



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

Author(s): Mylläri, Taina

Title: Konjunktiot ja syntaktinen kompleksisuus : konjunktioiden käyttö suomi toisena kielenä -teksteissä eri taitotasoilla

Year: 2022

Version: Published version

Copyright: © 2022 Puheen ja kielen tutkimuksen yhdistys ry

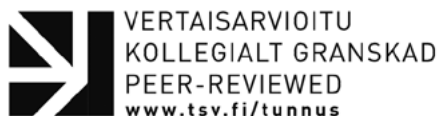
Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Mylläri, T. (2022). Konjunktiot ja syntaktinen kompleksisuus : konjunktioiden käyttö suomi toisena kielenä -teksteissä eri taitotasoilla. *Puhe ja kieli*, 42(3), 175-200.

<https://doi.org/10.23997/pk.126146>



KONJUNKTIOT JA SYNTAKTINEN KOMPLEKSISUUS: KONJUNKTIOIDEN KÄYTTÖ SUOMI TOISENA KIELENÄ -TEKSTEISSÄ ERI TAITOTASOILLA

Taina Mylläri, Kieli- ja viestintätieteiden laitos,
Jyväskylän yliopisto

Kielitaidon kehittymistä voidaan tutkia mittaamalla oppijankielen sujuvuutta, tarkkuutta ja kompleksisuutta. Kompleksisuus voidaan määritellä oppijankielen monipuolisuudeksi ja kehittyneisyydeksi, mutta syntaktista kompleksisuutta on kansainvälisessä toisen kielen oppimisen tutkimuksessa tyypillisesti mitattu alisteisiin lauserakenteisiin perustuvilla kvantitatiivisilla mittareilla. Tässä artikkelissa syntaktisen kompleksisuuden kehitystä tutkitaan tarkastelemalla rinnastus- ja alistuskonjunktioiden käyttöä suomi toisena kielenä -teksteissä Eurooppalaisen viitekehityksen eri taitotasolla. Aineistona käytetään Jyväskylän yliopiston Cefling-hankkeen oppijansuomen korpuksen tekstejä. Rinnasteisten ja alisteisten rakenteiden osuutta syntaktisen kompleksisuuden kehityksessä selvitetään tarkastelemalla rinnastus- ja alistuskonjunktioiden käyttöä suhteessa kielitaidon tasoon. Syntaktista kompleksisuutta oppijankielen monipuolisuutena selvitetään tarkastelemalla, mitä eri konjunktioita eri taitotasolla käytetään. Tulokset kyseenalaistavat käsityksen, että sivulauseet ilmaantuvat oppijankieleen vasta keskitasolla, sillä ne osoittavat, että suomi toisena kielenä -teksteissä käytetään sekä rinnastus- että alistuskonjunktioita jo taitotasolla A1. Lisäksi tulokset osoittavat, että konjunktioiden käyttö saattaa kehittyä eri tavalla aikuisten ja yläkouluikäisten teksteissä.

Avainsanat: konjunktioiden käyttö, suomi toisena kielenä, syntaktinen kompleksisuus

1 JOHDANTO

Toisen kielen taitoa ja sen kehittymistä voidaan tutkia esimerkiksi oppijankielen kielenpiirteiden muuttumisen tai toisen kielen käyttäjän kommunikatiivisten taitojen kehittymisen näkökulmasta. Molemmissa lähestymistavoissa kielitaidon ajatellaan usein kehittyvän

yksinkertaisista ja irrallisista ilmauksista monimutkaisempiin kokonaisuuksiin. Kun toisen kielen taidon kehitystä tarkastellaan kielenpiirteissä tapahtuvina muutoksina, apuna voidaan käyttää kompleksisuuden, tarkkuuden ja sujuvuuden käsitteitä (esim.

Kirjoittajan yhteystiedot:

Taina Mylläri

taina.myllari@jyu.fi

Ellis & Barkhuizen, 2005; Housen, Vedder & Kuiken, 2012). Näistä syntaktisen kompleksisuuden katsotaan usein kehittyvän niin, että alkeistasolla yksittäisiä merkityksiä ja lauseita aletaan yhdistää rinnastamalla, keskitasolla siirrytään rinnastamisesta sivulauseiden ja muiden alisteisten rakenteiden käyttöön ja ylimmillä taitotasolla lausekkeet pitenevät ja monipuolistuvat (esim. Kuiken & Vedder, 2019; Norris & Ortega, 2009).

Kielitaitoa kommunikatiivisesta näkökulmasta lähestyvässä Eurooppalaisessa viitekehyksessä (EVK, 2003) kielen osaaminen jaetaan taitotasoihin sen perusteella, miten kielenkäyttäjät selviytyvät erilaisista viestintätilanteista, ja taitotasolle tyypilliset kielelliset piirteet jäävät vähälle huomiolle (esim. Alanen, Huhta & Tarnanen, 2010; Martin, 2013). Konjunktioiden käyttö on yksi niistä harvoista kielenpiirteistä, jotka mainitaan EVK:n taitotasokuvauksissa: kirjoittamisen taitotasosteikon mukaan alkeistasolla A1 kielenkäyttäjät osaa kirjoittaa ”yksinkertaisia, irrallisia ilmauksia ja lauseita” ja seuraavalla tasolla (A2) hän osaa yhdistää niitä ”tavanomaisilla sidosanoilla, kuten sanoilla ’ja’, ’mutta’, ’koska’” (EVK 2003: 96). Aiempien tutkimustulosten mukaan sekä rinnastus- että alistuskonjunktiota kuitenkin käytetään jo toisen kielen oppimisen alkuvaiheessa (esim. Määttä, 2012; Vyatkina, 2012), eikä esimerkiksi rinnasteisten päälauseiden määrä virkkeessä tai sivulauseiden osuus kaikista lauseista välttämättä ole taitotasolla A1 merkittävästi pienempi kuin taitotasolla A2 (Myllylä, 2020a).

