



**TURUN
YLIOPISTO**

Viidesluokkalaisten sanastoresurssit mielipidekirjoituksissa

Kasvatustieteellinen tiedekunta
pro gradu -tutkielma

Laatija:
Ida Ojuva

14.2.2025
Turku

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

Pro gradu -tutkielma

Oppiaine: Kasvatustieteiden laitos

Tekijä: Ida Ojuva

Otsikko: Viidesluokkalaisten sanaston piirteet mielipidekirjoituksessa

Ohjaaja: Professori Sara Routarinne

Sivumäärä: 39 sivua

Päivämäärä: 14.2.2025

Tässä tutkimuksessa tutkittiin viidesluokkalaisten oppitunneilla kirjoittamia mielipidekirjoituksia (N= ja niiden sanastoa. Tarkoituksena oli selvittää, mitä sanoja viidesluokkalaisten osaa käyttää mielipidekirjoituksessa, sekä millaisia sanoja hän käyttää tekstissään.

Tutkimuksen aineisto oli 2013 kerätty Rohkaisukeskeinen kirjoittamisinterventio -hankkeen yhteydessä keräämät mielipidekirjoitukset (ks. esim. Routarinne, S. & Tenhola, E. 2015. Tekstilajin tajuja 5.-luokkalaisten kirjoitelmissa). Tässä tutkimuksessa aineiston muodostivat 44 mielipidekirjoitusta. Tutkimus on laadullinen tapaustutkimus, ja sen aineisto analysoitiin aluksi tekstien sanat lemmaamalla eli muuttamalla ne niin kutsuttuun sanakirjamuotoon. Viidesluokkalaisten käyttämien sanojen listaa frekvensseineen verrattiin Kotimaisten kielten keskuksen keräämään Suomen sanomalehtikielen – taajuussanastoon (CSC, 2004).

Ensimmäisten viiden sanan osalta taajuussanasto ja tutkittava aineisto vastasivat toisiaan. Yleisin sana aineistossa oli olla-verbi, mikä selittyy sen roolilla useissa eri kielioppirakenteissa. Kehittyneempiä sananmuodostuksia ja johdoksia oli vain muutamia. Tyypillisin ilmaus mielipidekirjoituksissa mielipiteen esittämiseksi oli ”mielestäni”-rakenne. Käytetyt ilmaukset olivat myös pääasiassa hyvin tavallisia ja kalpeita sanavalintoja. Sanavarastosta saatu kuva jäi pinnalliseksi aineiston pienuuden takia.

Avainsanat: Leksikko, sanaston kehitys, kirjoitustaito, kirjoitustaidon kehitys, kielen kehitys, leksikaalinen kompetenssi, mielipidekirjoitus

Sisällysluettelo

1	Johdanto	1
2	Lapsen kielellinen kehitys	5
2.1	Kielen varhainen kehitys	5
2.2	Kielen kehitys keskilapsuudessa	6
2.3	Sanaston kehitys	7
2.4	Kirjoitustaidon oppiminen ja kehitys	8
2.4.1	Genretaidot	9
2.4.2	Mielipidekirjoitus genrenä	10
3	Leksikko eli sanavarasto ja leksikaalinen tieto	11
3.1	Sanan määrittelyn haaste	12
3.2	Yksilön sanasto	13
3.3	Sanasto-opetusta ohjaavissa asiakirjoissa	14
3.4	Olemassa olevia testausmenetelmät	15
4	Tutkimusongelma	18
5	Menetelmä	19
5.1	Aineisto ja aineiston käsittely	19
5.2	Aineiston analyysi	20
6	Tulokset	23
6.1	Mitä sanoja viidesluokkalaisten käyttää mielipidekirjoituksissa?	23
6.1.1	Viidesluokkalaisten tekstit suhteessa sanomalehtikieleen	23
6.1.2	Sanaluokat	26
6.2	Millaisia sanoja viidesluokkalaisten käyttää mielipidekirjoituksessa ja miten?	28
6.2.1	Modaaliverbit	28
6.2.2	Epäfrekventit sanat	29
7	Pohdinta	34
7.1	Johtopäätökset	34
7.2	Luotettavuus	36
7.3	Jatkotutkimus	38
	Lähteet	39

1 Johdanto

Kielitaidossa tärkeän osa-alueen muodostaa yksilön osaama sanasto, joka kehittyy yhdessä yksilön kasvun myötä. Sanasto on avain kykyyn ilmaista itseään, omia ajatuksiaan sekä korkeamman tason ajattelutaitoa. Sana- ja ilmaisuvarasto kertoo kielitaidosta, ja nuorten lasten kohdalla vielä kielen kehityksestä. Sanasto ja kieli ovat tärkeitä jokaisessa koulun oppiaineessa, ja siksi siihen olisi hyvä kiinnittää enemmän huomiota opetustyössä.

Sanavaraston tutkiminen on jäänyt Suomessa melko vähäiselle huomiolle, eikä sen merkitystä esimerkiksi lukusujuvuuden osatekijänä nähdä. Tämä tutkimus pyrkii antamaan pienen katsauksen viidesluokkalaisten käytössä oleviin sanastollisiin resursseihin.

Kielen ja kirjallisen osaamisen taso on herättänyt paljon huolta viime vuosina (PIRLS, 2021 s.22; KARVI, 2021, s. 79; PISA22 s.31). Sanasto-osaamiseen keskittyminen voisi tuoda uutta näkökulmaa myös näiden taitojen tukemiseen.

Varhaisessa kielen oppimisen vaiheessa sanat ja ilmaisu opitaan vuorovaikutuksessa.

Sanavarasto onkin yksi ikkuna kielitaitoon. Se sisältää sanastotiedon lisäksi paljon tietoa kielen käytöstä sekä kieliopista. Sanaston tutkimus on kuitenkin haastavaa. Vaikka sana voi ensivaikutelmaltaan tuntua selkeältä käsitteeltä, se on oikeasti hyvin moniulotteinen rakenne. Sanastotutkijat eivät ole pystyneet määrittelemään *sanaa* tai sanan osaamista tyhjentävästi (esim. Masonen, 2013, s.189). Sanan määrittelyn haastavuutta avaan myöhemmissä luvuissa tarkemmin. Sanavarasto myös vaikuttaa suoraan lukusujuvuuteen, akateemisiin taitoihin ja siten koulumenestykseen (Treffers-Daller & Milton, 2013, s. 8; Honko, 2013, s. 21). Mitä korkeammalle koulutusasteelle mennään, sitä enemmän opiskeluun liittyy lukemista sekä kirjoittamista.

Opettajalla on hyvä olla käsitys oppilaiden sanavarastosta sekä sen kehittymisestä, koska on olemassa tutkimusviitteitä siitä, että sanavaraston kehitys olisi suoraan yhteydessä lukusujuvuuteen (Laufer, 2012, s.2). Lukemista ennustaa niin sanottu nopean nimeämisen taito, eli lapsi pystyy nopeasti nimeämään kuvattavan asian. Lukemista nopeuttaa, jos lapsen ei tarvitse lukea koko sanaa, vaan hän tunnistaa sen. Tunnistaminen kuitenkin vaatii sitä, että sana on ennestään jo tuttu. Jos tekstissä on paljon vieraita sanoja, tekstin ymmärtäminen jää taka-alalle ja yksittäisten sanojen tavaaminen ylikorostuu ja lukeminen hidastuu. (Heikkilä,

2015, s.30) Suppea sanasto hidastaa tekstin prosessointia, koska sanat täytyy tavata erikseen vuorollaan (Nation, 2001, s.2). Myös kokonaiskuvan muodostaminen vaikeutuu. Rungas vieras sanasto voi johtaa siihen, ettei sanojen merkityksen hahmottaminen kontekstivihjeiden avulla enää ole mahdollista (Nation, 2001, s.2).

Aiemmin sanastoa on tutkinut Honko. Hänen Unelmien päivä- tutkimuksessa (2013) myös huomattiin, kuinka 8–10-vuotiailla sanaston hallinnan taitoerot säilyvät läpi kouluvuosien ainakin siinä tapauksessa, jos sanaston kasvattamiseen ei kiinnitetä erityistä huomiota. Tässä tutkimuksessa aineiston kirjoittaneet oppilaat ovat kirjoittaneet itselleen vieraampaa tekstilajia, mikä vaatii käyttöön erilaisia tekstitaitoja.

Lukemisen sujuvuus puolestaan vaikuttaa positiivisesti sanavaraston kehittymiseen. (Honko, 2013, s. 65) Jos lukeminen on helppoa ja sujuu, on oppilaalla todennäköisesti enemmän motivaatiota jatkaa lukemista, jolloin hän törmää taas uusiin itselleen vieraisiin sanoihin. Sanaston yhteys lukusujuvuuteen ja siten akateemisiin taitoihin on aiemman tutkimuksen perusteella vahva. (esim. Lee, 2011) On tärkeää tietää, miten sanastotaidot eli taito ilmaista itseään sanoin, ja sanaston laatu kehittyvät, jotta näitä taitoja on mahdollista seurata ja tukea.

Erityisesti alkuopetuksen jälkeen olisi hyvä lukusujuvuuden ja nimeämisnopeuden lisäksi seurata sanavaraston, kielellisen työmuistin ja kuullun ymmärtämisen etenemistä (Heikkilä ym., 2019, s.57). Sanastotaitojen aktiivisemmän opettamisen lisäämiseksi on tärkeää, että sanastotaitojen ja sanaston koon merkitys tiedostetaan. Sen analyysillä voidaan myös tarkastella alakoululaisten kielen kehitystä (Honko, 2013, s.73).

Suomen kielen osalta Hongon (2013), sekä Pajusen ja Hongon (2021) tutkimukset ovat olleet siis tärkeitä. Niissä tutkimuksen kohteena on pääasiassa ollut kertova- ja kuvaileva kirjoitelma. Eri tekstilajien tuottamisvalmiuksien näkökulmasta sanastotaitoja ei alakoululaisilla ole juuri tutkittu eri tekstilajien tuottamisen kontekstissa. Kieli- ja käännöstieteiden laitokselta löytyy joitakin pro gradututkielmia, joissa on keskitytty ylä- ja alakoululaisten sanarakenteiden kompleksisuuteen (esim. Elomäki, 2011; Salmenoja, 2013). Myös varhaisen sanaston yhteyttä esilukutaitoon lukivaikeusriskilapsilla on pro gradututkielmassaan tutkinut Tentke (2022).

Tutkimuskentällä näyttää siis olevan selkeä aukko siinä, mitä alakouluikäisten sanaston käytöstä tiedetään kuvailevan kertomuskontekstin ulkopuolella. Sanasto-osaamisen merkitys näkyy lukusujuvuuden ja -innon yhteydessä. Tämä tutkimus pyrkii tarjoamaan näkökulmaa siihen, millaisia sanasto-osaamisen resursseja viidesluokkalaisella on käytössään mielipidekirjoituksen kontekstissa.

Kuitenkin on huomioitava, että sanaston kokoa on mahdotonta kartoittaa täydellisesti 'sanan' ja sanaston moniulotteisuuden takia. Tekstin analyysin avulla voi kuitenkin saada käsitystä siitä, millaisia sanoja oppilaalla on aktiivisessa käytössä. Mielipidekirjoitus genrenä vaatii erilaista sanasto-osaamista kuin esimerkiksi viestittely tai esseiden kirjoittaminen (Heikkinen ym., 2012, s.28).

Sanaston hallinta on myös yhteydessä tekstitaitojen kehittymiseen ja siten kirjalliseen kerrontaan (Pajunen ja Honko, 2021a, s. 65). Sanojen yleisyyden ja niiden käyttöyhteyden tunteminen auttaa opettajaa tavoitteiden asettamisessa sekä näiden tavoitteiden saavuttamisen seuraamisessa (Honko, 2013, s. 99). Kuitenkin laajaa tietoa tai tutkimusta leksikaalisesta kehityksestä – erityisesti suomalaisilla lapsilla – ei ole, mikä vaikeuttaa testaamisen tehtävänantojen suunnittelua, kuten Honko esitti (2013, s. 22). Kirjoitelmien sanaston tutkimus on yksi ikkuna siihen, millaisessa vaiheessa kielen oppiminen on. Myös sanaluokkien tarkastelu antaa kuvaa siitä, miten oppilas ajatuksiaan ilmaisee (Pajunen & Honko, 2021, s.13).

Termillä 'leksikko' viitataan yleensä kielen sanastoon ja sanavarastolla tarkoitetaan yksilön osaamia sanoja. Tässä tutkimuksessa kuitenkin myötäillään Hongon (2013) tapaa, jossa sanastoa, sanavarastoa, sekä leksikkoa käytetään rinnasteisina termeinä.

Oppilaiden kirjoitustaitoja, sekä lukutaitoa on tutkittu alakouluikäisten kohdalla paljon. (Pajunen & Honko, 2021a, s.14; Hiidenmaa, 2019, s.159; Kauppinen ym., 2015, s. 161) Kuitenkin käsitys siitä, mitä oppilaalta milläkin luokalla voi odottaa sanaston ja siten ilmaisujen tasolla on vähemmän tutkittu. Tämä johtunee siitä, että tutkimusta juuri suomalaisten lasten sanaston kehityksestä ja sanaston koosta ja laadusta on hyvin vähän (Pajunen & Honko 2021b, s. 65) Tässä tutkimuksessa tarkoitus on antaa käsitystä siitä, millaista sanavarastoa ja ilmaisutaitoa viidesluokkalaiselta on mahdollista odottaa mielipidekirjoituksessa, sekä testata sanaston tutkimusmenetelmää. Täysin vastaavanlaista

tutkimusta ei ole aikaisemmin Suomessa tehty. Kuten edellä kävi ilmi, sanasto on jäänyt suomalaisessa tutkimuksessa hyvin vähäiselle huomiolle.

Suomalaisten lasten keskilapsuuden sanaston kehittymisen tutkimuksessa tärkeän ja perinpohjaisen avauksen teki Honko. Hänen tutkimuksensa keskittyi S2-oppilaiden sanavarastoon, mutta tutkimukseen osallistui S2-oppilaiden lisäksi S1-verrokkeja. Hänen tutkimuksessaan analysoitiin ”Unelmien päivä”- tekstejä, joissa eri-ikäiset oppilaat kuvailivat unelmiensa päivää. Kuitenkin tässä tutkimuksessa aineistona käytetyt oppilaskirjoitelmat ovat kanta-aottavia mielipidetekstejä, joiden sanasto eroaa Unelmien päivä -tutkimuksen kertovasta tekstistä. Mielipidekirjoitus on viidesluokkalaisten vieraampi tekstilaji. Se myös rajoittaa ilmaisuvalintoja enemmän kuin Hongon (2013) tutkimuksen tuttu ja kertova teksti.

Mielipidekirjoituksessa käytetty ilmaisu vaatii tyyliltään vaikuttamaan ja vahvaan ilmaisuun pyrkivää argumentointia. Vakuuttavuutta lisää mielipidekirjoituksessa argumenttirakenteet, argumentoinnin tapa ja ilmaisun sävy (Aalto, 2020, s.18) sanastoa. Aineiston tekstit ovat viidesluokkalaisten oppitunnilla kirjoittamia tekstejä.

