisyyden piirteet vertaamalla poikkeamia Mantilan (1997) jaottelemiin

puhekielen piirteisiin ja Salosen (2012) tutkimuksen tuloksiin alakoululaisten kirjoitelmissaan käyttämistä puhekielen ilmauksista.

Mantila (1997: 11–12) hahmottelee nykypuhe-suomen piirteet viiteen ryhmään: yleiset ja neutraalit puhekielen piirteet, laaja-alaiset puhekielisyydet, yleistyvät puhekielisyydet, leimautuneet ja kartettavat aluepuhekielisyydet sekä elävät paikallisuudet. Luokittelun perusteina on käytetty piirteen alueellista edustusta ja nykykielisen variaation ominaispiirteitä. Unelmieni päivä -kirjoitelmat sisältävät jonkin verran kaikkia Mantilan nykypuhe-suomen piirteitä. Tämä tutkimus keskittyy varsinaisten oikeinkirjoitusvirheiden löytämiseen ja ryhmittelyyn, mutta mukana on muotoryhmiä, jotka muodollisesti voidaan tulkita kuuluviksi johonkin tutkimuksen virhekategoriiaan mutta jotka selvästi selittyvät myös puhekielisyyksiksi. Mantilan jaottelun mukaiset yleiset ja neutraalit sekä laaja-alaiset ja yleistyvät puhekielisyydet on rajattu tämän tutkimuksen virhekategorioiden ulkopuolelle. Sen sijaan mukaan on laskettu joitain alueellisesti leimautuneita puhekielen piirteitä, mutta ne on analyysissä pyritty erittelemään varsinaisista oikeinkirjoitusvirheistä. Sellaisia puhekielen piirteitä, joita ei Mantilan ryhmittelystä löydy, verrataan Mielikäisen (1991) selvitykseen nykypuhekielen piirteistä ja levinneisyydestä.

Tämä pro gradu -tutkielma on eräänlainen rinnakkaistutkimus Salosen pro gradu -tutkielmalle (2012) *Puhekielen piirteiden esiintyminen alakoululaisten kirjoitelmissa*. Salonen on tutkinut niin ikään koulukirjoitelmakorpuksesta alakoululaisten kirjoitelmissaan käyttämiä puhekielen ilmauksia. Salonen on ryhmitellyt käsittelemänsä puhekielen piirteet seuraavasti: äänne- ja muoto-opilliset puhekielisyydet (yksikön 1. ja 2. persoonan pikapuhemuodot nominatiivissa ja obliikvisijoissa, yksikön ja monikon 3. persoonan puhekieliset muodot, demonstratiivipronominien pikapuhemuodot, *olla-*, *tulla-* ja *mennä-*verbien pikapuhemuodot, *A:n* loppuheitto, *i:n* loppuheitto, jälkitavujen *i*-loppuisten diftongien jälkikomponenttien kato, sananloppuisen *n:n* kato, sananloppuisen *t:n* kato ja jälkitavujen *A*-loppuisten vokaaliyhtymien monoftongiutumisen), syntaktiset puhekielen piirteet (refleksiiviset ja ei-refleksiiviset omistuksen ilmaukset, passiivimuoto monikon 1. persoonan funktiossa, 3. infinitiivin illatiivin esiintyminen ilman *mA*-tunnusta ja monikon taivutuspäätteen puuttuminen) ja sanastolliset puhekielen piirteet (puhekieliset sanat ja slangi sekä interjektiot).

Äänne- ja muoto-opillisia puhekielen piirteitä esiintyy Salosen aineistossa melko vähän, kun taas passiivimuodon käyttö monikon 1. persoonan funktiossa on hänen aineistossaan laajimmin näkyvä puhekielen ilmiö ja sen puhekielisyydysprosentti (28 % tapauksista) on huomattavan suuri (Salonen 2012: 83, 115). Salosen tutkimuksen

alakouluikäiset kirjoittajat pyrkivät selvästi tuottamaan norminmukaista ja hyvää kirjoitelmien kieltä. Kirjoitelmien sanasto on yleiskielen mukaista ja poikkeamat siitä ovat pääosin jo hyvin yleisessä käytössä olevia arkisen puhekielen sanoja. Alemmilla luokka-asteilla kokemattomuus kirjoittamisessa ja kirjakielen tuottamisessa näkyy vielä voimakkaana, mutta ylemmille luokka-asteille siirryttäessä puhekielen piirteiden osuus vähenee ja kertoo selvästi siitä, että oppilaat alkavat sekä hallita norminmukaisen kirjoittamisen että tietoisesti pyrkiä siihen koulukontekstissa (Salonen 2012: 109). Puhekielen piirteiden runsaus on siis lähtökohdiltankin erilaista koulukontekstissa kuin tekstilajeissa, joissa puhekielisyyksiä esiintyy tavallisesti runsaasti.

Kaikki edellä mainitut Salosen tutkimat piirteet on jo luokiteltu puhekielen piirteiksi, mutta selvärajaista jaottelua varsinaisiin kirjoittaessa tehtyihin virheisiin ja puhekielen piirteisiin on mahdotonta tehdä. Varsinkin sananloppuisten kirjainten kadon luokittelu on erityisen ongelmallista, vaikka informantin kielellinen tausta tunnettaisiinkin. Aineiston analyysi tulee siis sisältämään jonkin verran päällekkäisyyksiä Salosen analyysin kanssa. Lisäksi käsitellään joitain sellaisia selvästi alueellisia puhekielisyyksiä, jotka menevät limittäin tähän tutkimukseen valittujen virhekategorioiden kanssa ja jotka Salonen on jättänyt oman tutkimuksensa ulkopuolelle. Salosen tutkimuksen tulosten perusteella voidaan olettaa, että alakoululaiset pyrkivät koulukirjoitelmissaan tuottamaan myöskin kieliopillisesti mahdollisimman virheetöntä tekstiä, joskin suomen kielen ongelmarakenteet tuottanevat vaikeuksia etenkin aineiston nuorimmille oppilaille ja näin ollen suurin osa poikkeamista lienee tahattomia normipoikkeamia.

### 3. Virheanalyysi Unelmieni päivä -kirjoitelmista

Tässä osassa tarkastellaan aineistosta poimittuja oikeinkirjoitusvirheitä kategoriakohtaisesti. Ensin esitellään virheiden ryhmittelyä ylä- ja alakategorioihin ja sen jälkeen virheanalyysia toteutetaan virhekatgoria kerrallaan. Virhekatgoriakohtaisia esiintymiä tarkastellaan luokkakohtaisesti ja sukupuolikohtaisesti, ja näin pyritään luomaan kuva siitä, mitä virheet kertovat alakoululaisten kirjoitustaidon kehitymisestä kyseisiin muuttujiin suhteutettuina. Lisäksi jokaisen virheen kohdalla on pyritty selittämään mahdollisia virheen taustalla olevia tekijöitä ja pohtimaan niiden yhteyttä suomen kielen ongelmarakenteisiin ja lapsen äänteellisen kehityksen ongelmiin.

#### 3.1 Virheiden ryhmittely

Tämän tutkimuksen virheiden ryhmittelyssä on lähdetty liikkeelle Ruoppilan, Römanin ja Västin (1969) fenomenologis-deskriptiivisestä virheanalyysista, eli luokitus on tehty esiintyneiden virheiden pohjalta. Virheryhmien nimeämisessä on kuitenkin otettu mallia myös Kuljun ja Mäkisen (2009: 513) kehittämästä sanarakenneanalyysista, sillä heidän luokituksessaan kiinnitetään huomiota erityisesti virheiden rakenteeseen. Kuljun ja Mäkisen analyysiin otettiin mukaan seuraavat 17 virhetyyppiä (kursiivilla on esitetty esimerkkejä Kuljun ja Mäkisen aineistosta virhetyyppien tunnistamisen helpottamiseksi): puuttuva kirjain geminaatasta (*apelsiini* »appelsiini»), puuttuva kirjain pitkästä vokaalista (*ratat* »rattaat»), muu puuttuva kirjain (*rendi* »trendi»), väärä kirjain (*gormula* »formula»), väärä geminaatta (*onkkimaan* »onkimaan»), väärä pitkä vokaali (*hyypäämisillään* »hyppäämisillään»), η-äänteen virhe (*ogkimaan* »onkimaan»), m/n-sekaannus (*moin* »noin»), rotaatio (*käärmeiben* »käärmeiden»), reversaali (*terndi* »trendi», *jouksee* »juoksee»), liika kirjain (*kiujas* »kiuas»), lisämerkkien virheet (*pöyta* »pöytä»), väärä sananloppu (*viidakkoveitsiini* »viidakkoveitsi»), puuttuva tavu (*säikäisi* »säikähtäisi»), väärä mielekäs sana (*pipo* »villapipoihini»), epämielekäs sana tai sanaraunio (*kupdala* »kaupungilla») ja puuttuva sana.

Ruoppilan tutkimusryhmän virheanalyysia mukaillen on tässä tutkimuksessa ryhmitelty poimitut virheet alakategorioihin niiden tyyppin mukaan. Näin syntyneet yläkategoriat ovat virheet sanarajoissa, puuttuvat kirjaimet, väärät kirjaimet, ylimääräiset kirjaimet ja sanan taivuttamiseen liittyvät virheet. Näihin kategorioihin valikoitui kaiken kaikkiaan 13 virhetyyppiä seuraavasti: Virheet sanarajoissa ovat yhdyssanan osien

kirjoittaminen erilleen (*sukellus vene* »sukellusvene») ja sanojen aiheeton yhteen kirjoittaminen (*veiverkot* »vei verkot»). Puuttuvien kirjainten virheisiin lukeutuvat puuttuva kirjain geminaatasta (*päätetiin* »päätettiin»), puuttuva kirjain pitkästä vokaalista (*huonesen* »huoneeseen») ja muut puuttuvat kirjaimet (*suikussa* »suihkussa»). Väärä kirjain -kategoriaan lukeutuvat väärät vokaalit (*leikittäisiin* »leikittäisiin», *jäätölöä* »jäätelöä») ja väärät konsonantit (*uinaan* »uimaan», *sänkyn* »sängyn»). Ylimääräinen kirjain -kategoriaan kuuluvat konsonantin geminoituminen (*mennen* »menen»), vokaalin pidentyminen (*makaroonia* »makaronia») ja muut ylimääräiset kirjaimet (*ouldolla* »oudolla»). Taivuttamiseen liittyviä virheitä ovat virheellinen taivutus (*asuisi* »asuisivat»), virhe symbolissa (2 »2:een/kahteen») sekä taivuttamisen yhteydessä käsiteltävät monitulkintaiset virheet, joita ei voida yksiselitteisesti ryhmitellä tiettyyn virhekkategoriaan kuuluviksi.

Tämän tutkimuksen virhekkategoriat eroavat siis jonkin verran Kuljun ja Mäkisen virheanalyysistä. Ensinnäkin tähän tutkimukseen on otettu mukaan Ruoppilan ym. (1969: 9) alkuperäisestä virheanalyysistä virhekkategoriat yhdyssanan osat erilleen ja sanat aiheettomasti yhteen, jotka taas Kulju ja Mäkinen jättivät omasta virheluokituksestaan pois. η-äänteen virhettä ei ole tässä tutkimuksessa eroteltu omaksi virheryhmäkseen, vaan se käsitellään sen mukaan, onko kyseessä esimerkiksi väärän tai puuttuvan konsonantin virhe. Samoin m/n-sekaannus käsitellään väärä konsonantti -virhekkategoriassa ja lisämerkkien virheet väärän vokaalin tai konsonantin yhteydessä sen mukaan, kummassa ne esiintyvät. Tämän tutkimuksen aineistossa varsinaisten puuttuvien tavujen määrä on niin vähäinen, että ne käsitellään muiden puuttuvien kirjainten yhteydessä. Kuljun ja Mäkisen väärä sananloppu -virhe lukeutuu tämän tutkimuksen virheellinen taivutus -kategoriaan, joka käsittelee taivuttamisessa esiintyviä ongelmia Kuljun ja Mäkisen vastaavaa virhetyyppiä laajemmin (ks. luku 3.6). Koska tämän tutkimuksen tarkoitus on analysoida Unelmieni päivä -kirjoitelmien sanatason virheitä, jäävät Kuljun ja Mäkisen virhetyypit väärä mielekäs sana ja puuttuva sana analyysin ulkopuolelle. Varsinkin puuttuva sana virhetyypinä on luonteeltaan sellainen, ettei sen tutkiminen valmiista kirjoitelmakorpuksesta ole mielekäästä eikä mahdollistakaan, vaan se sopii tutkittavaksi nimenomaan sanelukirjoituksen tai toistamiskirjoituksen arviointiin.

Yllämainittujen virheiden lisäksi kirjoitelmista löytyy Ruoppilan ym. luokituksen mukaisia virheitä, jotka kertovat sanan visuaalisen hahmottamisen ongelmista: reversaali ja rotaatio. Näitä virheryhmiä onkin pidetty tunnusomaisina lukihäiriöisille. Reversaali tarkoittaa sitä, että kirjainten järjestys on vaihtunut sanan sisällä (*juolu* »joulu»). Rotaatio tarkoittaa sitä, että kirjain on kääntynyt akselinsa ympäri (*tuoba* »tuoda»). Lisäksi sanaraunioiksi on tässä tutkimuksessa nimetty sellaiset sanat, joissa on enemmän kuin kolme virhettä, tai joita ei

pysty kontekstinkaan avulla palauttamaan aiottuun muotoonsa. Kaikkia näitä virhetyyppejä esiintyy aineistossa todella vähän; toisen luokan oppilaiden kirjoitelmissa 6 reversaalia, 3 rotaatiota ja 5 sanarauniota, neljännen luokan oppilaiden kirjoitelmissa 2 reversaalia, ei yhtään rotaatiota ja 6 sanarauniota sekä kuudennen luokan oppilaiden kirjoitelmissa 5 reversaalia, ei yhtään rotaatiota ja 4 sanarauniota. Frekvenssiltään näin suppeita virheryhmiä ei ole tässä tutkimuksessa relevanttia analysoida, sillä ne eivät tarjoa kattavaa tietoa kullakin luokka-asteella esiintyvien virheiden yleisyydestä tai saavutetusta kirjoitustaidosta.

Tutkimuksen virheanalyysiin ei ole laskettu mukaan virheitä, jotka esiintyvät vieraskielisissä sanoissa. Tällaisia sanoja esiintyy paljon etenkin nuorimpien oppilaiden kirjoitelmissa muun muassa lelujen, elokuvien ja pelien nimissä, jotka usein liittyvät nuorimpien oppilaiden unelmien päivään. Nämä virheet eivät kuitenkaan ole oleellisia, kun halutaan selvittää oppilaiden kirjoitustaitoa nimenomaan oman äidinkielen kannalta. Toisekseen aineiston nuorimmat oppilaat (toisen luokan oppilaat) eivät vielä opiskele vieraita kieliä, joten ei voida olettaa, että he osaisivat esimerkiksi englanninkielisiä sanoja kirjoittaa.

Aineistossa esiintyy jonkin verran päällekkäin meneviä virheryhmiä, toisin sanoen osa virheistä on sellaisia, että ne voitaisiin sijoittaa useampaan kuin yhteen virhekkategoriaan. Tällaisia virheitä ovat etenkin sellaiset, jotka voidaan tulkita selittyvän joko puuttuvilla tai väärillä kirjaimilla tai virheellisellä taivuttamisella. Etenkin sanan lopusta puuttuva kirjain on tyypillinen tällainen virhe, eikä sitä voida yksiselitteisesti kategorisoida jompaankumpaan ryhmään. Koska kaikki tällaiset virheet ovat toisaalta mahdollisesti virheellisellä taivuttamisella selittyviä, käsitellään niitä tässä tutkimuksessa erikseen omana ryhmänään juuri taivuttamiseen liittyvien virheiden yhteydessä luvussa 3.6.3.

Tässä tutkimuksessa käytetyn koulukorpuksen tekstiversio on siirretty ja koodattu Microsoft Excel -ohjelmaan. Excel-aineistossa kirjoitelmista näkyvät vain lauseen rajat eikä esimerkiksi pilkun paikkoja tai yhdysmerkkejä ole merkitty, joten välimerkkien tutkiminen tästä aineistosta ei ole mahdollista. Analyysissa huomioitujen virheiden kokonaismäärä on 2 588 kappaletta, mikä on noin 6,1 % tutkittujen kirjoitelmien kokonaissanamäärästä.

### 3.2 Virheet sanarajoissa

Virheet sanarajoissa ovat tyypiltään ensinnäkin sellaisia, että sanaväli kahden eri sanan välistä puuttuu, ja toiseksi sellaisia, että saman sanan eri elementtien väliin on virheellisesti jätetty väli. Viimeksi mainituissa on yleisimmin kyse yhdysnavirheestä, jossa yhdysnanan eri osat on kirjoitettu virheellisesti erikseen. Virheet sanarajoissa kertonevat varsinkin nuorimmilla

oppilaille vaikeudesta hahmottaa sanarajoja kirjoitettaessa, ja yksittäisissä aineiston kirjoitelmissa esiintyykin runsaasti sanojen peräkkäin kirjoittamista ilman sanavälejä. Tällaista runsasta sanojen yhteen kirjoittamista voidaan kutsua ketjuttamiseksi. Varsinkin tämän tutkimuksen vanhempien oppilaiden kirjoitelmissa virheet sanarajoissa ovat enemmänkin luonteeltaan yhdyssanavirheitä perinteisesti ymmärrettynä, toisin sanoen semanttisesti tai funktionaalisesti motivoituneita normivirheitä; esimerkiksi kuudennen luokan oppilaiden teksteissä on selvemmin kirjoitettu virheellisesti yhteen juuri samaan lausekkeeseen kuuluvia sanoja tai vastaavasti jätetty yhdyssanan osat erilleen.

Sanarajoissa esiintyvät virheet ovat yleisimpiä virheitä tässä aineistossa kaikilla luokka-asteilla: yhteensä niitä on aineistossa 1 008 kappaletta, mikä on 39 % tutkittujen virheiden kokonaismäärästä.

### 3.2.1 Erilleen kirjoitettu yhdyssana

Erilleen kirjoitettu yhdyssana on koko aineiston yleisin virhe 580 esiintymällään. Toisen luokan oppilaiden kirjoitelmissa yhteen kuuluvat sanan osat on kirjoitettu virheellisesti erilleen 194 kertaa, ja näin ollen se on toiseksi yleisin virhe toisen luokan oppilaiden kirjoitelmissa. Neljännen luokan oppilaiden kirjoitelmissa yhdyssanavirhe on yleisin virhetyyppi 181 esiintymällä, ja myös kuudennen luokan aineistossa yhdyssanan erilleen kirjoittaminen on analysoiduista virheistä yleisin 205 esiintymällä. Microsoft Excel -muotoisesta Koulukorpuksesta on poistettu kaikki merkit, myös yhdysmerkit. Jos oppilas on kirjoittanut yhdysmerkin sisältävän yhdyssanan oikein (esim. *uima-allas*), on se koodattu Exceliin ilman yhdysmerkkiä ja sanaväliä (*uimaallas*). Edellisen kaltaiset on siis tutkimuksessa tulkittu norminmukaisiksi yhdyssanoiksi. Jos oppilas ei olo osannut käyttää yhdysmerkkiä oikein vaan on jättänyt sen merkitsemättä (esim. *uima allas*), on muoto laskettu tähän virhekköategoriaan kuuluvaksi yhdyssanavirheeksi.

Virheryhmään kuuluvat nimensä mukaisesti virheellisesti erikseen merkityt yhdyssanat. Tämän ilmeisimmän virheen lisäksi tähän kategoriaan on laskettu mukaan tapaukset, joissa ei ole kyse varsinaisesta yhdyssanavirheestä, vaan yksittäisen sanan eri osat on merkitty erilleen kuten esimerkeissä 1 ja 2.

- (1) kello 2.00 lähti *simme* ulkomaille (4M)
- (2) viivytettiin *koko nainen* viikko hienousten keskellä (4F)

Näitä tapauksia voi selittää se, että suomen kielen sanat voivat erilaisten johtimien, liitteiden ja päätteiden ansiosta muodostua todella pitkiksi, eivätkä alakoululaiset vielä kykene hahmottamaan näin pitkiä sanoja vaan pyrkivät selkeyttämään niitä itselleen katkaisemalla ne jostain kohtaa. Toinen mahdollinen selitys on, ettei liitteitä ja päätteitä mielletä samaan sanaan kuuluviksi, vaan sanavartalo mielletään itsenäiseksi yksiköksi. Kolmas mahdollinen selitys liittyy rivinvaihtoon: oppilas on saattanut jakaa sanan kahdelle eri riville, ja jättänyt ensimmäisen osan jälkeen yhdysmerkin merkitsemättä, jolloin koodaaja on tulkinut sanan kahdeksi eri sanaksi.

Lähes poikkeuksetta tämän virhekattegorian virheet ovat kuitenkin sellaisia, joissa yhdyssanan yhdysosat, määrite- ja edusosa, on kirjoitettu erilleen. Tämä virhetyyppi on tyypillisimmin kompetenssivirhe, eli varsinainen virhe: oppilaan kielitaju ei siis kerro, että sanat tulisi kirjoittaa yhteen. Nuorimmat oppilaat eivät kirjoitelmissaan kovin kompleksisia yhdyssanoja käytä, joten heidän norminvastaiset yhdyssanansa ovat tyypillisimmin esimerkin 3 kaltaisia yhdyssubstantiiveja. Tyypillistä on myös, että oppilaat hahmottavat yhdyssanoiksi vain kaksiosaiset yhdyssanat. On mielenkiintoista havaita, että oppilaat hahmottavat sanoja pareittain, ja näin ollen useampiosaisen yhdyssanat on useimmiten katkaistu joko ensimmäisen osan jälkeen tai juuri ennen viimeistä osaa kuten esimerkissä 4.

- (3) Sitten haluaisin mennä kotiin ja syödä *lempi ruokaani* makaroonia Sitten haluaisin katsoa videota ja sitten haluaisin leikkiä *pikku sikkoni* kanssa (2F)
- (4) Hyvä ravihevonen tulikin se voitti 2 kertaa *suomen mestaruusravit* ja 3 kertaa tavalliset ravit se oli *suomenhevos ori* (4F)

Vanhempien oppilaiden kirjoitelmissa yksinkertaisimmat yhdyssubstantiivit eivät aiheuta yhtä paljon ongelmia kuin nuorimpien teksteissä. Erityisen alttiita erilleen kirjoittamiselle etenkin vanhempien oppilaiden kirjoitelmissa ovatkin kompleksisemmat rakenteet, esimerkiksi sellaiset yhdyssanat, joiden määriteosana toimiva nomini on jossain muussa sijassa kuin nominatiivissa kuten esimerkeissä 5 ja 6.

- (5) meidän perhe voittaisi *ulkomaan matkan* (6M)
- (6) Ennen *kotiin lähtöä* kävisimme katselemassa julkkisten koteja (4F)

Esimerkkien 5 ja 6 kaltaisia virheitä selittänee se, että määriteosan taipuminen saa sen vaikuttamaan nominatiivimuotoista määriteosaa itsenäisemmältä yksiköltä, jopa itsenäiseltä



sanalta. Suurin osa näistä virheistä esiintyy juuri genetiivimuotoisen määräiteosan yhteydessä. Muissa sijoissa virhettä selittänee määräiteosan taipumisen lisäksi se, että ne ovat usein *-inen*-adjektiivijohtoksia tai *O-* ja *U-*teonnimijohdoksia, joiden johtamattomat muodot siis kirjoitettaisiin erilleen (vrt. *tämä hetki* > *tämänhetkinen*, *lähteä kotiin* > *kotiinlähtö*).

Toinen virhealtis yhdyssananryhmä koostuu yhdyssanoista, joiden määräiteosa on numeraali kuten esimerkissä 7.

(7) nostin sängun selkänöjaa ja laitoin 50 tuumaisen plasmatelevisioon päälle (4M)

Myös numeraalin sisältävät yhdyssanat ovat usein johdosadjektiiveja (vrt. *viisi vuotta* > *viisivuotias*, *50 tuumaa* > *50-tuumainen*). Toiseksi numeroita sisältävien yhdyssanojen oikeinkirjoitukseen liittyy sellaisia sääntöjä (esim. yhdyssmerkin käyttäminen), joita oppilaat eivät joko ole vielä oppineet tai eivät osaa soveltaa.

Erilleen kirjoitettu yhdyssana on mielenkiintoinen virhetyyppi, sillä virheen kokonaismäärä pysyy hyvin samankokoisena läpi aineiston. Virheen verrattain suurta määrää vanhempien oppilaidenkin kirjoitelmissa selittää se, että sanamäärän kasvaessa ja kielen kompleksistuesssa kirjoitelmissa yhdyssanojen määrä suhteutettuna kirjoitelman sanojen kokonaismäärään kasvaa. Yhdyssanojen norminmukaiset ja norminvastaiset esiintymät on eritelty taulukossa 3.

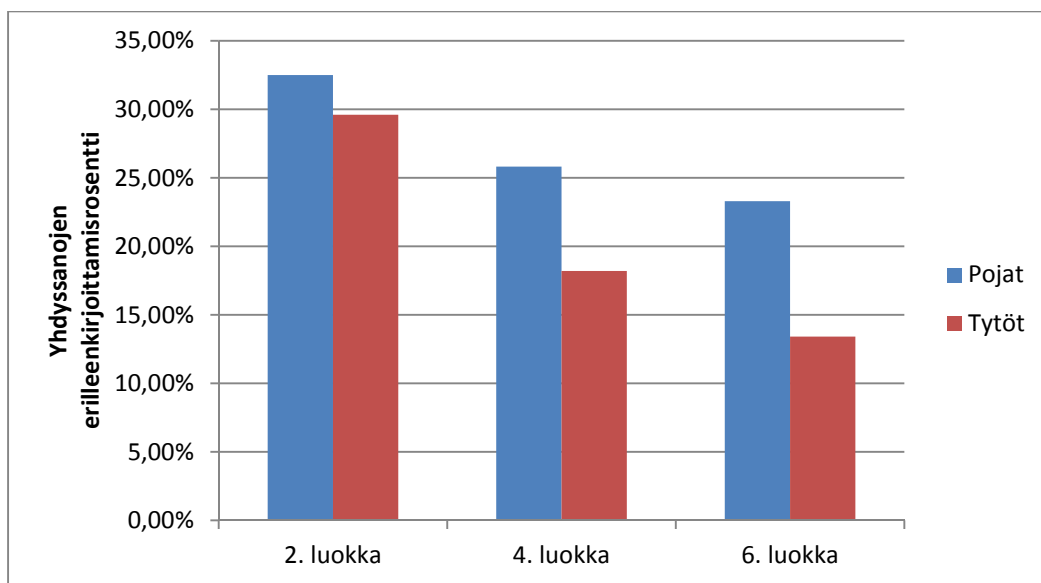
**Taulukko 3.** Yhdyssanojen virheellisten ja norminmukaisten esiintymien jakauma luokka-asteittain ja sukupuolittain.