Kompleksisuutta voidaan tarkastella kielisysteemin sisältämien osien ja niiden välisen kytkösten määränä ja monipuolisuutena (esim. Bulté & Housen, 2012). Tässä artik-

kelissa yhtä kompleksisuuden osa-aluetta, syntaktista kompleksisuutta, lähestytään tarkastelemalla konjunktiolla merkittyjä syntaktisia kytköksiä. Tavoitteena on täydentää aiemmassa tutkimuksessa (Myllylä, 2020a) saatua tietoa oppijansuomen syntaktisesta kompleksisuudesta EVK:n eri taitotasolla. Tutkimuskysymykset ovat:

1. Miten paljon konjunktiolla merkittyjä syntaktisia kytköksiä käytetään eri taitotasolla?
2. Miten konjunktioiden käyttö muuttuu taitotasolta toiselle?

Tutkimusaineisto koostuu aikuisten ja yläkouluikäisten kirjoittamista suomen kielenä (S2) -teksteistä, jotka edustavat kolmea eri tehtävätyyppiä, ja se kattaa kielitaidon taitotasot A1–C2. Konjunktiolla merkittyjen syntaktisten kytkösten määrää ja yleisyyttä kartoitetaan laskemalla konjunktioiden esiintymistaajuudet eri taitotasolla. Konjunktioiden käytön muutoksia kartoitetaan selvittämällä, millä taitotasolla yksittäiset konjunktiot ilmaantuvat ja yleistyvät S2-teksteissä, millaisia syntaktisia kielenyksiköitä rinnastuskonjunktioiden avulla yhdistetään ja miten suuri osa sivulauseista sisältää alistuskonjunktion. Artikkelissa keskitytään konjunktiolla merkittyjen syntaktisten kytkösten kuvaamiseen ryhmätasolla taitotasoin. Konjunktiota ja niiden käyttöä käsitellään kieliopin ilmiönä.

2 SYNTAKTINEN KOMPLEKSISUUS OSANA KIELITAITOA

Toisen kielen oppimisen tutkimuksessa syntaktisen kompleksisuuden määritelmien lähtökohtana on oppijankielen rakenteiden monipuolisuus, vaihtelu tai kehittyneisyys

(esim. Ellis & Barkhuizen, 2005; Housen, De Clercq, Kuiken & Vedder, 2019; Ortega, 2003). Tutkijoiden keskuudessa ei kuitenkaan ole yksimielisyyttä siitä, miten kompleksisuus näkyy kielellisessä tuotoksessa, mikä on sen suhde kielitaidon tasoon tai oppimisprosessin kestoon ja miten sitä parhaiten mitataan (esim. Housen ym., 2019; Ortega, 2003; Pallotti, 2009, 2015; Wolfe-Quintero, Inagaki & Kim, 1998). Tyypillisesti oppijankielen syntaktista kompleksisuutta on mitattu lauseisiin tai T-yksikköihin eli yhden päälauseen ja siihen liittyvien sivulauseiden muodostamaan kokonaisuuteen perustuvilla määrällisillä mittareilla, kuten T-yksikön pituus sanoina tai lauseina ja sivulauseiden osuus kaikista lauseista, joita on kuitenkin arvosteltu siitä, että ne yksinkertaistavat syntaktista kompleksisuutta keskittymällä lähinnä alisteisuuteen (engl. *subordination*) (esim. Bulté & Housen, 2012). Ne jättävät huomiotta esimerkiksi rinnastamisen (esim. Bardovi-Harlig, 1992) ja muut kuin lausetason yhdistämiskeinot (esim. De Clercq & Housen, 2017). Lisäksi mittareita on arvosteltu siitä, että ne tavoittavat vain kielenyksiköiden määrissä tapahtuvat muutokset eivätkä muutosten syitä, sillä keskenään hyvin erilaiset virkkeet tai T-yksiköt voivat tuottaa täysin samat määrälliset tulokset (esim. Biber, Gray & Poonpon, 2011; Rimmer, 2009).