Tulevissa luvuissa avataan aluksi hieman kielen kehitystä varhais-, sekä keskilapsuudessa. Tämän lisäksi kielen kehitystä tarkastellaan erityisesti sanaston oppimisen näkökulmasta, minkä jälkeen keskitytään kirjoittamisen oppimisen vaiheisiin. Luvussa 3 keskitytään sanan, sanaston sekä sanasto-osaamisen määrittelyyn. Samalla sivutaan opetusta ohjaavien asiakirjojen näkemystä sanasto-osaamisesta ja tarkastellaan kouluissa käytössä olevia testipattereita sanaston näkökulmasta. Luvussa 5 esitellään tutkimuksen aineisto, sekä menetelmä ja luvussa 6 avataan saatuja tuloksia. Lopuksi luvussa 7 tuloksia suhteutetaan aiempaan kirjallisuuteen, sekä pohditaan tutkimuksen luotettavuutta.

2 Lapsen kielellinen kehitys

2.1 Kielen varhainen kehitys

Edellytykset kielen oppimiselle alkavat jo sikiönä. Alkuun lapsi pyrkii kommunikoimaan elein, ilmein ja itkulla. Vähitellen hän oppii ensimmäiset sanansa. Nämä sanat ovat yleensä sellaisia, joita hoivaaja toistaa paljon. Sanojen oppimisen yhteydessä hän alkaa harjoitella lauseen muodostamista ja pidempien ilmausten rakentamista. (Buckley, 2003, s. 35, 79, 115, 147)

Lapsen kielen kehitys siis alkaa juuri kielen perussanastosta, sanonnoista ja fraaseista. (Pajunen & Honko, 2021a, s. 9). Siitä, montako sanaa vuodessa lapsi oppii, on hyvin vaihtelevaa tietoa (Treffers-Daller & Milton, 2013, s. 6). Yhtenä syynä lienee se, että eri kielissä sanasto toimii eri tavoilla. Lapsen ensisanat ja puhe kohdistuvat usein konkreettisiin ja läsnä oleviin asioihin, joita hän kuulee toistettavan monta kertaa päivässä. Vähitellen lapsi oppii erottamaan puhehetken ja puheen sisällön toisistaan eli hän oppii puhumaan asioista, joita ei tapahdu juuri puhehetkellä (Honko, 2013, s.72). Hän oppii siis kertomaan tarinoita.

Lapsen sanaston ja kielen kehitykseen vaikuttaa myös vahvasti se, millaisessa kieliympäristössä hän kasvaa (Masek ym., 2021, s.275). Kieliympäristöä muovaa se, miten lapsen hoivaaja käyttää kieltä lapsen läsnä ollessa. Esimerkiksi kuvaamalla koiraa tarkemmin (vrt. *suloinen vanha kissa – kissa*) hän voi syöttää lapselle useamman sanan ja opettaa kuvailutaitoja (Masek ym., 2021, s.275). Näin kieli opitaan vuorovaikutuksessa toisten kanssa (Pajunen, 2012, s.4).

Sanaston koko ja taidot käyttää sanastoa kehittyvät yleensä samassa tahdissa lapsen kognitiivisten taitojen kanssa (Honko, 2013, s. 7; Alatalo ym., 2018, s.66). Lapsi havainnoi ja kuuntelee tarkkaan ympärillä olevien aikuisten toimintaa ja puhetta yrittäen matkia tuotettuja ääniä (Tomasello, 2005, s. 39). Opittujen ja käytettyjen sanojen laatu riippuu siis siitä, missä kognitiivisen kehityksen vaiheessa lapsi on. Tässä tutkimuksessa hyödynnetään Piaget'n teorian vaiheita (ks. esim. Cerovac, M., & Keane, T., 2024)

Alkuun lapsi pyrkii kommunikoimaan ilmein ja elein, kun sanojen muodostaminen ei vielä suju (Tomasello, 2005, s.31). Lapsen puhe alkaa muodostua, kun hän siirtyy Piaget'n kuvaamasta sensomotorisesta vaiheesta esioperationaaliseen vaiheeseen. Lapsi on tässä vaiheessa hyvin minäkeskeinen ja pystyy ymmärtämään vain oman kokemuksensa ja

näkemyksensä, minkä takia myös hänen käyttämänsä sanat ovat hyvin konkreettisia ja kuvailua on vähän.

Lapsen sanasto kehittyy kokemuksen ja alussa erityisesti puheen pohjalta. Kommunikoinnin kannalta tärkeimmät ja samalla yleisimmät lekseemit opitaan yleensä ensimmäisenä (Pajunen & Honko, 2021b, s. 72). Kun lapsi oppii yksittäisiä sanoja, täytyy hänen alkaa harjoittelemaan myös niiden yhdistämistä kokonaisiksi lauseiksi (Tomasello, 2005, s. 40). Koulun alkaessa lapsi on yleensä siirtynyt kognitiivisessa kehityksessään konkreettisten operaatioiden vaiheeseen, jolloin hän pystyy jo rakentamaan sujuvia lauseita ja kokonaisuuksia, sekä irtautumaan puhehetkestä, sekä kehittämään tarinoita.

2.2 Kielen kehitys keskilapsuudessa

Tarinoinnista alkaa nopea kielellisen kehityksen vaihe, jossa kielelliset tiedot ja taidot kehittyvät ja rikastuvat (Pajunen & Honko, 2021a, s. 9). Kielellistä kehitystä nopeuttaa luku- ja kirjoitustaidon oppiminen, mutta ne eivät kuitenkaan ole välttämättömiä tämän kehityksen kannalta (Pajunen & Honko, 2021a, s. 10). Altistuminen kirjallisuudelle jo ennen lukemisen oppimista edesauttaa luetunymmärtämisen sekä mekaanisen luku- ja kirjoitustaidon kehittymistä (Mol & Bus, 2011, s. 285). Vanhempien lukemat iltasadut ja perheen yhteiset lukuhetket siis tukevat lapsen kielen ja sanaston kehitystä.

Kolmanteen luokkaan mennessä lähes poikkeuksetta jokainen oppilas on oppinut edes mekaanisen luku- sekä kirjoitustaidon, mikä tarkoittaa kielellisen kehityksen ja erityisesti sanaston koon räjähdysmäistä kasvua (Pajunen & Honko, 2021a, s. 17). Nyt oppilas voi kuluttaa ja tuottaa tekstiä itsenäisesti aivan uudella tavalla, jolloin sanavarasto kasvaa lukiessa ja tarve uusien sanojen käyttöön lisääntyy omien tekstien kirjoittamisen myötä.

Viimeisen ajattelun kehityksen vaiheen lapsi saavuttaa noin 12 vuoden iässä. Tätä vaihetta Piaget kuvaili formaalien operaatioiden vaiheeksi, missä lapsi pystyy hahmottamaan ja käsittelemään abstrakteja asioita sekä käsitteitä. Tämän pitäisi siis näkyä myös sanastossa ja sen kehittymisessä, sillä lapsen kokemusmaailma on taas muutoksessa (Pajunen, 2012, s.5).

Ajattelun kehityksen näkökulmasta tutkimuksen osallistujat ovat mielenkiintoisessa vaiheessa: viidennellä luokalla osa oppilaista on jo voinut siirtyä konkreettisista operaatioista formaaleihin operaatioihin, minkä pitäisi näkyä heidän kyvyssään perustella ja kuvailla mielipiteitään. Sanaluokista juuri kuvailevat adjektiivit, sekä adverbit tulevat viimeistään

tässä vaiheessa käyttöön (Pajunen & Honko, 2021a, s.19). Sanastoprofiileista pitäisi siis olla nähtävissä viitteitä oppilaan ajattelun kehityksestä. (Pajunen & Honko, 2021a, s.16)

2.3 Sanaston kehitys

Myös kielen oppimisen näkökulmasta sanastolla on suuri merkitys. Suuri osattujen sanojen määrä eli sanaston laajuus tukee ilmaisutaitoja. Kielioppi tarvitsee tuekseen laajaa sanastoa kehittyäkseen (Treffers-Daeller, 2013, s.3). Lapsen sanaston kehitys alkaa reseptiivisen sanaston kasvamisesta ja vasta myöhemmin lapsi alkaa itse käyttää osaamaansa sanastoa.

Oppiakseen sanoja ja niiden merkityksiä lapsen on osattava erotella kuulonvaraisesti sanoja aikuisen puheesta sekä yhdistää niitä viittauskohteisiin. Alle vuoden iässä lapsen sanastosta suurin osa on substantiiveja, koska verbien ja adjektiivien omaksuminen on hitaampaa. (Alatalo ym., 2018, s.66) Verbejä kuitenkin omaksutaan varhaiseen sanastoon hyvin nopeasti. Viimeisenä sanastoon omaksutaan pronomineja, konjunktioita, prepositioita, määrän ilmaisuja ja kysymyssanoja (Alatalo ym., 2018, s.70).

Puhe onkin usein kirjoitettua kieltä yksinkertaisempaa, sillä kirjoittaessa virkkeet pitenevät ja monipuolistuvat (Pajunen & Honko, 2021b, s. 63). Kirjoittaja joutuu miettimään sanavalintojaan tarkemmin, jolloin sanavaraston laajuudesta on hyötyä. Niinpä myös kirjoittaessa aluksi ilmaisut ovat hyvin lyhyitä ja ilmaisut helposti jopa töksähteleviä, koska kirjallisen ilmaisun taidot vielä kehittyvät ja sanavarasto on suppeahko (Pajunen, 2012, s.14). Lukemisen näkökulmasta laaja sanavarasto on hyödyllinen, koska lukeminen on nopeampaa, kun luettavat sanat ovat ennestään tuttuja. (Honko, 2013, s. 100)

Erytisesti alakouluiässä kirjalliset ilmaukset kompleksistuvat, kun sanapituus kasvaa ja sanarakenne monimutkaistuu (Honko, 2013, s. 17). Kompleksisessa sanassa sanan vartalo sisältää useita morfeemeja. Tällaisia ovat erimerkiksi johdokset eli kantasanaa johdetut monimutkaisemmat ilmaisut, kuten *jalka* sanasta johdettu *jalkine*, sekä yhdyssanat. (VISK §145). Sanaston kompleksistuminen tarkoittaa siis johdoksien, yhdyssanojen ja sanaliittojen hyödyntämistä, jolloin ilmaukset myös tarkentuvat tarkoitukseltaan ja sävyltään (Pajunen, 2014, s.14).

Yksittäinen oppilas saattaa osata useampia ja monipuolisempia ilmauksia, jotka eivät näy puhutussa kielessä. Lekseemit eli sanat kertovat myös syntaktisten taitojen kehityksestä (Pajunen & Honko, 2021b, s. 75). Vihjeitä syntaksin kehityksestä ovat esimerkiksi modaaliverbit, joita tässä tutkimuksessakin on tarkasteltu. Modaaliverbit ovat yleensä osana

verbiketjua, joiden määrä korreloi positiivisesti leksikon laajuuden kanssa (Honko, 2013, s.24). Ne esiintyvät yleensä osana verbiketjua (VISK §1564) tai nesessiivirakennetta (VISK §1565). Jotta mentaalinen kielioppi pääsee kehittymään, tarvitaan sanastotaitoa (Treffers-Daller & Milton, 2013, s. 3).

Kuten edellä esitettiin; Suomessa lasten sanaston kehitystä on tutkinut Honko Unelmien päivä-tutkimushankkeessaan (2013). Tutkimuksessa aineistoon kuului lähes 1900 tekstiä, joissa oli 230 000 sanaa. Tutkimukseen osallistui henkilöitä alakoululaisesta aikuisiin. Tutkimuksen päätavoite oli tutkia S2-oppilaiden sanastoa suhteessa S1-verrokkeihin.

Kirjoittamisen kannalta hallittu leksikko on osa toiminnallisia tekstitaitoja sekä keskeinen kommunikatiivisen kielitaidon kehittymisessä (Honko, 2013, s. 23). Hongon ja Pajusen tutkimuksessa on kuitenkin keskitytty lapsille tuttuun kertovaan tekstilajiin. Vähemmälle huomiolle on jäänyt se, millaista sanastoa yksilö hyödyntää muissa kirjoittamistehtävissä, joiden tavoitteet ja rakenteet ovat hänelle vieraampia. Sen takia tässä tutkimuksessa tarkastellaan sanastoa juuri toisenlaisessa tekstilajissa, eli mielipidekirjoituksessa.

2.4 Kirjoitustaidon oppiminen ja kehitys

Kirjoitustaitojen aktiivinen oppiminen alkaa viimeistään lapsen aloittaessa koulun. Alussa jo ennen luku- ja kirjoitustaidon opetusta lapsen taito havainnoida ympäristössä tuttuja kokosanoja ja logoja. Tätä vaihetta nimetään logografiseksi vaiheeksi (Lerikkanen, ym., 2013, s.12). Kirjoittamisessa tämä näkyy yrityksenä jäljentää kirjoittamista, vaikka oikeita kirjaimia ei vielä osata. Seuraavaksi lapsi osaa osittaisalfabeettisessa vaiheessa käyttää osaa sanaan liittyvistä kirjaimista. Alfabeettiseksi vaiheeksi kutsutaan vaihetta, jossa lapsi oppii äännekirjainvastaavuuden. Tästä lähtee luku- ja kirjoitustaidon todellinen kehitys. (Lerikkanen, ym., 2013, s.14)

Kirjoittamiseen valmistavia taitoja ovat karkea – ja hienomotoriikka. Ne koordinoivat käden suurempia ja sormen pienempiä tehtäviä. Tämän lisäksi vaaditaan visuaalisen hahmottamisen taitoja, eli taitoa ohjata kättä esimerkiksi oikealle riville. (Lerikkanen, ym.,2013, s.128–129)

Tämä myöhäisen kehityksen aika on sanavaraston kehityksen aikaa (Pajunen, 2012, s.14). Kirjoitettu kieli eroaa puhutun kielen rakenteista hyvin paljon (Beers & Nagy, 2011, s. 2). Ilmaisujen on oltava täsmällisempiä, koska mahdollisuutta korjausvuoroon, eli tarkenteiden lisäämiseen kuten keskustelussa, ei ole. Kirjoittamista ja kielentämistä voikin pitää kognitiivisena ongelmanratkaisuna (Pentikäinen, ym., 2017, s.164). Alkuopetuksessa

tavoitteena on, että oppilas saavuttaa tarkan lukutaidon ja sanaston oikeinkirjoitustaidon (Heikkilä ym., 2019, s.56).

Kirjoitustaidon oppiminen vaatii myös erilaisten kirjallisten genretyyppien oppimista, kun lapsi harjoittelee eri tekstilajien kirjoittamista ja kielimuodon hallintaa (Beers & Nagy, 2011, s. 2; Pentikäinen ym. 2017, s. 167). Tekninen taito sekä genretaitoja on hallittava aikuisena, koska kirjoittaminen on hyvin keskeinen työelämässä (Pentikäinen ym., 2017, s.160). Kirjallisen ilmaisun muodon hallinta on tärkeää myös siksi, että tekstilajeilla on erilaisia tehtäviä. Tekstilajin hallinta tarkoittaa myös sen tarkoituksen hallintaa (Pentikäinen ym., 2017, s. 173) ja tätä taitoa tarvitaan myös eri tekstilajien lukemisessa ja tulkinnessa.

2.4.1 Genretaidot

Kuten jo aiemmin todettiin, kirjoittaminen ei ole vain kirjoittamisen tekniikan ja kieliopin hallintaa. Tietyn tekstilajin kirjoittaminen vaatii myös tietoa eri genreistä ja niiden piirteistä (Heikkinen, ym., 2012, s. 29).