Luokka-aste	Norminmukainen M	Virheellinen M	Norminmukainen F	Virheellinen F	Yhteensä	Virheellisiä (%)
2.	181	87	255	107	630	30,8 %
4.	259	90	410	91	850	21,3 %
6.	388	118	561	87	1154	17,8 %
<b>Yhteensä</b>	828	295	1226	285	2634	22,0 %

Kuten taulukosta 3 ilmenee, yhdyssanojen kokonaismäärä lähes kaksinkertaistuu verrattaessa toisen ja kuudennen luokan kirjoitelmia keskenään. Samojen luokka-asteiden virheprosenttien vertailu taas osoittaa, että nuorimpien yhdyssanavirheiden osuus lähes puolittuu vanhimpien oppilaiden kirjoitelmissa. Samalla kun oppilaiden kirjoitelmissaan käyttämä kieli rikastuu, ja

he käyttävät yhdyssanoja enemmän, kehitty myös heidän sana- ja lausetajunsa, ja näin ollen norminvastaisten yhdyssanojen suhteellinen määrä vähenee.

Sukupuolittaisessa vertailussa tyttöjen ja poikien tekemien yhdyssanavirheiden kappalemäärät saattavat antaa harhaanjohtavan kuvan yhdyssanavirheiden tekijöistä. Virheellisten muotojen prosentuaalinen vertailu suhteessa norminmukaisiin yhdyssanoihin kuitenkin osoittaa, että pojat tekevät yhdyssanavirheitä tyttöjä enemmän kaikilla luokka-asteilla. Kaaviossa 1 on esitetty yhdyssanojen virheellisyysprosentit suhteessa norminmukaisiin yhdyssanoihin luokka-aste- ja sukupuolikohtaisesti.



**Kaavio 1.** Yhdyssanojen erilleen kirjoittaminen suhteessa norminmukaisiin yhdyssanoihin luokka-asteittain ja sukupuolittain.

Kuten kaaviosta 1 käy ilmi, toisen luokan poikien yhdyssanoista 32,5 % virheellisiä, kun samanikäisten tyttöjen virheprosentti on 29,6 %. Virheprosentti pienenee molemmilla sukupuolilla aina seuraavalle luokka-asteelle siirryttäessä, mutta tyttöjen yhdyssanavirheet vähenevät poikia enemmän. Neljännen luokan poikien yhdyssanoista 25,8 % on virheellisesti erikseen kirjoitettu, kun tyttöjen virheprosentti on jo huomattavasti pienempi, 18,2 %. Kaikkein suurin ero yhdyssanavirheissä on kuitenkin kuudennella luokalla, jolloin poikien yhdyssanoista 23,3 % on virheellisiä ja tytöillä liki kymmenen prosenttiyksikköä vähemmän, 13,4 %. Luokka-asteiden yhteenlasketuissa sukupuolikohtaisissa virheprosentteissa ero ei kuitenkaan ole yhtä radikaali kuin kuudesluokkalaisten aineiston: kaiken kaikkiaan aineiston yhdyssanoista pojat ovat tehneet virheitä 26,3 %:ssa ja tytöt 18,9 %:ssa.

Yhdyssanavirheitä tytöt ja pojat tekevät siis melko saman verran alemmilla vuosiluokilla, mutta tyttöjen poikia radikaalimpi yhdyssanavirheiden väheneminen alimmilta vuosiluokilta ylemmille siirryttäessä ilmenee juuri kuudesluokkalaisten aineistoja vertailtaessa. Tämän perusteella voidaan sanoa, että tyttöjen kielitaju saattaa kehittyä alakoulussa poikia nopeammin, ja näin he tunnistavat ja ymmärtävät yhdyssana-käsitteen merkityksen poikia paremmin ja osaavat soveltaa oppimiaan yhdyssanojen kirjoittamista ohjailevia sääntöjä poikia paremmin.

### 3.2.2 Puuttuva sanaväli

Puuttuva sanaväli virhekatgoria on koko aineiston toiseksi suurin virheryhmä 428 esiintymällään. Toisen luokan oppilaiden kirjoitelmissa on kirjoitettu kaksi yhteenkuulumatonta sanaa 261 kertaa yhteen. Kaiken kaikkiaan sanavälin puuttuminen on kaikkein yleisin virhe toisen luokan oppilailla. Virheen yleisyys selittyy sillä, että toisen luokan oppilaiden kirjoitustaito on vielä kovin kehittymätöntä. Sanarajojen hahmottaminen voi tuottaa varsinkin juuri kirjoittamaan opeteltaessa hankaluuksia. Lisäksi sanojen runsas yhteen kirjoittaminen voi johtua myös motoriikan kehittymättömyydestä. Sanavälin puuttumista selittää osaltaan myös se, että nimenomaan tätä virhettä on paljon yksittäisissä kirjoitelmissa; sanavälin puuttuminen on kimppuuntuva virhetyyppi, ja etenkin aineiston nuorimmat oppilaat saattavat ketjuttaa tekstiään runsaasti. Yhdessä kirjoitelmassa puuttuvia sanavälejä esiintyy todennäköisesti todella paljon kuten esimerkit 8 ja 9 osoittavat tai todella vähän jos lainkaan.

- (8) *äitisanoinminulle että mentäisii minunläksyjenjärkeenuinaan kunmetultiinuimastaniin kellooli puolikahdeksan joten olinukkumaanmeno aika (2F)*
- (9) *sitenhelähtivät kotiin jä meniväkatommaantelkaria ja sitehe söivä iltaruuan ja sitenhemmenivä nukummaan ja sitehesöivät aammuruuan ja menivä lenkille ja sitenhemmenivät kaveiilleen (2M)*

Paitsi sanarajan hahmottamisen vaikeudesta, puuttuva sanaväli saattaa johtua puhtaasti oppilaiden käsialan kehittymättömyydestä. Varsinkin nuorimpien oppilaiden käsiala saattaa olla sen verran harjaantumaton, että tekstiä lukevan voi olla haastavaa tulkita kirjainten välissä olevia välejä. Lapsia ei ole aineiston keruun yhteydessä ohjeistettu kirjoittamaan tietyllä käsialalla, ja etenkin nuorimmat ovatkin kirjoittaneet monella tyylillä. Sanarajat ovat

luonnollisesti yleensä ja kirjoittamistyylistä vaihdellen tulkinnallisia. Jokin sanaväli saattaa olla yksittäisen sanan kirjainten väliin jäänyttä väliä pienempi, vaikka oppilas varsin hyvin ymmärtäisi, mihin kohtaan sanaväli kuuluu.

Kirjoitustaidon kehittymättömyys näyttäisi selittävän toisen luokan oppilaiden kirjoitelmien runsasta yhteen kirjoittamista myös siinä mielessä, että aineistossa juuri tämä virheryhmä pienenee kaikkein radikaaleimmin ylemmille luokka-asteille siirryttäessä. Virheellisesti yhteen kirjoittaminen on nuorimmilla oppilailta hyvin epämääräistä, mutta myöhemmin ketjuttamista voidaan pitää kehityksellisesti heikkona piirteenä. Toisen luokan aineistossa tämä virhe esiintyi 261 kertaa, mikä on noin 2,5 % koko toisen luokan kirjoitelmien sanamäärästä. Vastaavasti neljäsluokkalaisten kirjoitelmissa on kaksi sanaa kirjoitettu virheellisesti yhteen 101 kertaa, mikä on noin 0,7 % neljäsluokkalaisten kirjoitelmien kokonaissanamäärästä, ja kuudesluokkalaisten kirjoitelmissa tämä virhe esiintyy 66 kertaa, mikä on noin 0,4 % kuudesluokkalaisten kirjoitelmien kokonaissanamäärästä. Sanavälin puuttuminen on toiseksi yleisin virhe neljäs- ja kuudesluokkalaisten kirjoitelmissa.

Sanavälin puuttuminen -ryhmän virheiden väheneminen siirryttäessä ylemmille luokka-asteille kuvaa alakoululaisten kirjoitustaidon kehittymistä ja erityisesti sanahahmotajun kehittymistä. Kun toisen luokan oppilaiden sanojen yhteen kirjoittaminen vaikuttaisi olevan melko mielivaltaista, niin neljännellä ja kuudennella luokalla sen voidaan ajatella olevan semanttisesti tai funktionaalisesti motivoitunutta. Vanhemmat oppilaat siis kirjoittavat selvemmin yhteen sellaisia sanoja, jotka he todella mieltävät yhdyssanoiksi. Tästä kertoo se, että ylempien vuosiluokkien oppilaat ovat kirjoittaneet virheellisesti määritteitä yhteen niiden edussanojen kanssa (10) varsinkin, jos ne ovat samassa sijassa (11–12), sekä adpositioita yhteen täydennystensä kanssa (13).

- (10) *Unelmienpäivänä* haluaisin olla jääkiekon smliaigan *paraspelaaja* ja minulla olisi *kultainen kypärä* (6M)
- (11) Tunsin olevani tänään *hyvässä kunnossa* (6M)
- (12) Koko perheeni ja minä olimme *hyvällä tuulella* (4M)
- (13) sinä saat tämän päivän aikaa olla *vedenalla* (4F)

Lisäksi erityisesti merkitykseltään abstraktit ja pituudeltaan lyhyet sanat ovat eräänlainen ongelmaryhmä, ja niitä ei ehkä aina hahmoteta itsenäisiksi sanoiksi. Tällaisia sanoja aineistossa edustavat esimerkiksi konjunktiot *kun* (14) ja *kuin* (15) erityisesti adverbin *niin*

kanssa yhteen kirjoitettuna (16) sekä postpositio *kanssa* ja sen erilaiset puhekieliset variantit (17).

- (14) *Hetikun* oli pimeä lähdettiin hautuumaalle kynttilöitä viemään (6M)
- (15) Mutta tämä unelma jää kyllä haaveeksi koska ei tällaista päivää voi tapahtua *muutakuin* unissa (6F)
- (16) minä haluaisin takaisin vanhaan kotiin melkein *niinkuin* elokuvassa *Villitusina* (4M)
- (17) Haluaisin myös olla *Roosankanssa* koulun suosituimpia tyttöjä Menisin *Roosankanssa* junalla taas Ouluun (4F)

Omistaminen koetaan myös selkeästi sanat tiukasti toisiinsa sitovaksi asiayhteydeksi, sillä hyvin usein on kirjoitettu virheellisesti yhteen sellaisia sanaliittoja, joiden ensimmäinen osa on genetiivissä kuten esimerkeissä 18 ja 19.

- (18) Minun unelmieni päivä oli eilen koska sain *särkänkausikortin* (4M)
- (19) Sitten me mentiin Vaasaan koska me haettiin *Vaasankissatalosta* kissa (4F)

Virheellisesti yhteen kirjoitetuista sanoista 264 kappaletta on poikien kirjoitelmista ja 164 kappaletta tyttöjen kirjoitelmista. Virheiden määrästä saa selkeämmän kuvan, kun sitä vertaa kirjoitelmien sanojen kokonaissanamäärään. Puutuva sanaväli -virheen jakautuminen eri luokka-asteiden ja sukupuolten välille selviää taulukosta 4.

**Taulukko 4:** Puuttuvat sanavälit sukupuolittain ja vuosiluokittain.

Luokka-aste	Kokonaissanamäärä M	Sanaväli puuttuu M	Virhe % M	Kokonaissanamäärä F	Sanaväli puuttuu F	Virhe % F	Kokonaissanamäärä virhe %
2.	3926	172	4,4 %	6482	89	1,4 %	2,5 %
4.	5522	59	1,1 %	8784	42	0,5 %	0,7 %
6.	6802	33	0,5 %	10774	33	0,3 %	0,4 %
<b>Yhteensä</b>	16250	264	1,6 %	26040	164	0,6 %	1,0 %

Kuten taulukosta 4 ilmenee, esiintyy poikien kirjoitelmissa kaikilla vuosiluokilla sanan virheellistä yhteen kirjoittamista kokonaissanamäärään suhteutettuna enemmän. Selvin ero on toisen luokan oppilaiden kirjoitelmissa: poikien virheprosentti on yli kolminkertainen tyttöjen

virheprosenttiin verrattuna. Neljännellä luokalla sukupuolten väliset erot hieman tasoittuvat, ja kuudennella luokalla tyttöjen ja poikien virheprosentti on jo lähes samansuuruinen.

Taulukosta 4 käy hyvin ilmi, että sanavälin puuttuminen on yleisintä nimenomaan nuorimmilla oppilailla. Toisen luokan oppilaiden kokonaisvirheprosentti on 2,5 %, mutta kuudennen luokan kirjoitelmissa vastaava luku on enää 0,4 %. Tämä kertonee ennen kaikkea sanarajan hahmottamisen ja käsialan kehittymisestä edetessä alakoulun alimmilta luokka-asteilta ylemmille. Se, että alakoulun ylemmillä luokilla kirjoitetaan sanoja semanttisesti tai funktionaalisesti motivoituneesti yhteen, on tulkinnallista; se voi kuvata oppilaiden herkkää korvaa sanojen semanttiselle tai funktionaaliselle yhteenkuuluvuudelle tai oikeinkirjoituskonventioiden tuntemattomuutta.

### 3.3 Puuttuva kirjain

Puuttuva kirjain voi olla puuttuva konsonantti geminaatasta, puuttuva vokaali pitkästä vokaalista tai yksittäinen puuttuva konsonantti tai vokaali, joka voi puuttua sanan alusta, keskeltä tai lopusta. Sanan lopusta kirjain voi puuttua monesta eri syystä; useissa tapauksissa kyse on loppuheitosta. Vaikka kaikkia puhekielen piirteitä ja mahdollisia murteen vaikutuksia koululaisten kirjoitelmista on mahdotonta erottaa, on tämän tutkimuksen virheanalyysissä pyritty mahdollisimman selvärajaiseen jaotteluun. Varsinkin puuttuvien kirjaimien ryhmässä on otettava huomioon monia tekijöitä. Tämän ryhmän ulkopuolelle on jätetty ensinnäkin sellaiset muodot, jotka Salonen (2012) on jo analysoinut puhekielen piirteiksi. Tällaisia muotoja ovat erilaiset pikapuhemuodot, *A*:n loppuheitto, *i*:n loppuheitto, jälkitavujen *i*-loppuisten diftongien jälkikomponentin kato sekä sananloppuisen *n*:n ja *t*:n kato. Lisäksi analyysin ulkopuolelle on rajattu myös frekvenssiltään vähäinen joukko selvemmin alueellisia puhekielen piirteitä, jotka myös Salonen on jättänyt omassa tutkielmassaan niiden vähäisyyden vuoksi analysoimatta. Analyysin ulkopuolelle jäävät tapaukset koskevat tämän tutkimuksen kategoriaa muu puuttuva kirjain, joten ulkopuolelle rajatut muodot eritellään kyseisen virhe kategorian yhteydessä tarkemmin.

Puuttuvan kirjaimen tapauksia on aineistossa yhteensä 651 kappaletta, mikä vastaa noin 25 %:a kaikista aineiston virheistä. Kuten virheiden ryhmittely -luvussa on kerrottu, on joidenkin virheiden ryhmittely yksiselitteisesti puuttuviksi kirjaimiksi mahdotonta; tämä tyyppi menee usein päällekkäin etenkin virheellinen taivutus -kategorian kanssa. Niinpä puuttuviin kirjaimiin lukeutuviksi virheiksi ei ole luokiteltu tapauksia, jotka voivat yhtä hyvin

selittyä virheellisellä taivuttamisella. Tämä monitulkintainen virhetyyppi käsitellään erikseen juuri virheellisen taivutuksen yhteydessä.

### 3.3.1 Puuttuva konsonantti geminaatasta

Puuttuva konsonantti geminaatasta on luonteeltaan äännepitouden hahmottamiseen liittyvä virhe. Varsinkin kirjoitustaidon alkutaipaleilla oppilaille saattaa olla hyvinkin haastavaa hahmottaa eroja konsonanttien äännepituuksissa, sillä eron tekeminen konsonantin äännepitouteen on huomattavasti vaikeampi havainnollistaa kuin vokaalien pituusero. Näin ollen aineiston nuorimpien oppilaiden kirjoitelmissa konsonantin puuttuminen geminaatasta on huomattavasti yleisempi virhe kuin vanhempien oppilaiden kirjoitelmissa. Toisen luokan oppilaiden kirjoitelmissa tämä virhe esiintyy 181 kertaa, neljännen luokan oppilaiden kirjoitelmissa 56 kertaa ja kuudennen luokan oppilaiden kirjoitelmissa 51 kertaa. Yhteensä puutteellisia geminaattoja on siis koko aineistossa 288 kappaletta. Pelkät kappalemäärät tämän virheen osalta eivät kuitenkaan anna tietoa oppilaiden kyvystä hahmottaa konsonanttien pituuksia, sillä eri-ikäisten oppilaiden aineistossa vaihtelee hyvin laajasti se, kuinka paljon geminaattoja (norminmukaisia ja virheellisiä) niissä ylipäätään esiintyy. Niinpä taulukossa 5 onkin esitetty kunkin luokka-asteen virheellisten geminaattojen määrä suhteutettuna norminmukaisiin.

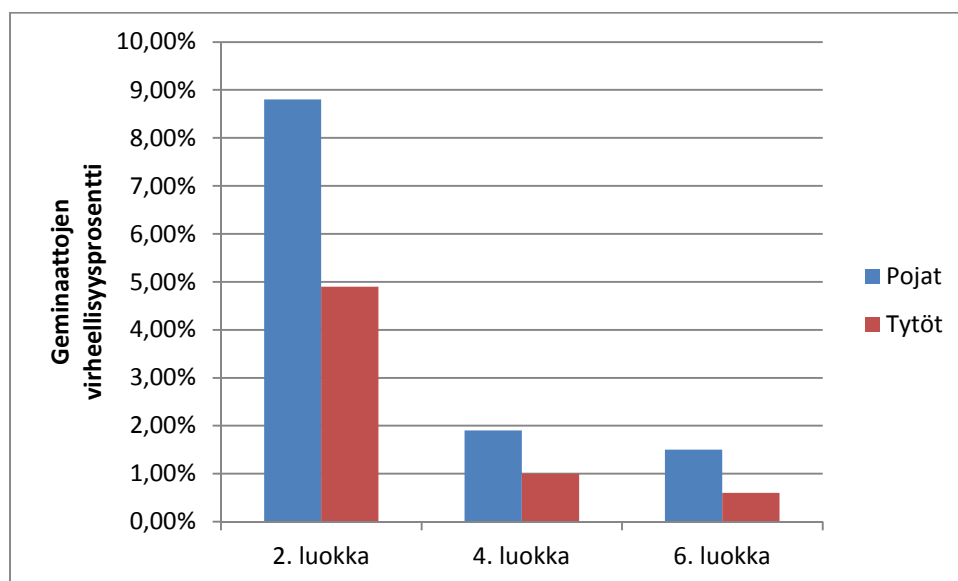
**Taulukko 5:** Konsonantin puuttuminen geminaatasta luokka-asteittain ja sukupuolittain.

Luokka-aste	Norminmukainen M	Virheellinen M	Norminmukainen F	Virheellinen F	Yhteensä	Virheet (%)
2.	990	96	1645	85	2816	6,4 %
4.	1508	29	2567	27	4131	1,4 %
6.	2125	32	3423	19	5599	0,9 %
<b>Yhteensä</b>	4623	157	7635	131	12546	2,3 %

Taulukkoon 5 on laskettu kunkin luokka-asteen geminaattojen virheellisyysprosentti, ja niiden luokka-astekohtainen vertailu osoittaa, että konsonantin puuttuminen geminaatasta on ylivoimaisesti yleisintä tämän aineiston nuorimpien oppilaiden kirjoitelmissa: toisen luokan oppilaiden geminaatoista peräti 6,4 % on puutteellisia, kun vastaava luku neljäsluokkalaisten geminaatoissa on enää 1,4 % ja kuudesluokkalaisten 0,9 %. Puuttuva konsonantti

geminaatasta -virheen väheneminen alemmilla vuosiluokilla ylemmille siirryttäessä kertoo kirjoitustaidon kehittymisestä, varsinkin, kun kyseessä on kirjoitustaitoa mittaava ongelmarakenne. Kirjoitustaidon kehittymisestä näyttäisi kertovan myös se, että ylemmillä vuosiluokilla tehdyt virheet konsonantin pituuden merkitsemisessä vaikuttaisivat olevan enemmänkin huolimattomuusvirheitä kuin osoitus osaamattomuudesta. Toisen luokan oppilaiden kirjoitelmissa geminaatan puuttuminen on melko johdonmukaista, mikä kertonee juuri vaikeudesta hahmottaa konsonantin äännepituuksia. Kuudennen luokan kirjoitelmissa tämä virhe ei ole enää yhtä johdonmukainen, vaan saman oppilaan kirjoitelmissa saattaa esiintyä puutteellisia geminaattoja norminmukaisten rinnalla. Satunnainen geminaattojen virheellisyys vanhempien oppilaiden kirjoitelmissa ohjaa niiden tulkitsemista satunnaisiksi lipsahduksiksi varsinaisen virheen sijaan.

Aineiston yhteensä 288 geminaatasta puuttuvasta konsonantista 157 kappaletta on poikien kirjoitelmista ja 131 kappaletta tyttöjen kirjoitelmista. Sukupuolittain ja luokka-asteittain lasketut geminaattojen virheellisyysprosentit esitellään kaaviossa 2.



**Kaavio 2.** Konsonantin puuttuminen geminaatasta -virheet suhteessa norminmukaisiin geminaattoihin luokka-asteittain ja sukupuolittain.

Kuten kaaviosta 2 ilmenee, konsonantin puuttuminen geminaatasta -virhe jakautuu melko epätasaisesti sukupuolten suhteen; jokaisella vuosiluokalla tyttöjen virheprosentti on poikien virheprosenttia pienempi. Huomattavin ero on aineiston nuorimpien oppilaiden kohdalla; toisen luokan poikien virheprosentti on lähes kaksinkertainen (8,8 %) tyttöjen virheprosenttiin



(4,9 %) verrattuna. Muilla luokka-asteilla sukupuolten erot geminaatista puuttuvien konsonanttien osalta hieman tasoittuvat, joskin myös neljännen ja kuudennen luokan poikien virheprosentti on tyttöihin verrattuna kaksinkertainen. Ylemmillä vuosiluokilla virheen frekvenssi on kaiken kaikkiaan kuitenkin niin suppea, ettei esiintymien kannalta ole mielekästä tehdä yleistäviä johtopäätöksiä.

Geminaattaa ja erityisesti klusiiligeminaattoja *kk*, *pp* ja *tt* pidetään yhtenä suomen kielen ongelmarakenteista, jotka usein aiheuttavat ongelmia varsinkin aloittelevalla kirjoittajalla (Ahvenainen & Holopainen 2005: 107). Onkin mielenkiintoista vertailla, missä konsonanteissa geminaatat toteutuvat norminmukaisesti ja minkä konsonantin kohdalla aineistossa on eniten virheitä. Taulukkoon 6 on koottu aineistossa esiintyvät norminmukaiset ja virheelliset geminaatat konsonanttikohtaisesti. Taulukossa ei ole mukana konsonantteja *d*, *h*, *j* ja *v*, sillä niiden geminoituminen suomen kielessä on hyvin marginaalista: geminaatta *dd* esiintyy vain vieras- ja slangisanoissa (*addiktio*, *spaddu*), *hh*, *jj* ja *vv* vain murteellisen geminaation tuottamina (ISK § 27).

**Taulukko 6.** Konsonanttikohtainen geminaattojen toteutuminen luokka-asteittain.

Geminaatta	Norminmukainen 2. lk	Virheelinen 2. lk	Norminmukainen 4. lk	Virheelinen 4. lk	Norminmukainen 6. lk	Virheelinen 6. lk	Yhteensä	Virhe %
<b>kk</b>	288	33	409	6	483	2	1221	3,4 %
<b>ll</b>	714	35	1136	18	1709	16	3628	1,9 %
<b>mm</b>	206	19	546	14	795	17	1597	3,1 %
<b>nn</b>	216	6	285	7	454	6	974	2,0 %
<b>pp</b>	84	12	118	1	148	4	367	4,6 %
<b>rr</b>	50	5	45	0	53	3	156	5,1 %
<b>ss</b>	388	4	557	3	714	2	1668	0,5 %
<b>tt</b>	689	67	979	7	1192	1	2935	2,6 %
<b>Yhteensä</b>	2635	181	4075	56	5548	51	12546	2,3 %

Suuri osa aineiston puuttuvista geminaatoista onkin juuri klusiileja, kuten taulukosta 6 käy ilmi. Tämä tosin saattaa selittyä sillä, että suomen kielessä juuri klusiilit muodostavat usein geminaatan astevaihtelun vahvana asteena. Esimerkit 20–22 ovat tyypillisiä puutteellisen geminaatan esiintymistapauksia, ja esimerkkien tapaan tyypillistä on, että samassa

kirjoitelmassa on paljon geminaatista puuttuvan konsonantin virheitä, varsinkin juuri aineiston nuorimpien oppilaiden kirjoitelmissa.

- (20) Olen karki landiassa näin siellä ison tikarin ja karkisauvan marjane karkin ja pienen mokin (2F)
- (21) siten hä hypasi ikkunasta hän luuli etä lento taito on tale/a (2M)
- (22) aamulla mentäis ulos ja käytäis kaupassa ja ostetais mulle vaateita ja site me mentäis syömään kotiin (2F)

Klusiiligeminaattojen lisäksi konsonantteja puuttuu aineiston kirjoitelmissa harvemmin dentaaligeminaatoista *ll* ja *ss*, jotka esiintyvät adessiivin (23), allatiivin (24) ja inessiivin (25) tunnuksissa. Lisäksi likvida- tai nasaaligeminaatta on usein astevaihtelun alaisen *t*-loppuisen konsonanttiyhtymän heikkoasteinen vastine, kuten esimerkeissä 26–27. Lisäksi nasaaligeminaatta *mm* puuttuu usein monikon kolmannen persoonan persoonapäätteestä kuten esimerkissä 28.

- (23) Olipa keran koira joka asui kadu/a Koira kulki kadui/a päivät ja varasti makkaroita (2F)
- (24) Mei/e tuli sisko jouluks käymään (2F)
- (25) Minun isosisko nukkui alakerrasa (2M)
- (26) Iso sisko tuo minulle koiran penun Lekin koko päivän koiran penun kansa (2F)
- (27) Kun siellä olemme muutaman satasen törsänneet farkkuihin siirymme H&Mlle ostelemaan paitoja (6F)
- (28) Kun pääsisime perille me saime ilmaiseksi ja hotellissa olisi iso TV (6M)

Konsonantikohtainen geminaattojen norminmukaisten ja virheellisten esiintymien vertailu osoittaa, että klusiiligeminaattojen lisäksi geminaatta *rr* on aineistossa melko riskialtis. Geminaatta *rr*:n norminmukaiset esiintymät (yhteensä 148 kappaletta) tosin ovat aineistossa huomattavasti muita konsonantteja harvinaisempia, joten pienikin virheiden määrä vaikuttaa nopeasti virheprosentin nousemiseen. Useimmiten norminmukaisena säilynyt geminaatta on *ss*, jonka virheprosentti on vain 0,5 %. *ss*:n säilymistä norminmukaisena selittänee se, että se on usein ensimmäinen alakoululaisille opetettava konsonantti, ja sen pituuseroja on hyvin helppo havainnollistaa ja kokeilla suhisemalla, toisin kuin esimerkiksi klusiilien. Lisäksi geminaatta *ss* sisältyy inessiivin sijapäätteeseen, mikä tekee siitä tunnusomaisen ja kenties helposti muistettavan kirjainyhdistelmän.