Nykyään syntaktista kompleksisuutta tutkitaankin yleensä moniulotteisena ilmiönä, jota tarkastellaan useilla rinnakkaisilla mittareilla (esim. Housen ym., 2019). Myös konjunktiota on tarkasteltu muiden syntaktisen kompleksisuuden mittareiden rinnalla. Alkeistason saksanoppijoiden rinnastus- ja alistuskonjunktioiden käyttöä tutkinut Vyatkina (2012, 2013) totesi, että ryhmätasolla rinnas-

tuskonjunktioiden käyttö väheni ja alistuskonjunktioiden käyttö lisääntyi kielitaidon kehittyessä mutta yksilötasolla kehitys saattoi poiketa ryhmätason kehityksestä. Sanoja ja lauseita kytkeviä ilmauksia yhtenä syntaktisen kompleksisuuden mittarina käyttäneet Benevento ja Storch (2011) puolestaan havaitsivat, että näiden ilmausten määrä ranskanoppijoiden teksteissä ei välttämättä kehity samalla tavalla kuin lauseiden määrä T-yksikössä: kolmella mittauskerralla kerätyissä teksteissä lauseiden määrä T-yksikössä oli suurimmillaan eri mittauskerralla kuin kielenyksiköiden kytkemiseen käytettyjen ilmausten määrä¹. Grant ja Ginther (2000) tutkivat englanti toisena kielenä -tekstien kielellisten piirteiden suhdetta kielitaidon arviointiin ja havaitsivat, että paremman arvosanan saaneissa teksteissä käytettiin enemmän konjunktiota ja sivulauseita kuin alemman arvosanan saaneissa. Taguchi, Crawford ja Wetzel (2013) puolestaan tutkivat yliopistotason vaihto-opiskelijoiden englannin kielen lähtötasotestien tekstejä ja havaitsivat, ettei alisteisuus välttämättä toimi kirjoittamisen taidon mittarina edistyneemillä taitotasolla, sillä pistemäärän perusteella kahteen ryhmään jaetuista teksteistä enemmän konjunktiota sisälsivät alemman pistemäärän saaneet tekstit, jotka myös joidenkin muiden tutkimuksessa käytettyjen syntaktisen kompleksisuuden mittareiden mukaan olivat hieman kompleksisempia kuin enemmän pisteitä saaneet tekstit. He toteavatkin, ettei lausetason kompleksisuus tai pelkkä alisteisten rakenteiden mittaaminen välttämättä anna oikeaa kuvaa kielitaidosta kaikissa konteksteissa ja ettei sivulauseiden

1 Tarkastelussa oli konjunktioiden lisäksi konnektiiveja, tosin konjunktiot *mais* (mutta) ja *parce que* (koska) Benevento ja Storch rajasivat pois niiden esiintymien suuren määrän vuoksi.

suuri määrä aina tee tekstistä parempaa vaan voi tuottaa liiallista kompleksisuutta, mikä saattaa hankaloittaa tekstin ymmärtämistä ja alentaa sen saamaa arvosanaa.

Syntaktisen kompleksisuuden ja kieli- tai kirjoitustaidon suhde ei olekaan yksiselitteinen, eikä kompleksisuuden lisääntyminen välttämättä merkitse parempaa kieli- tai kirjoitustaitoa (esim. Crossley & McNamara, 2014; Ortega, 2003; Pallotti, 2009; ensikielen osalta ks. esim. Jagaiah, Olinghouse & Kearns, 2020; Pajunen & Vainio, 2021a, 2021b). Esimerkiksi Lambert ja Kormos (2014) ovat todenneet, että harjaantuneet kirjoittajat saattavat käyttää aloittelijoita yksinkertaisempaa ilmaisua välittäessään monimutkaisia ajatuksia.

Aiemmissä kompleksisuuden tutkimuksissa on havaittu, että kompleksisuus voi kielitaidon eri tasoilla tai typologisesti erilaisissa kielissä ilmetä eri tavoin (esim. Bernardini & Granfeldt, 2019; Housen ym., 2019). Eroja voi olla siinä, miten kompleksisuus jonkin tietyn mittarin mukaan kehittyy (esim. Gyllstad, Granfeldt, Bernardini & Källkvist, 2014; Kuiken & Vedder, 2019) tai mitä pidetään kompleksisena. Esimerkiksi suomen kielen konjunktiot alkuiset sivulauseet tai epäsuorat kysymykset eivät välttämättä ole syntaksin näkökulmasta päälauseita kompleksisempia (Martin, Mustonen, Reiman & Seilonen, 2010). Alisteisuus ei myöskään ilmene kaikissa kielissä samalla tavalla (esim. Herlin, Visapää & Kalliokoski, 2014), eivätkä alisteiset rakenteet välttämättä muodosta yhtenäistä syntaktista kategoriaa (Cristofaro, 2014). Myös lähtökieli voi vaikuttaa kohdekielen syntaktiseen kompleksisuuteen (esim. Khushik & Huhta, 2020). Tässä artikkelissa keskitytään ryhmätason kehityskulkuihin ja ensikielen tai muiden aiemmin opittujen

kielten mahdollinen vaikutus on rajattu tutkimuksen ulkopuolelle.