Lapsen on siis sanaston, kieliopin ja motoriikan lisäksi hallittava eri tekstilajien skeemat ja tietää niiden tyypilliset piirteet, jotta hän osaa itse tuottaa vastaavanlaisen tekstin. Genre on muun muassa kielellisten, tekstuaalisten, retoristen, kognitiivisten ja sosiaalisten piirteiden kokonaisuus (Heikkinen, ym., 2012, s.65). Nämä eri piirteet lapsi oppii vuorovaikutuksessa eri tekstilajien kanssa. Hänen on kirjoitustapahtuman aikana osattava valita oikeat piirteet ja ilmaisut aina tekstilajikohtaisesti (Routarinne & Tenhola, 2015, s.174). Kuitenkin siitä, miten lapset oppivat eri genrejä kirjoittamaan, on suomen kielessä hyvin vähän tutkimusta.

Koulussa kirjoittamista opetetaan jo ensimmäiseltä luokalta asti. Opetussuunnitelmassa mainitaan myös eri tekstilajeihin tutustuminen osana suomen kielen ja kirjallisuuden opetusta (POPS14, s.106; 162–163). Mielipiteen ilmaisu on opetussuunnitelmassa myös alaluokilta asti. Ylemmillä luokilla aletaan esittämisen lisäksi myös vaatia perustelutaitoa. (POPS14, s.106; 162–163) Genrepedagogiikassa tarkoitus on siis kiinnittää oppilaiden huomio siihen, millaisia rakenteita ja sanoja eri tekstilajeissa käytetään (Rose, 2012, s.213; Enright, Toledo & Wright, 2023, s.44). Siinä tarkoituksena on kiinnittää oppijan huomio tekstin tarkoitukseen ja kohdeyleisöön, jotta viestiminen olisi mahdollisimman tehokasta ja tarkoituksenmukaista.

2.4.2 Mieliidekirjoitus genrenä

Mieliidekirjoituksessa kirjoittaja ilmaisee oman näkemyksensä jostain aiheesta. Teksteinä mieliidekirjoitukset ovat keskusteleivia sekä kantaaottavia. Siksi ne sisältävät yleensä erilaisia argumentointikeinoja ja niiden kieli on neutraalia kirjakieltä. Mieliidekirjoituksella voi avata uuden keskustelun tai vastata aiempaan kirjoitukseen. (Heikkinen, ym., 2012, s.192) Oppijan näkökulmasta mieliidekirjoituksessa harjoitellaan näkökulmanottokykyä sekä argumentointitaitoja.

Kaunokirjallisesta tekstistä mieliidekirjoitus eroaa myös siinä, että niissä useammin puhutellaan lukijaa (Routarinne & Tenhola, 2015, s.178) Nykyoppilaille mieliidekirjoitus voi olla haastava tekstilaji, koska sanomalehtien lukeminen on vähentynyt. He todennäköisesti yhdistävät mieliidekirjoituksen nettialstan kommentointiin. On tärkeää harjoitella mieliideen esittämistä myös kirjallisessa muodossa. Jos oppilas pystyy eläytymään perustelutilanteeseen, hänen perustelutaitonsa voivat parantua (Harjunen & Korhonen, 2011, s.403)

3 Leksikko eli sanavarasto ja leksikaalinen tieto

Leksikko tarkoittaa sanastotiedon varastoa. Tämä varasto on kaiken kielellisen toiminnan pohja ja välttämätön osa sekä merkittävä lauseenmuodostuksessa käytetty rakennuspalikka. (Honko, 2013, s. 64). Leksikaalinen tieto on hyvin monitasoista, sillä se kattaa lekseemien muodon, aseman, merkityksen, käyttötarkoituksen, sekä käyttörajoitukset (Nation, 1990, s. 200).

Jokaisella on olemassa oma mentaalileksikkonsa eli osattujen sanojen varasto, johon sanat tallentuvat. Sanat tallentuvat mentaalileksikkoon hyvin erilaisilla tavoilla. Yhdestä sanasta voi olla muistissa vain yksi taivutusmuoto ja toinen on vain osa jotain fraasia, josta irrallaan sanan käyttö ei onnistu. Leksikko kasvaa vuorovaikutuksessa toisten kielen käyttäjien kanssa (Nation, 2001, s.2).

Sanasto ei siis ole vain abstrakti joukko tai aakkostettu lista sanoja. Se yhdistää eri kielen osa-alueet käytössä. Perusalueena voi ajatella sanojen merkityksiä, mutta sanastossa sisällä on vahva yhteys muihin kielenhallinnan osa-alueisiin, mikä muodostaa kielen sanastollisen verkon. Sanasto on jatkuvassa vuorovaikutuksessa käyttäjän ympäristön kanssa. (Puro, 1999, s.6-7)

Mentaalileksikossa sanat ovat joko produktiivisia tai reseptiivisiä (James, 2009, s. 117). Produktiiviset sanat ovat sellaisia, joita kielen käyttäjä osaa käyttää oikein ja oikeassa tilanteessa. Reseptiiviset sanat käyttäjä tunnistaa, ja hän osaa päätellä niiden merkityksen kontekstista, mutta itse hän ei näitä sanoja tuota. Tämä jaottelu on kuitenkin jyrkkä, eikä aina riittävä, koska sanan osaamiseen liittyy muitakin osa-alueita: sanan eri merkitykset, johdokset, taivutussäännöt, kieliopilliset rakenteet, rekisteri eli käyttöympäristö, assosiaatiot ja kollokaatiot eli sanayhdistelmien vakiintuneita merkityksiä esimerkiksi voidaan käyttää ilmaisua *suuri rakkaus*, mutta ei ilmaisua *iso rakkaus* (Honko, 2013, s.81). Sanan osaaminen siis tapahtuu asteittain ja sanaa käyttämällä ja mahdollista käyttöympäristöä testaamalla (Salo & Mäntylä, 2017, s.67).

Tässä tutkimuksessa keskitytään juuri oppilaiden mentaalileksikon produktiiviseen osaan, koska tutkimuskohteena ovat oppilaiden kirjoittamat tekstit. Tarkemmin tutkimuksen kohteena ovat siis oppilaiden esille tuomat näytteet leksikostaan. Tämän lisäksi hieman kiinnostaa myös pragmatiikka; oppilaiden keinot suunnata kohti tavoittelemaansa tekstilajia.

3.1 Sanan määrittelyn haaste

Sanaston tutkimuksen kannalta haastavaa on kuitenkin se, että 'sanalle' ei ole tyhjentävää määritelmää, eikä sanan osaamiselle löydy kattavaa kuvausta (Masonen, 2013, s. 3). Usein se määritelläänkin tutkimuksessa aina uudelleen. Näiden haasteiden välttämiseksi Cruse (2004, s. 85) suosittelee prototyyppisen sanan kuvausta: Sen mukaan prototyyppisen sanan paikkaa ja asemaa voi lauseessa vaihtaa ja sen merkitys voi muuttua, kun lauseeseen lisätään sisältöä. Hänen mukaansa sanaa ei voi katkaista keskeltä tai sen osien järjestystä vaihtaa. Cruse (1986, s. 24) on myös tarjonnut toista määritelmää: leksikaalinen yksikkö, joka tarkoittaa pienintä semanttista yksikköä.

Sanaston rajauksessa ongelmia tuottaa myös kielisidonnaisuus. Sanojen taivutuksessa on paljon eroa eri kielten välillä, ja taivutusmuodot voivat olla hyvin erilaisia kantasanaan nähden (Honko, 2013, s. 52). Vaikka suomen kielessä hyvä-parempi-paras- tyyppinen taivutus onkin hyvin harvinainen, joissain kielissä se on tyypillisempi. Sanan taivutusmuotoja ei kuitenkaan tarvitse opetella jokaista erikseen, eikä jokainen muoto siis tallennu leksikkoon omana yksikkönään (Honko, 2013, s. 55). Leksikaaliset yksiköt eivät ole irrallisia, vaan ne liittyvät osaksi leksikon verkkoa (Honko, 2013, s. 58; Nagy & Scott, 2000). Leksikaalinen tieto siis rakentuu vanhan päälle ja on siten sidoksissa muihin sanoihin.

Crusen määritelmät pohjautuvat kuitenkin englannin kielelle, eikä niitä voi soveltaa suomen kieleen täysin. Sanan määrittely on hyvin kielisidonnaista ja riippuu siis siitä, minkä kielen näkökulmasta sanan käsitettä lähestytään (Honko, 2013, s.52). Termin sana haasteellisuudesta huolimatta se on yksi ensimmäisistä metakielellisistä termeistä, joita kouluopetuksessa lähdetään harjoittelemaan jo ensimmäisellä luokalla lukemaan opettelyn yhteydessä (POPS14, s.107).

Suomen kielessä sanojen rakenne on joustavampi. Kielitoimiston sanakirja määrittelee sanan kielen pienimmäksi itsenäiseksi merkityssisällölliseksi rakenneosaksi. Sana on yksi yhteneväinen kirjainjakso. Puheessa ja tekstiyhteydessä oleva sanan muoto on nimeltään sane. (VISK § 864) Kielitoimiston sanakirjan määritelmä ei kuitenkaan ole kovin kattava, sillä myös 'morfeemi' voidaan määritellä kielen pienimmäksi itsenäiseksi merkityssisällölliseksi rakenneosaksi, vaikka on olemassa myös sidonnaisia morfeemeja eli taivutuspäätteitä, joita ei esimerkiksi englannin kielessä ole. Englanti puolestaan hyödyntää funktiosanoja esimerkiksi prepositioita, joilla on vain kieliopillinen merkitys. Isossa suomen verkkokieliopissa (VISK)

sanat jaetaan perussanoihin, eli yksimorfeemisiin, jakamattomiin lekseemeihin, sekä kompleksisiin sanoihin, joita ovat johdokset ja yhdyssanat (VISK §145).

Tämän määrittelyn vaikeuden takia, sana määritellään tässä tutkimuksessa lekseemin avulla. Lekseemi on abstraktio, joka sisältää kaikki sanan taivutusmuodot, kuten sanakirjan hakusana. Sana on semanttinen yksikkö, joka sisältää kaikki taivutusmuodot, jotka tekstissä voivat esiintyä saneina. (Puro, 1999, s.8)

3.2 Yksilön sanasto

Yksilö käyttää myös hyvin erilaista kieltä ja sanastoa eri tilanteissa. Luokkatilanteessa oppija ei yleensä törmää kuin tiettyyn määrään sanoja, mikä johtuu sanojen eri yleisyysasteista (Laufer & Nation, 2012, s.163–164). Yleisten sanojen osaaminen ja oppiminen varsinkin vierasta kieltä opeteltaessa on tärkeää, koska ne tulevat kielen käyttäjää vastaan useammin ja kattavat suuren osan käytetystä kielestä (Honko, 2013, s. 99). Sanojen omaksuminen riippuu vahvasti siitä, miten paljon sanoja toistetaan (Milton, 2009, 26; Blom & Unsworth, 2010, s. 208).

Kielen sanaston oppimisessa on myös tilanteisen haasteen lisäksi määrällisiä ja laadullisia haasteita. Määrällinen ongelma liittyy siihen, että kieltä oppiessa oppija – oli kyseessä kotikieli tai toinen kieli – joutuu opettelemaan suuren määrän sanoja. Sanojen laadullinen haaste puolestaan on se, että sanoihin liittyy monia erilaisia ominaisuuksia kuten juuri käyttötarkoitus, -yhteys ja -tapa. (Laufer & Nation, 2012, s.163-164)

Sanasto-osaamisen vaatimukset myös muuttuvat usein oppiainekohtaisesti, kun eri oppiaineiden kohdalla tavoitteena on oppia sen omaa kieltä ja termistöä, joihin oppija ei arjessaan välttämättä muuten törmäisi. Tätä ilmiötä kuvaa Cummingsin termi CALP eli cognitive academic language proficiency, joka korostaa kielen oppimista erilaisissa akateemisissa konteksteissa (Cummins, 2008, s.72).

Erityisalojen sanastoa ei voi siis oppia kattavasti kielen tunnilla, vaikka oppikirjat pyrkivät tarjoamaan monipuolisia tekstejä (Kotilainen, ym., 2019, s.613). Nykyinen opetussuunnitelma ohjaa jokaisen aineen kohdalla opettajaa opettamaan alan erityissanastoa osana aineen opetusta (POPS 14, 28). Jotta sanan voi oppia ja sanavarastoaan kasvattaa,

täytyy sanoja päästä myös käyttämään (James, 2009, s.242), mikä ei välttämättä toteudu aina lukuaineiden tunneilla.

Sanaston laajuus antaa myös taidot taitavampaan kielenkäyttöön. Meara (1996, s.3) korostaa oppijoiden ollessa samalla tasolla muilla osa-alueilla ne, joiden sanavarasto on laajempi, erottuvat edukseen muista tasoisistaan.

3.3 Sanasto-opetusta ohjaavissa asiakirjoissa

Sanaston oppiminen on merkittävä osa kielitaidon oppimista. Erityisesti vieraan kielen opetuksessa usein keskitytään sanastotiedon kartuttamiseen. Usein saattaa jopa käydä niin, että opetuksessa keskitytään sanastoon ja sanalistoihin muiden osa-alueiden kustannuksella (Honko, 2013, s. 22). Kuitenkin ensikielen opetuksessa ei edes opetussuunnitelman tasolla mainita sanaston oppimista. Suomen kielen ja kirjallisuuden ensimmäisen ja toisen luokan tavoitteissa, sekä kolmas-kuudesluokkalaisten sisällöissä mainitaan sana- ja käsitevaraston laajentaminen osana monimuotoisten tekstien erittelyä ja tulkintaa, mutta suoraa sanasto-osaamista korostavaa tavoitetta (POPS 14, s.106; 162–163).

Sanasto tulee esiin taas suomi toisena kielenä oppimäärän tavoitteissa muutamassakin kohdassa, joissa juuri korostetaan kommunikatiivista osaamista kuten ensimmäisen ja toisen luokan tavoitteessa T2 ”harjaannuttaa ja rohkaista oppilasta kertomaan ajatuksistaan ja tunteistaan ja keskustelemaan erilaisista aiheista sekä vahvistamaan niissä tarvittavaa sanastoa ja ilmaisua” sekä kirjallisissa taidoissa tavoitteessa T6 ”ohjata oppilasta harjoittelemaan eri tekstilajien ja niissä tarvittavan sanaston ja kieliopillisten rakenteiden tuottamista”. (POPS14, s. 119–120)

Kolmas-kuudesluokkalaisten suomi toisena kielenä tavoitteissa niin ikään sanasto ja sen harjoittelu mainitaan tavoitteessa T8 ”auttaa oppilasta syventämään taitoaan suunnitella ja tuottaa tekstejä itsenäisesti ja ryhmässä sekä käyttämään monipuolisesti niissä tarvittavaa sanastoa ja kieliopillisia rakenteita” (POPS 14 s.186) Suomen kielen ja kirjallisuuden tavoitteessa T6 ”ohjata oppilasta päättelevään tekstien tulkintaan sekä laajentamaan sana- ja käsitevarantoaan”. (POPS14, s.162) Alaspesifin käsitteistön hallinta kuitenkin mainitaan jokaisessa oppiaineessa erikseen.

Eurooppalaisessa viitekehyksessäkin sanaston osaamista ei erotella omaksi taitoalueekseen muualla kuin A1 ja A2 tasoilla. Sanoja on tunnistettava ja ymmärrettävä, jolloin ne ovat esimerkki konkreettisesta kielitaidosta, mutta korkeammille tasoille siirryttäessä maininnat

sanastosta loppuvat. (EVK, 2008, s. 50–51) Ylemmillä tasoilla ei siis sanaston osaamisen erittelyä koeta tarpeellisenä, vaan kielen osaaminen näkyy muualla. Kuitenkin leksikaalinen kompetenssi määritellään ”kielen sanaston ja kieliopin elementeistä koostuvaksi sanavaraston tuntemukseksi ja kyvyksi käyttää sitä” myös eurooppalaisen viitekehyksen asiakirjoissa (EVK, 2008, s. 157).