### 3.3.2 Puuttuva vokaali pitkästä vokaalista

Puuttuva vokaali pitkästä vokaalista on myös äännepitouden hahmottamiseen liittyvä virhe. Vokaalien kohdalla pituuserojen havainnollistaminen on kuitenkin huomattavasti helpompaa kuin konsonanttien, joten aineistossa vokaalien pituuksissa on huomattavasti vähemmän virheitä kuin konsonanttien kohdalla. Vokaali puuttuu pitkästä vokaalista aineistossa 58 kertaa, kun konsonantti puuttuu geminaatasta lähes viisi kertaa useammin, 288 kertaa. Äännepitouuteen liittyvät virheet ovat enimmäkseen nuorimpien oppilaiden tekemiä, niin tässäkin virheryhmässä: 58 virheestä 39 kappaletta on toisen luokan oppilaiden tekemiä, 7 kappaletta neljäsluokkalaisten tekemiä ja 12 kappaletta kuudesluokkalaisten tekemiä.

Puuttuva vokaali pitkästä vokaalista -kategoria menee osittain päällekkäin väärä taivutus -virheen kanssa. Virheitä, joita ei voi yksiselitteisesti luokitella kuuluviksi puuttuvien vokaalien ryhmään, ei käsitellä tässä yhteydessä lainkaan, sillä virhekatteoria on kaiken kaikkiaan sen verran suppea, että kyseiset esiintymät aiheuttaisivat suurella todennäköisyydellä melko suuren vinouman tämän kategorian tuloksiin. Virheet, jotka voivat siis johtua joko puuttuvasta vokaalista pitkästä vokaalista tai virheellisestä taivutuksesta, käsitellään erikseen omana monitulkintaisten virheiden ryhmänä virheellisen taivuttamisen yhteydessä.

Samoin kuin geminaatasta puuttuvan konsonantin yhteydessä myöskään pitkästä vokaalista puuttuvan vokaalin yhteydessä pelkät virheiden kappalemäärät eivät riitä analyysin pohjaksi. Jotta oppilaiden kyvystä hahmottaa vokaalien pituuseroja saataisiin jonkinlainen kuva, tulee virheitä verrata norminmukaisiin, oikein merkittyihin pitkiin vokaaleihin, kuten taulukossa 7 on tehty.

**Taulukko 7.** Norminmukaiset ja virheelliset pitkät vokaalit luokka-asteittain ja sukupuolittain.

Luokka-aste	Norminmukainen M	Virheellinen M	Norminmukainen F	Virheellinen F	Yhteensä	Virheet (%)
2.	760	11	1376	28	2175	1,8 %
4.	1160	4	1789	3	2956	0,2 %
6.	1587	3	2341	9	3940	0,3 %
<b>Yhteensä</b>	3507	18	5506	40	9071	0,6 %

Taulukosta 7 ilmenee nopeasti, kuinka vokaalin puuttuminen pitkästä vokaalista on huomattavasti harvinaisempaa kuin konsonantin puuttuminen geminaatasta; kokonaisvirheprosentti konsonanttien kohdalla aineistossa on 2,3 %, kun se vokaalien kohdalla on vain 0,6 %. Tämä todistaa sen, että oppilaat erottavat vokaalien pituuserot huomattavasti paremmin kuin konsonanttien.

Toinen seikka, johon tämän virheen yhteydessä kiinnittää huomiota, on se, että pitkästä vokaalista toisen jättävät merkitsemättä etenkin nuorimmat oppilaat, samoin kuin konsonantin geminaatasta. Toisen luokan oppilaiden kokonaisvirheprosentti on 1,8 %, kun se kuudennen luokan oppilaiden kirjoitelmissa on enää 0,3 %. Toisen luokan oppilaiden kirjoitelmissa vokaalin pituuksissa on usein virheitä sellaisissa tapauksissa, joissa virheitä on runsaasti myös konsonanttien pituudessa kuten esimerkissä 29.

- (29) siteisä osti meillejätelöt site mementin hoteli huonesen ja syötin leivät ja mesaatin raha ja mentinkaupan (2F)

Esimerkin 29 kaltaisilla kirjoittajilla, joilla äännepitouden merkitsemisessä on johdonmukaisesti virheitä niin konsonanttien kuin vokaalienkin kohdalla, voidaan suurella todennäköisyydellä olettaa olevan ongelmia fonologisessa kehityksessä. Tosin kirjoituskokemus ja kielellinen kehitys ylipäättään saattavat vaikuttaa virheiden vähenemiseen, mikäli oppilaalla ei ole vakavaa kehityshäiriötä.

Neljännän ja kuudennen luokan oppilaiden kirjoitelmissa pitkästä vokaalista puuttuva vokaali on kaikkein kaikkiaan erittäin suppea virheryhmä (yhteensä vain 19 esiintymää), ja vanhempien oppilaiden kirjoitelmissa vokaalin pituuteen liittyvät virheet eivät ole yhtä johdonmukaisia kuin nuorimpien oppilaiden. Vokaalin pituuteen liittyvistä virheistä vanhempien oppilaiden kirjoitelmissa on vaikeampi sanoa, onko kyseessä varsinainen virhe vai tahaton lipsahdus kuten esimerkissä 30.

- (30) Pyytäisin äitiä lukemaan tarinoita omasta lapsudestani ja sen jälkeen olisin serkkun kanssa teitokonella kuuntelemassa musiikkia (6F)

Esimerkin 30 kaltaisissa tapauksessa tosin virheen tulkitsemista varsinaiseksi virheeksi puoltaa se, että samassa kirjoitelmassa esiintyy jonkin verran erilaisia äännepituusvirheitä: esimerkissä on pitkästä vokaalista puuttuvien vokaalien lisäksi myös konsonantin geminoitumistapaus *serkkun*, jossa astevaihtelun heikko aste on jäänyt toteutumatta. Toisaalta samassa tekstissä kuitenkin on myös täysin norminmukaisia pitkiä vokaaleja ja geminaattoja.

Vokaalin puuttuminen pitkästä vokaalista on hieman yleisempää tyttöjen kirjoitelmissa kuin poikien kirjoitelmissa. Tyttöjen kokonaisvirheprosentti on 0,7 %, ja poikien 0,5 %. Kuten aiemmin on todettu, on tämän virheen esiintyminen aineistossa kaiken kaikkiaan kuitenkin niin marginaalista, ettei näiden esiintymien pohjalta voida vetää yleistäviä johtopäätöksiä eri sukupuolten kirjoitustaidosta.

### 3.3.3 Muu puuttuva kirjain

Muu puuttuva kirjain on mikä tahansa yksittäinen vokaali tai konsonantti, joka sanasta puuttuu. Tähän ryhmään on laskettu mukaan myös puuttuvien kirjaimien muodostamat tavut, joiden osuus koko aineistossa on 22 kappaletta. Puuttuva kirjain -kategoria ja virheellinen taivutus -kategoria menevät useilta osin päällekkäin, sillä esimerkiksi genetiivin *n* voi puuttua sanan lopusta joko siitä syystä, että sana on merkitty virheellisesti nominatiiviin, tai siitä, että *n* on tippunut tahattomasti sanan lopusta pois esimerkiksi huolimattomuusvirheen seurauksena. Nimenomaan sananloppuisten kirjainten kato on ryhmä, jossa selittäviä tekijöitä voi olla useita. Tähän kategoriaan on laskettu mukaan vain sellaiset tapaukset, jotka ovat yksiselitteisesti puuttuvia kirjaimia, eivät siis esimerkiksi virheitä, jotka saattavat johtua myös virheellisestä taivuttamisesta. Mahdollisesti virheellisestä taivuttamisesta johtuvat puuttuvat kirjaimet kuuluvat monitulkintaisten virheiden kategoriaan, jota käsitellään omana ryhmänä virheellisen taivuttamisen yhteydessä. Virheitä, jotka siis saattaisivat selittyä virheellisen taivuttamisen lisäksi myös puuttuvalla kirjaimella, on aineistossa yhteensä 64 kappaletta.

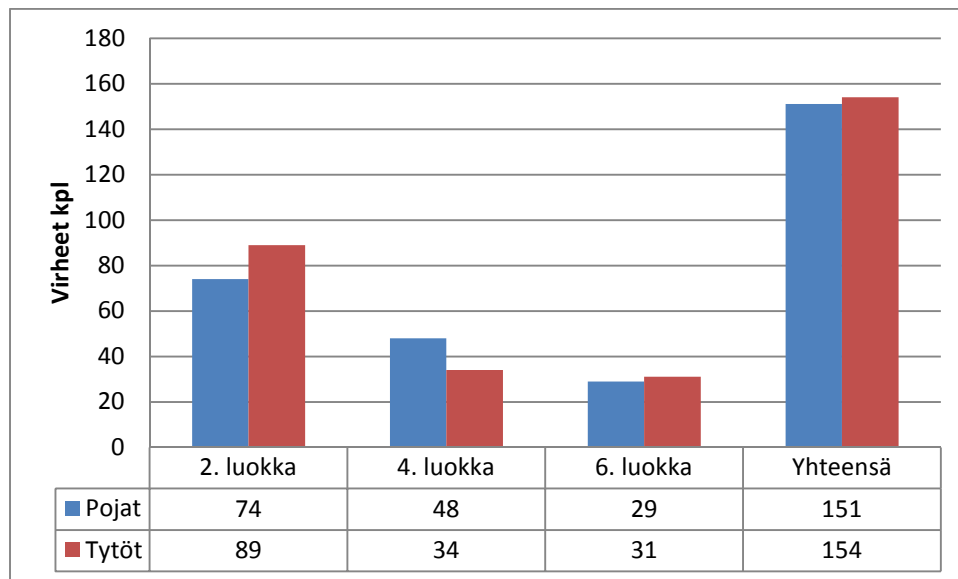
Tämän tutkimuksen virheanalyysi pohjautuu ensisijaisesti Ruoppilan, Römanin ja Västin (1969) virheanalyysiin. Vaikka Ruoppilan ym. virheanalyysiin pohjautuvilla testeillä on edelleen vankka asema kirjoitustaidon arvioinnissa, on se antamaa tietoa kriittisesti arvioitu muun muassa siksi, että se keskittyy arvioimaan kirjoitustaitoa liian yksiselitteisesti vain virheiden määrän kautta. Sen sijaan muun muassa Kulju ja Mäkinen (2009) ovat kehittäneet sanarakenneanalyysin, jossa virheiden määrien sijaan keskitytään siihen, millaisissa sana- ja tavuasemissa virheitä esiintyy ja mitkä rakenteet oppilaiden kirjoitelmissa säilyvät. Kuljua ja Mäkistä mukaillen tässä tutkimuksessa pyritään myös selittämään virheitä muun muassa sana- ja tavarakenteen kautta. Tämä koskee etenkin muu puuttuva kirjain -kategorian virheitä, sillä yksittäisten kirjainten puuttuminen liittyy hyvin usein tavoitesanan fonologiseen kompleksisuuteen.

Puuttuvien kirjainten kohdalla voi usein olla kyseessä myös puhekielinen tai murteellinen ilmaisu; puhekielen yleinen piirre on, että sanan lausumista yksinkertaistetaan usein lyhentämällä, ja näin sanan yleiskielisestä versiosta puuttuu joitain kirjaimia sen puhekielisessä variantissa. Tämän virheryhmän ulkopuolelle on jäänyt joukko yleisinä puhekielen piirteinä pidettäviä muotoja, jotka esitellään seuraavaksi. Ensinnäkin tutkimuksen ulkopuolelle on rajattu kaikki Salosen (2012) analysoimat puhekielisyydet: Hänen mukaansa pikapuhemuodot toteutuvat alakoululaisten kirjoitelmissa *minä*-pronomineissa, demonstratiivipronomineissa sekä *olla-* *mennä-* ja *tulla-*verbeissä. *A:n* loppuheitto toteutuu inessiivissä ja essiivissä, *kanssa*-postpositioissa, adverbeissa *siellä*, *vielä*, *ajoissa* ja *yhdessä*, *kyllä*-sanassa, konjunktioissa *mutta* ja *että*, pronomineissa *meillä*, *missä* ja *samassa*, kolmannessa infinitiivissä sekä yksittäisissä sanoissa *laskettelust* ja *ehdokkaat*. *i:n* loppuheitto toteutuu konditionaalissa, imperfektissä, translatiivissa sekä numeraaleissa *yksi*, *kaksi*, *viisi* ja *kuusi*. Jälkitavujen *i*-loppuisten diftongien jälkikomponenttien kato toteutuu aineistossa *inen-*loppuisissa nomineissa, *oin-*loppuisissa adverbeissa sekä *takaisin*-adverbissä, *oitta-* ja *ai-*verbeissä ja lisäksi konditionaalissa. Sananloppuisen *n:n* kato toteutuu adverbeissa *sitten*, *niin*, *paljon*, *vähän*, *varmaan*, *kotiin*, *suoraan*, *jotain*, *kahdestaan*, *kokonaan*, *nopeaan*, *vastakkain* ja *sisään*, liitepartikkeleissa *-kin* ja *-kAAn*, konjunktioissa *kun* ja *kuin*, postpositioissa *jälkeen* ja *kanssaan*, passiivin lopussa sekä kolmannen infinitiivin illatiivin lopussa. Sananloppuisen *t:n* kato taas toteutuu aktiivin 2. partisiipin nominatiivimuodoissa sekä *nyt*-adverbissa.

Tähän tutkimukseen valitun puuttuvan konsonantin piirrejoukon ulkopuolelle on rajautunut myös frekvenssiltään vähäinen joukko verbejä ja nomineja, joista yleiskielen *d* on kadonnut (*oottaa*, *nähä*, *yhellä*, *kaheksa*, *ees*). Lisäksi tarkastelun ulkopuolelle on jäänyt *tt:n* tai *t:n* esiintyminen *ts:n* vastineena (*mettä*, *katto*, *ite*). Yleiskielen *d:n* kato ja *ts:n* murteelliset vastineet on kontekstiin suhteutettuna mielekästä luokitella mieluummin puhekielisiksi tai murteellisiksi, joten ne on syytä pitää erillään varsinaisista virheistä. Mielikäisen (1991: 10–11) mukaan *ts:n* *tt*:lliset tai *t*:lliset variantit ovat laajalevikkisiä ja varsinkin läntisissä murteissa vallitsevia ja nuortenkin puheeseen siirtyviä piirteitä.

Yksitulkintaisia puuttuvia kirjaimia on aineistossa yhteensä 305 kappaletta, ja kun tähän lasketaan mukaan tapaukset, jotka voivat johtua myös virheellisestä taivuttamisesta, luku on 369. Luokkatasokohtaisesti nämä virheet jakautuvat seuraavasti (sulkujen sisään merkityssä luvussa mukana vertailun vuoksi myös mahdolliset virheelliset taivutukset): toisen luokan oppilaiden kirjoitelmissa 163 (195), neljännen luokan oppilaiden kirjoitelmissa 82 (106) ja kuudennen luokan oppilaiden kirjoitelmissa 60 (68). Virheiden määrästä huomaa, että

vanhempien oppilaiden kirjoitelmista puuttuu yksittäisiä kirjaimia huomattavasti vähemmän kuin nuorempien. Myös virheiden kaksitulkintaisuus vähenee luokka-asteelta toiselle siirryttäessä, mikä kertoo kielen ja kerronnan selkiytymisestä ylipäätään. Tämän virhekategorioiden taulukoissa esiintyvät luvut ovat vain yksiselitteisesti puuttuvista kirjaimista koostuvia. Kaaviossa 3 on esitetty muiden puuttuvien kirjainten jakautuminen eri luokka-asteiden ja sukupuolten välillä.



**Kaavio 3.** Muut puuttuvat kirjaimet luokka-asteittain ja sukupuolittain.

Kaaviosta 3 käy ilmi, että muu puuttuva kirjain -virhe on kaikkien yleisin aineiston nuorimmilla oppilailla: kaikista 305 puuttuvasta kirjaimesta yhteensä 163 kappaletta on toisen luokan oppilaiden kirjoitelmista. Puuttuvien kirjainten määrä vähenee selvästi siirryttäessä ylemmälle luokka-asteelle, ja neljäsluokkalaisten kirjoitelmissa niitä on 82 ja kuudesluokkalaisten 60 kappaletta. Sukupuolittainen virhejakauma näyttäisi melko tasaiselta, mutta pelkät virhemäärät ovat tässä kohtaa myös hieman harhaan johtavia, sillä luokka-astekohtaiset ja sukupuolikohtaiset aineistot ovat hyvin erikokoisia. Muita puuttuvia kirjaimia tarkastellaan tästä eteenpäin kolmena alaryhmänä: laajemmin puuttuvina konsonanteina ja puuttuvina vokaaleina sekä lopuksi suppeammin puuttuvina tavuina.

Puuttuvat konsonantit on jaettu edelleen yksinäiskonsonantteihin, jotka puuttuvat sanan lopusta tai sanan sisältä. Puuttuva sanansisäinen yksinäiskonsonantti merkitsee tässä aineistossa yksinäiskonsonanttia, joka puuttuu joko sanan alusta tai keskeltä, toisin sanoen

kahden vokaalin välistä Sanansisäiset konsonantit voivat puuttua myös konsonanttiyhtymistä. Konsonanttiyhtymät, joista konsonantti puuttuu, taas voidaan jakaa tavunsisäisiin ja tavunrajaisiin konsonanttiyhtymiin. Lisäksi on eritelty erikseen sellaiset *j*-kirjaimet, jotka puuttuvat kirjoitelmissa tekijännimistä. Tarkempi puuttuvien konsonanttien asema jakauma selviää taulukosta 8.

**Taulukko 8.** Puuttuvien konsonanttien esiintyminen eri sana- ja tavuasemissa luokka-asteittain.

Luokka-aste	Sananloppuinen konsonantti	Tavunsisäinen konsonanttiyhtymä	Tavunrajainen konsonanttiyhtymä	Yksinäis-konsonantti	<i>j</i> tekijänimistä	Yhteensä
2.	37	15	19	9	8	88
4.	7	2	7	10	17	43
6.	5	1	8	2	6	22
<b>Yhteensä</b>	49	18	34	22	30	153

Puuttuvien konsonanttien asemien vertailu osoittaa, että konsonantti puuttuu useimmiten sanan lopusta; tämä tyyppi on toisen luokan oppilaiden kirjoitelmissa yleisin syy konsonantin puuttumiseen, mutta määrä pienenee radikaalisti ylemmille vuosiluokille siirryttäessä. Sananloppuinen puuttuva konsonantti on useimmiten *n*: jopa 35 tapauksessa 49:stä on kyse sananloppuisen *n:n* puuttumisesta. Tätä selittää ensinnäkin se, että suomen kielen monet tunnukset ja päätteet ovat *n*-loppuisia. Etenkin aineiston nuorimpien oppilaiden kirjoitelmissa näiden tunnusten ja päätteiden merkitseminen on usein hieman vaillinaista kuten esimerkeissä 31 ja 32.

(31) Kun minä menin katsomaan *kuuse* alle siellä oli kauhea kasa lahjoja (2F)

(32) Kun me oltiin perillä oli ihana tulla *siistii* kotiin kun ennen lähtöä me siivottiin (2F)

Toinen selittävä tekijä on se, että puheessa sananloppuinen *n* katoaa hyvin usein. Tämän virhekategorioiden ulkopuolelle on jo jätetty laaja joukko nimenomaan sananloppuisen *n:n* katotapauksia ensinnäkin siksi, että ne on Salosen pro gradu -tutkielmassa (2012) jo tulkittu puhekielen piirteiksi ja toisekseen siksi, että sananloppuisen *n:n* kato voi selittyä myös virheellisellä taivutuksella. Tähän tutkimukseen on kuitenkin laskettu mukaan sellaiset sananloppuiset *n:n* kadot, joita Salonen ei ole analysoinut omassa tutkimuksessaan ja jotka



eivät voi johtua virheellisestä taivuttamisesta. Etenkin joukko *n*-loppuisia adverbeja on hyvin altis nuorimpien oppilaiden kirjoitelmissa sananloppuisen *n:n* kadolle kuten esimerkissä 33.

(33) Äiti kertoi että lähemme *tänää* kylpylään ja olemme yötä hotllissa (2F)

Kyseessä on siis useimmiten puhekielen vaikutuksesta syntynyt virhe, mikä kertoo siitä, että etenkin aineiston nuorimmilla oppilailla on pientä horjuvuutta norminmukaisen ja norminvastaisen kielimuodon erottamisessa.

Toinen sanan lopusta helposti oppilaiden kirjoitelmissa katoava konsonantti on *t*. Sananloppuisten konsonanttien 49 esiintymästä on kyseessä *t:n* kato 12 kertaa. Samoin kuin sananloppuisen *n:n* katoa, myös *t:n* katoa selittää se, että jotkin suomen kielen päätteet ja tunnuksen ovat *t*-loppuisia. Etenkin monikon 3. persoonan persoonapäätteestä *t* on altis katoamaan toisen luokan oppilaiden kirjoitelmissa kuten esimerkki 34 osoittaa.

(34) *hepäättivä* mennä kumituskartanoon he löysivät älämölönpoistoaineen sitten *hepalasiva* kotiin (2M)

Kuten esimerkki 34 osoittaa, sananloppuinen *t* voi kadota kirjoitelmissa sellaisesta ympäristöstä, jossa se samassa asemassa on myös säilynyt. Oppilas siis todennäköisesti tietää, että sanan loppuun kuuluu kirjain *t*, koska hän ei jätä sitä järjestelmällisesti kaikista tapauksista pois. Tämä kertoo siitä, että sananloppuisen *t:n* kato on luonteeltaan todennäköisesti huolimattomuusvirhe kuin esimerkiksi sananloppuisen *n:n* kato.

Kuten edellä on todettu, sananloppuisten kirjainten katoa voivat selittää monenlaiset seikat. Ensinnäkin on oppilas saattanut taivuttaa sanan väärin tai ei ole tiennyt, miten sana kuuluisi taivuttaa (35). Toiseksi voi kyseessä olla puhekielinen tai murteellinen ilmaus, jossa sana on lyhentynyt lopusta päin (36). Tätä tulkintaa tukee se, että lyhentyneitä sanoja esiintyy etenkin kirjoitelmissa, joissa on muutenkin runsaasti puhekielisiä ilmauksia. Kaikkein uskottavin selitys sananloppuisten kirjainten puuttumisen kannalta kuitenkin näyttäisi olevan huolimattomuusvirhe (37), sillä näiden virheiden esiintyminen tietyssä ympäristössä tietyssä kirjoitelmassa ei ole kovin johdonmukaista.

(35) Katsoin Amazing *race* ja Selviytyjät (4M)

(36) ehkä juteltais koska nähään seuraavan *kerra* (6F)

(37) Ja stten kun herään *saa iso* hodarin ja kylmää kokakolaa (2F)

Esimerkkien 35 ja 37 kaltaisia tapauksia ei siis kuitenkaan ole laskettu mukaan lopulliseen puuttuvien kirjainten virhekatteeseen niiden monitulkintaisuuden vuoksi. Kuten esimerkeistäkin voi huomata, on näiden kohdalla kuitenkin useimmiten todennäköisemmin kyse juuri puuttuvasta kirjaimesta kuin virheellisestä taivutuksesta tai sen puutteesta.

Luokka-asteittainen puuttuvien konsonanttien paikkojen vertailu osoittaa, että aineiston nuorimpien oppilaiden kirjoitelmista puuttuva konsonantti on hyvin usein konsonanttiyhdyntä. Tavunsisäisten ja tavunrajaisien esiintymien välillä ei ole merkittävää eroa. Konsonanttiyhdyntät mielletään yhdeksi suomen kielen ongelmarakenteeksi, joten ei ole yllättävää, että niiden merkitsemisessä on ongelmia etenkin aineiston nuorimmilla oppilailta.

Tavunsisäiset konsonanttiyhdyntät voidaan jakaa tavunalkuisiin ja tavunloppuisiin konsonanttiyhdyntäisiin. Tavunalkuiset konsonanttiyhdyntät, joista konsonantti puuttuu, esiintyvät yleensä vain sananalkuisessa asemassa kuten esimerkissä 38. Tavunloppuiset konsonanttiyhdyntät ovat osa laajempaa tavunrajaisista konsonanttiyhdyntästä, jonka loppu sijaitsee seuraavan tavun puolella. Tavunloppuisista konsonanttiyhdyntästä puuttuukin usein konsonanttiyhdyntän ensimmäinen (yksinäis)konsonantti kuten esimerkissä 39.

(38) haluaisin robo raktorin kauko ohjattavan ja ison kasan lunta (2F)

(39) Pääsimme hädintuskin sisälle kun Misku hyökkäsi meidän kippuun (2F)

Tavunrajaisissa konsonanttiyhdyntäisissä on hieman enemmän virheitä kuin tavunsisäisissä konsonanttiyhdyntäisissä kaikilla luokka-asteilla: tavunsisäisistä konsonanttiyhdyntäisistä konsonantteja puuttuu 18 kappaletta, kun tavunrajaisista konsonanteista niitä puuttuu yhteensä 34 kappaletta. Myös tavunrajaisista konsonanttiyhdyntäisistä puuttuvien konsonanttien määrä kuitenkin pienenee vertailtaessa aineiston nuorimpien ja vanhimpien oppilaiden virhemääriä keskenään.

(40) menimme myös syömään ja olime pajon ulkona ja myöskin puistossa ja tultiin vihdoikin kotiin (2M)

Sanasta puuttuva konsonantti on harvemmin yksinäiskonsonantti. Vanhempien oppilaiden kirjoitelmissa yksinäiskonsonanttien puuttuminen sanasta ei vaikuta mitenkään johdonmukaiselta, vaan hyvin satunnaiselta. Tällaiset virheet ovat siis todennäköisimmin huolimattomuusvirheitä. Jonkin verran yksinäiskonsonantteja puuttuu myös sellaisista sanoista, joista konsonantin lisäksi puuttuu myös muita kirjaimia kuten esimerkissä 41. Tällaiset virheet saattavat olla niin ikään huolimattomuusvirheitä, mutta ne saattavat liittyä

myös kirjoitustaidon kehittymisessä ilmeneviin ongelmiin. Myös aineiston nuorimpien oppilaiden kirjoitelmista puuttuu jonkin verran yksinäiskonsonantteja, mutta näiden esiintymien kohdalla on hyvin yleistä, että samassa sanassa on useampia virheitä kuten esimerkissä 42. Nuorimmat oppilaat kirjoittavat hitaammin ja näin ollen huolellisemmin kuin vanhemmat oppilaat, joten yksinäiskonsonantin puuttuminen huolimattomuusvirheen seurauksena ei ole yhtä todennäköinen selitys heidän kohdallaan, joten nämä virheet kertovat todennäköisesti kirjoitustaidon kehittymättömyydestä ja vaikeudesta hahmottaa sanassa olevia äännteitä.