Oppijansuomen syntaktista kompleksisuutta on tutkittu erilaisten määrällisten mittareiden avulla. Alisaari (2016) tutki taitotason A2 suomenoppijoiden T-yksiköiden sanamäärää² eikä löytänyt siinä tilastollisesti merkitsevää kehitystä neljän viikon kurssin aikana. Kompleksisuuden ja tarkkuuden kehitystä suomea toisena kielenä ja vieraana kielenä opiskelevien teksteissä tutkinut Tilma (2014) havaitsi kompleksisuuden lisääntyvän molemmissa ryhmissä, kun mittareina käytettiin virketyyppejä sekä virkkeiden ja lauseiden sisältämien morfeemien määrää. Oppijansuomen kompleksisuutta ja tarkkuutta dynaamisten systeemien teorian (DST) näkökulmasta tutkineet Spoelman ja Verspoor (2010) puolestaan havaitsivat, että yhdeltä oppijalta kolmen vuoden aikana kerättyssä aineistossa kompleksisuus lisääntyi epälineaarisesti, kun sitä tarkasteltiin sanojen, substantiivilausekkeiden ja virkkeiden kompleksisuutta kuvaavien suhdelukujen avulla. Tämän artikkelin aineistossa tehdyssä määrällisessä tutkimuksessa (Mylläri, 2020a) S2-tekstien syntaktista kompleksisuutta ja sen kehitystä tarkasteltiin seitsemän kansainvälisessä toisen kielen oppimisen tutkimuksessa yleisesti käytetyn määrällisen mittarin avulla. Tutkimuksen tulokset osoittivat, että ryhmätasolla sanojen määrä T-yksikössä, lauseiden määrä T-yksikössä ja sivulauseiden osuus kaikista lauseista kasvoivat ja rinnasteisten päälauseiden määrä virkkeessä pieneni kielitaidon kehittyessä, kun aineiston tehtävyytyyppejä tarkasteltiin rinnakkain. Mittareiden tavoittama kehitys

2 Alisaari käytti T-yksikön sanamäärää sujuvuuden mittarina (vrt. Wolfe-Quintero ym., 1998).

ei kuitenkaan ollut lineaarista ja varianssi taitotasojen sisällä oli suurta, mikä osittain selittää tutkimuksen määrällisten mittareiden heikkoa kykyä erotella taitotasoja.

Myös konjunktioiden käyttöä oppijansuomessa on tutkittu. Alkeistason suomenoppijoiden sanastoa tutkinut Määttä (2012) totesi, että jo alkuvaiheessa teksteissä käytettiin samoja konjunktioita, jotka esiintyivät oppikirjassa: *ja*, *kun* ja *mutta*. Alakouluikäisten S1- ja S2-oppilaiden leksikaalista osaamista tutkinut Honko (2013) puolestaan havaitsi, että ylemmillä luokilla konjunktioiden käyttö monipuolistui ja niiden määrä yleisimpien sanojen joukossa lisääntyi, minkä lisäksi hän havaitsi virkerakenteiden kompleksistuvan ja monipuolistuvan alakoulun aikana. Pajunen & Vainio (2021b) ovat tarkastelleet S1-tekstien alisteisia elementtejä suhteessa kaikkiin lauseita yhdistäviin elementteihin ja todenneet, että kun tekstit jaettiin kahteen ryhmään niiden T-yksiköiden keskimääräisen sanamäärän (MLTU) tai kompleksisuusindeksin (CI) mediaanien ylä- ja alapuolelle, alisteisten elementtien osuus oli suurempi teksteissä, joiden MLTU tai CI asettuivat ryhmän mediaanin arvon yläpuolelle. Lisäksi he totesivat, että kehityksellisesti relatiivilauseet lisääntyvät ja adverbiaalilauseet vähenevät.

Tässä tutkimuksessa syntaktista kompleksisuutta lähestytään konjunktioiden syntaktisen käytön avulla, ja tarkasteluun otetaan mukaan sekä rinnasteisuus että alisteisuus. Ilmiötä tarkastellaan ryhmätasolla suhteessa kielitaidon taitotasoon. Tarkastelu rajataan vain konjunktioihin ja niillä ilmaistuihin kytköksiin. Käsiteltävät konjunktiot ja niiden jako rinnastus- ja alistuskonjunktioihin esitellään seuraavassa luvussa.

3 AINEISTO JA MENETELMÄT

Tutkimusaineisto sisältää 667 tekstiä (48 876 sanaa) aikuisilta S2-kirjoittajilta ja 411 tekstiä (16 590 sanaa) yläkouluikäisiltä S2-kirjoittajilta sekä 453 tekstin (19 826 sanan) vertailuaineiston S1-nuorilta (taulukko 1). Se on koostettu Jyväskylän yliopiston Cefling-hankkeen³ S2-aineistosta ja S1-vertailuaineistosta, joka on Cefling-hankkeessa kerätty yläkouluikäisiltä S1-kirjoittajilta samoilla tehtävillä kuin S2-nuorten aineisto. Tehtävätyyppinä ovat epämuodollinen viesti (esim. sähköposti ystävälle tai opettajalle), muodollinen viesti (esim. reklamaatio verkkokauppaan) ja mielipideteksti annetusta aiheesta (esim. kännykän käytöstä koulussa). Alkuperäiset tekstit on kirjoitettu käsin rajoitetussa ajassa ja ilman apuvälineitä. S2-tekstit on Cefling-hankkeessa sijoitettu Eurooppalaisen viitekehysten mukaisille taitotasolle A1–C2 tavalla, jota voidaan pitää varsin luotettavana sekä tilastollisin että laadullisin menetelmin tarkasteltuna (Huhta, Alanen, Tarnanen, Martin & Hirvelä, 2014). Aikuisten aineistossa on tekstejä kaikilta taitotasoilta, nuorten aineistossa tasoilta A1–B2, tosin nuorten aineiston taitotasolla B2 on vain pieni määrä tekstejä. S1-tekstejä ei ole arvioitu taitotasolle. (Jantunen & Pirkola, 2015; Mustonen, 2015.) Aineisto on aiemman syntaktisen kompleksisuuden tutkimuksen yhteydessä koodattu xml-muotoon ja siihen on merkitty virke- ja lauserajat sekä lauseiden keskinäiset suhteet. Teksteistä on poistettu sanasta sanaan tehtävänannoista kopioidut otsikot sekä erityisesti viesteille tyypilliset tervehdykset ja lähettäjän

3 Cefling = Linguistic Basis of the Common European Framework for L2 English and L2 Finnish. (<http://www.jyu.fi/hytk/fi/laitokset/kivi/tutkimus/hankkeet/paattyneet-tutkimushankkeet/cefiling/suom>)

tai vastaanottajan yhteystiedot silloin, kun ne ovat sisältäneet vain verbittömiä rakenteita. (Mylläri, 2020a; aineiston segmentoinnista tarkemmin Mylläri, 2020b.)