Sanasto-osaaminen mainitaan Perusopetuksen opetussuunnitelmassa suomen kielen ja kirjallisuuden oppimäärässä ohimennen osana monilukutaidon tavoitteita. Sen sijaan suomi toisena kielenä -tavoitteissa sanasto-osaamiseen ohjataan kiinnittämään enemmän huomiota. Tämä voi mahdollisesti johtua siitä, että sanasto-osaamista pidetään niin itsestään selvänä, ettei sitä erikseen korosteta tavoitteissa. Kuitenkin kielen ja kirjallisuuden osaamisen taso on hyvin heikkoa, ettei tällaista oletusta voi enää perustellusti tehdä (PIRLS, 2021 s.22; KARVI, 2021, s. 79; PISA22 s.31),

3.4 Olemassa olevia testausmenetelmiä

Suomalaisessa koulujärjestelmässä on systemaattista oppilaiden osaamisen testausta, vaikka valtakunnallisia, standardoituja kokeita ei ole ennen ylioppilaskirjoituksia. Tavallisten, opettajan pitämien kokeiden lisäksi oppilailla teetetään moniammatillisessa yhteistyössä monenlaisia standardisoituja testejä, joiden tarkoituksena on seurata lapsen kehitystä omaan tasoonsa sekä ikätasoon nähden. Näiden testien tulosten perusteella pyritään tunnistamaan oppilaan mahdollisen tuen tarve.

Seuraavaksi esitellään vain muutamia käytössä olevia testistöjä, joita teetetään suurimmalle osalle oppilaista. Opettajat voivat myös hieman valikoiden päättää, mitä tarjolla olevista testistöistä käytetään. Sanastotaitoihin keskittyviä tai edes huomioivia testistöjä ei oikeastaan löytynyt.

Kielen kehityksen alkuvaiheessa 6–24 kuukauden iässä lapsille teetetään Sanaseula-testistö. Tässä vaiheessa testataan ymmärretyn ja ilmaistun sanaston kehitystä 9–18 kuukauden ikäisillä ja ilmaistun sanaston kokoa 18–24 kuukauden iässä. Testin tarkoitus on seuloa ajoissa esiin lapset, joilla saattaa olla puheen kehityksen viivästyä. (Stolt & Vehkavuori, 2018)

Lukemisen ja matematiikan oppimis- ja arviointiympäristö (LUKIMAT) on testipatteri, jota käytetään oppilaan oppimisen arviointiin. Testin teettää usein erityisopettaja, minkä jälkeen tuloksia ja tarvittavia jatkotoimenpiteitä pohditaan moniammatillisesti luokanopettajan ja erityisopettajan johdolla. Tässä katsauksessa keskitytään esittelemään vain luku- ja kirjoitustaidon arvioinnin osaan. LUKIMAT-testauksen tarkoitus on antaa tietoa ja välineitä lukemisen, kirjoittamisen ja matematiikan tuen tarpeen tunnistamiseen ja oppimisen seurantaan. Testi on suunnattu esi- ja alkuopetusikäisille.

Lukemisen osalta testi arvioi ensimmäisellä luokalla fonologisia taitoja, sekä kokoavaa lukemista ja toisella luokalla teknistä lukutaitoa ja oikeinkirjoitusta. Testi tehdään oppilaiden kanssa kolmena eri ajankohtana ensimmäisellä ja toisella vuosiluokalla: syksyllä, talvella ja keväällä. Toisella luokalla lukutaitoa ja kirjoitustaitoa arvioidaan lukutehtävillä sekä oikeinkirjoitusta testaavilla sanelutehtävillä. Kaikki oppilaat tekevät tehtävät itsenäisesti, ja opettaja arvioi jokaisen testin erikseen kooten pisteet omalle taulukolleen.

Lapsen pisteitä verrataan laajaan ikäverrokeista kerättyyn aineistoon persentiilien eli prosenttipisteiden avulla. Tällöin oppilaan suoritus on mahdollista suhteuttaa toisiin samanikäisiin. (Eklund ym., 2013; Eklund ym., 2013)

Vaikka testissä arvioidaankin lukutaitoa sekä sanojen tunnistamista ja oikeinkirjoitusta, ei se testaa oppilaan sanavarastoa tai sen kehitystä. Testi tunnustaa sanojen tunnistustaidon tärkeyden lukusujuvuuden edistämisen kannalta, mutta ei keskity siihen.

Allu-testi on tarkoitettu alakoulun kaikille vuosiluokille 1–6. Sen tarkoitus on kartoittaa oppilaan lukivaikeuksia, teknistä lukutaitoa ja luetun ymmärtämistä. Allu-testin avulla on mahdollista seurata oppilaan kehitystä vuosiluokalta toiselle sekä tunnistaa lisätukea tarvitsevat oppilaat. Testin avulla on mahdollista suhteuttaa oppilaan osaamista ikätasoonsa tai arvioida, minkä luokkatason osaamista yksittäisen oppilaan suoritus vastaa. Testi ei kuitenkaan perustu suomen kielen ja kirjallisuuden opetussuunnitelmaan tai kirjasarjoihin. Testin teko on aikataulutettu maaliskuulle, eli sen tarjoama taitotaso vastaa jokaisen vuosiluokan maaliskuun taitoja. Tehtävät tehdään digitaalisessa ympäristössä. Alkuluokilla testiin kuuluu neljä osaa: sanojen tavuttaminen, sanojen erottaminen toisistaan, kirjoitusvirheiden etsiminen ja epäsanojen etsiminen. Isommilla oppilailla testiin kuuluu luetun ymmärtämisen osio. Tehtävien suorittamisessa oppilaalla on aikaraja, joka vaihtelee

luokka-asteen mukaan. (Lindeman, 1998) Testeissä näkyy sanastotaitojen testaaminen, mutta oppilaan leksikon kokoa tai kasvua ei tässä testistössä seurata.

YTTE-testi arvioi luetun ja kuullun ymmärtämisen sekä lukemisen sujuvuuden taitoja 2.–3.vuosiluokalla. Materiaali on tehty käyttöön opettajille, erityisopettajille, psykologeille ja muille lukutaidon kanssa työskenteleville ammattilaisille. YTTE-testistössä oppilaita testataan yksitellen. Luetunymmärtämistä testataan tarinoilla ja selittäville asiategsteillä. Ennen lukemista varmistetaan oppilaan ennakkotiedot aiheesta, koska se voi vaikuttaa luetunymmärtämiseen. Lukemisen jälkeen oppilas vastaa suullisesti kysymyksiin. (Kajamies, ym., 2008)

Testistö tunnistaa esitiedon auttavan lukemisen ymmärtämisessä ja lukusujuvuudessa. Oppilaalle tuttuja sanoja voi siis löytyä tekstistä, mikä helpottaa tekstin lukemista.

Sanaseula -testauksen jälkeen ei siis testauksessa enää suoraan keskitytä sanaston oppimiseen ja kasvattamiseen. Sen jälkeen testistö alkaa painottamaan lukemisen oppimisen valmiuksia. Myöhemmissä testeissä korostuu lukujusuvuus ja -tarkkuus sekä oikeinkirjoitus.

Vaikuttaa siis siltä, että sanasto-osaamiseen ei ole kiinnitetty erityistä huomiota testien kehittälyvaiheessa. Tämä voi osaltaan johtua siitä, että sanaston testaaminen on hankalaa, kuten myös Honko (2013, s.20) tuo esiin. Toisaalta ehkä sanasto-osaamiseen olisi tulevaisuudessa kiinnitettävä enemmän huomiota myös testauksessa.

4 Tutkimusongelma

Suomessa lukutaidon ja lukusujuvuuden tukemisessa sanaston näkökulma on jäänyt vähemmälle huomiolle eikä tutkimusta sanaston kehityksestä ja sen etenemisestä juuri ole tehty kehittyvästä kielitaidosta (katso luku 2.3). Tämä näkyy esimerkiksi standardoitujen testien painoalueissa ja sanastollisen osaamisen huomiotta jättämisessä (katso luku 3.4).

Sanaston osaaminen korostuu kuitenkin kaikilla kielitaidon alueilla: puheessa, kuuntelemisessa, lukemisessa ja kirjoittamisessa (katso luku 3.1). Tekstejä usein tutkitaan kirjoitustaitojen ja genretaitojen näkökulmasta. Tässä tutkimuksessa tarkoitus on muodostaa käsitys viidesluokkalaisten produktiivisesta sanastosta, jota he osaavat käyttää uudessa kirjallisessa kontekstissa ja siitä, miten he ilmaisevat itseään kirjallisesti mielipidetekstissä. Tavoitteena on siis yrittää löytää vastauksia kysymykseen:

1. Mitä sanoja viidesluokkalainen käyttää mielipidekirjoituksessa?
2. Millaisia sanoja viidesluokkalainen käyttää ja miten?

5 Menetelmä

5.1 Aineisto ja aineiston käsittely

Aineistoon kuului monia eri tekstilajeja edustavia tekstejä, joita kirjoitettiin viitenä eri kirjoituskertana. Koko aineistoon kuului monia eri oppilastekstejä, jotka edustivat eri tekstilajeja muun muassa tarinoita, ohjeita, tietotekstejä ja jännityskertomuksia. Tekstein kerääminen aloitettiin viidennellä luokalla ja keruuta jatkettiin vielä, kun oppilaat olivat jo 6. luokalla.

Tekstit kerättiin eteläsuomalaisesta alakoulusta lukuvuoden aikana. Aineisto on vuonna 2013 Routarinteen ja Absetzin (2013) keräämä Rohkaisukeskeinen kirjoitusinterventio- hanke ja hankkeen ”Rokki” -aineisto, joka koostuu erilaisista viidesluokkalaisten teksteistä, joita oppilaat olivat kouluvuoden aikana kirjoittaneet.

Alkuperäisessä aineistossa tarkasteluluokkia oli kolme, mutta tässä tutkimuksessa keskitytään vain kahden luokan teksteihin, koska kolmas ryhmä ei ollut kirjoittanut mielipidekirjoitusta milläkään kirjoituskerralla. Kirjoitelmien aiheisiin tai oppilaille annettaviin ohjeistuksiin eivät Routarinne ja Absetz olleet antaneet tarkempia neuvoja, vaan antoivat opettajien ohjata kirjoitustilannetta ja -tehtäviä heille itselleen sopivalla tavalla.

Tutkimukseen valikoitui kahden luokan kirjoittamat mielipidetekstit, joita on yhteensä 44 kappaletta. Mielipidetekstit valittiin tutkimuskohteeksi, koska ne ovat oppilaille vieraampi tekstilaji. Se myös antaa oppilaille mahdollisuuden osoittaa enemmän erityistä sanasto-osaamista, kuin kertova teksti. Luokalta E tekstejä oli 23 ja luokalta G 21. Luokkien kirjaimet on salaustarkoituksessa vaihdettu. Oppilaita oli siis kahdelta eri luokalta samasta koulusta.

Molempien luokkien mielipidetekstit ovat neljännen kirjoituskerran tuotoksia eli ne on kirjoitettu viidennen luokan lopussa. Oppilaat ovat tällöin olleet 11–12-vuotiaita. He ovat saaneet päättää kirjoituksensa aiheen itse ja otsikoineet tekstinsä valitsemansa aiheen mukaan. Ennen tekstien kirjoittamista opettajat ovat kuitenkin hieman johdatelleet aiheeseen keskusteluttamalla luokkaa aiheista, joista heillä on mielipiteitä.

Mielipidetekstejä on jokaiselta oppilaalta yksi eli yhteensä 44 tekstiä. Tekstien pituudet vaihtelivat G luokalla 44 sanasta 145 sanaan, kun taas E luokan tekstit vaihtelivat 53 ja 195

sanan välillä. Jokaisella tekstillä on oma kirjain ja numero koodi, joista ei voi tunnistaa kirjoittajia.

Eettisestä näkökulmasta aineisto ei vaatinut tutkimusluvan hakua, sillä se oli koodattu jo valmiiksi niin, ettei tässä vaiheessa ollut mahdollista tunnistaa edes kouluja tarkasti. Anonymiteettiä lisäsi vielä keräysajankohdasta kulunut aika. Tutkimuslupa on aikanaan hankittu koululta aineiston keruun aikaan ja vanhempia informoitu tehdystä tutkimuksesta. Oppilailla on myös ollut mahdollisuus kieltäytyä osallistumisesta. Siitä, onko aineistossa koko luokan tekstit, ei enää tässä vaiheessa ole tietoa. Aineistosta oli käytettävissä enää vain digitoitu ja koodattu aineisto. Alun perin tekstit ovat kuitenkin olleet käsin kirjoitettuja.

Aineisto luovutettiin vasta salassapitosopimuksen allekirjoittamisen jälkeen ja sitä käsiteltiin yksityisessä ympäristössä, missä kukaan ulkopuolinen ei päässyt siihen käsiksi. Tekstejä, sekä sanalistoja säilytettiin Turun yliopiston omalla salatulla palvelimella, joka on tarkoitettu juuri tutkimusaineistojen ja muiden opiskelijoiden salausta tarvitsevien tiedostojen säilyttämiseen. Raportointi on noudattanut vastuullisuuden ja huolellisuuden periaatteita.

Tutkimus on monimenetelmällinen, mutta kallistuu enemmän laadulliselle kannalle. Tämä johtuu siitä, että korpuksen eli tekstiaineiston koko on niin pieni, että tutkimus on enemmän tyypiltään tapaustutkimus eikä kattavia yleistyksiä voinut tehdä. Kyseessä onkin siis enemmän tapaustutkimus, joka kuitenkin antanee viitteitä siitä, minkälaisia sanastollisia taitoja oppilas ottaa käyttöön uuden ja vieraamman tekstilajin harjoittelussa.

5.2 Aineiston analyysi

Ennen analyysin aloittamista aineisto lemmattiin eli sanastettiin. Tieteen termipankki määrittelee lemman eli lekseemin sanan kaikki taivutusmuodot käsittäväksi abstraktioksi (s.v. lemma, viitattu 19.5.2024). Tutkimuksessa käytettiin siis sanan kielen jäsenistömuotoa eli niin sanottua sanakirjamuotoa, joka yhdistää saman sanan eri taivutusmuodot kuten esimerkiksi *kattila* – *kattilan* samaan kategoriaan *kattila*.

Jokaiselle sanalle määritettiin lekseemi ja sanaluokka. Lemmaus tehtiin Turun yliopiston NLP-ryhmän parserilla. Se on luonnollisen kielen analysointiin tarkoitettu alusta, joka on erikoistunut kielellisiin ominaisuuksiin kuten esimerkiksi kielioppiin ja sanojen perusmuotojen hakuun.

Monissa tapauksissa saneella oli useampia lekseemivaihtoehtoja (esimerkiksi sane *asua* voi olla verbi *asua* tai partitiivi substantiivista *asu*). Parseri tunnisti melko hyvin kontekstisidonnaiset merkitykset, mutta jokainen sane tarkastettiin erikseen niin, että se sopi kontekstiin.

Epäselvissä tapauksissa lemmaus tarkastettiin näissä tapauksissa syöttämällä ensin tekstin lause, jossa monitulkintainen sane esiintyi, yksinään parseriin. Lopuksi sana syötettiin vielä yksin, jos tulos oli edelleen epävarma. Näin pyrittiin minimoimaan niin tietokoneen tekemät virheet kuin inhimilliset erehdyksetkin.