- (41) Sitten huudettiin *ensimmälle* palkinto pallille Juoksu kilpailun voittaja on allekirjoittaneen nimi (4F)
- (42) Aamulla kun *häerin* (po. heräsin) puin päälle ja veinkoiran ulos (2F)

Varsinkin vanhempien oppilaiden kirjoitelmissa taas puuttuva konsonantti on usein tekijäsanoista puuttuva *j*, joka saattaa äännettäessä liudentua *i*:ksi kuten esimerkeissä 43 ja 44 näyttää tapahtuneen. Kyseessä on todennäköisesti puhtaasti ääntämisestä johtuva virhe, joten *j*:n merkitsemättä jättäminen on hyvin erityyppinen virhe kuin esimerkiksi konsonantin puuttuminen konsonanttiyhtymästä.

- (43) Unelma päivä olisi päivä Tanskassa jonkun Elokuvan koekuvauksissa *Näytteliänä* (6F)
- (44) Katson ikkunasta ja näen *palvelioideni* putsaamassa uimaallasta Yksi *palvelioistani* tuo minulle kaikkea hyvää tarjottimella Syön aamiaiseni ja sanon *palvelialleni* että Laita pore amme valmiiksi (4M)

Virheitä tekijännimissä esiintyy melko johdonmukaisesti kaikilla aineiston luokka-asteilla, ja myös tyypillistä on, että samassa kirjoitelmassa kaikki tekijännimet on kirjoitettu virheellisesti *j*:ttömiksi kuten esimerkissä 44. Tämä kertoo myös siitä, että oppilas todennäköisesti luulee, että hänen käyttämänsä sananmuoto olisi norminmukainen.

Sanasta puuttuvat muut vokaalit on jaettu tässä tutkimuksessa sanansisäisiksi lyhyiksi vokaaleiksi, sananloppuisiksi vokaaleiksi ja diftongeista puuttuviksi vokaaleiksi. Sanansisäinen lyhyt vokaali voi siis puuttua vokaaliyhtymästä eli tavunrajaisesta vokaalijonosta tai kahden konsonantin välistä. Tarkempi puuttuvien vokaalien sana- ja tavuasemainen erittely ilmenee taulukosta 9.

**Taulukko 9.** Sanasta puuttuvien vokaalien esiintyminen eri sana- ja tavuasemissa luokka-asteittain.

Luokka-aste	Sanansisäinen lyhyt vokaali	Vokaali diftongista	Sananloppuinen vokaali	Yhteensä
2.	29	32	7	68
4.	15	16	2	33
6.	8	20	1	29
<b>Yhteensä</b>	52	68	10	130

Taulukosta 9 ilmenee, että myös vokaalien kohdalla virheiden määrä vähenee huomattavasti siirryttäessä alemmilta vuosiluokilta ylemmille. Sanansisäisten lyhyiden vokaalien puuttuminen on hyvin samantapaista kuin sanansisäisten yksinäiskonsonanttien puuttuminen; hyvin usein sanasta, josta puuttuu sanansisäinen yksinäisvokaali, puuttuu myös muita kirjaimia, ja teksti sisältää muutenkin runsaasti virheitä, mitä havainnollistaa esimerkki 45.

- (45) mä ja mun sisko nukutin siellä ja syötin ja käutin kaupois lähetetin kortia *lasktin* vesi liukumäestä (2F)

Sanansisäinen lyhyt vokaali ei siis juurikaan esiinny ainoana virheenä sanassa, saati kirjoitelmassa, ellei se todennäköisesti ole huolimattomuusvirhe. Lyhyen vokaalin merkitseminen ei olekaan varsinainen ongelmarakenne, joka itsessään kertoisi jotain oppilaiden kirjoitustaidosta. Tällaisia ongelmarakenteita sen sijaan ovat diftongit, joista puuttuvat vokaalit muodostavatkin aineiston suurimman puuttuvien vokaalien ryhmän kaikilla luokka-asteilla. Vokaalin puuttuminen diftongista on kuitenkin huomattavasti yleisempää toisen luokan oppilaiden kirjoitelmissa kuin vanhempien oppilaiden; toisen luokan aineistossa puutteellisia diftongeja esiintyy 32 kertaa, mikä on kaksinkertainen määrä neljäsluokkalaisten virheiden (16) määrään verrattuna. Kuudesluokkalaisten kirjoitelmissa vaillinaisia diftongeja esiintyy 20 kappaletta. Diftongien ongelmallisuudesta kertoo se, että on hyvin tyypillistä, että samassa kirjoitelmassa esiintyy useampia virheellisiä diftongeja kuten esimerkit 46 ja 47 osoittavat.

- (46) syötyämme lähdimme Särkänniemeen lopupiväksi kävimme melkein jokaliteessa (po. joka laitteessa) (2F)
- (47) Unelmieni päivä olisi se että päivä olisi oikein pitkä – – kun paisime takasin vahtaisime rahat euroksi ja olisimme TRUE rikkaita olisi hieno talo oma minigolf rata kellaris keilahalli ja perhe olisi muutenkin rikas eika velosta olisi haittaa (6M)

Virheen toistuminen saman kirjoitelman sisällä kertoo siitä, että oppilaalla on ongelmia nimenomaan kyseisen rakenteen hallinnassa. Näin ollen diftongien merkitsemistä voidaankin pitää yhtenä kirjoitustaitoa ja sen kehittymistä mittaavana piirteenä. Koska virheet diftongeissa vähenevät radikaalisti verrattaessa aineiston nuorimpien oppilaiden kirjoitelmia vanhempien oppilaiden kirjoitelmiin, voidaan olettaa, että kirjoitustaito ja kyky hahmottaa sanassa esiintyviä äänneitä kehittyvät iän ja kirjoituskokemusten myötä.

Vokaaleja puuttuu sanan lopusta hyvin harvoin, ja 10 esiintymästä 7 on toisen luokan oppilaiden kirjoitelmista. Toisen luokan oppilaiden kirjoitelmien puuttuvista sananloppuisista vokaaleista lähes kaikki esiintyvät ympäristössä, jossa sanat on kirjoitettu virheellisesti yhteen ja sanavälin lisäksi ensimmäisestä sanasta puuttuu siis myös viimeinen vokaali kuten esimerkeissä 48 ja 49.

(48) Olisimme *kaikk*pelihahmoja (2M)

(49) sitten mä sain sieltä paukkuksesta sellasen koiran pennun *joksyö* kulhosta ja näyttää kieltä (2F)

Sananloppuinen vokaali ei siis ole yksinään esiintyvä virhetyyppi, vaan se liittyy yleensä juuri sanarajan hahmottamisen ongelmaan.

Joistain sanoista puuttuu useampia kirjaimia, joskus jopa kokonaisia tavuja. Puuttuviksi tavuiksi ei ole laskettu selkeitä kongruenssivirheitä, joissa siis monikon kolmannen persoonan päätte on jätetty merkitsemättä (esim. *ne esiintyisi* »he esiintyisivät«), vaan nämä tapaukset on tulkittu yksiselitteisesti virheellisen taivutuksen ryhmään kuuluviksi. Kaikista muu puuttuva kirjain -kategorian virheistä 22 on puuttuvan tavun virheitä. Niistä 7 kappaletta esiintyy toisen luokan oppilaiden kirjoitelmissa, 6 kappaletta neljäsluokkalaisten kirjoitelmissa ja 9 kappaletta kuudesluokkalaisten kirjoitelmissa. Kyseessä on siis frekvenssiltään hyvin pieni virheryhmä, mutta siitä on syytä tehdä muutama huomio: Ensinnäkin puuttuvat tavut puuttuvat yleisesti pitkistä ja rakenteellisesti haastavista sanoista sellaisissa teksteissä, joissa oikeinkirjoitusvirheitä on muutenkin verrattain runsaasti. Toiseksi on erittäin yleistä, että sanasta puuttuva tavu muistuttaa läheisesti jotain toista sanassa olevaa tavua tai esiintyy jopa sellaisenaan tavoitellussa sanamuodossa useammin kuin kerran.

(50) Ja ampua illalla 5 metriä pitkän uuden vuoden raketin kuuden metrin *sytyt langalla* ja lähteä sillä hämeen koskelle (2M)

- (51) Lätisimme aamulla voittamalla perheen *lomatalla* lämpimään maahan ja menisimme sinne laivalla (6M)

Molemmat esimerkit ovat tyypillisiä puuttuvien tavujen tapauksia ja ilmentävät niiden esiintymisympäristöjen vaikutusta virheisiin. Puuttuvat tavut eivät esiinny juurikaan ainoana virheinä yksittäisessä sanassa saati koko kirjoitelmassa. Lisäksi pitkien sanojen tavuomissio on yksi lapsen äänteellisen kehityksen fonologisista prosesseista, jolla kompleksisia rakenteita pyritään yksinkertaistamaan.

Muiden puuttuvien kirjainten yleinen eri tavu- ja sanarakenteissa esiintymisen tarkastelu osoittaa, että kirjaimia puuttuu usein konsonanttiyhtymistä ja diftongeista. Konsonanttiyhtymät ja diftongit ovatkin Ahvenaisen ja Holopaisen (2005: 107) mukaan suomen kielen ongelmarakenteita, joissa kirjoitettaessa tehdään paljon virheitä. Näin ei olekaan yllättävää, että juuri näissä ongelmarakenteissa esiintyy etenkin nuorimpien oppilaiden kirjoitelmissa paljon virheitä. Samalla nämä havainnot tukevat Kuljun ja Mäkisen (2009) oletuksia sanahahmon vaikutuksesta oppilaiden tekemiin virheisiin.

### 3.4 Väärä kirjain

Väärän kirjaimen virheet on jaettu kahteen ryhmään: väriin konsonantteihin ja väriin vokaaleihin. Lisäksi aineistossa on yhteensä 8 tapausta, joissa kyse ei ole kyse vokaalin korvautumisesta toiselle vokaalilla tai konsonantin korvautumisesta toisella konsonantilla, vaan konsonantti on korvattu vokaalilla tai vokaali konsonantilla. Nämä ovat lähes poikkeuksetta i/j-vaihteluja, ja niitä käsitellään väärän vokaalin yhteydessä, sillä on yleisempää, että j-kirjain on korvattu i-kirjaimella kuin toisin päin. Vääriä kirjaimia on aineistossa kaiken kaikkiaan 418 kappaletta, mikä on noin 16 % aineiston kaikista virheistä.

#### 3.4.1 Väärä konsonantti

Väärät konsonantit on tässä tutkimuksessa analysoitu sointivirheiksi, piirtovirheiksi, m/n-sekaannuksiksi tai konsonanttiassimilaatioiksi, joka tarkoittaa yksittäisen konsonantin korvautumista toisella, sanassa muutoinkin esiintyvällä konsonantilla. η-äänteen virheitä ei ole kategorisoitu perinteisestä virheanalyysistä poiketen omaksi virhetyypikseen, sillä niitä

esiintyy tässä aineistossa verrattain vähän, vaan ne on luokiteltu väärin konsonanttien yhteydessä sointivirheiksi.

Vääriä konsonantteja esiintyy aineistossa kaiken kaikkiaan 175 kertaa: toisen luokan oppilaiden kirjoitelmissa 52 kertaa, neljännen luokan oppilaiden kirjoitelmissa 46 kertaa ja kuudennen luokan oppilaiden kirjoitelmissa 77 kertaa. Sukupuolittain väärät konsonantit jakautuvat siten, että poikien kirjoitelmista niitä on poimittu 98 kappaletta ja tyttöjen kirjoitelmista 77 kappaletta.

Sointivirheitä on eniten aineiston nuorimmilla kirjoittajilla. Virheet  $\eta$ -äänteessä on rinnastettu väärin konsonanttien osalta sointivirheisiin (*sänkyn* »sängyn»), vaikka mukana on muutama sellainen tapaus, joissa ei ole kyse puhtaasti sointivirheestä (*kuniggas* »kuningas»). Viimeksi mainitun kaltaisia  $\eta$ -äänten virheitä on koko aineistossa vain kaksi kappaletta, joten niiden rinnastaminen ensiksi mainitun kaltaisiin sointivirheisiin on järkevää. Sointivirhe on usein seurausta toteutumattomasta astevaihtelusta (52) tai vierasperäisten äänteiden korvaamisesta tutummalla omaperäisellä äänteellä (53). Jonkin verran varsinkin nuorempien oppilaiden teksteissä sointivirhe on kääntynyt ikään kuin väärin päin, ja soinniton klusiili on korvattu soinnillisella klusiililla (54). Tällaiset tapaukset on tässä analyysissä luokiteltu myös kuuluviksi sointivirheisiin.

(52) sen jälkeen menimme äidin ja isän kanssa katulle ottamaan aurinkoa (2F)

(53) meille ei osteta kissaa koska minun kummitätini on allerkinen kissoille (2M)

(54) Mentäis lappiin laskettelemaa Ja saisajaa lastemootorigelkalla (2M)

Sointivirheiden yleisyys nuorimmilla oppilailta johtunee kielitajun kehittymättömyydestä. Vieraat äänteet eivät normaalisti kuulu vielä pienen lapsen kielitajuun, vaan puheessakin ne korvataan yleensä tutulla äänteellä. Niiden merkitseminen kirjoitettuun kieleen vaatinee parempaa käsitystä suomen kielen äännejärjestelmästä.

Piirtovirhe konsonantissa ilmenee tässä aineistossa l/t-sekaannuksena. Useimmiten kyse on siis siitä, että t:n poikkiviiva on jäänyt merkitsemättä, ja näin lopputulos muistuttaa l-kirjainta (55).

(55) Eräs ystävämme haki meidät bussiasemalle puoli kahdeksal/a Matka Helsinkiin kes/i  
56 tuntia (6M)

Kaikilla vuosiluokilla ilmenee myös runsaasti m/n-sekaannusta. Melko yleistä on, että *m* on korvautunut *n*:llä adjektiivin komparatiivimuodossa (56). Virhe saattaa osittain liittyä myöskin käsialan kehittymättömyyteen, sillä käsialasta riippuen ei *m:n* ja *n:n* ero ole välttämättä selvä vaan yleensä ja kirjoittamistyylistä vaihdellen tulkinnallisia. Tällaiselta vaikuttavat etenkin virheet kirjoitelmissa, joissa m/n-sekaannusta on johdonmukaisesti läpi koko tekstin mutta muita virheitä ei juurikaan ole (57). Lisäksi m/n-sekaannus johtuu osittain selvästi sanan ääntöasusta esimerkiksi nasaaliassimilaation vuoksi kuten esimerkissä 58.

- (56) Illalla meri näytti vielä kauniinmalta kuin päivällä (4F)
- (57) Kahden viikon päästä lähtisinne takaisin Suomeen ja joutuisinne istua taas noin yhdeksän tuntia (6M)
- (58) kumpu se olisi totta (2F)

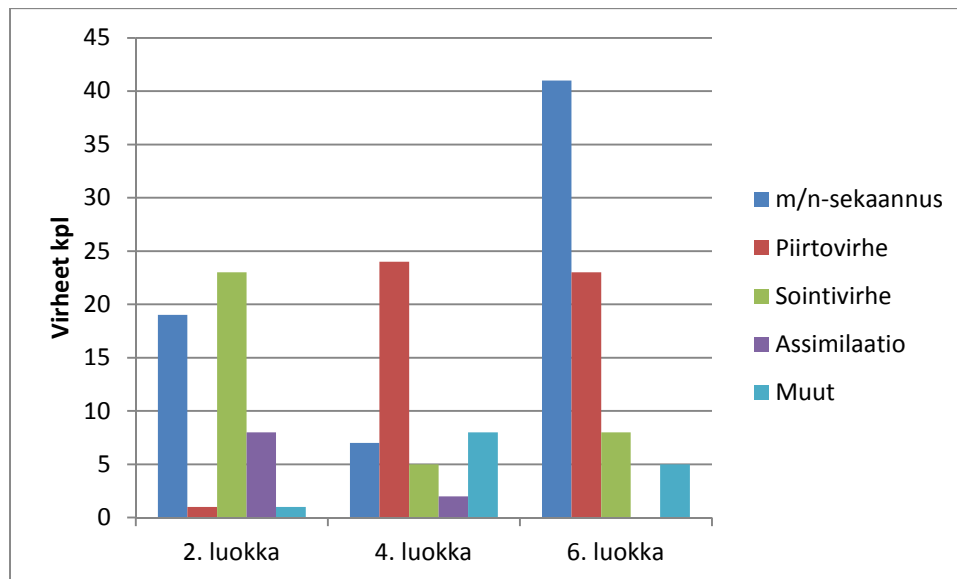
Konsonanttiassimilaatiossa norminmukainen konsonantti on korvautunut toisella, sanassa edempänä tai jäljempänä esiintyvällä konsonantilla kuten esimerkissä 59.

- (59) sitten haluaisin leikkiä pikku *sikkoni* kanssa pitkään ja saada barbin (2F)

Konsonanttiassimilaatio on nuorimpien oppilaiden kirjoitelmissa yleisintä juuri konsonanttiyhtymissä niin, että konsonanttiyhtymä korvautuu loppujen lopuksi geminaatalla, mikä tarkoittaa tavoitellun konsonanttiyhtymän yksinkertaistamista. Konsonanttiassimilaatio on yksi lapsen kielen fonologisista prosesseista, joilla hän pyrkii helpottamaan muutoin vaikeiden sanojen ääntämistä. Kirjotukseen nämä assimiloituneet muodot ovat voineet siirtyä myös fonologisen prosessin tuloksena oppilaan kirjoittaessa sanaa äänne äänteiltä. Ne voivat johtua myös tarpeesta yksinkertaistaa sanaa visuaalisesti: kahden saman merkin tuottaminen kahden eri merkin asemesta yksinkertaistaa sanan ulkomuotoa.

Luokka-asteittain on mielenkiintoista vertailla, mistä väärät konsonantit johtuvat. Kuten kaaviosta 4 selviää, toisen luokan oppilaiden väärät konsonantit ovat suurimmaksi osaksi sointivirheitä, neljäsluokkalaisten piirtovirheitä ja kuudesluokkalaisten m/n-sekaannuksia.





**Kaavio 4:** Väärien konsonanttien selitykset luokka-asteittain.

Syyt väärän konsonantin merkitsemiseen vaihtelevat luokka-asteittain erittäin paljon. Sointivirhe on selvästi yleisin toisen luokan oppilaiden kirjoitelmissa: yhteensä 23 kappaletta kaikista 52 väärästä konsonantista selittyy sointivirheellä, kun sitä esiintyy neljäs- ja kuudesluokkalaisten kirjoitelmissa huomattavasti vähemmän. Piirtovirheitä taas toisen luokan oppilaat eivät juurikaan tee, mutta aineiston neljäsluokkalaisten 46 väärästä konsonantista 24 on piirtovirheitä ja kuudesluokkalaisten 77 väärästä konsonantista piirtovirheellä selittyy 23 kappaletta. m/n-sekaannus on huomattavasti yleisempi toisen luokan oppilailta kuin neljännen luokan oppilailta. Epäodotuksenmukaista sen sijaan on, että m/n-sekaannus on kuudesluokkalaisten yleisin väärän konsonantin selitys. Kuudesluokkalaisten runsaan m/n-sekaannusten määrän tosin selittää yksittäinen kuudesluokkalaisten pojan kirjoitelma, jossa m/n-sekaannuksia on yhteensä 29 kappaletta, eli lähes kolme neljäsosaa koko kuudennen luokan kirjoitelmien m/n-sekaannuksista. Kyseisessä kirjoitelmassa m/n-sekaannus saattaa selittyä hankalasti tulkittavalla käsialalla, tai vaikeudella hahmottaa *m*- ja *n*-kirjaimet visuaalisesti tai äänteellisesti toisistaan eri kirjaimiksi. Tässä tekstissä *m*-kirjaimet on hyvin usein korvattu *n*-kirjaimella tai geminaatta mm:stä puuttuu toinen kuten esimerkissä 60. Lisäksi tekstissä on jonkin verran muitakin virheitä muun muassa äännepituuksissa ja sanarajoissa.

- (60) kiertää *ynpäri* *anerikkaa* ja pysähtyä välillä skeitata *ynpäri* paikpelupeja ja sää tääsielä ja kun *olisinne* kierrelleet *ynpäri* *anerikkaa* niin *tapaisime* sielä huippu skeittaaaja Ja *lähtisimme* heidän kanssa skeitataa maan ja säättään sinne kun *palaisinne* lentokentälle ja ja *lentäisime* lomalle johonkin päin ehkä thaimaaseen lonalle (6F)

Jos ei lasketa mukaan tämä yhden kirjoitelman m/n-sekaannuksia, kuudesluokkalaisten m/n-sekaannusten määrä laskee 12:een. Näin ollen myös m/n-sekaannus olisi yleisempi virhe aineiston nuorimmilla kirjoittajilla kuin ylemmillä vuosiluokilla. Tällöin kuudesluokkalaistenkin yleisimmäksi väärän konsonantin selittäjäksi nousisi piirtovirhe.

Väriä konsonanttien luonteen muuttuminen alimmilta luokka-asteilta ylemmille edetessä kertoo kahdenlaisesta kehityssuunnasta alakoululaisten kirjoitustaidoissa: Ensinnäkin, sointivirheiden väheneminen ylemmille vuosiluokille siirryttäessä kertoo suomen kielen äännetajun kehittymisestä; tietoisuus astevaihtelusta sekä vahvoista ja heikosta äännepareista parantuu. Toiseksi, piirtovirheiden huomattavan lisääntymisen ylemmille luokille edetessä voidaan tulkita tarkoittavan huolimattomuusvirheiden lisääntymistä. Toisen luokan oppilas keskittyy jokaisen kirjaimen merkitsemiseen vanhempia oppilaita tarkemmin, sillä koko aakkosjärjestelmä on heille verrattain uusi. Sen sijaan kokeneempi, alakoulun viimeisillä luokilla oleva oppilas saattaa keskittyä enemmän itse sisällön tuottamiseen ja kirjoittaessaan nuorempia nopeammin tekee helposti huolimattomuusvirheitä eikä kiinnitä yhtä tarkasti huomiota yksittäisten kirjainten muotoon.

### 3.4.2 Väärä vokaali

Väärät vokaalit ovat yksittäisiä norminvastaisia vokaaleja, jotka esiintyvät joko lyhyissä vokaaleissa, vokaaliyhtymissä tai diftongeissa. Tämän ryhmän ulkopuolelle ovat jääneet selvimmän yleiset ja neutraalit, levikiltään laaja-alaiset puhekielen ja murteen vaikutuksesta syntyneet muodot. Salonen (2012) on jo omassa tutkimuksessaan samasta aineistosta analysoinut jälkitavujen *A*-loppuisten vokaaliyhtymien (*eA*, *OA*, *UA* ja *iA*) monoftongiutuneet muodot, jotka ovat etenkin nuorten käyttämä yleistävä puhekielen piirre (Salonen 2012: 67). Kyseiset muotoryhmät on siis jätetty tämän tutkimuksen ulkopuolelle. Sen sijaan tähän ryhmään on laskettu mukaan etenkin länsimurteille tyypilliset diftongin avartumistapaukset, joita aineiston virheistä on 14 kappaletta, sekä lounais- ja savolaismurteissa esiintyvät diftonginreduktiot, joita aineiston virheistä on 6 kappaletta. Diftongin avartuminen ja diftonginreduktio ovat selvemmin alueellisia murrepiirteitä kuin esimerkiksi analyysin ulkopuolelle jäänyt *A*-loppuisten vokaaliyhtymien monoftongiutuminen, ja niitä pidetään

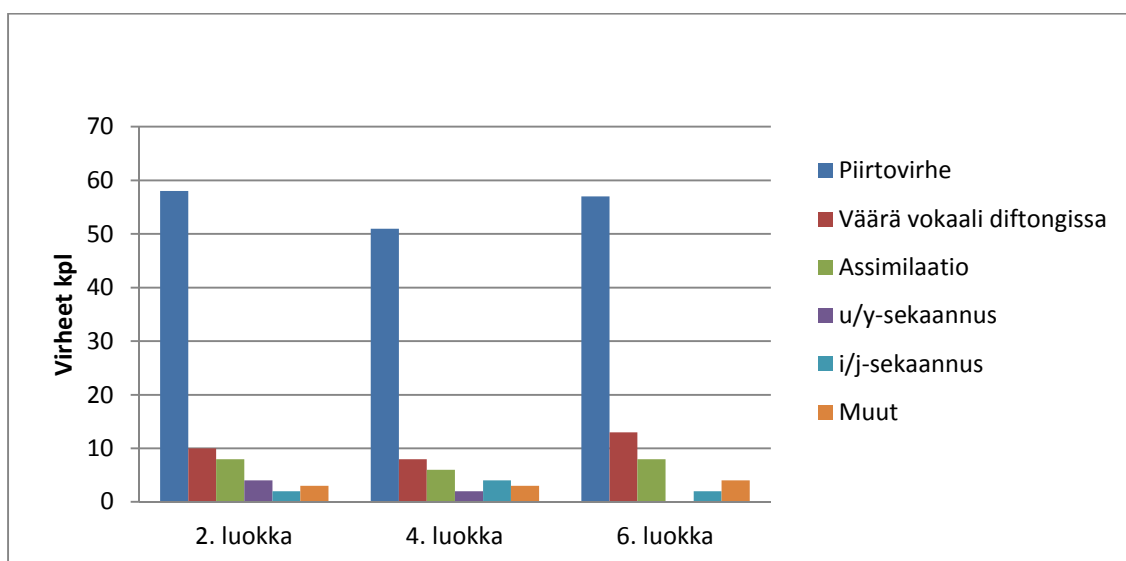
myös alueellisesti hyvin leimaavina piirteinä eivätkä ne lukeudukaan Mantilan (1997) tai Mielikäisen (1991) jaotteluiden mukaisiin nykypuhekielen piirteisiin. Koska kyseisiä muotoja ei lasketa yleistyvien tai laaja-alaisten puhekielisyyspiiriin kuuluviksi, niitä voidaan pitää selvemmin norminvastaisina kuin yleisiä puhekielen piirteitä. Niinpä ne on sisällytetty tämän tutkimuksen analyysiin. Molemmat muotoryhmät ovat myös aineiston kannalta frekvenssiltaan niin pieniä, etteivät ne aiheuta tuloksissa vinoutumia.

Väriä vokaaleja on aineistossa yhteensä 247, joista tosin 4 saattaa selittyä myös virheellisellä taivuttamisella (61). Näissä tapauksissa siis yhtä vokaalia muuttamalla muuttuu myös sanan sijamuoto, joka on lausekontekstin kannalta virheellinen.

(61) peikko ei uskaltanut lentää lentokoneelle Koska ihmiset saattaisivat nähdä peikon (2F)

Mahdollisella virheellisellä taivutuksella selittyvien väärin vokaalien määrää ei ole laskettu mukaan varsinaiseen väärin vokaalien virhekatteeseen, vaan nämä 4 tapusta käsitellään taivuttamiseen liittyvien virheiden yhteydessä monitulkintaisina virheinä luvussa 3.6.3.