Tässä tutkimuksessa eri tehtävätyypit on yhdistetty. Konjunktioita ja niiden käyttöä tarkastellaan suhteessa kielitaidon taitotasoon ja erilaiset kirjoittajakohtaiset taustatekijät, kuten ensikieli tai Suomessa vietetty aika⁴, sekä niiden mahdollinen vaikutus tekstien syntaktisiin piirteisiin on rajattu tarkastelun ulkopuolelle. Aikuisten ja nuorten

tekstejä tarkastellaan kuitenkin erikseen, sillä aiemmissa Cefling-aineistoa käyttävissä tutkimuksissa (Kajander, 2013; Reiman, 2011; Seilonen, 2013) on havaittu, että joissakin kielenpiirteissä ja niiden muuttumisessa taitotasolta toiselle on eroja aikuisten ja nuorten välillä. S1-vertailuaineiston tekstit on tässä tutkimuksessa yhdistetty yhdeksi ryhmäksi (ensikielisen vertailuaineiston tarpeellisuudesta ks. esim. Pallotti, 2009).

TAULUKKO 1. Aineiston teksti- ja sanamäärät kirjoittajaryhmittäin ja tehtävätyypeittäin (EM = epämuodolliset viestit, M = muodolliset viestit, MP = mielipidetekstit).

Kirjoittajaryhmä / S2-taitotaso	Tekstejä				Sanoja			
	EM	M	MP	Yhteensä	EM	M	MP	Yhteensä
Aikuiset								
A1	39	22	50	111	1 582	752	2 261	4 595
A2	39	27	37	103	1 533	1 494	2 272	5 299
B1	41	42	43	126	2 453	2 290	5 142	9 885
B2	39	34	35	108	2 206	1 983	4 166	8 355
C1	26	45	46	117	1 528	3 337	5 876	10 741
C2	14	58	30	102	970	5 152	3 879	10 001
S2-nuoret								
A1	25	33	32	90	677	860	775	2 312
A2	79	40	39	158	2 879	1 489	1 589	5 957
B1	64	40	40	144	2 971	1 980	2 232	7 183
B2	12	7	.	19	677	461		1 138
Vertailuryhmä								
S1-nuoret	163	162	128	453	6 186	6 931	6 709	19 826

⁴ Aineiston S2-kirjoittajat ovat äidinkieleltään heterogeeninen ryhmä, eri ensikieliä on yli 20 (Jantunen & Pirkola, 2015). Kirjoittajat ovat voineet oppia suomea toisena tai vieraana kielenä tai molempina.

Tarkasteltaviksi konjunktioiksi valittiin Ison suomen kieliopin (VISK, 2004) pykälissä 813–816 käsitellyt konjunktiot, jotka jaoteltiin syntaktisen käytön mukaan rinnastus- ja alistuskonjunktioihin (taulukko 2). Vertailukonjunktio *kuin* laskettiin alistuskonjunktioiksi, vaikka sillä on myös rinnastuskonjunktin piirteitä (VISK § 819). Konjunktioiden esiintymät etsittiin siten, että aineistosta kerättiin ensin koneellisesti kaikkien lauseiden ensimmäiset sanat, joiden joukosta tarkasteltavaksi valitut konjunktiot poimittiin käsin. Näin mukaan saatiin myös kirjoitusasultaan odotuksenmukaisesta poikkeavat lauseenalkuiset esiintymät. Esimerkiksi sanat *etta*, *ette* ja *silla* sekä puhekieliset muodot *et*, *ko*, *ku*, *kuha(n)* ja *jossei* otettiin mukaan tarkasteluun, jos ne kontekstin perusteella olivat tulkittavissa konjunktioiksi. Näin löytyneiden eri kirjoitusasujen kaikki esiintymät kerättiin aineistosta koneellisesti virkekonteksteineen, jolloin mukaan saatiin myös muut kuin lauserajalla olevat esiintymät.

Tarkastelun ulkopuolelle rajattiin tapaukset, joissa sanaa käytettiin muuten kuin konjunktiona. Löydettyjen esiintymien joukosta poistettiin käsin kysymyssanana käytetty *koska*, adverbina käytetty *vaan* (esim. *missä vaan*), sekä sanat *et* (< *ei*), *mut* (< *minut*) ja *sillä* (< *se*), kun niitä käytettiin kieltoverbin tai pronominin taivutusmuotoina. Lisäksi pois rajattiin *vaikka* ja *vaikkapa*, kun niitä käytettiin muuten kuin konjunktiona (esim. *tavataan vaikka keskiviikkona*). Konjunktiot koodattiin aineistoon, ja esiintymien määrät laskettiin koneellisesti. Parikonjunktio *sekä – että* koodattiin ja laskettiin yhdeksi rinnastuskonjunktioiksi. Lopuksi tarkistettiin koneellisesti, ettei aineistossa ole sellaisia konjunktioita, jotka on mainittu VISK:n pykälissä 813–816 mutta joita ei esiintynyt lauserajoilla. Esiintymien koneellisessa poiminnassa ja laskemisessa hyödynnettiin tätä tutkimusta varten kirjoitettuja Python-skriptejä.