Korpuksen lemmatut tekstit koottiin sana kerrallaan sanalistauksiksi. Näistä lekseemeistä koostuneista teksteistä sanat kerättiin sanastolistoiksi. Ensimmäisessä vaiheessa tekstit purettiin yksiköinä ja jokaiselle lemmalle laskettiin sen jälkeen frekvenssi sen mukaan, kuinka monta kertaa se esiintyi yhdessä tekstissä. Tämän jälkeen lemmat aakkostettiin ja yksittäisten tekstien lemma-listat yhdistettiin luokkakohtaisiksi listauksiksi, joissa näkyi luokan tekstien lemموjen yhteenlasketut frekvenssit. Nämä kaksi listausta yhdistettiin vielä molemmat luokat kattavaksi listaksi frekvensseineen. Listat siis liitettiin peräkkäin samaan tiedostoon.

Frekvenssien avulla oli mahdollista analysoida syntyneitä sanastoprofiileja. Frekvenssit ovat tärkeä indikaattori käyttöpohjaisessa kielenoppimisen teoriassa (Honko, 2013, s. 21), koska se on selkeärajainen keino luokitella sanoja (Honko, 2013, s. 97). Sanan frekvenssi tekstissä voi myös kuvata hyvin selkeästi sanan vaikeutta tai harvinaisuutta (Milton, 2009, s. 22, 35, 242), koska sanojen oppiminen ja osaaminen riippuu sanojen toistamisesta (Milton, 2009, s. 26; Blom & Unsworth, 2010, s. 208). Tutkimus keskittyi pelkästään sanastoon ja käytettyihin ilmauksiin, vaikka sanasto ei suljakaan kielioppia, lauserakenteita tai ilmaisun sujuvuutta pois, vaan ne liittyvät hyvin tiivisti yhteen sanaston kanssa. Rajaus tehtiin metodologisen valinnan – frekvenssien tarkastelun – vuoksi.

Tässä tutkimuksessa päädyttiin kokoamaan yhden lekseemin taivutusparadigmaan eli kaikki sanan mahdolliset taivutusmuodot omiksi sanaperheikseen *aika-aikaa-ajan* ja niin edelleen luettiin siis kaikki saman perheen *aika* jäseniksi. Kuitenkin saman kannan johdokset huomioitiin omina lekseemineen, ja siten ne saivat oman frekvenssiluvun ja mahdollisen sanaperheen kuten esimerkiksi sanan *aika* johdos *aikainen* luettiin omaksi sanaperheekseen.

Lemmauksen ja listojen kokoamisen jälkeen siirryttiin tulkitsemaan tekstin lekseemien esiintyvyyksiä ja etsittiin poikkeavia tai harvinaisia ilmauksia. Aineiston pienen koon takia monen sanan frekvenssi jäi yhteen, jolloin nämä sanat käytiin vielä erikseen läpi. Niitä

verrattiin Kotimaisten kielten keskuksen kokoamaan Suomen lehtikielen- taajuussanastoon (CSC, 2004). Harvinaisten sanojen määrä oli alhainen, ja tutkijan oli mahdollista tarkastella mielipidekirjoituksissa esiintyvien harvinaisempien lekseemien käyttöä. Samassa yhteydessä etsittiin myös sanoja, jotka poikkeavat kompleksisuudeltaan viidesluokkalaisen ilmauksista.

Myös useimmiten esiintyviä lekseemejä verrattiin samaan taajuussanastoon, jotta selviäisi, miten lähellä aineisto on suurempia korpuksia. Tällä tavalla varmistettiin myös tulosten vertailukelpoisuutta. Kaikkia lemmoja verrattiin taajuussanastoon, mutta erityisesti huomiota kiinnitettiin kahteenkymmeneen ensimmäiseen lemmaan molemmissa listoissa, sekä niihin, joita ei taajuussanastosta löytynyt.

Lemmat, joiden frekvenssi oli yksi ja joita ei löytynyt taajuussanastosta tarkasteltiin vielä tarkemmin. Niille etsittiin tavallista esiintymisympäristöä. Niiden merkitystä ja aineistossa näkyvää esiintymisympäristöä verrattiin myös Suureen kielioppiin, sekä Kielitoimiston sanakirjaan. Modaaliverbeihin kiinnitettiin erityistä huomiota, koska ne sisältävät paljon syntaktista, eli lauseopillista tietoa

6 Tulokset

6.1 Mitä sanoja viidesluokkalaisten käyttää mielipidekirjoituksissa?

6.1.1 Viidesluokkalaisten tekstit suhteessa sanomalehtikieleen

Aineistoissa käytettiin yhteensä 3710 lekseemiä eli sanaa, joista eri lekseemejä oli 974. Määrät jakautuivat luokkien välille niin, että E- luokalla oli 2064 lekseemiä, kun kaikki lekseemit laskettiin yhteen. Keskenään erilaisia lekseemejä oli 655. G-luokalla käytettiin 1664 lekseemiä, joista erilaisia oli 550.

Taulukko 1: 5. luokkalaisten kirjoitelmien kaksikymmentä yleisintä sanaa ja niiden frekvenssit

	(N=3710)	
<i>olla</i>	268	7,22
<i>ja</i>	124	3,34
<i>ei</i>	112	3,01
<i>se</i>	107	2,88
<i>että</i>	52	1,40
<i>jos</i>	51	1,37
<i>mieli</i>	46	1,23
<i>joka</i>	39	1,05
<i>niin</i>	38	1,02
<i>pitää</i>	38	1,02
<i>koira</i>	36	0,97
<i>minä</i>	36	0,97
<i>kun</i>	34	0,91
<i>mutta</i>	34	0,91
<i>kakka</i>	28	0,75
<i>voida</i>	28	0,75
<i>hän</i>	25	0,67
<i>hyvä</i>	23	0,61
<i>kaikki</i>	23	0,61
<i>tehdä</i>	23	0,61

Taulukko 2: Suomen lehtikielen taajuussanasto, Kotimaisten kielten keskus

<i>sana</i>	<i>n</i> (N= 43 999 826)	%
<i>olla</i>	2716396	6,17
<i>ja</i>	1566108	3,55
<i>ei</i>	593462	1,34
<i>se</i>	538609	1,22
<i>että</i>	443301	1,00
<i>joka</i>	417984	0,95
<i>vuosi</i>	344927	0,78
<i>hän</i>	302803	0,68
<i>myös</i>	267016	0,60
<i>saada</i>	248176	0,56
<i>mutta</i>	217016	0,49
<i>tämä</i>	209851	0,47
<i>voida</i>	197072	0,44
<i>tulla</i>	192327	0,43
<i>suomi</i>	181128	0,41
<i>tehdä</i>	171406	0,38
<i>kun</i>	165584	0,37
<i>pitää</i>	161237	0,36
<i>mukaan</i>	157837	0,35
<i>uusi</i>	149954	0,34

Tarkasteltaessa molempien luokkien viidesluokkalaisten tekstejä eniten käytetty lekseemi oli *olla*-verbi. Tämä selittyy sillä, että suomen kielessä menneen ajan liittomuodot pluskvamperfekti esimerkiksi *oli pukemassa* (MT2) ja perfekti esimerkiksi *on taso laskenut* (KP22) muodostetaan apuverbin *olla*-avulla eri taivutuksilla. *Olla*-verbi esiintyy myös

useissa erilaisissa verbiketjuissa ja liittomuodoissa. *Olla*-verbi kattaa kirjoitelmien tekstien lekseemeistä noin 7 %. Merkitykseltään se on hyvin kalpea. *Olla*-verbi esiintyy usein myös osana predikatiivilauseita, joissa kuvataan jotain entiteettiä tai ilmiötä esimerkiksi *Se* (, että auttaisimme vanhainkodin vanhuksia) *olisi reilua!* (LK29). Sillä on myös roolinsa eksistentiaalilauseissa (VISK §893). kuten esimerkiksi *Vuonna 2006 oli mielestäni parhaat lastenohjelmat* (KP22). Myös sanomalehtisanastossa *olla*-verbi oli listan kärjessä eli se ei kuvaa mitään tarkkaan ja tarvitsee aina adjektiivin tarkentamaan, jos sillä kuvataan olotilaa

Toiseksi eniten kirjoitelmissa oli käytetty *ja*- ja *ei*-lekseemejä. Sana *ja* esiintyi noin 3 prosenttia kuten myös *ei*. Verbin *ei* yleisyys on *olla*-verbin tavoin ainakin osittain selitettävissä sen roolissa liittomuodoissa. Rinnastuskonjunktio *ja* puolestaan kertoo päälauseiden rinnastamisesta sekä mahdollisista luetteloista, joissa rinnastetaan samantasoisia kielen ilmiöitä. Allekkain kirjoitetut luettelot tai listat eivät kuitenkaan yleensä ole tyypillinen mielipidekirjoituksen piirre (Heikkinen ym., 2012 s. 714). Niitä ei esiintynyt teksteissä edes kolmiosaisena argumentoinnin keinona, mikä olisi sopinut tekstilajiin. Konjunktio kertoo teksteissä esiintyvistä rinnastamisesta. Lasten puheessa myös *ja sitten* on kerronnassa toistuva ilmaus, joka kertoisi puhekielisistä ilmauksista. Kuitenkin vain neljässä tekstissä käytettiin ilmausta *ja sitten*.

Konjunktion *jos* yleisyys kertoo tekstilajin tajusta. Se on sävyltään suostutteleva (KS v.s. *jos* viitattu 26.1.2025), mikä sopii mielipidekirjoituksen genreen. *Jos* on aineiston kuudenneksi käytetyin sana. Se kertoo myös alisteisten lauseiden määrästä, mikä viittaa kirjoitustaitojen ja kieliopin hallintaan, koska alisteisuus sivulauseessa on kompleksisempi rakenne kuin rinnastettu lause. Taajuussanastossa kuudenneksi käytetyin sana oli, pronomini *joka*.

Tekstien yleisimmät sanat vastasivat Suomen kielen lehtikielen-taajuussanastoa viiden ensimmäisen sanan osalta. Viisi ensimmäistä ilmausta ovat kuitenkin melko värittömiä ja pääasiassa rakenteiden osina käytettyjä palasia. Ne eivät siis ole kovin mielenkiintoisia.

Viidesluokkalaisten kirjoitelmissa lekseemi *mieli* kattoi 1,2 % kaikista sanoista eli se esiintyi 46 kertaa, mikä olikin seitsemänneksi useimmin esiin noussut lekseemi. *Mieli* esiintyi pääasiassa muodossa *mielestäni*, jolla kirjoittajat olivat korostaneet omaa mielipidettään. *Mielestäni* ilmaisua voi pitää tämän aineiston kohdalla kiteytyneenä ilmaisuna, koska *mieli* ei esiinny missään muussa muodossa teksteissä. Oppilaat käyttävät ilmaisua harjoittellessaan

mielipidekirjoituksen genren skeemaa. Kolmessakymmenessä kahdessa (32) tekstissä kirjoittaja oli käyttänyt *mielestäni* rakennetta. Kuten näissä muutamassa esimerkissä on tehty: ”*Mielestäni koiran kakat kadulla ovat epäsiivistyksellisiä. (RP21)*”, sekä ”*Mielestäni kaikkien koulujen pitäisi tarjota yhtä laadukasta ja monipuolista opetusta jotta tällaista ei tapahtuisi. (HR23)*”.

Usein teksti jopa aloitettiin ilmauksella *mielestäni*, mikä viittaa siihen, että kirjoittaja on alusta asti halunnut korostaa tekstin olevan hänen mielipiteensä, jonka hän yrittää seuraavaksi perustella. Teksteistä seitsemän aloitettiin virkkeellä, jossa esiintyi *mielestäni*.

Suostuttelutarkoituksessa käytettiin myös sanoja *'pitää'* (1 %), *'jos'* (1,3 %) ja . Kuten esimerkissä 2:

Teidän pitäisi muuttaa töihin ja hankkia aivot! (AJ11)

Ennen aikaan läksyihin ja kotitöihin, ulkona olemiseen ja lukemiseen käytettiin lähes kaikki vapaa-aika, ja jos läksyjä ei tehnyt, se oli lähes kamalaa, mutta nykyään sekä vanhemmat, että lapset eivät piittaa, jos läksyjä ei tee, sillä vain pelit on tärkeitä. (HM26)

Sana *'kaikki'* oli käytössä myös lukijoihin vetoamisessa, sekä perusteluissa. *Kaikki* on yksi universaaleista kvanttoripronomineista eli se viittaa koko puheena olevaan joukkoon (VISK §750). Usein nuoret ja lapset voivat käyttää sanaa *kaikki* perustelutarkoituksessa eli yleisyyteen vetoamisen keinona. Esimerkissä 3 kuitenkin *kaikki* on käytössä tarkoituksena vedota lukijoihin:

Joten nyt te kaikki inhottavat ihmiset jotka käytätte kettuja kaulahuiveina ajatelkaa niitä pieniä ketunpoikasia joiden äidi tai isät tai molemmat roikkuvat nyt kauloissanne.

Lekseemi *niin* on perustelutarkoitusten lisäksi painottamassa adjektiiveja. *Niin* käyttäminen korostamismerkityksessä on kuitenkin puhekielinen ilmaisu. Sen käyttäminen viestii vielä kielitaidon kehittymisestä ja yleiskielen normien hallinnan harjoittelusta. *Niin* on tavan adverbi, joka tarkoittaa *'sillä tavalla'*, *'silleen'*, tai *'siten'* (VISK 1160). Sen peruskäyttö ei

siis liity intensiteetti-ilmauksiin eli sen perustarkoitus eli ole korostaa kuten *Jos turkisten käyttö on niin tärkeä tapa...* (KJ2)

Myös yksi suosituimmista aiheista on nähtävissä taajuuksista. Kirjoitelmien sanoissa nousevatkin *koira*, sekä *kakka* yleisimpien kahdenkymmenen sanan joukkoon. Koirankakat kadulla olivat yhteensä seitsemän eri tekstin aiheena. Teksteissä se esiintyi 28 kertaa. Taajuussanastossa *kakka* ei esiintynyt kertaakaan. *Koira* sen sijaan oli sanastossa sijalla 954 eli sen osuus oli 0,12 % sanastosta.

6.1.2 Sanaluokat

Aineistoa tarkasteltiin myös sanaluokkien näkökulmasta, koska ne tarjoavat tietoa siitä, miten ilmaisut on teksteissä rakennettu. Lyhenne aux-viittaa (ks. taulukko 3: Aineistossa esiintyvät sanaluokat luokittain) verbiketjun osana olevaan verbiin, esimerkiksi verbiketjussa *on haluttu*, *olla*-verbi on luokiteltu aux-kategoriaan ja *haluta*-verbi verbeihin.

Taulukko 3: Aineistossa esiintyvät sanaluokat luokittain

<i>sanaluokka</i>	<i>E-luokka (n)</i>	<i>E-luokka %</i>	<i>G-luokka (n)</i>	<i>G-luokka %</i>	<i>N</i>	<i>%</i>
<i>substantiivi</i>	505	24,68	468	28,13	973	26,22
<i>verbi</i>	347	16,95	254	15,26	601	16,19
<i>adjektiivi</i>	194	9,48	188	11,3	382	10,29
<i>pronomini</i>	281	13,73	193	11,6	474	12,77
<i>konjunktio</i>	119	5,81	119	7,15	238	6,41
<i>adverbi</i>	259	12,65	190	11,42	449	12,1
<i>aux</i>	219	10,70	172	10,34	391	10,53
<i>adpositio</i>	18	0,88	20	1,2	38	1,02
<i>numeraali</i>	15	0,73	33	1,98	48	1,29
<i>propn (nimi)</i>	38	1,86	18	1,08	56	1,5
<i>int</i>		0	1	0,06	1	0,02

Aineistossa substantiiveja oli huomattavasti enemmän, kuin muita sanaluokkia, mikä on hyvin tavallista, koska asioista ei voi puhua ilman substantiivien osaamista (Puro, 1999, s.13).