Luokkatasokohtaisesti väärät vokaalit jakautuvat seuraavasti: toisen luokan oppilaiden kirjoitelmissa 85 kappaletta, neljännen luokan oppilaiden kirjoitelmissa 74 kappaletta ja kuudennen luokan oppilaiden kirjoitelmissa 84 kappaletta. Yksiselitteisistä 243 väärästä vokaalista 132 kappaletta on poikien kirjoitelmista ja 111 tyttöjen kirjoitelmista. Väärän vokaalin merkitsemistä voidaan selittää monenlaisilla tekijöillä, ja kaaviosta 5 käykin ilmi, miten virheelliset vokaalit mahdollisesti selittyvät luokkatasokohtaisesti.



**Kaavio 5:** Väärin vokaalien selitykset luokka-asteittain.

Niissä tapauksissa, joissa oikea vokaali on korvattu väärällä vokaalilla, on useimmiten kyse piirtovirheestä. Piirtovirheellä tarkoitetaan a/ä- tai o/ö-sekaannuksia, joissa siis ä:n ja ö:n pisteet ovat jääneet merkitsemättä tai vastakohtaisesti a:n ja o:n päälle on virheellisesti merkitty pisteet (*päästa* »päästä», *kamerä* »kamera»). Kuten kaaviosta 3 ilmenee, huomattavan suuri osa vääristä vokaaleista selittyy piirtovirheellä kaikilla luokka-asteilla: toisen luokan oppilaiden 85 väärästä vokaalista 58 johtuu piirtovirheestä, neljäsluokkalaisten 74:stä 51 ja kuudesluokkalaisten 84:stä 57. Piirtovirheiden määrä pysyy lähes samana kaikilla luokka-asteilla, mikä kertonee siitä, ettei kyseessä ole varsinaisesta kirjoitustaidon kehittymisestä kertova piirre. Piirtovirheet ovatkin tyypillisiä huolimattomuusvirheitä, mistä kertoo myös se, että piirtovirheitä esiintyy usein myös sanoissa, joissa ä tai ö on myös norminmukaisesti merkittynä mukana kuten esimerkissä 62.

(62) *Mina halusin lahteä ulkomaille* (2F)

Jotkin tapaukset selittyvät myös piirtovirheeseen rinnastuvalla u/y-sekaannuksella, jossa siis y:n ”alakoukku” on jäänyt merkitsemättä, ja näin ollen lopputulos muistuttaa lähinnä u-kirjainta kuten esimerkissä 63.

(63) *SuhtyMäpäivänäni haluaisin saada lahjanani arkullise rahaa jaxkflanggelin* (2M)

Esimerkin 63 kaltaisista tapauksista ilmenee hyvin u/y-sekaannusten ja piirtovirheiden samankaltaisuus, sillä myös u/y-sekaannuksia esiintyy sanoissa, joissa on mukana myös norminmukaisia muotoja. Tätä virhettä esiintyy toisen ja neljännen luokan oppilaiden kirjoitelmissa, tosin molemmilla hyvin marginaalisena määränä. Kuudesluokkalaisten kirjoitelmissa u/y-sekaannuksia ei ole lainkaan. Vanhimpien oppilaiden kirjoitustaito lienee jo niin rutiininomaista, että riski u/y-sekaannukseen on huomattavasti pienempi kuin nuoremmilla oppilailla.

Toiseksi yleisin väärän vokaalin esiintymisympäristö on diftongeissa. Väärät vokaalit diftongeissa johtuvat usein etenkin länsimurteille tyypillisestä diftongin avartumisesta, jolloin diftongin puolisuuppea jälkikomponentti on korvautunut väljällä vokaalilla kuten esimerkissä 64.

(64) *minä ja siskoni Riina olimme lähdössä Ruatsiin seuraavana aamuna aikaisin Emme meinanneet saada unta kun jännitti niin paljon mennä laivalla Kello oli neljä aamu yällä piti lähteä Vaasaan että kerkeää laivaan* (4F)

Diftongin avartumisesta johtuvia vääriä vokaaleja on aineistossa yhteensä 14 kappaletta, ja ne esiintyvät lähes poikkeuksetta juuri länsimurteiden piiriin kuuluvien oppilaiden kirjoitelmissa.

Väärä vokaali diftongissa voi johtua myös diftongin oikenemisestä, jolloin diftongi on korvautunut pitkällä vokaalilla. Tämä prosessi voi johtua puhtaasti siitä, että oppilaalla on vaikeuksia merkitä (ja kenties ääntää) diftongia sen siirtymä-äänteisyyden vuoksi kuten todennäköisesti esimerkissä 65. Lisäksi kyseessä voi olla lounais-, savolais- ja paikoin pohjoispohjalaisillekin murteille tyypillinen diftonginreduktio (ISK § 23), jossa suppenevan diftongin jälkikomponentti *i*, *u* tai *y* esiintyy väljentyneenä, ja lopputuloksena on pitkä vokaali. Diftonginreduktiosta on todennäköisesti kyse esimerkissä 66.

(65) minulla olisi monta matkaluukullista satasen seteleitä (2M)

(66) pussailtais ja ehkä halaaltais ja ehkä juteltais – – Meidän päivän risteily alkaa loppua eikä me olla muuta paljota tehty ku pussaaltu ja halailtu (6F)

Esimerkissä 66 diftonginreduktioksi tulkitsemista horjuttaa se, että redukoituneiden diftongien lisäksi kirjoitelmassa esiintyy täysin vastaavissa ympäristöissä norminmukaisia diftongeja. Toisaalta kyseisen oppilaan koulu sijaitsee pohjoispohjalaismurteiden alueella, joten diftonginreduktio olisi hänen murteelleen luontainen. Murrepiirteiden leimallisuudesta kertonee siis sekin, ettei se esiinny johdonmukaisena piirteinä edes sellaisten oppilaiden kirjoitelmissa, joiden luontaiseen puhemuotoon se saattaa kuulua.

Kolmanneksi yleisin väärän vokaalin selitys on vokaaliassimilaatio, eli se, että sanassa edempänä tai jäljempänä esiintyy vokaali, jolla sanan tavoitteenmukainen vokaali on virheellisesti korvaantunut kuten esimerkissä 67.

(67) välipalaksi minä saan jäätölöä (2F)

Esimerkin 67 kaltaista tapausta voidaan kutsua myös vokaaliharmoniaksi; siinä norminmukainen vokaali *e* on korvattu muutoinkin sanassa esiintyvällä vokaalilla *ö* ja näin sanan ääntäminenkin yksinkertaistuu. Myös tämä ryhmä on hyvin suuruudeltaan hyvin samankokoinen kaikilla luokka-asteilla, mutta etenkin vanhemmilla oppilailla sen voidaan olettaa johtuvan enemmänkin huolimattomuusvirheestä kuin varsinaisesta vokaaliassimilaatiosta.

Osa vääristä vokaaleista (8 kappaletta) on *i/j*-sekaannuksia. Tähän ryhmään lukeutuvat siis sekaannukset kumpaankin suuntaan, jolloin myös *i* on voinut korvaantua *j*:llä, joten aina ei ole kyse varsinaisesti väärästä vokaalista. *i/j*-sekaannukset muodostavat kuitenkin yhteensä

pienen virhekategoriain, ja johdonmukaisuuden vuoksi esiintymät on mielekästä käsitellä yhtenä ryhmänä. i/j-sekaannukset johtunevat suurimmaksi osaksi *i:n* ja *j:n* ääntämisen samankaltaisuudesta, ja onkin hyvin todennäköistä, että oppilas todella ääntäisi sanan niin kuin on sen kirjoittanut (68).

(68) Sitten seuraavana kesänä Liinu puraisi postelioonia jalkaa (4M)

Muihin mahdollisiin selityksiin lukeutuu joitain sellaisia tapauksia, joissa väärän vokaalin merkitseminen liittyy sanan taivuttamisen vaikeuteen (69). Lopuille tämän kategorian virheille ei näyttäisi olevan mitään erityistä selitystä, vaan vokaali on korvautunut väärällä tuntemattomasta syystä (70).

(69) Ja sitten purattiin ostoskassit (2M)

(70) Ja sitten samalla menä keksin sille nimen siitä tuli PukuLiinu (4M)

Esimerkin 69 kaltaiset virheet johtunevat sanan taivutusparadigman monimutkaisuudesta. Suomen kielessä sanoja taivutettaessa tapahtuu monenlaisia äänneellisiä muutoksia, ja oppilaalle on haasteellista hallita etenkin itselle vieraiden sanojen paradigmoja. Esimerkin 69 *purattiin*-muoto lienee syntynyt analogian kautta purkaa-verbin monikäyttöisemmistä muodoista (esim. aktiivin preesens).

### 3.5 Ylimääräinen kirjain

Ylimääräinen kirjain voi merkitä virhettä äännepitäisyydessä: konsonantti on voinut geminoitua virheellisesti tai vastaavasti vokaali pidentyä. Sanassa voi myös esiintyä yksittäisiä ylimääräisiä konsonantteja ja vokaaleja sanan alussa, keskellä tai lopussa. Jos ylimääräisiä kirjaimia on peräkkäin useampi, ne voivat muodostaa myös kokonaisen ylimääräisen tavun. Kaiken kaikkiaan ylimääräiset kirjaimet ovat usein sanan ääntämisen aiheuttamia virheitä, ja niillä voidaankin pyrkiä yksinkertaistamaan sanan ääntämistä. Ylimääräisten kirjainten virheitä on aineistossa yhteensä 233 kappaletta, mikä vastaa 9 % tutkittujen virheiden kokonaismäärästä.

### 3.5.1 Konsonantin geminoituminen

Konsonantin geminoituminen on äännepitouden hahmottamiseen liittyvä virhe. Sen voidaan katsoa olevan vastakkainen prosessi konsonantin geminaatosta puuttumiselle. Konsonanti kuitenkin geminoituu aineistossa huomattavasti harvemmin kuin jää puuttumaan geminaatosta: kun puutteellisia geminaattoja oli aineistossa 288 kappaletta, geminoituneita konsonanteja on vajaa puolet tästä, 123 kappaletta. Onkin paljon todennäköisempää, että geminaatosta toinen jää merkitsemättä kuin että sen kirjoittaisi virheellisesti kahteen kertaan. Konsonanti on virheellisesti geminoitunut toisen luokan oppilaiden kirjoitelmissa 64 kertaa, neljäsluokkalaisilla 35 kertaa ja kuudesluokkalaisilla 24 kertaa. Konsonantin geminoitumistapaukset jakautuvat sukupuolten välille melko tasaisesti: poikien kirjoitelmissa niitä on yhteensä 59 kappaletta ja tyttöjen 64. Luvut antavat suuntaa tarkastelulle, mutta niissä on huomioitava kirjoittajien välinen suuri hajonta. Kaikilla luokka-asteilla tyttöjen kirjoitelmat ovat sanamäärältään poikien kirjoitelmia pidempiä. Konsonantin geminoitumistapaukset on esitetty luokka-asteittain ja sukupuolittain taulukossa 10.

**Taulukko 10.** Konsonantin geminoituminen luokka-asteittain ja sukupuolittain.

Luokka-aste	Pojat	Tytöt	Yhteensä
2.	26	38	64
4.	22	13	35
6.	11	13	24
<b>Yhteensä</b>	59	64	123

Konsonantin geminoitumista selittää useissa tapauksissa sanan ääntöasu; joihinkin tapauksiin on varmasti vaikuttanut murteellinen ääntöasu kuten esimerkissä 71, joitain selittää sanan vierasperäisyys. Vierassanojen ääntämys saattaa poiketa niiden vakiintuneesta kirjoitusasusta äänneiden pituuden osalta. Monilla puhujilla vierassanojen yksinäisklusiilit ja -frikatiivit esiintyvät geminaattoina likvidan (72), nasaalin (73), pitkän vokaalin tai diftongin jäljessä (ISK § 8).

(71) Unelmieni päivää ei ole vielä tullut (4F)

(72) Lähdimme kesäloman lopussa etelä Portugaliin (6M)

(73) Kun olimme Helsingin olympiastadionilla (4M)

Joissain geminoitumistapauksissa on kyse toteutumattomasta astevaihtelusta, jolloin heikkoasteisen yksinäiskonsonantin on korvannut geminaatta (74). Varsinaisia astevaihtelusääntöjä ei alakoulussa opeteta, mutta kielellisen tietoisuuden kehittyessä lapsi taitaa ne tiedostamattaan. Hyvin usein konsonantin geminoitumisessa on selvästi kyse rajageminaatiosta, jota useat aikuisetkaan eivät kirjoituksessa hallitse (75).

(74) Äitin pikkuserkkun nimi on Aleksandra (2F)

(75) me matkustaisimme perheen kanssa jonnekkain ulkomaille lomalle (4M)

Toisen luokan oppilaiden kirjoitelmissa konsonantti on geminoitunut huomattavasti useammin kuin vanhempien oppilaiden. Nuorimpien alaluokkalaisten tuntuu olevan vaikeaa hahmottaa äänteiden pituuksia, mikä varmasti vaikuttaa myös tämän virheen yleisyyteen. Melko yleistä on, että konsonantti on geminoitunut sanassa, jonka yleiskielisessä versiossa on oikeastikin jonkin toisen konsonantin muodostama geminaatta kuten esimerkissä 76. Joissain tapauksissa molempien konsonanttien pituus on virheellinen, niin että geminaattaa merkitään virheellisesti yksinäiskonsonantilla ja vastaavasti yksinäiskonsonanttia geminaatalla kuten esimerkissä 77. Tällaiset virheet eivät välttämättä kerro niinkään vaikeuksista hahmottaa äännepituuksia vaan vaikeudesta sanahahmon visuaalisessa hahmottamisessa.

(76) voittaissimme kaikki kolme peliä 3-0 (2M)

(77) Tänään menen katsomaan jääkiekko ote//un Ilves-tappara (4M)

Neljäsluokkalaisten kirjoitelmissa konsonantti on geminoitunut 35 kertaa ja kuudesluokkalaisten kirjoitelmissa 24 kertaa. Yhteensä aineiston virheistä 123 kappaletta on siis konsonantin virheellisiä geminoitumisia. Virheiden yleisyyden luokkatasokohtaisesta vertailusta selviää, että konsonantin geminoituminen vähenee voimakkaasti alemmilta vuosiluokilta ylemmille edettäessä. Tämä johtuu varmasti kielitajun kehitymisestä ja erityisesti äännepituuksien hahmottamisen kehitymisestä. Lisäksi vierasperäisten sanojen oikeinkirjoitukseen aletaan kiinnittää opetuksessakin enemmän huomiota alakoulun ylemmillä luokilla, joten oppilaat tulevat tietoisemmiksi ääntämisen ja kirjoitusasun eroista.



### 3.5.2 Vokaalin pidentyminen

Vokaalin pidentyminenkin on äännepitouden hahmottamiseen liittyvä virhe. Virheenä se on hyvin samankaltainen kuin konsonantin geminoituminen ja vastakkainen vokaalin puuttumiselle pitkästä vokaalista. Kuten pitkien vokaalien kohdalla on todettu, on vokaalin pituus huomattavasti helpompi huomata ja hahmottaa kuin konsonantin pituus, ja siitä kertoo myös vokaalin pidentymisesiintymien pienempi frekvenssi konsonantin geminoitumiseen nähden: konsonantti on aineistossa geminoitunut 123 kertaa, mutta vokaali on pidentynyt vain 45 kertaa. Pidentyneen vokaalin luokka-asteittaiset ja sukupuolittaisen esiintymät käyvät ilmi taulukosta 11.

**Taulukko11.** Vokaalin pidentyminen luokka-asteittain ja sukupuolittain.

Luokka-aste	Pojat	Tytöt	Yhteensä
2.	16	9	25
4.	5	3	8
6.	6	6	12
<b>Yhteensä</b>	27	18	45

Toisen luokan oppilaiden kirjoitelmissa vokaali on pidentynyt 25 kertaa ja neljännen luokan oppilaiden kirjoitelmissa vain 8 kertaa. Kuudesluokkalaisten aineistossa kyseinen virhe esiintyy 12 kertaa, eli koko aineistossa yhteensä 45 kertaa. Tyttöjen ja poikien tekemien virheiden määrien vertailu osoittaa, että poikien kirjoitelmissa vokaali pidentyy varsinkin aineiston nuorimmilla oppilailla huomattavasti useammin kuin tyttöjen kirjoitelmissa. Tosin virhemäärät ovat kaiken kaikkiaan niin pieniä, ettei niiden osalta kannata sukupuolikohtaista vertailua juurikaan suorittaa.

Oppilaat ovat usein merkinneet lyhyen vokaalin sijasta pitkän vokaalin todennäköisesti sanan ääntöasun mukaisesti. Tällaisia tapauksia on paljon ennen kaikkea vierasperäisissä sanoissa, joiden yleiskielinen kirjoitusasu poikkeaa niiden ääntöasusta (78).

(78) Jedeillä on aseenaan *laasermiekka* (2M)

Vokaaleissa ääntämyksen ja kirjoitusasun pituusero ei ole yhtä tavallista kuin konsonanttien kohdalla, mikä myös selittää vokaalin pidentymisen suhteellisen pientä määrää konsonantin geminoitumiseen nähden.

Toinen mahdollinen selitys vokaalin pidentymiseen on se, että samassa sanassa esiintyy muitakin pitkiä äänteitä, jolloin toisen äänteen pituus on visuaalisen yhdenmukaistamisen tuotoksena pidentänyt virheellisesti toisenkin äänteen kuten esimerkissä 79.

(79) Kun olisin syönyt menisin uimaan *aaltaaseen* (4M)

Esimerkin 79 kaltaista vokaalin pidentymistä sanassa, jossa kaikki muutkin vokaalit esiintyvät pitkinä, voidaan pitää eräänlaisena äänneassimilaation ilmentymänä, jossa oppilas on pyrkinyt yhdenmukaistamaan sanan äännepituuksia.

Kaiken kaikkiaan vokaalin pidentyminen on verrattain pieni virhekatteoria. Vaikka vokaalin pidentyminen on konsonantin geminoitumisen tapaan äännepituuksien hahmottamiseen ja merkitsemiseen liittyvä ongelma, selittyy se enneminkin juuri joidenkin lyhyiden vokaalien ääntymisenä puheessa pitkinä, kuin varsinaisella äännekirjainvastaavuuden horjumisella.

### 3.5.3 Muu ylimääräinen kirjain

Muu ylimääräinen kirjain -ryhmä sisältää muut yksittäiset aineiston kirjoitelmissa ylimääräisinä esiintyvät kirjaimet. Yhteensä aineistossa on ylimääräisiä kirjaimia 65 kappaletta, joista 12 kappaletta on ylimääräisen tavun virheitä. Muista ylimääräisistä kirjaimista 26 on poikien kirjoitelmista ja 39 tyttöjen kirjoitelmista. Tyttöjen poikia suurempi virhemäärä tosin tasoittuu, kun otetaan huomioon erot tyttöjen ja poikien tekemien kirjoitelmien pituuksissa. Taulukosta 12 selviää ylimääräisten kirjainten laadut tarkemmin.

**Taulukko 12.** Muun ylimääräisen kirjaimen laatu luokka-asteittain.

Luokka-aste	Ylimääräinen konsonantti	Ylimääräinen vokaali	Ylimääräinen tavu	Yhteensä
2.	7	17	5	29
4.	16	3	3	22
6.	7	3	4	14
<b>Yhteensä</b>	30	23	12	65

Yksittäiset ylimääräiset konsonantit saattavat olla selkeästi sanan ääntämistä helpottavia siirtymä-äänteitä, tyypillisimmin joko *v*- tai *j*-kirjaimen muodossa. *i*:n tai *U*:n sisältävää vokaaliyhtymää äännettäessä sen tavurajalla voi esiintyä konsonanttisia siirtymä-äänteitä, ja nämä ovat osittain siirtyneet myös kirjoitettuun kieleen. Tutkimuksen aineistossa *v*-siirtymä-ääne esiintyy tyypillisimmin *kauan*-adverbissa kuten esimerkissä 80 ja *j*-ääne *aikoa*-verbissä kuten esimerkissä 81.

(80) Olen unelmoinut jo kauvan omasta hevosesta (4F)

(81) Jari oli lämmittämässä saunaa sillä ajomme mennä uimaan (4F)

Konsonanttiset siirtymä-äänteet selittävät suurimmaksi osaksi yksittäisiä ylimääräisiä konsonantteja: yhteensä 30 esiintymästä 12 selittyy siirtymä-äänteillä. Muut kuin siirtymä-äänteellä selittyvät ylimääräiset konsonantit esiintyvät enimmäkseen aineiston nuorimpien oppilaiden kirjoitelmissa ja siirtymä-äänteet enemmän vanhempien. Vanhempien oppilaiden ylimääräiset konsonantit ovat näin ollen tarkoituksenmukaisempia ja harkitumpia kuin nuorimpien oppilaiden.

Yksittäistä ylimääräistä vokaalia sanassa ei tunnu selittävän selkeästi mikään tietty sanan piirre; se saattaa esiintyä hyvin satunnaisesti millaisessa äänneympäristössä tahansa. Toisen luokan oppilaiden kirjoitelmissa ylimääräinen vokaali näyttäisi hakeutuvan joko konsonanttiyhtymän väliin (82) tai toisen vokaalin viereen muodostaen näin diftongin sen kanssa (83).

(82) sen jälkeen lähidimme mökille (2M)

(83) mene sisälle koska he rakensivat taillin (2F)

Edellisistä esimerkeistä konsonanttiyhtymän väliin sijoitettu ylimääräinen vokaali kertoo konsonanttiyhtymien merkitsemisen vaikeudesta varsinkin nuorimmilla oppilailla. Juuri konsonanttiyhtymät muodostavatkin yhden suomen kielen ongelmarakenteista, jotka Ahvenaisen ja Holopaisen (2005: 107) mukaan tuottavat erityisiä vaikeuksia lukivaikeuksiselle lapselle. Äänteen lisääminen eli additio taas on yksi lapsenkielen fonologisista prosesseista, joilla hän pyrki helpottamaan sanan lausumista.

Ylimääräisiä tavuja on aineistossa yhteensä 12 kappaletta. Ylimääräiset tavut ovat lähes poikkeuksetta sellaisia, että sanassa esiintynyt yksittäinen tavu on jostain syystä merkitty kokonaan tai osittain kahteen kertaan kuten esimerkeissä 84 ja 85.

(84) Söin hampurilaista ja cococolaa pullosta *viimeitteitteeksi* söin hedelmän (4F)

(85) *saisisin* nukkua pitkään aamulla ja meidän perhe voittaisi ulkomaan matkan (6M)

Tavujen reduplikaatio voi liittyä vaikeuteen hahmottaa ja tuottaa pitkiä, johdettuja ja taivutettuja sanoja. Oppilaan voi olla vaikea hahmottaa, mitä kaikkea kuuluu sanavartaloon ja mitä siihen tulee johtamisen ja taivuttamisen yhteydessä lisätä. Toisaalta tavureduplikaatio voi kertoa ongelmasta fonologisessa strategiassa, jolloin sanan tuottamisessa ilmenevä ongelma johtuu siitä, että sen jakaminen pienempiin yksiköihin on haastavaa.

### 3.6 Taivuttamiseen liittyvä virhe

Sanassa oleva virhe voi liittyä sanan taivuttamiseen. Virheellinen taivutus on virhetyyppi, jossa sana on yksinkertaisesti kontekstiin nähden väärässä muodossa. Tällaisia virheitä ovat muun muassa kongruenssivirheet sekä valenssi- ja rektiovirheet. Lisäksi tähän kategoriaan on laskettu mukaan sellaiset virheet, joissa sana on korvattu toisella sanalla, jonka kantasana on sama kuin tavoitellun sanan (*menemästä* »menemisestä», *hälyttämisen* »hälyttimen»). Virheellinen taivutus voi tarkoittaa myös epäonnistunutta taivutusta, jossa sanamuoto on ollut periaatteessa oikea, mutta taivuttamisen toteutus epäonnistunut (*syökseä* »syöstä», *lapseja* »lapsia»). Salonen (2012) on pro gradu -tutkielmassaan analysoinut Koulukorpuksen puhekielisyyksiä, ja hän on kategorisoinut kongruenssivirheet yhdeksi puhekielen piirteeksi. Kongruenssivirheet ovat mukana myös tässä tutkimuksessa, sillä ne ovat hyvin oleellinen osa virheellistä taivuttamista, ja muodostavat oivan vertailujoukon kirjoitelmissa tehtyihin muihin taivutusvirheisiin.

Toinen taivuttamiseen liittyvä virheryhmä on virhe symbolissa. Tähän kategoriaan on poimittu aineiston numerot, lyhenteet ja muut merkit, joihin ei ole osattu liittää taivutustunnusta oikein (*nhllän* »NHL:n»). Kolmanneksi virheellisen taivuttamisen yhteydessä tarkastellaan myös monitulkintaista virhejoukkoa, johon lukeutuvat virheet eivät ole yksiselitteisesti luokiteltavissa vain yhteen virhekategoriaan. Yhteistä näille virheille on, että niiden kaikkien yksi mahdollinen selitys on juuri virheellinen taivutus. Useimmiten muu mahdollinen virhe on puuttuva kirjain tai väärä kirjain, ja useimmissa tapauksissa se on myös todennäköisempi virhe.

Kolmanneksi virheellisen taivutuksen yhteydessä käsitellään monitulkintaisten virheiden joukko, joita ei ole voitu yksiselitteisesti sijoittaa vain yhteen virhekategoriaan. Yhteistä kaikille näille virheille on se, että ne saattavat selittyä virheellisellä taivutuksella, ja

sen vuoksi ne juuri taivuttamiseen liittyvien virheiden yhteydessä käsitelläänkin. Virheellistä taivuttamista todennäköisempi selitys näille monitulkintaisille virheille kuitenkin on puuttuvan kirjaimen virhe. Taivuttamiseen liittyviä virheitä (monitulkintaiset virheet mukaan laskettuna) on aineistossa yhteensä 278 kappaletta, mikä on noin 11 % tutkittujen virheiden kokonaismäärästä.

### 3.6.1 Virheellinen taivutus

Virheellinen taivutus -kategorian virheet ovat sellaisia, että kirjoitelman yksittäinen sana on lauseyhteyteen tai muuhun kontekstiin nähden virheellisesti taivutettu. Virheellisyys voi tarkoittaa sitä, että taivutusmuoto on norminvastainen tai muodollisesti epäonnistunut. Norminvastaisia taivutuksia ovat esimerkiksi kongruenssivirheet, rektio- ja muut suuntaisen täydennyksen virheet sekä yksikkö/monikko-sekaannukset. Sanan muoto voi olla myös virheellinen niin, että taivutettaessa syntynyt muoto ei ole norminmukainen sana lainkaan. Tällaiset virheet näyttäisivät johtuvan sanan kompleksisesta taivutusparadigmasta; suomen kielessä sanavartaloissa tapahtuu taivutettaessa monenlaisia muutoksia, jotka voivat varsinkin nuorimmille oppilaille olla todella haasteellisia.