TAULUKKO 2. Tutkimuksessa tarkasteltavat konjunktiot.

Aineistossa esiintyvät konjunktiot		Aineistossa esiintymättömät konjunktiot
Rinnastuskonjunktiot	Alistuskonjunktiot	
-kä, eli, elikä, ja, mutta (<i>mut</i>), muttei, sekä, sillä, tai, vaan, vai; sekä – että, joko – tai, joko – taikka, niin – kuin	ellei, että (<i>et</i>), ettei (<i>etteikö</i>), jos (<i>jossei</i>), jotta, jottei, koska, kuin, kun (<i>ku</i> , <i>ko</i>), kunhan, kunnes, mikäli, vaikka, vaikkapa; ennen kuin, niin kuin, kun taas	jahka, jollei, joskin, josko, kunpa, saati, sun, taikka, ynnä, vaikkakaan, vaikkakin

Konjunktioiden esiintymistä tarkastellaan ensisijaisesti ryhmitasolla ja aineiston tilastollisessa kuvailussa kunkin taitotason tekstejä käsitellään tekstimassana. Koska tekstien lukumäärä ja sanamäärä vaihtelevat taitotasolta toiselle, taitotasojen ja kirjoittajaryhmien rin-

nakkaisessa tarkastelussa käytetään konjunktioiden normalisoituja esiintymistaajuuksia suhteessa tuhanteen sanaan. Eri konjunktioiden absoluuttiset ja normalisoidut esiintymistaajuudet on esitetty liitteen 1 taulukoissa 1A (aikuisen aineisto) ja 1B (nuorten aineisto

sekä S1-vertailuaineisto). Aineistossa esiintyvien konjunktioiden valikoiman tarkastelussa käytetään myös konjunktioiden suhteellisia frekvenssejä ($f\%$), jotka on laskettu jakamalla yksittäisen konjunktion esiintymien määrä kullakin taitotasolla kaikkien saman taitotason konjunktioesiintymien kokonaismäärällä.

Taitotasojen tilastollista vertailua varten laskettiin myös tekstikohtaiset konjunktioiden esiintymistäajuudet, jotka normalisoitiin ryhmätason tarkastelusta poiketen suhteessa sataan sanaan, koska kaikilla taitotasolla samoin kuin S1-aineistossa tekstien keskimääräinen pituus on alle sata sanaa. Vertailussa käytettiin ei-parametrisiä testejä, sillä aineisto ei Shapiro-Wilkin ja Levenen testien mukaan täyttänyt parametristen tilastollisten testien taustaoletuksia jakauman normaaliuden ja varianssin homogeenisuuden osalta. Tilastollisina testeinä käytettiin Kruskal-Wallis testia ja parittaisissa vertailuissa Mann-Whitneyn U-testiä vastaavaa Wilcoxonin järjestyslukujen summan testiä, jonka p-arvoihin tehtiin Bonferroni-korjaus. Tilastollisen merkitsevyyden rajana käytettiin arvoa $p < 0,05$. Tilastolliset analyysit tehtiin R-ohjelmiston versiolla 3.4.4 RStudio:n versiolla 1.1.456 avulla.

Tarkastelu aloitetaan rinnastus- ja alistuskonjunktioiden esiintymistä eri taitotasolla. Tässä yhteydessä testataan, onko taitotasojen välillä tilastollisesti merkitseviä eroja kaikkien konjunktioiden sekä erikseen rinnastus- ja alistuskonjunktioiden normalisoiduissa esiintymistäajuuksissa. Sen jälkeen selvitetään, mitä kielenyksiköitä rinnastuskonjunktiolla yhdistetään ja miten suuri osa aineiston sivulauseista sisältää alistuskonjunktion. Lopuksi kartoitetaan kullekin taitotasolle tyypillistä konjunktioiden varantoa selvittämäl-

lä, mitä konjunktioita eri taitotasolla esiintyy, sekä laskemalla yksittäisten konjunktioiden suhteelliset frekvenssit ja se, miten monessa eri tekstissä ne esiintyvät (vrt. Michel, 2013).