Verbejä tekstissä on toiseksi eniten. Lausetta, eikä sitä myöden tekstiä voi kirjoittaa ilman verbejä. Verbien määrä antaa siis viitteitä lauseiden määrästä. Suoraan yhteydessä se ei ole, koska osa verbeistä on voinut olla osana verbiketjua. Taulukon aux-merkintä ilmaisee liittomuodossa osana ollutta – yleensä *olla*- tai kieltoverbiä – lekseemiä.

Adjektiivit nousivat aineistosta, koska niitä käytetään tekstissä ilmaisemaan mielipidettä tai arviota asiasta. Niiden avulla voidaan luoda erilaisia sävyjä ilmaisuun. Kielen rikkaus näkyy myös adjektiivien määrässä (Honko, 2013, s.96). Tekstissä esiintyi 382 adjektiivia, joista 25 oli adjektiivi *hyvä*, joka on hyvin kalpea ilmaus. Seuraavaksi käytetyimpiä olivat *oma*, *pieni*, *iso*, *kiva* ja *vanha*. Nämä kaikki ovat hyvin värittömiä, mikä kertoo siitä, että kielen käytön taito ei ole vielä kehittynyt. Ilmaisut ovat vielä hyvin tyypillisiä lapsille. Toisaalta ne ovat myös hyvin tavallisia sanoja. Seuraavassa luvussa esittelen aineistosta nousseita harvinaisia ja ilmaukseltaan rikkaampia ilmauksia, muun muassa adjektiiveja.

Pronomineja aineistossa oli 12,77 %. Esimerkiksi tekstissä SJ12 pronomineja oli n. 12%, mikä oli suurin pronominiinien osuus kaikista teksteistä. Siitä huolimatta pronominiinien osuus tämän tekstin osalta vastasi kaikkien tekstien yhteenlaskettua pronominiinien määrää. Teksti argumentoi Real Madridin ja ”Barcan” eli Barcelonan jalkapallojoukkueiden välisistä eroista ja Real Madridin paremmuudesta. Teksti on myös yksi aineiston pisimmistä teksteistä, eli siinä on yhteensä 195 sanaa.

... Cristiano Ronaldo joka on kaikkien aikojen eniten maaleja tehnyt pelaaja ja tehnyt ne tosi pienessä ajassa. joten ei ole ihme että Ronaldo voitti myös La Ligan maalikuninkuuden kaudella 2010-2011. 41:dellä maalilla hän teki koko Ligan maali ennätyksen. Real Madrid on joukkue jonka vahvuus on hyökkääminen... SJ12

Esimerkkitekstissä eniten käytetyt pronominit ovat *hän*, *se* ja *joka*, kun kirjoittaja viittaa joukkueeseen ja joukkueen pelaajiin. Kirjoittaja on myös ymmärtänyt relatiivipronominin *joka* roolin sivulauseen rakentamisessa, mutta tekstiä tarkastellessa käy ilmi, ettei hän kuitenkaan ole tunnistanut tai osannut rakentaa sujuvia sivulauseita. Rakenne ”*joten ei ole ihme*” on kuitenkin hyvin taitavasti rakennettu merkityssuhteiden ilmaus, mikä kertoo syy-seuraussuhteiden päättelyn taidosta.

6.2 Millaisia sanoja viidesluokkalainen käyttää mielipidekirjoituksessa ja miten?

6.2.1 Modaaliverbit

Huomioitavaa on myös, että viidesluokkalaisten teksteissä modaaliverbejä esiintyy vähän suhteessa siihen, miten yleisiä ne ovat taajuussanastossa (ks. taulukko 2 s. XX). Modaaliverbit ilmaisevat mahdollisuutta ja välttämättömyyttä. Tällaisia verbejä ovat esimerkiksi *voida*, *saattaa*, *taitaa* ja *mahtaa*), sekä välttämättömyyttä ilmaisevat *täytyä*, *pitää*, *tarvita* ja *joutua* (VISK §1562). Modaaliverbit esiintyvät yleensä osana verbiketjua ja muodostavat kompleksisia lauserakenteita. Niiden käyttäminen kertoo siis oppilaan syntaktisesta eli lauseopillisesta osaamisesta.

Modaaliverbit, joita tutkituissa teksteissä käytettiin usein olivat *pitää* (1%) ja *voida* (0,07%). Myös *olla*- verbiä voidaan käyttää modaalisessa merkityksessä, vaikka se ei suoraan modaaliverbi olekaan (VISK1581).

Taajuussanastossa modaaliverbejä oli useampia. Käytetyin oli *saada*. Muita taajuussanaston kahdenkymmenen käytetyimmän sanan joukossa olevaa modaaliverbiä olivat *voida*, *olla* ja *pitää*, mutta näitäkin modaaliverbejä voi käyttää ei modaalisesti. Myös verbi *tulla* mainittiin, joskaan se ei aina ilmaise modaalisuutta, vaikka sillä on sellainenkin merkitys. Teksteissä *tulla* ei kuitenkaan esiintynyt *täytyä* merkityksessä.

Kehittyvästä kieli- ja kirjoitustaidosta kertoo myös pronominin *minä* yleisyys aineistossa. Kirjoittajat eivät siis vielä välttämättä kykene asioiden käsittelyyn toisesta näkökulmasta, mikä on tyypillistä vielä tässä ajattelun kehityksen vaiheessa (Pajunen, 2012, s.4), kuten esimerkissä PM 13:

...Jos olisin eläin, niin en minä haluaisi, että juuri minulle viritettäisiin kaikennäköisiä hiirenloukkuja tai, että joutuisin eroon perheenjäsenistäni, sillä minut myydään...

PM13

Tässä esimerkissä kirjoittaja perustelee mielipidettään tavalla, jolla aikuiset usein selvittävät lasten riitoja, eli ohjaamalla pohtimaan, miltä toisesta tuntuu. Hän käyttää itseään perusteluna sille, miksi toisen pitäisi muuttaa mielipidettään.

6.2.2 Epäfrekventit sanat

Johdokset ja abstraktit sanat ovat yleensä harvinaisempia kuin konkreettiset ilmaukset (Pajunen, 2012, s.32). Teksteissä oli monia sanoja, jotka esiintyvät vain kerran kaikissa teksteissä. Epätavalliset ja epäfrekventit sanat löydettiin katsomalla lekseemien frekvenssejä ja harvinaisuutta tarkistettiin vertaamalla näitä sanoja taajuussanastoon.

Teksteissä esiintyi kuitenkin merkkejä myös siitä, miten oppilaat olivat oppineet suomen kielen sananmuodostuksen periaatteen.

Eräässä tekstissä oppilas oli itse johtanut uuden sanan todennäköisesti hyödyntämällä sananmuodostuksen periaatteita:

Mielestäni koirankakat kadulla ovat epäshivistyksellisiä. (TP12)

Oppilas osoittaa, että hän ymmärtää epä- liitteen muuttavan sanan alkuperäisen tarkoituksen kielteiseksi. Kontekstista on myös pääteltävissä, että oppilas hallitsee verbistä johdetun substantiivin *shivistys*, joka on jo itsessään kantasanaansa (<*shivistaa*<*siveä*) abstraktimpi. Tästä hän on johtanut adjektiivin *shivistyksellinen*. Näin oppilas osoittaa kykenevänsä jo abstraktimpaan sanaston käsittelyyn. Ilmaisua ei kuitenkaan ole oikeaa kieltä, vaan hänen itsensä keksimä epäsanana. Sen tarkoitus on kuitenkin selvä ja osoittaa kielellä leikkimisen taitoa.

Toinen esimerkki sananmuodostuksen ymmärtämisen harjoittelusta on tekstissä SM16:

Useimmissa kouluissa esiintyy koulukiusausta, mutta siihenkin on puututtu. Koulukiusaus on epäreilua kiusattavaa kohtaan ... Koulukiusauksesta ei joskus edes välttämättä tiedetä... (SM16)

Kontekstin perusteella oppilas on tarkoittanut koulukiusaamista. Koulukiusaamisesta puhutaan kouluissa säännöllisin väliajoin, eli termin pitäisi olla oppilaalle tuttu. Hän on kuitenkin aivan otsikosta alkaen toistanut saman virheen koulukiusauksesta. *Kiusaus* voi tarkoittaa tietynlaista ruokaa kuten esimerkiksi lohikiusaus tai se voi olla sanan *houkutus* synonyymi.

Muita viidesluokkalaisen sanaston näkökulmasta erikoisempia johdoksia ja ilmaisuja olivat *hännänhuippu*, *uppo-outo*, *höperö* ja *talouskomissaari*. *Hännänhuippu* esiintyy usein saduissa ja tarinoissa, eikä juuri näy puhutussa saatikka virallisessa kielessä. Yksi esimerkki sen esiintymisestä on Seitsemän veljeksien pilkkalaulu: *Sitten tulee hännänhuippu, pikku Eero liukas liukku*. Tämän lisäksi ilmaisun voi kuulla esimerkiksi urheiluselostuksessa.

Viidesluokkalaisen tekstissä se ilmenee vastaavassa kontekstissa.

Suosittuja manga-sarjoja suomennetaan aivan liian hitaasti, kuten Kuroshitsuji ja Soul Eater jotka tehtiin ajat sitten, mutta suomennetaan vasta nyt.

Kaiken hännänhuippu on Naruto, joka tuli vuonna -99, mutta alettiin suomentaa vasta vuonna 2008 eli 9 vuotta väliä. (EM4)

Rakenteena *kaiken hännänhuippu* ei siltikään ole sujuva. *Hännänkuippu* ei kollokoi muun ympäristön kanssa. Kuitenkin kontekstissa sillä ilmaistaan naurettavan myöhässä olemista, jolloin se on ymmärrettävästi käytetty.

Samana henkilön kirjoittamassa tekstissä esiintyy myös adjektiiviksi *uppo-outo*. *Hännänhuipun* tavoin se on tyypillisempi saduissa ja kertomuksissa. *Uppo-outo* ei ole nykyaikana puhekielinen ilmaus, eikä se myöskään esiinny virallisessa kielessä, vaikka se on yleiskielinen ilmaus (KS, s.v.*uppo-outo* viitattu 10.2.2025). Tekstissä adjektiivia käytetään ymmärrettävästi:

Sen sijaan suomennetaan uppo-outoja sarjoja, joista kukaan ei tiedä mitään. (EM4)

Kielitaitoa osoittaa myös adjektiivin oikeinkirjoitus. Kirjoittaja on ymmärtänyt yhdyssanojen kirjoitussäännöt, sekä tunnistanut tarpeen väliviivalle. *Uppo-outo* on myös semanttisesti ymmärretty oikein, vaikka sen merkitys tarkennetaankin vielä sivulauseella.

Yhdessä teksteistä esiintyy myös johdos adjektiivista *hellävarainen - hellävaraisesti*, jota ei löytynyt taajuussanastosta tai muista teksteistä. Arjessa myös shampoota tai pesuohjelmia voidaan kuvailla hellävaraiseksi. Myös Kielitoimiston sanakirja (s.v. *hellävarainen* viitattu: 10.2.2025) *Hellävarainen* on kirjoitettu oikein. Se antaa *hellävaraisen* määritelmän

yhteydessä esimerkin shampoosta. Se kuitenkin on kaksi kertaa samassa sanaympäristössä, mikä todennäköisesti on ympäristö, jossa kirjoittaja sen on kuullut eli toimintaohjeena.

Mielestäni eläimiä pitäisi käsitellä hellävaraisesti. ... En kirjoittanut tästä aiheesta vain siksi, että olen eläinrakas, vaan siksi, että minulla on ollut lemmikki, jota jotkut ihmiset eivät osanneet käsitellä hellävaraisesti.

(PM13)

Talousskomissaari on terminä etäinen suhteessa viidesluokkalaisten kokemusmaailmaan, minkä takia se nousi esiin analyysissä. Sanan tunteminen kertoo yksittäisen oppilaan mielenkiinnonkohteista sekä ehkä kieliympäristöstä kotona. Aineistossa se oli osana eduskunnan riitoja käsittelevässä tekstissä. Myös eduskunnan riidat ovat aika kaukana keskiverto neljäsluokkalaisten kokemuspiiriä.

Hyvänä esimerkkinä se kun Timo Soini (PS) haukkui talousskomissaari Olli Rehniä, isoa vastuuta kantavia henkilöitä pitää tukea eikä haukkua. (LA3)

Oppilas on selkeästi tuntenut politiikan kenttää sekä poliitikkojen riidoista kirjoittamisen muodon, koska hän on osannut laittaa Timo Soinin puolueen lyhenteen sulkeisiin tämän nimen perään. Merkintätavasta voi olettaa, että oppilas lukee uutisia jonkin verran ja hänen teksti- ja sanavarastonsa rikkaus heijastuu hänen kirjoittamiseensa.

Mielenkiinnon herättää myös aineistossa esiintyvä *höperö*, joka on puhekielessä melko harvinainen sana, eikä toisaalta virallisen kielen ilmauskaan. Se on myös tutumpi tarinakontekstista, kuten esimerkiksi sarjakuvista, eikä vastaa viidesluokkalaisten tyypillistä ilmaisua.

Esimerkiksi: Jos kauppaan tulisi joku aivan vanha ja höperö mies ja kysyisi että saisinko parhaimman ja tarkimman haulikon? (LO17)

Sävyltään kuvailevampi adjektiivi *iljettävä* löytyi myös aineistosta. *Iljettävä* esiintyy usein sarjakuvissa ja tarinoissa. Se on myös kärjistävä ilmaus, mikä sopii vaikuttamaan pyrkivään tekstiin. Taajuussanastossa se ei esiintynyt. Sanomalehtikieli on kuitenkin melko väritöntä ja *iljettävä* on hyvin värikäs ilmaus.

Sekä koiran kakka tien varsissa on todella iljettävän näköistä. (TP12)

Esimerkissä adjektiivi iljettävä esiintyy päälauseessa, joka on aloitettu rinnastuskonjunktioilla *sekä*. Lausetta ei kuitenkaan oikeasti rinnasteta mihinkään, vaan konjunktioita edeltää piste. Kontekstin perusteella voisi olettaa, että lause rinnastuu edelliseen päälauseeseen, mutta oppilas ei vielä hallitse pilkkusääntöjä.

Aineistossa esiintyi vain kerran verbi *kusettaa*. Vaikka ilmaus on aineistossa harvinainen, on se puhekielessä hyvin tyypillinen nuorten käyttämä puhekielinen ilmaus. Yleensä puheessa sillä tarkoitetaan huijaamista tai harhaan johtamista. Nyt se kuitenkin esiintyi verbin *pissattaa* synonyymina. Molemmat käyttötarkoitukset ovat yleisesti käytössä (KS s.v. *kusettaa* viitattu: 10.2.2025), mutta ilmaisu on puhekielinen.

Kaikki urpot kusettavat ja kakattavat koiransa talvella lumeen ja potkivat lunta päälle kuvitellen, että kakat häviävät johonkin. (RP 21)

Kuitenkaan puhekielinen verbi ei sovi yleiskielisen mielipidekirjoituksen tyyliin. Silti se on yksi harvoista kuvailevista verbeistä, joita teksteissä esiintyy.

Toinen teksteistä noussut kuvaileva verbi oli *hutaista*. Sekin on tyyliiltään hyvin puhekielinen, mutta antaa tekstile silti enemmän väriä, kuin suurin osa käytetyistä verbeistä.