Yksiselitteisiä virheellisiä taivutuksia aineistossa on 150 kappaletta. Toisen luokan oppilaiden kirjoitelmissa virheitä on 65 kappaletta, neljännen luokan oppilaiden kirjoitelmissa 46 kappaletta ja kuudennen luokan oppilaiden kirjoitelmissa 39 kappaletta. Sukupuolittain virheet jakautuvat siten, että poikien tekemiä on 90 kappaletta ja tyttöjen 57 kappaletta, eli sanan norminmukaisessa taivutuksessa pojat tekevät selvästi tyttöjä enemmän virheitä. Taulukossa 13 on eritelty tarkemmin virheellisen taivutuksen virheitä muotoryhmittäin ja luokka-asteittain.

**Taulukko 13.** Virheellisten taivutusten selitykset luokka-asteittain.

Luokka-aste	Kongruenssi-virhe	Rektio-virhe	Infinitiivi-virhe	Yksikkö/monikko-sekaannus	Muotoon liittyvä	Muut	Yhteensä
2.	28	13	3	6	8	7	65
4.	30	1	9	1	3	2	46
6.	15	5	8	0	6	5	39
<b>Yhteensä</b>	73	19	20	7	17	14	150

Kuten taulukosta 13 ilmenee, on virheellinen taivutus useimmiten luonteeltaan kongruenssivirhe; lähes puolet kaikista virheellisistä taivutuksista selittyy kongruenssivirheellä. Subjektin ja predikaatin välisessä lukukongruenssissa on monenlaista vaihtelua. Useimmissa puhutun kielen varianteissa monikkokongruenssi puuttuu kolmannesta persoonasta kokonaan tai ainakin vaihtelee inkongruenssin kanssa puhujittain tai tapauskohtaisesti. Tämä heijastuu myös kirjoitetun kielen käytäntöön vaihteluna. (ISK § 1280.) Koska tämän tutkimuksen aineisto koostuu rakenteeltaan hyvin perinteisistä koulukirjoitelmista, inkongruenssitapaukset ovat rakenteeltaan sellaisia, että monikon 3. persoonan yhteydessä on merkitty monikkomuotoinen subjekti mutta yksikkömuotoinen predikaatti, kuten esimerkeissä 86 ja 87.

- (86) Aina heti heti kun tulee nälkä syödään makaroonilaatikkaa vielä että uimaseuran *reenit olisi* joka päivä.niin kuin isoveljellä (2M)
- (87) Minä ja kaksi ystävääni menimme tavarataloon. *Heidän nimensä oli* Nicklas ja Eetu (4M)

Subjektin ja predikaatin lukukongruenssi jääkin usein toteutumatta erityisesti luonteeltaan puhekielisissä kirjoitelmissa. Samoin inkongruenssia tuntuisi esiintyvän erityisesti niissä tapauksissa, joissa henkilöihin viittaava persoonapronomini *he* on korvautunut puhekielisellä demonstratiivipronominilla *ne*, kuten esimerkissä 88.

- (88) minä menin katsomaan sitä ja sielä asuyi tyttö ja poika *ne söi* tikaria (2F)

Tarkasteltaessa kirjoitelmissa esiintyviä kongruenssivirheitä, on syytä ottaa huomioon sekä kongruentit että inkongruentit tapaukset, jotta virheiden määrästä saataisiin selkeämpi kuva. Niinpä vertailua varten aineistosta on laskettu myös tapaukset, joissa monikon 3. persoonassa olevan subjektin ja predikaatin kongruenssi toteutuu. Taulukossa 14 on esitetty näin syntyneet vertailuluvut luokka-aste- ja sukupuolikohtaisesti.

**Taulukko 14.** Monikon 3. persoonan kongruentit ja inkongruentit tapaukset luokka-asteittain ja sukupuolittain sekä luokittainen virheellisyysprosentti.

Luokka-aste	Yleiskieliset M	Inkongruentit M	Yleiskieliset F	Inkongruentit F	Yhteensä	Inkongruenssi %
2.	18	12	23	16	69	40,6 %
4.	29	18	58	12	117	25,6 %
6.	35	4	124	11	174	8,6 %
<b>Yhteensä</b>	82	34	205	39	360	20,3 %

Yleiskielisiin esiintymiin suhteutettuna eniten virheellisiä esiintymiä on toisella luokka-asteella, jolla inkongruentit tapaukset muodostavat yli 40 % kaikista esiintymistä. Ylemmille luokille siirryttäessä inkongruenttien esiintymien suhteellinen määrä yleiskielisiin verrattuna alkaa hiljalleen laskea, ja neljännellä luokka-asteella vastaava luku onkin enää noin 26 % ja kuudennella luokalla inkongruentit tapaukset muodostavat enää alle 9 % kokonaisuudesta. Vaikka inkongruenttien tapausten suhteellinen määrä väheneekin melkoisesti alempiin luokka-asteisiin verrattuna, on 9 %:n virheellisyysprosentti kuudennella luokka-asteella siitä huolimatta melkoinen verrattuna moneen muuhun tästä aineistosta tutkittuun piirteeseen. Voidaankin epäillä monikon 3. persoonan inkongruenssin olevan sellainen puhekielen piirre, jota vanhemmatkaan kirjoittajat eivät tunnista norminvastaiseksi, vaan luulevat muodon olevan huoliteltuun yleiskieleen kuuluva.

Kongruenssivirheen lisäksi väärä taiputus voi olla luonteeltaan rektio- tai valenssivirhe. Rektiolla tarkoitetaan, että tietyn sanan täydennys on tiettyssä sijassa riippumatta siitä, mikä on täydennyksen kategoria tai semanttinen luonne. Esimerkiksi verbin *tykätä* substantiivitäydennys on vaihtoehdoita elatiivissa (Tykkään hänestä) ja postposition *takana* täydennys genetiivissä (tuon puun takana). Rektioksi ei kutsuta tapauksia, joissa sana edellyttää täydennykseltään pelkästään tiettyä suuntaisuutta, ei tiettyä sijamuotoa. Tällöin täydennyksellä on erilaisia sija- ja sanaluokkavaihtoehtoja. Rektioon perustuvan sijan ja semanttisesti määräytyvän sijan välillä ei ole jyrkkää eroa. (ISK § 1225.) Rektio-ryhmään on tässä tutkimuksessa kuitenkin laskettu mukaan myös tapaukset, jotka eivät ole varsinaisia rektiovirheitä, vaan enemmänkin semanttisia suuntaisuusvirheitä. Rektio-virheiksi on siis tässä luettu kaikki nimitäydennykset, jotka eivät ole edussanan vaatimassa muodossa. Aineiston virheistä löytyy myös rektiovirheiden lisäksi virheitä suuntaisen täydennyksen sijamuodoissa. Yhteensä tällaisia virheitä on 19 kappaletta, joista suurin osa, 13 kappaletta, esiintyy toisen luokan oppilaiden kirjoitelmissa, kuten esimerkki 89.

- (89) Hän unelmoi jokapäivä Joskus *koiraa* joskus *kissaa* (po. unelmoida + elat.) (2F)

Infinitiivi-virheet taas ovat infinitiivitäydennyksiä, jotka eivät ole pääverbin vaatimassa muodossa. Tyypillinen infinitiivivirhe on A-/MA-infinitiivisekaannus. Valinta A- ja MA-infinitiivin välillä on abstraktimmissa tapauksissa rektion luonteista mutta vaihtelee jonkin verran. Verbi *alkaa* saa yleiskielessä A-infinitiivin, verbit *joutua*, *sattua*, *kyetä* ja *pystyä* MA-infinitiivin. Puhutussa kielessä infinitiivin valinta voi olla päinvastainenkin (ISK § 470). Esimerkin 90 kaltainen *alkaa*-verbin saama täydennys on melko yleinen puhekielen piirre, ja esimerkin 91 kaltainen infinitiivitäydennys on myös puhutussa kielessä yleistyvää.

- (90) Kun vien tavarat sisälle ja päästän tiikeriraitaisen kissan tikrun ulos sekä *alan purkamaan* tavaroitani (6F)
- (91) tämä toive on hartain toiveeni On että *pystyisin olla* sille isäntänä ja että se tottelisi minua koska se tarkoittaa että *pystyy olla* isäntänä koiralla (4F)

Yksikkö/monikko-sekaannuksia ei aineistossa ole montaa ja nekin ovat yhtä lukuun ottamatta kaikki toisen luokan oppilaiden kirjoitelmista (92).

- (92) saisi olla maailmalla kiertää ja kaartaa ja olla maailmanparas ja voittaa osa kilpailuja ja saisi pyttyjä ja *kaulamitalia* (2M)

Muoto-virheeseen sisältyvät niin mielekkäät väärät sanat kuin epämielekkäät taivutuksetkin. Väärä mielekäs sana on norminmukainen lekseemi, joka on kuitenkin kontekstiin nähden väärä. Väärän mielekkään sanan kantasana on sama kuin aiotun oikean sanan. Tällaisia virheitä on vain toisen luokan oppilaiden kirjoitelmissa. Väärille mielekkäille sanoille on tyypillistä, että ne ovat useamman johtimen sisältäviä johdoksia, mikä tietenkin selittää sekaannuksia sanoissa. Kantasanaan on liitetty joko väärä johdin (kuten esimerkissä 93) tai jokin johdin on voinut jäädä merkitsemättä.

- (93) museo oli kiinni Siellä oli vain vartioita koira katsoi joka ikisen nurkan Kun koira oli kyllästynyt niin koira vahingossa astui *hälyttämisen* (po. hälyttimen) päälle niin vartiat huomasi koiran niin alkoi takaa ajo (2F)

Epämielekäs taivutus tarkoittaa sanaa, jonka taivuttamisessa on virhe eikä taivuttamisen tuloksena syntynyt sana ole norminmukainen. Tällaisia muotoja saavat sanat, joiden taivutusparadigman sisällä sanavartalolle tapahtuu runsaasti muutoksia kuten esimerkeistä 94



ja 95 käy ilmi. Esimerkissä 94 on epämielekkään taivutuksen lisäksi myös väärän mielekkään väärän sanan virhe.

- (94) Jos saisin joskus omia kilttejä *lapseja* niin olisi ehkä hauskaa opetalla (po. opettaa) niille matikkaa ja muuta (2F)
- (95) Perillä etsin ensiksi hotellin ja *purkasin* kassini lepäsin myös (6F)

Kaiken kaikkiaan muodollisesti virheellisiä taivutuksia ei ole aineistossa montaa, ja virheet vaikuttavat epäjohdonmukaisilta ja inhimillisiltä, sillä usein ne noudattavat jonkin toisen sanan taivutusparadigmaa.

### 3.6.2 Virhe symbolissa

Virhe symbolissa voi ilmetä, kun lyhenteen, numeron tai muun merkin perään tulisi liittää taivutustunnus. Varsinainen virhe on useimmiten kokonaan puuttuva taivutustunnus tai tunnuksen merkitseminen virheellisesti. Virhe symbolissa on siis eräänlainen virheellisen taivutuksen alatyyppejä, joten niitä on mielekästä tarkastella taivuttamiseen liittyvien virheiden yhteydessä omana ryhmänään. Kyseessä ole frekvenssiltään kovin suuri virhekattegoria, eikä alakoulussa ylipäättään juurikaan opeteta taivutustunnusten ja päätteiden lisäämistä lyhenteiden, lukujen ja muiden merkkien perään, vaan niitä opetetaan järjestelmällisesti vasta yläkoulussa.

**Taulukko 15.** Virheet symboleissa luokka-asteittain ja sukupuolittain.

Luokka-aste	Pojat	Tytöt	Yhteensä
2.	9	8	17
4.	8	4	12
6.	3	4	7
<b>Yhteensä</b>	20	16	36

Virheitä symboleissa ilmenee aineistossa yhteensä 36 kappaletta, ja niistä 20 on poikien kirjoitelmissa ja 16 tyttöjen kirjoitelmissa. Luvut antavat suuntaa tarkastelulle, mutta niissä on huomioitava kirjoittajien välinen suuri hajonta. Kaikilla luokka-asteilla tyttöjen kirjoitelmat ovat sanamäärältään poikien kirjoitelmia pidempiä. Symboleita esiintyy oppilaiden kirjoitelmissa etenkin lelujen, tv-ohjelmien, kauppojen ja pelien nimissä. Yleistä on, että

symboli on jätetty kokonaan taivuttamatta, eli taivutustunnusta ei ole symbolin perään merkitty kuten esimerkissä 96.

(96) Haluaisi pelata play station2 koko päivän (2M)

Virheistä symboleissa 15 kappaletta on esimerkin 96 tapaisia, eli sellaisia, joissa symbolia ei ole taivutettu lainkaan. Loput 21 tapausta koostuvat virheistä, joissa taivutustunnus tai pääte on yritetty lisätä symbolin perään, mutta näin syntynyt muoto on virheellinen useimmiten ääntämistä mukaillen kuten esimerkissä 97.

(97) Tämän seurauksena pääsen maailmahuipulle *nhllään* joukkueeseen boston (2M)

Esimerkin 97 kaltaiset virheet symboleissa kertonevat oppilaan vaikeudesta hahmottaa symbolin vartalon ja taivutustunnusten välistä rajaa. Esimerkin 97 *nhllään* lausutaan [änhooällään], ja hyvin tyypillistä virheille symboleissa onkin juuri tällainen merkitseminen, joissa itse symboli on oikein merkitty, mutta vartalon ja taivutustunnuksen väliin on lisätty ääntämistä helpottavia kirjaimia.

Vaikka kyseessä onkin verrattain hyvin pieni virhekatgoria, se kertoo kirjoitustaidon kehittymisestä, sillä virheiden yleisyys pienenee huomattavasti verrattaessa eri luokka-asteiden esiintymien määriä keskenään: toisen luokan oppilaiden kirjoitelmissa virheitä symboleissa on 17 kappaletta, neljännen luokan oppilaiden kirjoitelmissa 12 kappaletta ja kuudennen luokan oppilaiden kirjoitelmissa enää 7 kappaletta. Tyttöjen ja poikien välillä symbolin merkitsemisessä ei ole juurikaan eroa: virheistä 20 on poikien kirjoitelmista ja 16 tyttöjen kirjoitelmista.

### 3.6.3 Monitulkintaiset virheet

Aineistossa on runsaasti monitulkintaisia virheitä, jotka voivat virheellisen taivutuksen lisäksi johtua puuttuvista, ylimääräisistä tai vääristä kirjaimista. Näitä virheitä ei siis voi yksiselitteisesti sijoittaa vain yhteen virhekatgoriaan niiden monitulkintaisuuden vuoksi, joskin jotkin virheet ovat kontekstinsa puolesta sellaisia, että niiden voidaan arvella johtuvan enemminkin muusta kuin taivuttamiseen liittyvästä virheestä. Tällaisia virheitä on aineistossa yhteensä 92 kappaletta. Tarkempi monitulkintaisten virheiden luokka-asteittainen ja sukupuolittainen jakauma käy ilmi taulukosta 16.

**Taulukko 16.** Monitulkintaiset virheet luokka-asteittain ja sukupuolittain.

Luokka-aste	Pojat	Tytöt	Yhteensä
2.	15	30	45
4.	13	16	29
6.	7	11	18
<b>Yhteensä</b>	35	57	92

Kuten taulukosta 16 ilmenee, monitulkintaisten virheiden määrä vähenee huomattavasti siirryttäessä alimmilta luokka-asteilta ylemmille; toisen luokan oppilaiden kirjoitelmissa monitulkintaisia virheitä on 45 kappaletta, kun vastaava virhe esiintyy kuudesluokkalaisten kirjoitelmissa enää 18 kertaa. Monitulkintaiset virheet ovat tytöillä hieman yleisempiä kaikilla luokka-asteilla. Tosin täytyy muistaa, että tyttöjen kirjoitelma-aineistot ovat myös poikia suuremmat kaikilla luokka-asteilla, joten virheiden määrät tasoittuvat aineistojen kokoihin suhteutettuina.

Monitulkintaiset virheet ovat useimmiten tapauksia, joissa yhden tai useammat kirjaimen puuttuessa sanasta lopputuloksena on norminmukainen sana, joka on lauseyhteyteen tai muuhun kontekstiin nähden väärän muotoinen. Useimmiten tällaiset tapaukset syntyvät, kun sanan lopusta jää kirjaimia pois, sillä suomen kielessä on runsaasti yksikirjaimisia päätteitä ja tunnuksia, joilla siis käytännössä koko sanan muoto muuttuu. Kaikkien monitulkintaiset virheiden yksi mahdollinen selitys on siis virheellinen taivutus. Taulukosta 17 käy ilmi monitulkintaisten virheiden muut mahdolliset virhekategorioidet.

**Taulukko 17.** Monitulkintaisten virheiden laadut luokka-asteittain.

Luokka-aste	Konsonantti puuttuu lopusta	Vokaali puuttuu lopusta	Pitkä vokaali puuttuu	Tavu puuttuu	Ylim. kirjain lopussa	Väärä vokaali	Konsonantti puuttuu	Yht.
2.	23	5	6	4	4	3	0	45
4.	21	1	5	2	0	0	0	29
6.	4	1	7	1	2	1	2	18
<b>Yhteensä</b>	48	7	18	7	6	4	2	92

Kuten taulukosta 17 ilmenee, ylivoimaisesti yleisin monitulkintaisten virheiden selitys on yksittäisen kirjaimen sanan lopusta puuttuminen; yhteensä sanan lopusta puuttuvia konsonantteja ja vokaaleja on 55 kappaletta, mikä on yli puolet kaikista monitulkintaisten virheistä. Sellaisissa monitulkintaisten virheissä, joissa virheellisen taivutuksen lisäksi toinen

mahdollinen virhe on yksittäinen sanan lopusta puuttuva kirjain, ei todennäköisesti ole kyse virheellisestä taivuttamisesta, vaan nimenomaan yksittäisen kirjaimen sanan lopusta puuttumisesta. Tätä tulkintaa tukee se, että tällaiset virheet eivät esiinny kirjoitelmissa johdonmukaisesti, vaan vaikuttavat tahattomasti sanan lopusta pois jääneiltä kuten esimerkeissä 98 ja 99.

(98) keksin sille nimen siitä tuli PukuLiinu ja *pari* päivän päästä se oli Buhku (4M)

(99) Kesälomalla kävin isin kanssa kalastusretkellä isi sai ison hauen Ja *pieni* ahvenia (2M)

Tapauksissa, joissa sananloppuinen *n* puuttuu illatiivista, voi olla kyse myös puhekielisestä muodosta, sillä illatiivin päätteestä voi puheessa jäädä loppu-*n* ääntämättä (100).

(100) sen jälkeen menimme *Amerikkaa* ja kun sieltä tulimme nukuin hyvin (6M)

Yksittäisiä sananloppuisia kirjaimia puuttuu sellaisista kirjoitelmista, joista puuttuu jonkin verran myös muita yksittäisiä kirjaimia, joten myös tämä ohjaisi ajattelemaan monitulkintaiset, mahdollisesti kirjaimen puuttumisella sanan lopusta selittyvät virheet muuksi kuin virheelliseksi taivutukseksi. Huomionarvoista on myös, että nimenomaan sellaisten monitulkintaisten virheiden, jotka saattavat selittyä myös yksittäisen kirjaimen sanan lopusta puuttumisella, ryhmä pienenee kaikkien radikaaleimmin alemmilla luokkasteilta ylemmille siirryttäessä, joten se kertonee myös kirjoitustaidon kehitymisestä.

Pitkän vokaalin puuttuminen sanasta on hyvin samantyyppinen virheryhmä kuin yksittäisen sananloppuisen kirjaimen puuttuminen. Näissä tapauksissa virheellisen taivutuksen selitys tuntuu myös melko epätodennäköiseltä virheiden satunnaisuuden vuoksi. Puuttuva pitkä vokaali aiheuttaa aiotun illatiivin muuttumisen genetiiviksi, joka siis on lauseyhteyden perusteella tulkittava virheelliseksi muodoksi kuten esimerkissä 101.

(101) menimme uimaan uima altaallemutta en tietenkään mennyt *syvän* altaaseen vaan menin pikuveljeni kanss matalaan altaaseen (2F)

Esimerkin 101 kaltaisissa tapauksissa vaikuttaa melko selvältä, että kyse on puuttuvasta kirjaimesta eikä virheellisestä taivutuksesta, sillä oppilas on osannut taivuttaa vastaavissa ympäristöissä sanat täysin oikein (*altaaseen, matalaan altaaseen*).

Monitulkintaiset virheet, jotka saattavat selittyä puuttuvilla tavuilla, ovat useimmiten syntyneet niin, että sanan lopusta puuttuu useampia kirjaimia, jotka siis yhdessä muodostavat

kokonaisen tavun. Näitä virheitä on aineistossa todella vähän, ja nekin selittyvät todennäköisemmin huolimattomuusvirheellä kuin virheellisellä taivutuksella kuten esimerkissä 102.

(102) Siitä tuli minulle hyvä kaveri ja riehakas Se vaan ekana yö ulisi paljon ja oli villi (4M)

Esimerkin 102 kaltaisissa tapauksissa vaikuttaa selvältä, että kyse on sanan lopusta puuttuvista kirjaimista eikä virheellisestä taivutuksesta, sillä samassa lausekkeessa esiintyvä sana *ekana* on osattu taivuttaa essiivissä täysin oikein.

Monitulkintaisista virheistä hyvin pieni osa selittyy mahdollisesti ylimääräisellä sananloppuisella kirjaimella, väärällä vokaalilla tai puuttuvalla yksinäiskonsonantilla; yhteensä näitä tapauksia on vain 12 kappaletta koko aineistossa. Ylimääräinen sananloppuinen kirjain on useimmiten verbin loppuun lisätty *n*, jolloin yksikön kolmas persoona on muuttunut yksikön ensimmäiseksi persoonassa kuten esimerkissä 103.

(103) Eräänä päivänä hän *menin* puistoon nallensa kanssa Sitten hän kuvitteli että hänen nallensa olisi koira (2F)

Esimerkin 103 kaltaiset tapaukset johtuvat kaikkein todennäköisimmin virheellisestä taivutuksesta. Unelmien/Unelmieni päivä aiheena saa oppilaat usein tuottamaan kirjoitelmia, joissa he kertovat yksikön ensimmäisessä persoonassa omasta unelmien päivästä tai kirjoittavat ikään kuin itsestään yksikön kolmannessa persoonassa. Välillä samassa kirjoitelmassa eri kerronnan muodot ja persoonat saattavat mennä sekaisin, ja oppilas on saattanut vaihtaa kerronnan näkökulmaa kesken tekstin. Tällaisissa tapauksissa saattaa siis olla horjuntaa myös yksittäisten verbien lauseyhteyteen sopimisessa.

Väärä vokaali -virhe saattaa aiheuttaa adessiivi/allatiivi-sekaannuksia kuten esimerkissä 104.

(104) kun possu oli käynnyt muta kylvys niin possu meni kaverilleen ja kun possu oli kaverilleen niin he leipoivat possu piiraan (2M)

Esimerkin 104 kaltaisissa tapauksissa virheelliseen allatiivimuotoon on todennäköisesti vaikuttanut edellisessä lauseessa esiintyvä allatiivi kaverilleen, joten kyseessä voi olla myös sana-assimilaatio.

Monitulkintaisissa virheissä puuttuvan yksinäiskonsonantin tapauksia on vain 2 kappaletta, jotka molemmat ovat kuudesluokkalaisten kirjoitelmista. Molemmissa tapauksissa on kyseessä puuttuva genetiivin tunnus *n*, joka on jäänyt puuttumaan sanan lopusta ennen liitepartikkelia kuten esimerkissä 105.

(105) Niillä ei ees melkeen riittäneet *ruokaakaan* rahaa (6F)

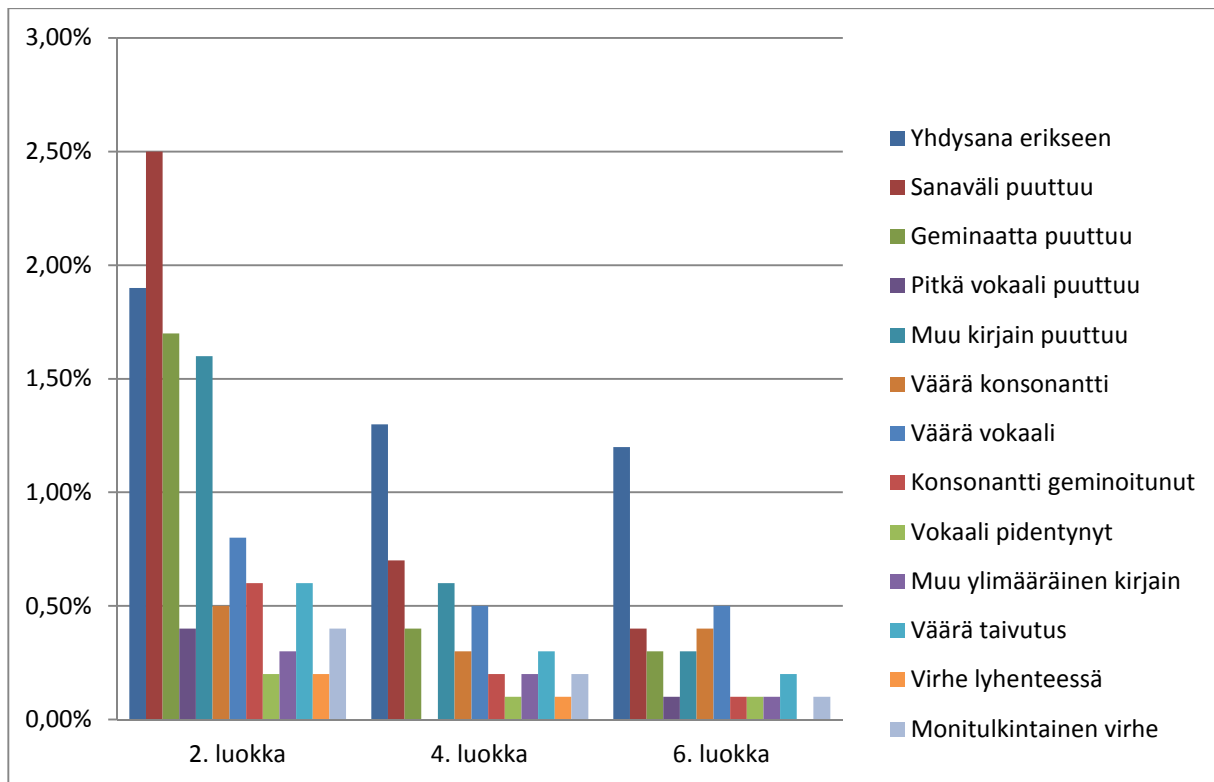
Esimerkissä 105 virheellistä taivutusta todennäköisemmältä selitykseltä tuntuisi juuri puuttuvan konsonantin virhe, joka puolestaan saattaa johtua sanan rakenteellisesta kompleksisuudesta liitepartikkelin *-kaan* vuoksi.