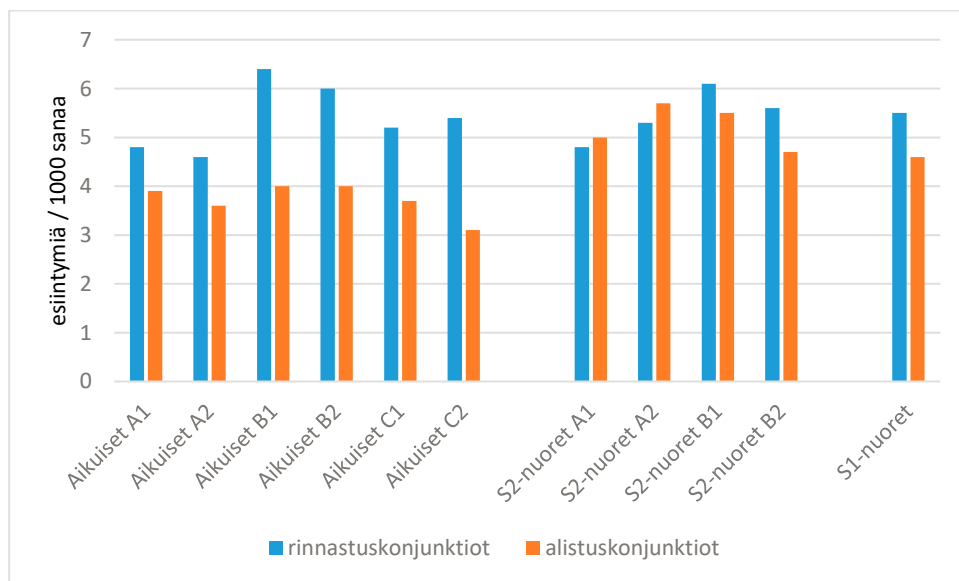
4 TULOKSET

Sekä rinnastuskonjunktioita että alistuskonjunktioita käytetään aineiston kaikilla taitotasolla. Jo taitotasolla A1 aikuisten aineistossa 89 % ja nuorten aineistossa 86 % teksteistä sisältää vähintään yhden konjunktion ja noin puolessa teksteistä (aikuisten aineistossa 51 prosentissa ja nuorten aineistossa 46 prosentissa) on käytetty sekä alistus- että rinnastuskonjunktioita molempia vähintään kerran. Aineiston alimmilla taitotasolla on myös täysin konjunktiottomia tekstejä: taitotasolla A1 aikuisten aineistossa 12 (11 %) ja nuorten aineistossa 13 (14 %), taitotasolla A2 aikuisten aineistossa 7 (7 %) ja nuorten aineistossa 3 (2 %). Lisäksi aikuisten aineiston taitotasolla B1 on yksi teksti, jossa ei käytetä konjunktioita. Vaikka ilmiö näyttää olevan tyypillinen S2-aineiston alimmille taitotasolle, myös S1-nuorten vertailuaineistossa on 20 täysin konjunktiotonta tekstiä (4 % kaikista S1-teksteistä). Tyypillisesti kirjoittajat käyttävät sekä rinnastus- että alistuskonjunktioita.

Aikuisten aineistossa sekä rinnastus- että alistuskonjunktioiden normalisoitu esiintymistäajuus on suurimmillaan taitotasolla B1 ja B2, nuorten aineistossa puolestaan alistuskonjunktioiden taitotasolla A2 ja B1 ja rinnastuskonjunktioiden taitotasolla B1 (kuvio 1). Rinnastuskonjunktioiden normalisoitu esiintymistäajuus on lähes kaikilla taitotasolla suurempi kuin alistuskonjunktioiden eli rinnastuskonjunktiot ovat yleisempiä kuin alistuskonjunktiot. Ainoa poikkeus tästä ovat S2-nuorten aineiston taitotasot A1 ja

A2, joilla alistuskonjunktioiden normalisoitu esiintymistaajuus on hieman suurempi kuin rinnastuskonjunktioiden. Tältä osin aikuisten ja nuorten aineistot poikkeavat toisistaan. S1-vertailuaineistossa rinnastuskonjunktiot

ovat yleisempiä kuin alistuskonjunktiot, mikä vastaa yleissuomesta aiemmin saatuja tuloksia (Saukkonen, Haipus, Niemikorpi & Sulkala, 1979, s. 10).



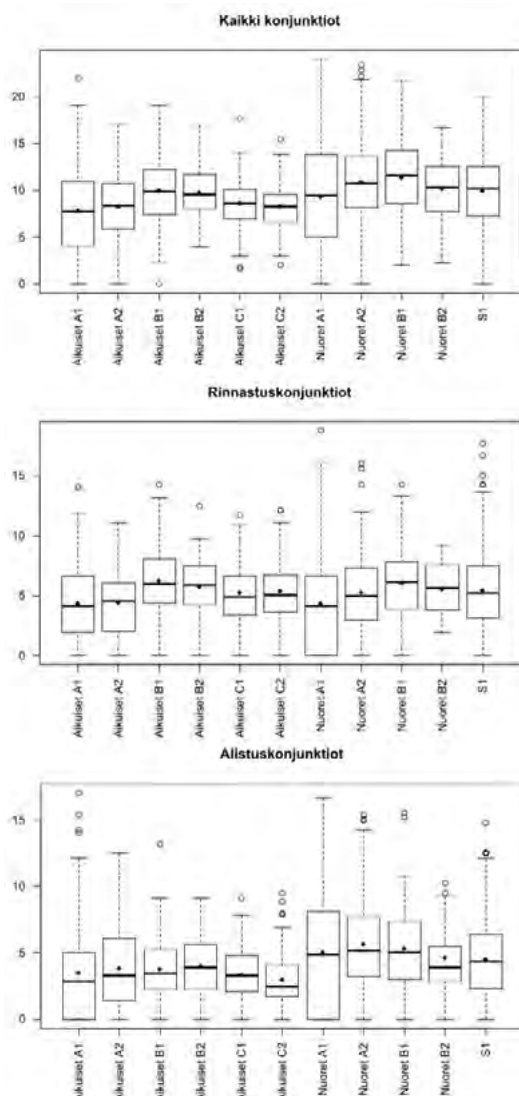
KUVIO 1. Rinnastus- ja alistuskonjunktioiden normalisoidut esiintymistaajuudet.