Onko kukaan koskaan lisäkseni miettinyt, etteivät nykynuoret enää käytä aikaa kotitöihin, ja läksytkin tehdään hutaisemalla -> Tai ei sitten ollenkaan. (HM26)

Esimerkistä nousee myös oppilaan käyttämä kuvallinen ilmaisu. Sana *tai* aloittaa rinnasteisen lauseen, joka tässä tapauksessa on sivulause toistamatta jääneen *että*-konjunktin takia. Pilkun sijaan hän on piirtänyt nuolen sivulauseeseen. Ilmeisesti tarkoitus on kuvata liikettä kahden vaihtoehdon välillä. Yleiskieliseen tekstiin se ei kuitenkaan sovi, vaan on tyyliiltään enemmän tekstiviestin kaltainen. Ylimääräistä kirjoittamista pyritään välttämään, jotta viestin saisi nopeammin eteenpäin.

Epäfrekventit sanat näissä aineistoissa eivät lähtökohtaisesti ja yksiselitteisesti liittyneet mielipidekirjoituksen genreen, eivätkä ne siten osoittaneet välttämättä tekstilajin hallintaa. Ne olivat kuitenkin harvinaisempia sanoja viidesluokkalaisen sanastossa, mikä osoittaa kielen hallintaa ja ilmaisuvarojen rikkautta. Kuitenkin ilmaisultaan sävykkäämmät adjektiivit, kuten *uppo-outo* ja *iljettävä*, ovat vaikuttamaan pyrkivälle mielipidekirjoitukselle hyvin tyypillisiä.

Esimerkiksi teksti Eduskunnan riidoista (LA3), puhui *talouskomissaarista*. Harvinaisesta, poliittiseen kieleen kuuluvasta lekseemistä huolimatta hän puhui myös *haukkumisesta* loukkaamisen sijaan. *Haukkua* tässä yhteydessä on hyvin puhekielinen ilmaus, vaikka se onkin kuvaileva verbi. Mielipidettä ilmaistaan kuvaamalla tilanteen olevan *ärsyttävä*, mutta asiaa ei perustella enempää. Tekstissä on myös rakenne *se kun*, joka yhtä lailla on dialoginen rakenne, joka voi esiintyä mediatekstissä.

Erityisesti sanasto-osaamisen näkökulmasta esiin nousi teksti koodilla EM4. Kirjoittaja on käyttänyt kahta tyypillisestä viidesluokkalaisen sanastosta eroavaa ilmausta: *uppo-outo* ja *hännänhuippu*. Teksti oli kirjoitettu Manga-sarjakuvien suomentamisen tarpeesta. Aihevalinnasta voi päätellä, että kirjottaja lukee melko paljon ja erityisesti Manga-sarjakuvia. Lukuharrastuneisuus näkyy myös hänen ilmaisuvalinnoistaan. Teksti on kuitenkin suhteellisen lyhyt, vain 82 sanaa. Vaikka yleisesti on havaittu, että tekstin pituus on alakouluiässä yleensä yhteydessä myös kielelliseen osaamiseen.

7 Pohdinta

Tämän tutkimuksen tulokset antavat näkökulmaa siihen, millaisia kielellisiä resursseja viidesluokkalainen pystyy hyödyntämään itselleen vieraamman tekstilajin yhteydessä. Aineisto on kuitenkin niin pieni, että kyseessä on tapaustutkimus, joka antaa vain viitteitä siitä, mitä suuremmilta ryhmiltä voi odottaa. Tutkimuksessa hyödynnettiin uudenlaista tutkimusmenetelmää, jolloin tutkimus oli myös kokeilu siitä, kannattaako sanaston kehitystä tutkia kirjoitelmissa käytetyn sanaston näkökulmasta.

Työnsä näkökulmasta opettaja saa tämän tutkimuksen tuloksista konkreettisempia lähtökohtia oppilaiden kielellisen kehityksen sekä erityisesti sanavaraston arviointiin arjessa, jolloin hän voi antaa tarvittaessa tarjota täsmällisempää tukea sanaston kasvattamisessa, sekä lukusujuvuuden tukemisessa. Tulevaisuudessa olisi hyvä, jos myös sanavaraston kehittymiselle tehtäisiin omia testipattereita.

Sanan ja sanaston moniulotteisuuden ja määrittelyn haasteen takia tutkimuskysymyksen muotoilu oli haastavaa. Tämän aineiston perusteella vaikuttaa siltä, että esiin nousi melko toistuvia piirteitä kuten *mielestäni* -rakenne.

7.1 Johtopäätökset

Harvasta tekstistä oli erotettavissa selviä merkkejä erityisen taitavasta sanasto-osaamisesta, kun kirjoitettavana tekstinä on mielipidekirjoitus, mutta muutama teksti viittasi heikompaan osaamiseen. Harvinaiset ilmaukset tekstissä eivät lähtökohtaisesti ilmaiseet tekstilajin hallintaa, mutta silti antoivat viitteitä sanasto-osaamisesta. Yksi kielen rakenne, jolla mielipidetekstiä lähdettiin rakentamaan oli *mielestäni*-ilmaus. Vain kahdessatoista tekstissä *mielestäni* -rakennetta ei oltu käytetty ollenkaan.

Tutkimuksessa keskityttiin erityisesti viidesluokkalaisen käyttämiin ilmauksiin ja kuvauksiin. Huomio kiinnittyi erityisesti adjektiiveihin ja adverbeihin, koska ne ilmaisevat kehittyneempää kielitaitoa (Honko, 2013, s. 96). Harvinaisimmat lekseemit aineistossa olivat juuri adjektiiveja. Harvinaiset adjektiivit olivat lähtökohtaisesti myös pitkiä sanoja ja rakenteeltaan monimutkaisia. Siitä huolimatta suurin osa adjektiiveista oli hyvin tavallisia ja värittömiä ilmauksia. Alakouluikä on vaihe, jolloin sanapituus kasvaa ja sanarakenne

monimutkaistuu (Honko, 2013, s.17). Havainto vastasi siis aiempaa tutkimusta. Koulun tarkoitus on myös laajentaa osaamista perussanoista ja kehittää ilmausta kohti abstraktimpaa sanastoa. Johdokset ja yli 5 merkin lemmat viestivät sanastonkäyttötaidosta. Tällaisia ilmauksia oli kuitenkin hyvin vähän käytössä. Teksteistä oli myös havaittavissa kieliopillisen osaamisen piirteitä. Lyhenne aux- kuvasi aina verbiketjun osana ollutta verbiä – yleensä ollatai kieltoverbiä – jolloin saadaan kuvaa siitä, että oppilaat ovat hallinneet myös kompleksisempia lauserakenteita.

Vaikeassa tekstilajissa sanasto voi olla kalpeampaa. Kompleksisesti kielenkäytöstä aineistossa löytyy kuitenkin merkkejä. Näitä merkkejä ovat esimerkiksi modaaliverbien runsas käyttö, omien sanojen rakentaminen eli kielellinen luovuus (*epäsivistyksellinen*) ja pidemmät, kuvailevat adjektiivit.

Tekstit kuvasivat myös hyvin sitä, miten oppilas harjoittelee tekstilajitaitoja. Tekstit oli kirjoitettu pääasiassa kirjoittajan itsensä näkökulmasta, eikä lukijaa puhuteltu kuin muutamissa teksteissä. Mielipidetekstin tarkoitus on kuitenkin vaikuttaa lukijaa, jolloin lukijan puhuttelun puute antaa viitteitä egosentrisyydestä, mikä on ajattelun kehitykselle tässä vaiheessa vielä tyypillistä Piaget'n teorian mukaan.

Useiden tekstien tarkoitus oli siis ilmaista oma mielipide, eikä kirjoittaa mielipidekirjoitusta, mikä kertoo vielä tarpeesta harjoitella mielipidekirjoitusta genrenä. Myös opetussuunnitelma ohjaa mielipiteenilmaisun harjoitteluun alkuopetuksessa (POPS14, s. 106). Ylemmillä luokilla tavoitteessa mainitaan myös mielipiteen ilmaisu osana oppilaan oman viestijäkuvan rakentumista (POPS14, s.163).

Ilmaisut olivat osassa tekstejä vielä töksähteleviä tai puheenomaisia, mikä on tutkitusti tyypillistä kirjoitustaitojen vasta kehittyessä (Pajunen & Honko, 2021b, s.63). Vaikka teksteissä ei toistunut lasten tarinoille tyypillinen ”ja sitten”-rakenne, ei esimerkiksi sivulauseita vielä osattu erottaa päälauseista, vaan ne saattoivat esiintyä irrallisina. ”ja sitten”-rakenteen puuttuminen on kuitenkin osoitus kehittyvästä tekstilajin tajusta, mikä viittaa ajattelun kehitykseen.

7.2 Luotettavuus

Tarkasta sanaston laadusta ja sanavaraston koosta on vaikea saada kuvaa. Koska tekstilaji määrittää vahvasti sitä, millaisia ilmauksia ja sanoja käytetään (Pentikäinen ym., 2017, s. 165), eivät tekstit voi antaa täydellistä kuvaa. Oppilaat kirjoittivat kuitenkin heitä itseään kiinnostavista aiheista, jolloin on syytä olettaa, että käytössä oli heille luontevaa ja tuttua sanastoa.

Harvinaiset sanat olivat suurilta osin tarinalliselle tekstille tyypillisiä ilmauksia, mikä on hyvin ymmärrettävää. Silti alaspesifin sanaston hallinta jää piiloon, mikä on harmi, koska juuri alaspesifi sanasto mainitaan opetussuunnitelman perusteissa jokaisen reaaliaineen kohdalla, sekä yleisessä arvopohjassa (POPS 14, s.28). Harvassa tekstissä oli edes mahdollisuutta alaspesifille sanastolle. Vain muutama teksti oli kirjoitettu koirankakkoja ja elokuvien ikärajoja haastavammista aiheista, kuten eduskunnan riidoista tai ilmastonmuutoksesta.

On myös muistettava, että aikuisten sanomalehtikieli on hyvin erilaista verrattuna lasten käyttämään kieleen tai lasten kirjoittamien mielipidekirjoitusten kieleen. Taajuussanasto (CSC, 2004) ei ollut täysin vastaava vertailukohta. Parempia vaihtoehtoja ei kuitenkaan ollut.

Aineistoa tarkastellessa on kuitenkin huomioitava, että aineisto on hyvin pieni korpus eli tekstiaineisto, jolloin se ei ole suoraan yleistettävissä. Korpuksen koosta huolimatta yhtäläisyyttä laajaan taajuussanastoon löytyi sen verran, että tulokset kannatti raportoida. Aineistossa näkyy myös oppilaita eniten kiinnostaneet aiheet, koska esimerkiksi *kakka* oli ylikorostunut koirankakkojen keräämisen puolesta kirjoitettujen tekstien määrän takia.

Myös Hongon (2013) tutkimuksessa kymmenen yleisimmän lekseemin joukkoon kaikissa ikäluokissa kuuluivat *ja*, *olla*, ja *se*. Tässä tutkimuksessa *minä* jäi sijalle yksitoista, eikä päivä näkynyt ollenkaan kahdenkymmenen ensimmäisen lekseemin joukossa. Myös Hongon aineistosta näkyy siis Unelmien päivä -kirjoitustehtävä (Pajunen & Honko, 2021b, s.74). Vaikka kirjoitustehtävä oli hänen tutkimuksessaan erilainen, myös yhtäläisyyksiä tämän tutkimuksen sanastoon löytyi.

Oppilaat olivat saaneet valita kirjoitustensa aiheet itse, mutta luokissa oli käsitelty mahdollisia aiheita aluksi keskustellen. Koirankakat ovat varmasti puhuttaneet, koska kirjoituskerta sijoittuu kevääseen, jolloin lumet sulavat ja kakat paljastuvat.

Se, että aineisto on kymmenen vuotta vanhaa tarkoittaa, että tulokset kuvaavat ilmaisutaitoja kymmenen vuoden takaa. Todennäköisesti tuoreemmalla aineistolla tehty tutkimus olisi nostanut esiin erilaista sanastoa, mutta *mielestäni* – rakenne on niin vahva, että se todennäköisesti toistuisi edelleenkin. Kuitenkin nykykartoitusten valossa voi olla, että kuva nykyisestä sanastosta olisi vielä suppeampi (PIRLS, 2021 s.22; KARVI, 2021, s.79)

Tutkimusmenetelmä sopi aineiston tarkasteluun ja antoi tutkimuskysymykseen vastaavaa tietoa. Tutkimuksessa täytyi soveltaa ja yhdistää kasvatustieteelle ja kielentutkimukselle tyypillisiä analyysimenetelmiä. Joustavuus kuitenkin näkyi valinnan vahvuutena, koska suunnan löytäminen tutkimusvaiheessa oli haastavaa. Tämän takia analyysivaihe näyttäytyi epäjohdonmukaisena, koska aineisto sisälsi niin monia eri näkökulmia.

Menetelmän valintaa haastoi kuitenkin suomenkielisen, täysin vastaavan tutkimuksen puuttuminen. Vaikka vastaavaa tutkimusta ei ole aikaisemmin tehty, aikaisempaan tutkimukseen perehtyminen osoitti työn merkityksen. Lasten sanastosta ja ilmaisuvalinnoista on hyvin vähän tietoa.

Vaikka kategorisoinnin joustavuus antoikin mahdollisuuden tarkastella aineistoa monesta eri näkökulmasta, se myös vaikeuttaa tutkimuskysymyksen rajaamista sekä tutkimuksen kannalta merkittävien seikkojen huomioimista. Koska tutkimukselle ei ollut selkeää ohjetta siitä, millä kategorioilla oli väliä, vaan kategoriat on päätetty analyysivaiheessa, tutkimuksen objektiivisuus on saattanut heikentyä. Samalla sanoista on voinut kadota niiden eri merkityssisällöt.

Sanastossa yksittäisellä sanalla ja sen ilmaisulla voi olla monta eri kategoriaa. Tässä tutkimuksessa käytetyillä kategorioilla, eli sanaluokissa on voinut tulla virheluokitteluja. Esimerkiksi sana 'asua' voi olla verbi asua tai substantiivin asu partitiivimuoto. Aineistosta on mahdollisuus silloin saada tilannekuva, jota ei välttämättä voi yleistää muuhun kuin vain tutkimuksen tekstin aikalaisteksteihin.

Johdosten lukeminen omaksi sanaperheekseen oli myös toimiva ratkaisu, sillä johdosten merkityserojen tarkentaminen onkin yksi koulun sanastotyön alueista. Sen merkitystä ei pidä vähätellä.

Koko aineisto analysoitiin, eikä analyysin ulkopuolelle tarvinnut jättää sanoja. Eteneminen aineiston purussa oli johdonmukaista, mutta syvempää analyysia tehtiin poukkoillen sanasta toiseen, jolloin huolellisuus on saattanut kärsiä.

Olla-verbi muodosti haasteen tutkimuksen analyysin näkökulmasta. Se on hyvin yleinen verbi, mutta yksinään käytettynä se on hyvin mitäänsanomaton. *Olla*-verbin yleisyys kuitenkin johtuu myös siitä, että se on monikäyttöinen ja osana monia eri rakenteita (VISK §456). Yleisyys ei sen tapauksessa kerrokaan kielenkäytön suurpiirteisyydestä tai yksipuolisuudesta. *Olla*-verbillä voidaan rakentaa myös hyvin kompleksisia lauserakenteita. Toisaalta sanasto sisältää monia eri kielen osa-alueita, joiden tutkiminen kerralla on hyvin työlästä yhdelle ihmiselle, jos tavoitteena olisi havainnoida syntaksia, semantiikkaa tai muita sanasto-osaamisen osa-alueita.