Tämän luvun analyysi osoittaa, että suurin osa monitulkintaisista virheistä selittyy todellisuudessa jolloin muulla virheellä kuin virheellisellä taivutuksella. Virheiden tarkastelu niiden todennäköisemmistä virhekategoriosta erillään on kuitenkin perusteltua, jotta välttyttäisiin ylitulkitsemasta oppilaiden Unelmieni päivä -kirjoitelmissa tekemiään virheitä.

#### 4. Lopuksi

Tutkimuksen tavoitteena on ollut rakentaa virheanalyysin keinoin käsitys siitä, kuinka norminmukaista alakoululaisten kirjoitelmissaan käyttämä kieli on. Tutkimukseen valittiin tarkasteluun sanatasolla esiintyviä, lähinnä fonologisia ja morfologisia piirteitä. Tutkimuksen ulkopuolelle jäi kuitenkin virheitä kahdesta syystä: Ensinnäkin, tutkimuksessa käytetyn Koulukorpuksen tekstiversio on siirretty ja koodattu Microsoft Excel -ohjelmaan, johon tehtiin myös tämän tutkimuksen aineiston koodaus. Excel-aineistossa kirjoitelmista näkyvät vain lauseen rajat eikä esimerkiksi pilkun paikkoja tai yhdysmerkkejä ole merkitty, joten välimerkkien tutkiminen tästä aineistosta ei ole mahdollista. Koska käytössä on siis ollut sähköiseen muotoon muutettu korpus, ei ole myöskään ollut mielekästä tarkastella piirteitä, jotka ilmenevät puhtaasti oppilaan käsinkirjoitetusta tekstistä. Näin ollen esimerkiksi alkukirjainvirheet ovat myös rajautuneet tutkimuksen ulkopuolelle. Toiseksi, tutkimuksen ulkopuolelle on jäänyt joukko sellaisia virheitä, joiden esiintyminen suhteutettuna koko aineiston kokoon on hyvin marginaalinen. Tällaisia virheitä ovat muun muassa reversaali ja rotaatio. Edellä mainittuja piirteitä tutkimalla olisi voitu saada toisenlaisia tuloksia.

Joillekin virhekategorioiden on laskettu norminmukaisten vastineiden mukaan vertailuluvut, ja näin ollen on voitu luoda kattavampi käsitys virheiden yleisyydestä aineistossa. Kaikille virheryhmille tällainen norminmukaisten esiintymien laskeminen ei kuitenkaan ole mahdollista (esim. yksittäiset puuttuvat kirjaimet ja yksittäiset ylimääräiset kirjaimet), joten näistä kategorioista on voitu esittää vain virheiden lukumäärät ja verrata niitä suhteessa kirjoitelmien sanamääriin. Kaavio 6 esittää kaikkien tutkielmaan valittujen virhekategorioiden laajuudet suhteutettuna vastaavaan aineiston kokoon. Vaikka osa näin syntyneistä suhteellisista vertailuluvuista onkin hyvin pieniä, saadaan näin luotua käsitys eri virheiden yleisyyksistä ja niiden vaihteluista eri luokka-asteilla. Virheet ovat kaaviossa tutkimuksen analyysiosan mukaisessa järjestyksessä.



**Kaavio 6.** Eri virheiden esiintymät aineistojen kokonaissanamääriin suhteutettuna luokka-asteittain.

Kuten kaaviosta 6 ilmenee, suurin osa tutkituista virheistä muodostaa aineiston kokonaissanamäärään suhteutettuna todella pienen prosentin. Yli kokonaisen prosenttiyksikön muodostavat kaikilla luokka-asteilla vain yhdyssana erikseen -virhekatgoria. Tämän lisäksi toisen luokan aineistossa on kolme sellaista virheryhmää, joiden esiintymät vastaavat yli yhtä prosenttiyksikköä toisen luokan kirjoitelmien kokonaissanamäärästä: puuttuva sanaväli, puuttuva geminaatta ja muu puuttuva kirjain. Oikeinkirjoitusvirheiden suhteellisesta vähenemisestä sen sijaan kertovat yhteenlasketut luokka-asteikohtaiset kokonaisvirheprosentit: toisen luokan oppilaiden virheet vastaavat 11,7 % heidän aineistonsa kokonaissanamäärästä, neljäsluokkalaisten 4,9 % ja kuudesluokkalaisten 3,8 %. Oikeinkirjoitusvirheet siis vähenevät radikaalisti alakoulun alimmilta luokilta ylemmille siirryttäessä, mutta kehitys hidastuu huomattavasti alakoulun ylimmillä luokilla. Tämän suuntainen kehityslinja kertoo ensinnäkin siitä, kuinka horjuvaa oikeinkirjoitustaito vielä alakoulun alimmilla luokilla on ja kuinka nopeasti se kehittyy oppilaan kirjoitustaidon vakiintuessa. Oppilaskohtaiset erot oikeinkirjoitustaidoissa ovat alakoulun alimmilla luokilla suurimmillaan, ja juuri alimmilla luokilla kirjoitustaito myös kehittyy nopeiten. Toiseksi voidaan sanoa, että neljäs- ja



kuudesluokkalaisten oppilaiden varsinaisessa oikeinkirjoitustaidossa ja sanatason tarkkuudessa ei ole radikaaleja tasoeroja, vaan kirjoitustaidon kehittyminen ilmenee alakoulun ylimmillä vuosiluokilla muunlaisilla muuttujilla kuten kirjoitelmien sanamäärän kasvamisella (Pajunen 2012) ja kertomusrakenteen kehittymisellä (Vesala 2012).

Ahvenaisen ja Holopaisen (2005: 107) mukaan kieleemme kuuluvia sanoja samoin kuin lainasanoja voidaan ryhmitellä rakenteen mukaisesti helppoihin ja vaikeisiin sanoihin. Lukivaikeuksisten oppilaiden ongelmat näkyvät juuri vaikeiden rakenteiden hallinnassa. Sanastotasolla voidaan suomen kielen ongelmarakenteet Ahvenaisen ja Holopaisen mallin mukaan ryhmitellä seuraavasti: pitkä vokaali ja diftongi (erityisesti reversoituvat yhdistelmät ei/ie, uo/ou), tavunloppuinen -h-, geminaatta (huomioitava erityisesti klusiiligeminaatat kk, pp, tt), konsonanttikasaumat (erityisesti reversoituvat yhdistelmät ts/st, sk/ks), η-äänne, vierasperäiset konsonantit ja rakenteet ja kasautuvat rakenteet. Tässä tutkimuksessa virheanalyysin keinoin Koulukorpuksesta tehdyt havainnot noudattavat Ahvenaisen ja Holopaisen suomen kielen ongelmarakenteet -ryhmittelyä, sillä etenkin aineiston nuorimmilla kirjoittajilla puuttuvia ja vääriä kirjaimia esiintyy runsaasti juuri pitkissä äänneissä, diftongeissa ja konsonanttiyhtymissä. Lisäksi tulokset vahvistavat sekä Ruoppilan ym. (1969) että Kuljun ja Mäkisen (2009) virheanalyysien tuloksia, sillä molemmissa tutkimuksissa puuttuvat kirjaimet ovat oppilaiden yleisimpiä kirjoitusvirheitä. Varsinkin aineiston nuorimmilla oppilailla runsasta puuttuvien kirjainten määrää esimerkiksi diftongeissa, konsonanttiyhtymissä ja äännepituuksissa selittänevät ongelmat fonologisessa koodaamisessa ja äänne-erottelussa. Mahdollista lukivaikeutta ei voida myöskään poissulkea, mutta ilman laajempaa tietämystä oppilaiden lukitaidoista ja kehityksestä, ei myöskään voida yhden kirjoitelman perusteella oppilaalle lukivaikeutta diagnosoida. Lisäksi etenkin nuorimpien oppilaiden kirjoitelmissa on huomattavan suuria pituus- ja tasoeroja jo senkin takia, että lasten kirjoitustaito on vasta löytymässä, ja tasoerot tasoittuvat iän myötä.

Ruoppilan tutkimusryhmän tuloksista hieman poiketen tämän analyysin selvästi suurimman virheryhmän muodostavat sanarajan hahmottamiseen liittyvät virheet yhdyssanan osat erilleen ja sanat aiheettomasti yhteen. Virheet sanarajoissa kattavat kaikista tässä tutkimuksessa analysoiduista virheistä 39 %, kun samat virheryhmät Ruoppilan ym. (1969: 18–19) tutkimuksessa vastaavat keskimäärin 15 %:a heidän analysoimistaan virheistä. Tämän tutkimuksen verrattain runsasta virheellistä yhteen ja erilleen kirjoittamista tosin selittävät yksittäiset runsaasti ketjuttamista sisältävät toisen luokan oppilaiden kirjoitelmat. Näiden tutkimustulosten valossa on kuitenkin syytä pohtia, mitä tällaiselle kehitykselle voidaan tehdä ja kuinka oppilaiden sanatarkkuutta voitaisiin jatkossa paremmin kehittää.

Vaikka kaaviossa 6 esitetyt aineiston määrään suhteutetut virheprosentit vaikuttavat pääosin hyvin pieniltä, on virheiden tarkastelu suhteessa norminmukaisiin esiintymiin huomattavasti informatiivisempaa. Tällaista norminmukaisten vastineiden tarkastelua virheiden rinnalla on tutkimuksessa toteutettu yhdyssanojen, geminaattojen, pitkien vokaalien sekä subjektin ja predikaatin lukukongressin toteutumisien kohdalla. Taulukossa 18 on esitetty edellä mainittujen virheiden määrät norminmukaisine vastineineen luokkatasokohtaisesti.

**Taulukko 18.** Norminvastaiset ja norminmukaiset muotoryhmät luokka-asteittain.

	<b>Yhdyssanat</b>	<b>Geminaatat</b>	<b>Pitkät vokaalit</b>	<b>Lukukongruenssi</b>
<b>2. luokka norminmukainen</b>	436	2635	2136	41
<b>2. luokka virheellinen</b>	194	181	39	28
<b>2. luokka virhe %</b>	30,8 %	6,4 %	1,8 %	40,6 %
<b>4. luokka norminmukainen</b>	669	4075	2949	87
<b>4. luokka virheellinen</b>	181	56	7	30
<b>4. luokka virhe %</b>	21,3 %	1,4 %	0,2 %	25,6 %
<b>6. luokka norminmukainen</b>	949	5548	3928	159
<b>6. luokka virheellinen</b>	580	51	12	15
<b>6. luokka virhe %</b>	17,8 %	0,9 %	0,3 %	8,6 %
<b>Kokonaisvirhe %</b>	22,0 %	2,3 %	0,6 %	20,3 %

Taulukkoon 18 on valikoitunut virheitä kahdesta syystä: Ensinnäkin, kyseiset virheet muodostavat enemmistön koko tutkittujen virheiden joukosta eli ovat kaikkein yleisimpiä virheitä. Toiseksi näille virheille on ylipäätään ollut mahdollista laskea norminmukaiset vastineet, toisin kuin suurimmalle osalla tutkimuksessa tutkituista virheistä. Taulukon 18 virheiden tarkastelu osoittaa, että kyseiset virheet eivät ole ainoastaan lukumääriltään suuria, vaan kaikkien norminmukaisiin vastineisiin verrattu määrä on myös huomattavan suuri.

E erityisen suuret virheprosentit ovat yhdyssanojen sekä predikaatin ja subjektin lukukongruenssin suhteen: kaikista yhdyssanoista virheellisiä on toisen luokan oppilaiden kirjoitelmissa lähes 31 %, neljäsluokkalaisten kirjoitelmissa noin 21 % ja kuudesluokkalaisten kirjoitelmissa lähes 18 %. Luvuista huomaa, että yhdyssanojen suhteellinen virheellisyys kuitenkin vähenee luokka-asteelta toiselle siirryttäessä, mutta virheprosentti on kaikilla

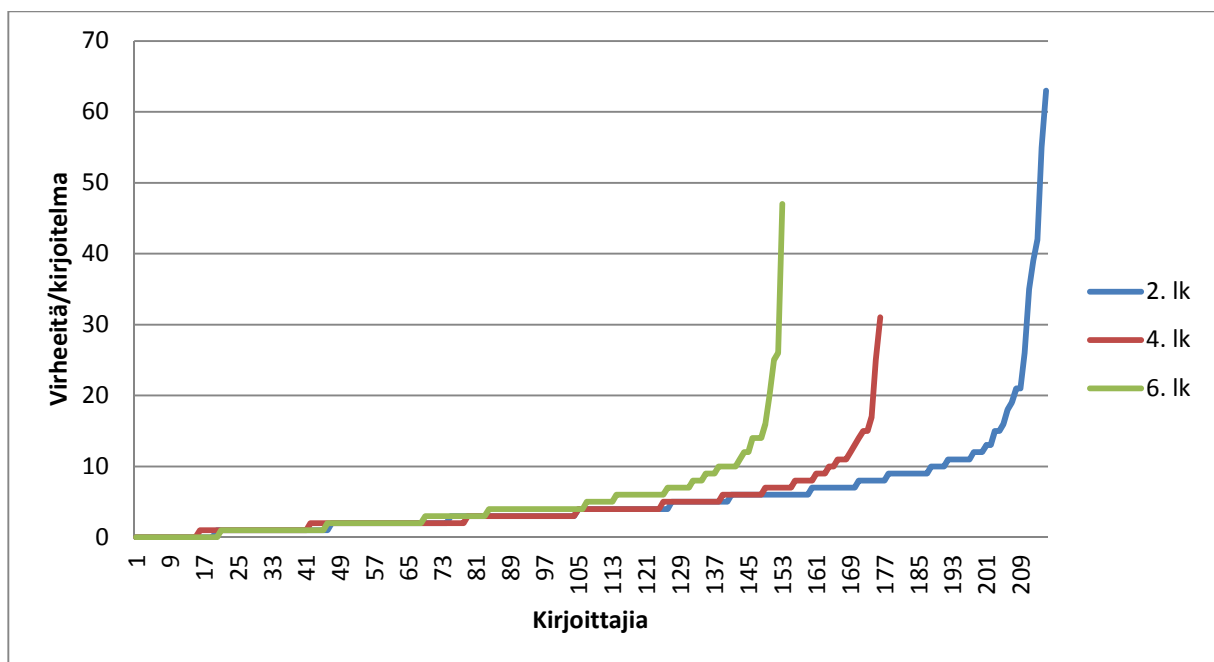
luokka-asteilla edelleen huomattava. Vaikeuden yhdyssanojen merkitsemisessä liittyvät ensinnäkin siihen, että suomen kielessä on joukko sanojen yhteen kirjoittamista ohjaavia sääntöjä, eikä näitä sääntöjä varsinkaan alakoulun ensimmäisillä luokilla järjestelmällisesti opetella. Toiseksi oppilaan sanatajun ja lausetajun kehittymättömyys voi johtaa siihen, ettei hän ymmärrä kahden tai useamman sanan muodostamaa kokonaisuutta uudeksi merkitykseksi eikä näin ollen osaa soveltaa tätä kielioppisääntöä. Kolmanneksi voidaan ajatella, että varsinkin aineiston vanhimpien oppilaiden kirjoitelmissa yhdyssanan erilleen merkitsemiseen on voinut vaikuttaa muiden heidän opiskelemiensa kielten erilaiset yhdyssanasäännöt. Esimerkiksi englannin kielessä kahta sanaa ei kirjoiteta välttämättä yhteen, vaikka ne muodostaisivat yhdessä eri merkityksen kuin erikseen kirjoitettuina (esim. suom. *pikku + sisko* = *pikkusisko*, vrt. engl. *little sister*).

Myös lukukongruenssin virheprosentit on huomattavan suuret kaikilla luokka-asteilla: toisen luokan norminmukaisiin vastineisiin verrattu virheluku on yli 40 %, neljännen luokan oppilaiden noin 26 % ja kuudennen luokan oppilaiden noin 9 %. Monikon 3. persoonan inkongruenssi on piirre, jonka voi lukea kuuluvaksi yleisiin ja neutraaleihin puhekielisyyspiirteihin. Vaikka inkongruenttien tapausten suhteellinen määrä väheneekin melkoisesti alempiin luokka-asteisiin verrattuna, on 9 %:n virheellisyysprosentti kuudennella luokka-asteella siitä huolimatta melkoinen verrattuna moneen muuhun tästä aineistosta tutkittuun virhetyyppiin. Inkongruenttien tapausten suhteellisen suuri määrä vielä ylimmilläkin luokka-asteilla kertoo kuitenkin siitä, että kyseessä on helposti puheesta kirjoitukseen siirtyvä puhekielen piirre, jonka välttäminen hyvään kirjoitelman kieleen pyrittäessä saattaa aiheuttaa hankaluuksia, vaikka normitetun kielen säännöt pääosin olisivat jo kirjoittajan hallinnassa. Voidaankin epäillä monikon 3. persoonan lukukongruenssin olevan sellainen piirre, joka ei ole vanhemmillekaan kirjoittajille kovin vakiintunut verrattuna vaikkapa äännepituuksien hahmottamiseen ja merkitsemiseen, joiden luokittaisen esiintymisen kehityskulku on kovin toisenlainen.

Geminaattojen ja pitkien vokaalien toteutuminen on nuorimpia todennäköisempää aineiston vanhimmilla oppilailta, ja näissä virhekatgorioissa kehityssuunnat ovat hyvin selvät: toisen luokan oppilaiden kirjoitelmissa on äännepituuksien hahmottamisesta kertovia ongelmia moninkertainen määrä vanhempiin oppilaisiin verrattuna. Konsonanttien pituudet ovat selvästi vaikeampia hahmottaa kuin vokaalien. Tätä selittänee ainekin se, että vokaaleja on helpompi ääntää ja näin konkretisoida äännepituuksia kuin konsonantteja. Puutteellisten geminaattojen määrä on toisen luokan oppilaiden kirjoitelmissa 6,4 % norminmukaisiin verrattuna, kun puutteellisten pitkien vokaalien määrä on vain 1,8 %. Erot ovat suuret

kahdella muullakin luokka-asteella, tosin virheprosentit ovat huomattavasti pienemmät: neljäsluokkalaisten geminaatoista 1,4 %:sta puuttuu konsonantti ja pitkistä vokaaleista 0,2 %:sta puuttuu vokaali. Kuudesluokkalaisten geminaatoista vain 0,9 % on puutteellisia ja pitkistä vokaaleista 0,3 %. Geminaattojen ja pitkien vokaalien toteutuminen ovat erityisesti fonologisen tietoisuuden kehittymistä mittaavia piirteitä. Äännepituuksien erottaminen ja oikea merkitseminen kertoo myös siitä, että oppilaan fonologisen koodaamisen taito on nopeutumassa ja automatisoitumassa.

Kirjoitelmien pituuden ja siten myös virheiden määrien hajonta on suurta jokaisella luokka-asteella. Sen vuoksi virheiden esiintymistä yksittäisissä kirjoitelmissa on myös mielekästä tarkastella yksilövariaation kautta. Kaavioissa 7 on esitetty kaikki analysoidut virheet kirjoitelmakohtaisesti niin, että kirjoitelmat on järjestetty vähävirheisimmästä runsasvirheisimpään sen mukaan, kuinka monta virhettä niissä esiintyy.



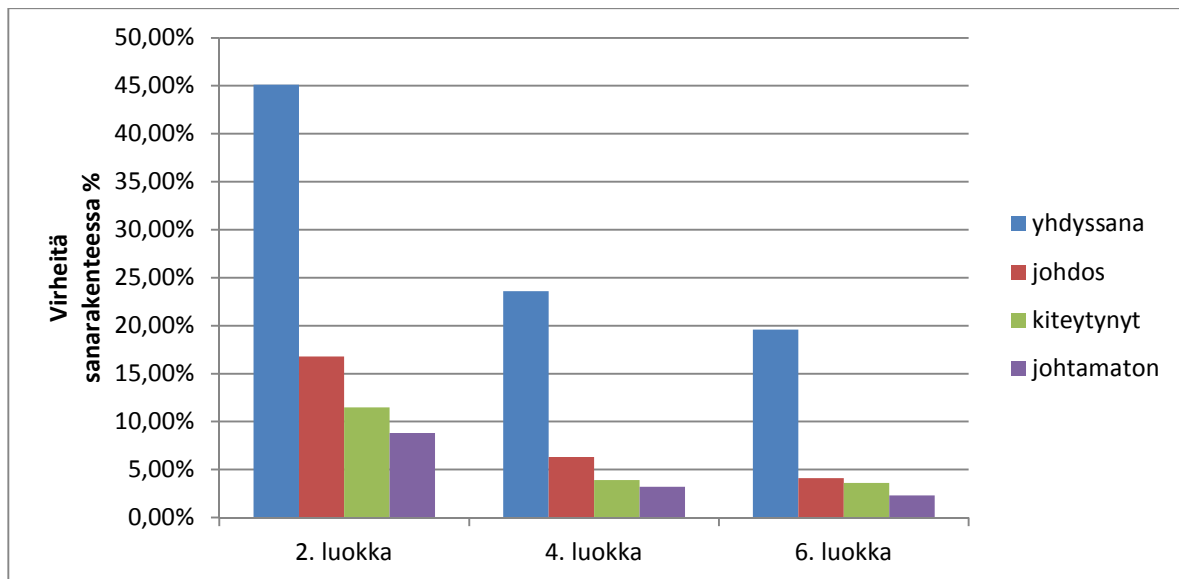
**Kaavio 7.** Kirjoitelmissa esiintyvien virheiden määrä yksilötasolla luokka-asteittain

Kaaviosta 7 käy ilmi, että suurimmassa osassa kirjoitelmia on virheitä melko vähän; alle kymmenen virhettä sisältäviä kirjoitelmia on 90 % kaikista kirjoitelmissa. Runsasvirheisiä kirjoitelmia on vähän, ja ero vähävirheisten ja runsasvirheisten kirjoitelmien välillä on radikaali. Siitä kertoo kaavio 7 kaikkien luokka-asteiden virhemäärien kuvaajien muoto.

Käyrien eksponentiaalinen nousu selittyy runsasvirheisillä kirjoitelmilla kaikilla luokka-asteilla. Toisen luokan runsasvirheisimmässä kirjoitelmassa esiintyy yhteensä jopa 63 virhettä, neljännen luokan 31 virhettä ja kuudennen luokan runsasvirheisimmässä kirjoitelmassa on 47 virhettä. Toisaalta erittäin runsasvirheisten kirjoitelmien osuus kaikista kirjoitelmista jää melko pieneksi: kymmenen tai enemmän virheitä sisältävien kirjoitelmien osuus on vain noin 10 %.

Koulukorpushankkeen yhteydessä ei ole kerätty tietoja oppilailla mahdollisesti todetuista luku- tai kirjoitusvaikeuksista tai muista mahdollisista kehityshäiriöistä. Tällaisten taustatietojen valossa tämän tutkimuksen tulosten tarkastelun yhteydessä olisi mahdollista luokitella tarkemmin niitä piirteitä, jotka mahdollisesti kertovat oppilaan lukiongelmista. Yksilövariaation virheanalyysipohjainen ja virhekategoriakohtainen tarkastelu tästä aineistosta olisi toki mahdollista, ja se onkin yksi mahdollinen jatkotutkimuksen aihe. Kuitenkin yksilövariaatiosta saatavat tulokset voivat tuottaa vain arvailuja oppilaan varsinaisista kirjoitustaidoista, jos hänen taustoistaan ja kehityksestään ei ole tietoa saatavilla. Yksittäisen oppilaan yksittäisen kirjoitelman tasoon kun vaikuttaa hyvin monenlaisia tekijöitä, kuten kirjoitustilanne ja senhetkinen keskittymiskyky.

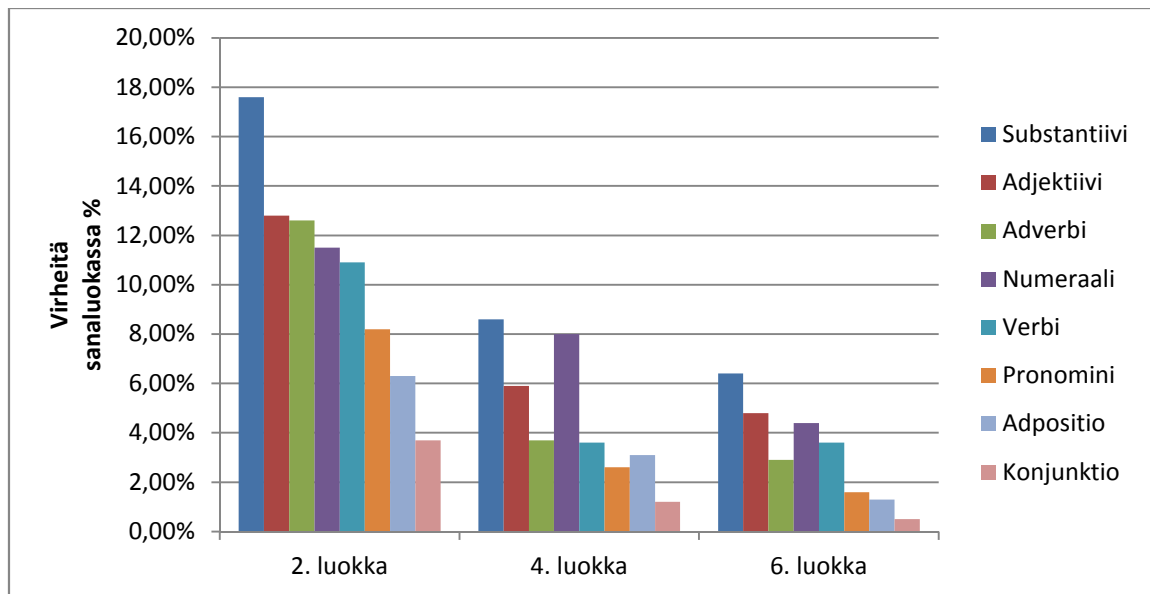
Tämän tutkimuksen analyysi pohjautuu ensisijaisesti Ruoppilan, Römanin ja Västin. (1969) perinteiseen virheanalyysiin, mutta mukailee Kuljun ja Mäkisen (2009) sanarakenneanalyysia niin, että tässä tutkimuksessa pyritään myös osin selvittämään virheiden sana- ja tavuusemia ja näin selittämään virheitä. Kuljun ja Mäkisen sanarakenneanalyysissa selviää, että erilaisissa rakenteissa tehdään erilaisia virheitä niin määrällisesti kuin laadullisestikin. Jotta tämän tutkimuksen tuloksia voitaisiin verrata Kuljun ja Mäkisen sanarakenneanalyysiin, on kaaviossa 8 esitetty analysoidut virheet sanarakennekohtaisesti. Kaaviosta 8 on jätetty pois aineiston alkuperäiset sanarakenteet symboli, vierassana ja sanaliitto sillä niissä esiintyy virheitä niin vähän, ettei virheitä ole mielekästä verrata muissa sanarakenteissa esiintyviin virheisiin.



**Kaavio 8.** Virheiden esiintyminen eri sanarakenteissa prosentteina esitettynä luokka-asteikohtaisesti.

Pelkkien virheiden määrien vertailu tuottaisi tässä kohtaa harhaanjohtavaa tietoa, sillä kirjoitelmat kaikilla luokka-asteilla koostuvat pääosin johtamattomista perussanoista; tutkimusaineiston sanoista lähes 74 % on johtamattomia perussanoja. Kaaviossa 8 esitetyt prosentit kuvaavat sitä, kuinka monta prosenttia kunkin sanarakenteen sanasta on sellaisia, joissa virheitä esiintyy. Kaaviosta huomaa, että yhdyssanat ovat ylivoimaisesti virheellisin sanarakenne kaikilla luokka-asteilla. Yhdyssanoissa esiintyvät virheet myös vähenevät eniten verrattaessa toisen ja neljännen luokan prosentteja keskenään. Toiseksi eniten virheitä on kaikilla luokka-asteilla johdoksissa ja vähiten johtamattomissa perussanoissa. Nämä tulokset vahvistavat ainakin välillisesti Kuljun ja Mäkisen (2009) oletuksia sanahahmon vaikutuksesta oppilaiden tekemiin virheisiin. Yhdyssanat ja johdokset ovat johtamattomia perussanoja kompleksisempia rakenteita ja näin ollen alttiimpia virheille.

Tutkimuksessa analysoituja virheitä on mielenkiintoista tarkastella myös sanaluokittain. Kaaviossa 9 on esitetty kunkin sanaluokan prosentuaalinen virheellisyys, eli kuinka monta prosenttia kunkin sanaluokan sanoista sisältää virheitä. Kaaviosta 9 on jätetty pois alkuperäisaineiston sanaluokista interjektiot, sillä sekä niiden virheettömät että virheelliset esiintymät ovat niin suppeita, ettei niitä ole mielekästä verrata muihin sanaluokkiin.



**Kaavio 9.** Virheiden esiintyminen eri sanaluokissa prosentteina esitettynä luokka-asteikohtaisesti.

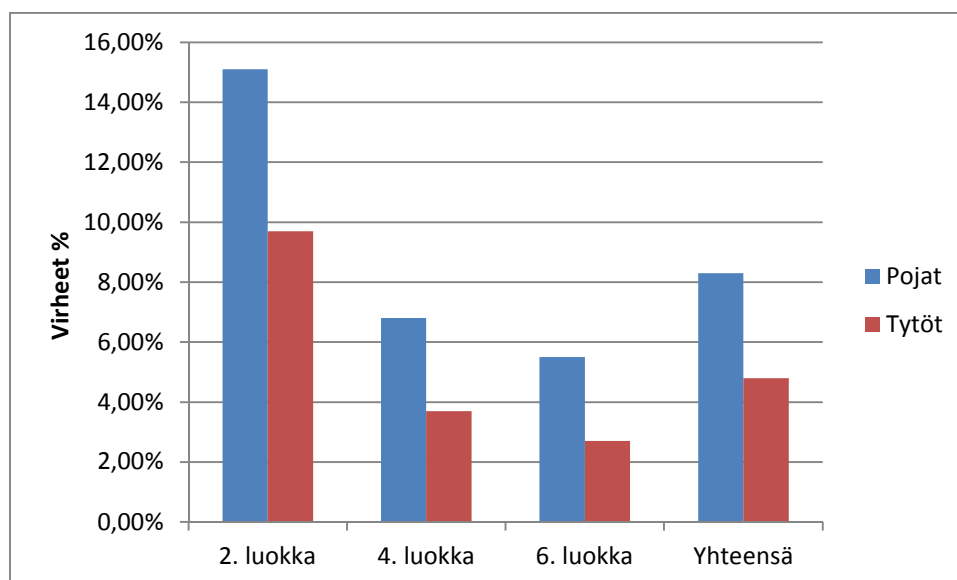
Kuten kaaviosta 9 käy ilmi, on substantiivi kaikkein virhealttein sanaluokka kaikilla luokka-asteilla: toisen luokan oppilaiden substantiiveista liki 18 % sisältää virheen, neljäsluokkalaisten substantiiveista lähes 9 % ja kuudesluokkalaisten substantiiveista yli 6 % on virheellisiä. Substantiivien muihin sanaluokkiin verrattuna suurta virheprosenttia selittää se, että sanaluokista juuri substantiivit ovat sanarakenteeltaan kompleksisimpia. Kaaviosta 8 kävi ilmi, että yhdyssanat ovat ylivoimaisesti eniten virheitä sisältävä sanarakenne, ja kaikista aineiston yhdyssanoista 88 % onkin substantiiveja. Vastaavasti taas johtamattomat perussanat sisältävät suhteessa muihin sanarakenteisiin vähiten virheitä, ja kaavion 9 osoittamista vähävirheisimmistä sanaluokista, konjunktioista ja adpositioista 87 % on johtamattomia perussanoja. Näin ollen sanaluokkakohtainen virheiden vertailu osaltaan myös vahvistaa Kuljun ja Mäkisen (2009) oletuksia sanarakenteen vaikutuksesta oppilaiden tekemiin kirjoitusvirheisiin.

Yksi tämän tutkimuksen tavoitteista on tuottaa tietoa tyttöjen ja poikien oikeinkirjoitustaidoista ja näin tarkastella, millaisia kehityskulkuja kummallakin sukupuolella ilmenee. Vertailun vuoksi on kaikki tutkimuksessa tutkitut virheet laskettu yhteen ja esitelty näin saadut virheiden kokonaismäärät luokka-asteittain ja sukupuolittain taulukossa 19.

**Taulukko 19.** Kaikkien tutkittujen virheiden esiintymät luokka-asteittain ja sukupuolittain.

Luokka-aste	Pojat	Tytöt	Yhteensä
2.	592	628	1220
4.	374	325	699
6.	375	294	669
<b>Yhteensä</b>	1341	1247	2588

Jo pelkästään virheiden kokonaislukumääristä voi huomata alakoululaisten kirjoitustaidon kehittämisessä kahdenlaisia kehityslinjoja: Ensinnäkin, virheiden kokonaislukumäärä vähenee alemmalta luokka-asteelta ylemmälle siirryttäessä. Toiseksi, poikien yhteenlaskettu kokonaisvirhemäärä on tyttöjä suurempi. Lukumäärällisesti tyttöjen kirjoitelmissa on hieman enemmän virheitä kuin poikien kirjoitelmissa toisen luokan aineistossa, neljännellä ja kuudennella luokalla poikien virhemäärä sen sijaan on suurempi. Luvut antavat suuntaa tarkastelulle, mutta niissä on huomioitava kirjoittajien välinen suuri hajonta. Kaikilla luokka-asteilla tyttöjen kirjoitelmat ovat sanamäärältään poikien kirjoitelmia pidempiä. Edellä mainitut kehityslinjat korostuvat, kun kokonaisvirhemäärät suhteutetaan kunkin luokka-asteen aineiston kokonaissanamäärään. Kaaviossa 10 virheiden määrät onkin suhteutettu luokka-asteiden ja sukupuolten kirjoitelmien kokonaissanamääriin ja näin laskettu kyseisiin muuttujiin verrattavat virheellisyysprosentit.



**Kaavio 10.** Kaikkien tutkittujen virheiden esiintymät aineistojen kokonaissanamääriin suhteutettuna luokka-asteittain ja sukupuolittain.



Kaaviossa 10 korostuvat sekä luokka-asteiden että sukupuolten väliset erot. Luokka-asteikohtaiset erot ovat huomattavasti selvemmat toisen ja neljännen luokan välillä kuin neljännen ja kuudennen luokan välillä: toisen luokan kokonaissanamäärään suhteutettu virheprosentti on lähes kolminkertainen sekä tytöillä että pojilla verrattuna neljäsluokkalaisten vastaavaan lukuun. Sen sijaan neljäsluokkalaisten kokonaisvirheprosentti on vain noin yksi prosenttiyksikköä suurempi kuin kuudesluokkalaisten virheprosentti molemmilla sukupuolilla. Nämä havainnot siis vahvistavat aiemmin esitettyä ajatusta siitä, että alakoululaisten kirjoitustaito kehittyy nopeammin alakoulun alemmilla luokilla kuin ylemmillä.

Kaavion 10 virheprosentit myös osoittavat, että virheiden kirjoitelmiin kokonaissanamääriin suhteutettu osuus ei ole kovin merkittävä: toisen luokan oppilaiden virheet vastaavat keskimäärin noin 12 %:a, neljäsluokkalaisten noin 5 %:a ja kuudesluokkalaisten noin 4 %:a luokka-asteiden kirjoitelmien kokonaissanamäärästä. Varsinkin neljäs- ja kuudesluokkalaisten virheprosentit ovat siis todella pieniä, mikä kuvastaa heidän oikeinkirjoitustaitonsa vakiintumista. Toisen luokan verrattain suurempi virheprosentti taas johtunee enimmäkseen kirjoituskokemuksen puutteesta ja runsaista kirjoittajakohtaisista eroista.

Sukupuolten väliset erot korostuvat selvästi kaaviossa 10: kaiken kaikkiaan poikien kokonaissanamäärään suhteutettu virheprosentti on lähes kaksinkertainen tyttöjen vastaavaan lukuun verrattuna. Molemmilla sukupuolilla virheiden suhteellinen määrä kyllä vähenee luokka-asteelta toiselle siirryttäessä, mutta huomionarvoista on, että tyttöjen virheluvut suhteutettuna poikien virhelukuihin ovat kaikilla luokka-asteilla huomattavasti pienemmät. Kehityssuunta kulkee niin, että toisen luokan poikien kokonaissanamäärään suhteutettu virheprosentti on yli 15 % ja tyttöjen alle 10 %, eli poikien virheprosentti on noin kolmanneksen tyttöjen virheprosenttia suurempi. Neljännelle luokalle siirryttäessä poikien virheluku laskee lähes 7 %:iin ja tyttöjen alle 4 %:iin. Kuudennella luokalla luvut ovat taas edelliseen verrattain pienemmät, mutta sukupuolten välinen ero on selvästi nähtävissä edelleen: poikien virheprosentti on vielä kuudennella luokalla (5,5 %) suurempi kuin tyttöjen neljännen luokan prosentti (3,7 %). Tyttöjen kehityskulku kiihtyy kuudennelle luokalle siirryttäessä niin, että heidän kokonaissanamäärään suhteutettu virheprosenttinsa on enää 2,7 %, eli alle puolet kuudesluokkalaisten poikien virheprosentista.

Tyttöjen parempaa luku- ja kirjoitustaitoa on perinteisesti pyritty tulkitsemaan ennen kaikkea tyttöjen nopeamman kehityksen avulla (ks. mm. Mäki 2002), mutta on viitattu myös persoonallisuuspierre-eroihin, jotka ovat yhteydessä kirjoitustehtävissä menestymiseen.

Selittävien tekijöiden erittely on hankalaa, koska ne todennäköisesti ovat vuorovaikutusyhteydessä. Yleisesti on todettu, että pojilla esiintyy enemmän kirjoitusvaikeuksia kuin tytöillä, joten poikien tyttöihin verrattuna suurempaa virheiden määrää voidaan selittää myös sillä, että pojilla on kaiken kaikkiaan enemmän sellaisia lukiongelmia, jotka ilmenevät heidän kirjoitustaitonsa heikkoutena.

Tyttöjen ja poikien välisistä koulumenestyseroista on tehty havaintoja etenkin äidinkielen ja kirjallisuuden oppiaineessa viime aikoina runsaasti. Esimerkiksi Lappalainen (2008: 87) selittää tyttöjen huomattavasti poikia parempaa kirjoitustaitoa sillä, että valtaosa pojista ei pidä äidinkieltä ja kirjallisuutta kiinnostavana oppiaineena. Tytöt muun muassa lukevat poikia enemmän vapaa-aikanaan, mikä osoittaa ja tukee kiinnostusta kyseiseen oppiaineeseen. Poikien ja tyttöjen erilaisia kiinnostuksenkohteita tulisi Lappalaisen mukaan ottaa nykyistä paremmin huomioon oppiaineen opiskelua suunniteltaessa esimerkiksi harjoittelemalla entistä tehokkaammin ja monipuolisemmin kirjoittamista oppilaiden mielekkäinä pitämässä tosielämän tilanteissa ja luonteissa yhteyksissä opiskeluun ja harrastuksiin.

Samantapaisia kehityskulkuja sukupuolten välillä on tehty muihinkin muuttujiin verrattuna Unelmieni päivä -kirjoitelmista. Tyttöjen kirjoitelmat (sekä lyhyet että pitkät) ovat kaikilla luokka-asteilla pidempiä kuin poikien ja sanamäärien vaihtelu on luokittain suurempi kuin pojilla (Pajunen 2012: 8). Unelmieni päivä -kirjoitelmien tasoa on arvioitu jakamalla ne erilaisiin kerronnan tasoluokkiin: kertomuksiksi, skripteiksi, kertomusskeemoiksi ja listoiksi. Tytöt kirjoittivat enemmän kertovia tekstejä kuin pojat: erityisesti kaikkein onnistuneimmat kertovat tekstit, kertomukset, ovat valtaosin tyttöjen kirjoittamia (Vesala 2012: 77). Pajusen (2012) mukaan tyttöjen kirjoitelmien taso on kaikilla mittareilla mitattuna poikia edellä koko alakoulun ajan, eivätkä suomalaispojat näytä saavuttavan normeja, jotka kansainvälisessä tutkimuksessa on oletettu kyseiselle ikäkaudelle.

Pajusen (2012: 28) mukaan kansainväliseen tutkimukseen verrattuna etenkin poikien huono suoriutuminen on huolestuttavaa. Hän arvelee tilanteen heijastavan muutoksia paitsi suomalaisessa yhteiskunnassa, perheessä ja koulujen antamassa opetuksessa myös harrastuksissa. Ulkoisten muutosten vaikutus viittaa siihen, että kielen ja sen osa-alueiden oppimisessa vuorovaikutus aikuisten kanssa on ratkaisevan tärkeää. Tämä johtopäätös ei tue niitä näkemyksiä, joiden mukaan lasten kielellinen kehitys perustuu yksin kielikyvyn synnynäisyyteen tai on yksin kehityksellistä. Tämän tutkimuksen tulokset puolustavatkin enemmän Tolchinskyn (2004) teoriaa siitä, että lapsen myöhempi kielenkehitys on vahvasti sidoksissa häntä ympäröiviin kulttuuriin tekijöihin ja niiden asettamiin vaatimuksiin.

Tulevaisuudessa tähän asiaan tulisikin kiinnittää erityistä huomiota, jotta sukupuoliin kohdistuvat erilaiset odotukset ja paineet sekä kiinnostuksen väheneminen äidinkielen ja kirjallisuuden oppiainetta kohtaan eivät kasvattaisi eroja entisestään.

Tässä tutkielmassa on haluttu kartoittaa alakoululaisten kielen norminmukaisuutta ja sanatason tarkkuutta toisella, neljännellä ja kuudennella luokka-asteella ja seurata, miten norminmukaiseen yleiskieleen pyrkivät koululaiset tehtävästään suoriutuvat. Tutkimuksessa tutkitut virheet muodostavat koko aineiston kokoon verrattain melko pienen joukon: virheiden määrä vastaa 6,1 %:a aineiston kokonaissanamäärästä. Voidaan siis sanoa, että alakoululaiset pyrkivät kirjoitelmissaan tuottamaan pääosin mahdollisimman norminmukaista ja huoliteltua yleiskieltä, vaikka alimmilla luokilla aloittelevien kirjoittajien kirjoitustaidoissa on huomattavan suuria eroja. Lisäksi aloittelevalle kirjoittajalle on koulukirjoitelmissa käytettävä yleiskieli tuntemattomampi kuin hänen puhetilanteissa käyttämänsä ja kuulemansa kieli. Niinpä koulukirjoitelman laatiminen vaatii myös tietoisuutta eri kielimuodoista ja taitoa soveltaa näitä tietoja tarkoituksenmukaisesti eri konteksteissa.

Luokka-asteiden 2–6 virheellisyyslukuja vertailtaessa huomataan, että oikeinkirjoitustaito kehittyy toisen ja neljännen luokan välissä huomattavasti mutta neljännen ja kuudennen luokan välissä erot tasoittuvat. Lähes kaikkien 13 tutkitun piirteen kokonaissanamäärään suhteutettu virheprosentti laskee alemmilla luokka-asteelta ylemmälle siirryttäessä. Erityisen paljon laskevat puuttuvan sanavälin ja puuttuvien kirjainten virheet. Puuttuvien sanavälien vähenemiseen liittyy monenlaisia tekijöitä: Ensinnäkin, oppilaiden motoriikka ja tekninen kirjoitustaito kehittyvät, jolloin kirjoituksesta tulee visuaalisesti selkeämpää; eri sanat siis erottuvat selkeästi toisistaan käsialan kehittyessä. Toiseksi oppilaiden sanataju kehittyy, ja he ryhtyvät hiljalleen ymmärtämään sana-yksikön merkityksen. Puuttuvien kirjainten väheneminen johtuu todennäköisesti kirjoitusprosessissa käytettävän väylän muuttumisesta fonologisesta suoraksi. Kokeneemman kirjoittajan kirjoitustaito on jo niin automatisoitunut, ettei hänen tarvitse käyttää äänne-erottelua. Poikkeuksia kirjoitustaidon kehittymistä implikoivasta kehityssuunnasta ovat ainoastaan virheet, jotka eivät ole kompetenssivirheitä, eli varsinaisia virheitä, vaan tahattomia lipsahduksia. Selvimmin tällaisia lipsahduksia ovat piirtovirheet, joita siis esiintyy neljäs- ja kuudesluokkalaisilla huomattavasti enemmän kuin aineiston nuorimmilla oppilailla. Virheiden luonteita tarkasteltaessa voidaan siis todeta, että varsinaiset virheet vähenevät mutta satunnaisten lipsahdusten riski kasvaa luokka-asteelta toiselle siirryttäessä.

Virheryhmiä, joiden esiintymät vähenevät radikaalisti luokka-asteelta toiselle siirryttäessä, voidaan siis pitää kirjoitustaidon kehittymisestä kertovina piirteinä.

Huomattavan osan aineiston virheistä kaikilla luokka-asteilla muodostavat yhdyssanavirheet ja kongruenssivirheet. Nämä virheet ovat tyypillisiä kielioppivirheitä, ja niitä voidaan selittää kahdella tavalla: Ensinnäkin, yhdyssanavirheet kertonevat sanatajun eli morfologisen tietoisuuden kehittymättömyydestä ja siitä, ettei yhdyssanojen kirjoittamista ohjaavien erityisiä kielioppisääntöjä osata soveltaa kirjoituksessa. Toiseksi kongruenssivirheet ovat tulkittavissa myös puhekielen piirteiksi. Näin ollen niiden runsas esiintyminen Unelmieni päivä -kirjoitelmissa kertoo, että kyseessä on helposti puheesta kirjoitukseen siirtyvä piirre, jonka välttäminen hyvään kirjoitelman kieleen pyrittäessä saattaa aiheuttaa hankaluuksia, vaikka normitetun kielen säännöt pääosin olisivat jo kirjoittajan hallinnassa. Näitä piirteitä esiintyy näkyvästi vielä ylimmilläkin luokka-asteilla, ja ne toiminevat tulevaisuudessa äidinkielen opetuksen haasteena standardin mukaista kirjoitettua kieltä opeteltaessa. Haasteita lisää se, että kirjalliseen kulttuuriimme syntyy kaiken aikaa uusia tekstilajeja, joissa kieliopillisten sääntöjen noudattaminen ei ole ei yhtä säännönmukaista tai edes vaadittavaa kuin kirjoitetussa kielessä perinteisesti ja entistä nuoremmat lapset kommunikoivat ja viettävät yhä enemmän aikaa näiden uusien medioiden ja tekstien parissa.

## Lähteet

- AHVENAINEN, OSSI – HOLOPAINEN, ESKO 2005: *Lukemis- ja kirjoittamisvaikeudet. Teoreettista taustaa ja opetuksen perusteita*. Jyväskylä: Kirjapaino Oma Oy.
- BAKKER, DIRK J. 1979: Hemispheric differences and reading strategies: Two dyslexias? – Springer US (toim.), *Annals of dyslexia* s. 84–100. Bulletin of Orton Society, osa 29, 1. painos.
- CORDER, S. PIT 1976: *Miten kielitiedettä sovelletaan*. Suomalaisen Kirjallisuuden Seuran Toimituksia 321. Helsinki: Suomalaisen Kirjallisuuden Seura.
- EHRI, LINNEA 1989: *The development of spelling knowledge and its role in reading acquisition and reading disability*. Journal of Learning Disabilities, 22 (6) s. 356–365.
- HOLOPAINEN, LEENA – SAVOLAINEN, HANNU 2006: Nuorten lukemisen ja kirjoittamisen vaikeudet. – Marjatta Takala & Elina Kontu (toim.), *Luki-vaikeudesta luki-taitoon* s. 203–229. Helsinki: Yliopistopaino.
- HONKO, MARI (tekeillä): *Frekvenssipohjaisten mittarien validiteetti kouluikäisten S1- ja S2-oppijoiden sanastollisen taidon arvioinnissa*.
- ISK = HAKULINEN, AULI – VILKUNA, MARIA – KORHONEN, RIITTA – KOIVISTO, VESA – HEINONEN, TARJA RIITTA – ALHO, IRJA 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- KETONEN, RITVA 2010: *Dysleksiariski oppimisen haasteena. Fonologisen tietoisuuden interventio ja lukemaan oppiminen*. Jyväskylä studies in education, psychology and social research 404. Jyväskylä: Jyväskylän yliopisto
- KULJU, PIRJO – MÄKINEN, MARITA 2009: *Sanelukirjoituksen arviointi – tyyppivirheistä sanarakenteiden fonologiseen analyysiin [The assessment of dictation exercises – from specific mistakes to the phonological analysis of word structures]*. – Virittäjä 4/2009 s. 508–532.
- KUORILEHTO, MIIA 2003: *Television Chat-ohjelmien puhekielisyydet ja kielelliset erityispiirteet*. Pro gradu -tutkielma. Tampereen yliopiston kieli- ja käännöstieteiden laitos.
- LAPPALAINEN, HANNU-PEKKA 2008: *On annettu hyviä numeroita. Perusopetuksen 6. vuosiluokan suorittaneiden äidinkielen ja kirjallisuuden oppimistulosten arviointi 2007*. Opetushallituksen oppimistulosten arviointi 3/2008. Helsinki: Edita Prima Oy.

- LUMILA, LIISA 2000: *TEXTAILLAAN!!!ÜÜÜ. Nuorten tekstiviestien kielen äänne- ja muoto-opilliset puhekielisyydet sekä erityispiirteet*. Pro gradu -tutkielma. Oulu: Oulun yliopiston suomen ja saamen kielen ja logopedian laitos.
- MANTILA, HARRI 1997: *Pohjoissuomalaisen puhekielen sosiolingvistinen variaatio* s.1–23. Oulun yliopiston suomen ja saamen kielen ja logopedian laitoksen julkaisuja 8. Oulun yliopistopaino.
- MIELIKÄINEN, AILA 1991: *Murteiden murros. Levikkikarttoja ja nykypuhekielen piirteistä*. Jyväskylän yliopiston suomen kielen laitoksen julkaisuja 36. Jyväskylä: Jyväskylän yliopiston monistuskampus.
- MÄKI, HANNA 2002: *Elements of spelling and composition. Studies on predicting and supporting writing skills in primary grades*. Turun yliopiston julkaisuja B: 255. Turku: Turun yliopisto
- OPETUSHALLITUS 2004: *Perusopetuksen opetussuunnitelman perusteet 2004*. Vammala: Vammalan Kirjapaino Oy.
- PAJUNEN, ANNELI 2012: *Kirjoittamistaitojen kehitys 8–12-vuotiailla. Alakoululaisten unelmakirjoitelmat*. – Virittäjä 1/2012 s.4–23.
- PAJUNEN, ANNELI – PURANEN, MAIJA – YLI-PAAVOLA, ANNUKKA 2007–2012: *Koulukorpus* (osia Koulukorpukselta). Tampereen yliopisto.
- PONSILA, MARJA-LIISA 2011: *Fonologisten taitojen yhteys lukemaan ja kirjoittamaan oppimiseen*. – Kaisa Launonen & Anna-Maija Korpijaakko-Huuhta (toim.), *Kommunikoinnin häiriöt. Syitä, ilmenemismuotoja ja kuntoutuksen perusteita* s. 77–94. Palmenia-sarja nro 12. Tallinna Raamatutrukikoda: Gaudeamus Helsinki University Press Oy Yliopistokustannus, HYY Yhtymä.
- RUOPPILA, ISTO – RÖMAN, KYLLIKKI – VÄSTI, MAIRE 1969: *KTL:n diagnostisia kirjoituskokeita peruskoulun II ja III luokille: Kokeiden standardointi ja kirjoitusvirheiden faktorianalyttinen tutkimus*. Kasvatustieteiden tutkimuslaitoksen julkaisuja 50. Jyväskylä: Jyväskylän yliopisto.
- SALONEN, IINA 2012: *Puhekielen piirteiden esiintyminen alakoululaisten kirjoitelmissa*. Pro gradu -tutkielma. Tampereen yliopiston kieli-, käännös- ja kirjallisuustieteiden yksikkö.
- SAVINAINEN-MAKKONEN, TUULA – KUNNARI, SARI 2004: *Systemaattisen kehityksen ja fonologisen viimeistelyn kaudet*. – Sari Kunnari & Tuula Savinainen-Makkonen (toim.), *Mistä on pienten sanat tehty. Lasten äänteellinen kehitys* s. 99–124. Helsinki: WSOY.

- STENBERG, PIA-RIITTA 2010: *Yhdyssanatajuna kehittyminen ja lingvistinen tiedonrakentelu peruskoulusta yliopistoon*. Pro gradu -tutkielma. Itä-Suomen yliopiston suomen kielen ja kulttuuritieteiden oppiaineryhmä.
- TAKALA, MARJATTA 2006: Lukemaan opettaminen. – Marjatta Takala & Elina Kontu (toim.), *Luki-vaikeudesta luki-taitoon* s. 13–36. Helsinki: Yliopistopaino.
- TAKALA, SAULI 1990: Lukeminen prosessinäkökulmasta. – Pirjo Linnakylä & Sauli Takala (toim.), *Lukutaidon uudet ulottuvuudet* s. 43–57. Jyväskylän yliopisto. KTL:n julkaisusarja B:61.
- TAKALA, SAULI – VÄHÄPASSI, ANNELI 1983: *On the specification of the domain of writing. Kohti kirjoittamisen kuvailua ja erittelyä*. Reports from the Institute for Educational Research, 333/1983. Jyväskylän yliopiston kasvatustieteiden laitos.
- THUNEBERG, HELENA 2006: Näkökulmia alakoulun luki-opetukseen. – Marjatta Takala & Elina Kontu (toim.), *Luki-vaikeudesta luki-taitoon* s. 117–201. Helsinki: Yliopistopaino.
- TOLCHINSKY, LILIANA 2004: The nature and scope of later language development. – Berman, Ruth Aronson (toim.), *Language development across childhood and adolescence*. Amsterdam, NLD: John Benjamins Publishing Company.
- VESALA, ANNA 2012: *Kerronnan keinot alakoululaisten kirjoitelmissa*. Pro gradu -tutkielma. Tampereen yliopiston kieli-, käännös- ja kirjallisuustieteiden yksikkö.