7.3 Jatkotutkimus

Jatkossa olisi mielenkiintoista tehdä vastaavanlainen tutkimus, jossa verrataan oppilaiden lukuharrastuksen laatua eli sitä, millaisia tekstejä oppilas arjessaan kuluttaa ja kielellisen kehityksen yhteyttä toisiinsa. Tässä tutkimuksessa nousi jo viitteitä niiden yhteydestä. Samalla olisi mahdollista nähdä, onko kymmenessä vuodessa tapahtunut muutosta oppilaiden kielellisen kehityksen etenemisessä ja sanastotaidoissa.

Myös uudemman aineiston keruu ja vertailu tässä tutkimuksessa käytettyyn aineistoon tarjoaisi mielenkiintoisen mahdollisuuden tarkastella sanastotaitojen yleistä kehitystä kymmenen vuoden aikana. Hyödyllistä dataa voisi tuottaa myös tutkimus, jossa yhdistetään passiivisen sanavaraston tutkimus esimerkiksi monivalintatehtävä sekä vastaavanlainen tuottavan tekstin tutkimus, jolloin sanavarastoa voisi kuvata laajemmin ja pureutua opetussuunnitelmassa mainittuihin eri alojen sanastoihin.

Lähteet

- Aalto, I. (2020). Yhdeksäsluokkalaiset ottavat kantaa luonnonsuojeluun ja ympäristön tilaan. Teoksessa. Hellgren, J. & Kauppinen, M. (toim.), *Kirjoitan, siis ajattelen : näkökulmia kirjoittamisen opetukseen* (pp. 13-27). Kansallinen koulutuksen arviointikeskus. Artikkelikokoelma / Kansallinen koulutuksen arviointikeskus, 2.
https://karvi.fi/app/uploads/2020/09/KARVI_Artikkelikokoelma_2_2020.pdf
- Absetz, P., & Routarinne, S. (2013) Writing performan and Writing Self-Efficacy: The Case of Finnish Fifth graders. Teoksessa Fisher, R., Howard, L. & Monteith, K. (toim.) *Why do we write as we write?*
- Alatalo, P., Munck, P., Lapinleimu, H., & Stolt, S. (2018). Pikkukeskosena syntyneiden lasten ymmärretyn sanaston kehitys vuoden iässä ja sen yhteys kognitiiviseen kehitykseen kaksivuotiaana. *Puhe Ja Kieli*, 38(2), 65–82. <https://doi.org/10.23997/pk.75366>
- Beers, S. F., & Nagy, W. E. (2011). Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing*, 24(2), 183–202.
- Buckley, B. (2003). *Children's Communication Skills: From Birth to Five Years* (1st ed.). Routledge.
<https://doi.org/10.4324/9780203865750>
- Blom, E., & Unsworth, S. (2010). Comparing L1 children, L2 children and L2 adults. Teoksessa. E. Blom & S. Unsworth (toim.), *Experimental methods in language acquisition research* (Vol. 27) 201-222.
- Cerovac, M., & Keane, T. (2024). Early insights into Piaget's cognitive development model through the lens of the Technologies curriculum. *International Journal of Technology and Design Education*.
<https://doi.org/10.1007/s10798-024-09906-5>
- Cruse, Alan (2004). *Meaning in the Language. An Introduction to Semantics and Pragmatics. Second Edition*. Oxford Textbooks in Linguistics. <https://philpapers.org/archive/CRUMIL-2.pdf>
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.
- Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. *Encyclopedia of language and education*, 2(2), 71-83.

Daller, Helmut, James Milton, Janine Treers-Daller (2007). *Modelling and Assessing Vocabulary Knowledge*. Cambridge Applied Linguistics Series.

CSC - IT Center for Science (2004). *Frequency Lexicon of the Finnish Newspaper Language*. Kielipankki. Retrieved from <http://urn.fi/urn:nbn:fi:lb-201405272>

Flores, N. (2024). Rebranding CALP. *Journal of Curriculum and Pedagogy*, 21(3), 388–391. <https://doi.org/10.1080/15505170.2024.2380327>

Ece Demir-Lira, Ö., Applebaum, L. R., Goldin-Meadow, S., & Levine, S. C. (2019). Parents' early book reading to children: Relation to children's later language and literacy outcomes controlling for other parent language input. *Developmental Science*, 22(3), e12764-n/a. <https://doi.org/10.1111/desc.12764>

Enright, E.A., Toledo, W., & Landau Wright, K. (2023). Advancing civics-specific disciplinary writing in the elementary grades. *Journal of Writing Research*, 15(1), 41-71. <https://doi.org/10.17239/jowr-2023.15.01>.

Eklund, K., Salmi, P., Polet, J. & Aro, M. (2013). LukiMat - Oppimisen arviointi: Lukemisen ja kirjoittamisen tuen tarpeen tunnistamisen välineet 1. luokalle. Tekninen opas. (Saatavilla osoitteessa <http://www.lukimat.fi/lukimat-oppimisen-arviointi/materiaalit/tuen-tarpeen-tunnistaminen/1lk/lukeminen/tekninen-opas>) (viitattu 2.2.2024)

Eklund, K., Salmi, P., Polet, J. & Aro, M. (2013). LukiMat - Oppimisen arviointi: Lukemisen ja kirjoittamisen tuen tarpeen tunnistamisen välineet 2. luokalle. Tekninen opas. (Saatavilla osoitteessa <http://www.lukimat.fi/lukimat-oppimisen-arviointi/materiaalit/tuen-tarpeen-tunnistaminen/2lk/lukeminen/tekninen-opas>) (viitattu 2.2.2024)

Elomäki, K. (2011). Kirjoitetun kielen rakenteellinen kompleksisuus kolmas- ja neljäsluokkalaisten kirjoitelmissa. Pro gradu -tutkielma. Kieli-, käännös- ja kirjallisuustieteiden yksikkö. Tampereen Yliopisto

EVK = Eurooppalainen viitekehys. Kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys. 2008. 1.–3. painos. Helsinki: WSOY

Rose, D., Gee, J. P., & Handford, M. (2012). Genre in the Sydney school. Teoksessa J. P. Gee & M. Handford (Toim.), *The Routledge handbook of discourse analysis* (1. painos, s. 209–225). Routledge.

POPS14 = Opetushallitus. Perusopetuksen opetussuunnitelman perusteet 2014. (OPS 14) oph.fi

Harjunen, E., & Korhonen, R. (2011). Keksitään kirjoittamista alakoulussa: herutellaan perusteluita. *Virittäjä*, 115(3). 395 – 405

Heikkilä, R. (2015). Rapid automatized naming and reading fluency in children with learning difficulties. University of Jyväskylä.

Heikkilä, R., Puttonen, J. & Siiskonen, T. (2019). Virtaa ja välineitä luku- ja kirjoitustaidon arviointiin. *NMI Bulletin Vol. 29, No.1*. Niilo Mäki Instituutti.

Heikkinen, V; Voutilainen, E; Lauerma, P; Tiililä, U. & Lounela M. (toim.) (2012): Genreanalyysi – tekstilajitutkimuksen käsikirja. *Kotimaisten kielten keskuksen julkaisuja 169*. Helsinki: Gaudeamus.

Hiidenmaa, P. (2018). Yhä moniulotteisempi lukutaito. *Virittäjä*, 122(2), 159-160.

<https://doi.org/10.23982/vir.70975>

Honko, M. (2013). Alakouluikäisten leksikaalinen tieto ja taito : toisen sukupolven suomi ja S1-verrokkit. Tampere.

Honko, M., & Pajunen, A. (2021a). Myöhempi kielenkehitys: sanatiedosta rakenne- ja tekstitietoon. teoksessa A. Pajunen & M. Honko (toim.) *Suomen kielen hallinta ja sen kehitys: Peruskoululaiset ja nuoret aikuiset. 8-60*, Helsinki: Suomalaisen Kirjallisuuden Seura. <https://doi.org/10.21435/skst.1472>

Honko, M., & Pajunen, A. (2021b).: Unelmieni päivä- sanasto kehityksellisestä näkökulmasta. Teoksessa A. Pajunen & M. Honko (toim.) *Suomen kielen hallinta ja sen kehitys: Peruskoululaiset ja nuoret aikuiset*. 61-107. <https://doi.org/10.21435/skst.1472>

Häyrinen, T., Serenius-Sirve, S. & Korkman, M. (1999). Lukilasse. Lukemisen, kirjoittamisen ja laskemisen seulontatestistö peruskoulun ala-asteen luokille 1–6.

VISK = Iso suomen verkko kielioppi. Hakulinen, A., Vilkuna, M., & Koivisto, V. (2008). (toim.) Iso suomen kielioppi. Helsinki: Suomalaisen kirjallisuuden seura <https://scripta.kotus.fi/visk/etusivu.php> (viitattu 14.9.2024)

KS = *Kielitoimiston sanakirja*. Kotimaisten kielten keskus Verkkoversion päivitys 19.3.2024 (viitattu 10.2.2024)

Kauppinen, M., Pentikäinen, J., Hankala, M., Kulju, P., Harjunen, E., & Routarinne, S. (2015). Systemaattinen katsaus perusopetusikäisten kirjoittamisen opetusta ja osaamista koskevaan tutkimukseen. *Kasvatus*, 46(2), 160-175.

Kajamies, A., Poskiparta, E., Annevirta, T., Dufva, M. & Vauras, M. (2003). YTTTE. Luetun ja kuullun ymmärtämisen ja lukemisen sujuvuuden arviointi. Turku: Turun yliopisto, Oppimistutkimuksen keskus (OTUK)

Kotilainen, L., Kurhila, S., & Kalliokoski, J. (toim.) (2019). *Kielenoppiminen luokan ulkopuolella*. (Tietolipas; Nro 262). Suomalaisen Kirjallisuuden Seura.

KARVI = Ukkola, A., & Metsämuuronen, J. (2021). Matematiikan ja äidinkielen ja kirjallisuuden osaaminen kolmannen luokan alussa. (Vol 20). Kansallinen koulutuksen arviointikeskus.

Laufer, B., I. S. P. Nation (2012). Vocabulary. teoksessa. Susan M. Gass, Alison Mackey (toim.). *e Routledge Handbook of Second Language Acquisition*. Lon don: Routledge, 163–176

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32(1), 69–92. <https://doi.org/10.1017/S0142716410000299>

Lerikkanen, M.-K., Vahtera, L., & Sanoma Pro, kustantaja. (2013). Lukemaan oppiminen ja opettaminen esi- ja alkuopetuksessa (1.-3. p.).

Lindeman, J. (1998). Ala-asteen lukutesti *ALLU*. Turku: Turun yliopisto, Oppimistutkimuksen keskus.

Masek, L. R., Ramirez, A. G., McMillan, B. T., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Beyond counting words: A paradigm shift for the study of language acquisition. *Child Development Perspectives*, 15(4), 274-280.

Masonen, V. (2013). Leksikaalinen kompetenssi osana kielitaitoa. *Lähivõrdlusi – Lähivertailuja Vol. 23*. s.187-209

Mol, S. E., & Bus, A. G. (2011). *To Read or Not to Read: A Meta-Analysis of Print Exposure From Infancy to Early Adulthood*. *Psychological Bulletin*, 137(2), (s.267–296).

<https://doi.org/10.1037/a0021890>

Meara, P. (1996). e Dimensions of Lexical Competence. *Performance and competence in second language acquisition* <https://www.lognostics.co.uk/vlibrary/meara1996a.pdf>

Milton, J. (2009). Measuring Second Language Vocabulary Acquisition. *Second Language Acquisition* 45. Bristol: Multilingual Matters

<https://ebookcentral.proquest.com/lib/kutu/reader.action?docID=543908>

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284).

Nation ISP. (2001). *Learning Vocabulary in Another Language*. Cambridge.

Pajunen, Anneli. (2012). Kirjoittamistaitojen kehitys 8–12-vuotiailla Alakoululaisten unelmakirjoitelmat. *Virittäjä* (Vol.1) s. 4–32

PISA = Hiltunen, J., Ahonen, A., Hienonen, N., Kauppinen, H., Kotila, J., Lehtola, P., ... & Vettenranta, J. (2023). PISA 2022 ensituloksia. *Opetus- ja kulttuuriministeriön julkaisuja 49*.

PIRLS = Leino, K., Sirén, M., Nissinen, K., & Puhakka, E. (2023). Puoli tuntia lukemista: Kansainvälinen lasten lukutaitotutkimus (PIRLS 2021). (Koulutuksen tutkimuslaitos, Tutkimuksia).

Pirskanen, H. & Kurki, S. (2014). Kokemuksia genrepedagogiikasta ja Reading to learn -ohjelmasta. Teoksessa Shore, S., & Rapatti, K. (toim.) *Tekstilajitaidot : lukemisen ja kirjoittamisen opetus koulussa*. Äidinkielen opettajain liitto.

Puro, Tarja. (1999). Sanastollinen tieto ja suomen kielen oppikirjojen sanasto. *Virittäjä* 1 Vol.1 (s. 2–26)

Pentikäinen, J., Routarinne, S., Hankala, M., Harjunen, E., Kauppinen, M., & Kulju, P. (2017). Oikeinkirjoituksesta monilukutaitoon : suomalainen kirjoittamisen opetus ennen ja nyt. Teoksessa V. Korhonen, J. Annala, & P. Kulju (toim.), *Kehittämisen palat, yhteisöjen salat : näkökulmia koulutukseen ja kasvatukseen* (s.57-180). <http://urn.fi/URN:ISBN:978-952-03-0500-0>

Routarinne, S. & Tenhola, E. (2015). Tekstilajin tajua 5.-luokkalaisten kirjoitelmissa. Teoksessa M. Kauppinen, M. Rautiainen & M. Tarnanen (toim.) *Elävä ainepedagogiikka: Ainedidaktiikan symposium Jyväskylässä 13.–14.2.2014 Ainedidaktisia tutkimuksia 9*, (s. 172–191)

Salmenoja, E. (2013). Rakenteellinen kompleksistuminen yläkouluikäisten kirjoitelmissa. Pro gradu-tutkielma. Kieli-, käännös-, ja kirjallisuustieteiden yksikkö. Tampereen yliopisto.

Salo, O-P. & Mäntylä K. (2020). Yksittäisistä sanoista sanaketjuihin – näkökulmia sanaston opetukseen, opiskeluun ja arviointiin. Teoksessa. Hildén, E. & Hahl, K. (toim.) *Kielididaktiikan katse tulevaisuuteen: Haasteita, mahdollisuuksia ja uusia avauksia kielten opetukseen*. Ainedidaktinen seura. <http://hdl.handle.net/10138/312321>

Stolt, S., & Vehkavuori, S. (2018). SANASEULA – Lapsen varhaisen sanaston kehityksen seulontamenetelmä. Niilo Mäki Instituutti.

Tentke, T. (2022). Varhaisen sanavaraston koon yhteys esilukutaitoihin perinnöllisen lukivaikeusriskin lapsilla. Pro gradu- tutkielma. Psykologia, Lääketieteellinen tiedekunta. Helsingin yliopisto.

Tomasello, M. (2005). *Constructing a language: A usage-based theory of language acquisition*. Harvard university press.

Tuomi, Jouni, ja Anneli Sarajärvi. Laadullinen tutkimus ja sisällönanalyysi. Uudistettu laitos. Helsinki: Kustannusosakeyhtiö Tammi, 2018. (e-kirja).

Treffers-Daller, J. & Milton, J. (2013). Vocabulary size revisited: the link between vocabulary size and academic achievement. *Applied Linguistics Review*, 4 (1). 151-172.
<https://doi.org/10.1515/applirev-20130007>