Kaikkien konjunktioiden, rinnastuskonjunktioiden ja alistuskonjunktioiden normalisoituja esiintymistaajuuksia (kuvio 2) verrattiin tilastollisin menetelmin. Aikuisten aineistossa taitotasojen välillä on Kruskal-Wallis testin mukaan tilastollisesti merkitseviä eroja kaikkien konjunktioiden ($Khiin\ neliö = 35,29; df = 5; p < 0,001$), rinnastuskonjunktioiden ($Khiin\ neliö = 36,17; df = 5; p < 0,001$) ja alistuskonjunktioiden ($Khiin\ neliö = 16,15; df = 5; p = 0,006$) normalisoidussa esiintymistaajuudessa. Kaikkien konjunktioiden esiintymistaajuus on suurimmillaan B-tasolla, ja erot sekä alimpiin että ylimpiin taitotasoihin ovat Wilcoxonin järjestyslukujen summan testin mukaan tilastollisesti merkitseviä. Taitotaso

B1 eroaa tilastollisesti merkitsevästi taitotasosta A1 ($W = 4933; p = 0,001$), A2 ($W = 4983; p = 0,038$), C1 ($W = 5657; p = 0,026$) ja C2 ($W = 4521; p = 0,002$), taitotaso B2 puolestaan taitotasosta A1 ($W = 4194,5; p = 0,002$), C1 ($W = 5657; p = 0,025$) ja C2 ($W = 3752; p = 0,001$). Rinnastuskonjunktioiden normalisoidussa esiintymistaajuudessa on aikuisten aineistossa tilastollisesti merkitseviä eroja tasojen A ja B välillä. Taitotaso A1 eroaa tilastollisesti merkitsevästi taitotasosta B1 ($W = 4653; p < 0,001$) ja B2 ($W = 4290; p = 0,004$), samoin taitotaso A2 taitotasosta B1 ($W = 4332; p < 0,001$) ja B2 ($W = 3963; p = 0,005$). Aikuisten aineistossa alistuskonjunktioiden normalisoidussa esiintymistaajuus

juudessa ainoa tilastollisesti merkitsevä ero on taitotasojen B2 ja C2 välillä ($W = 7071$; $p = 0,006$). Nuorten aineistossa sekä kaikkien konjunktioiden että rinnastuskonjunktioiden normalisoitu esiintymistäajuus kasvavat taitotasolta A1 taitotasolle B1 ja alistuskonjunktioiden taitotasolta A1 taitotasolle A2. Kruskal-Wallis testin mukaan nuorten aineistossa on tilastollisesti merkitseviä eroja taitotasojen välillä kaikkien konjunktioiden ($Khiin\ neliö = 10,14$; $df = 3$; $p = 0,017$) ja rinnastuskonjunktioiden ($Khiin\ neliö = 16,98$;

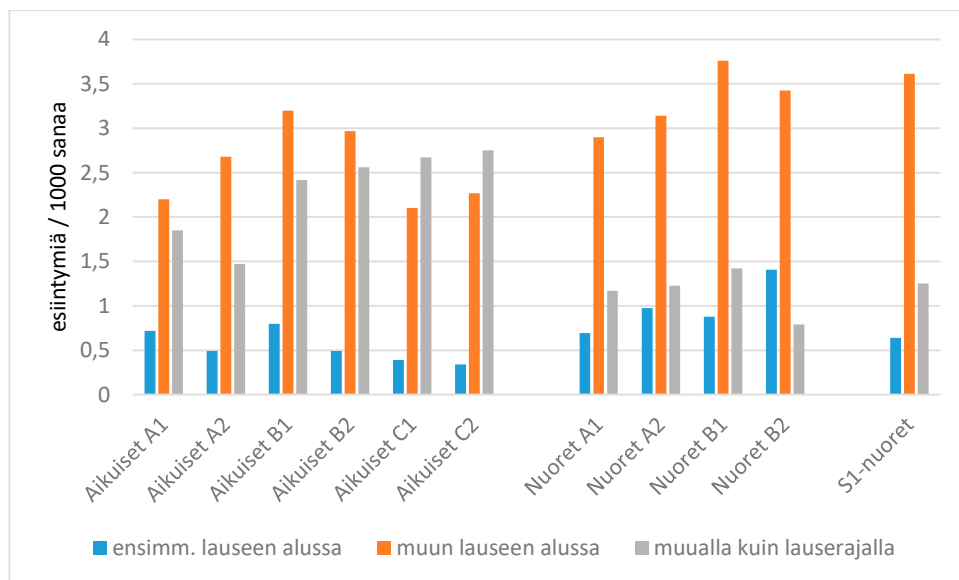
$df = 3$; $p < 0,001$) normalisoidussa esiintymistäajuudessa. Parittaisten vertailujen perusteella taitotasojen A1 ja B1 välillä on tilastollisesti merkitsevä ero sekä kaikkien konjunktioiden ($W = 5023$; $p = 0,023$) että rinnastuskonjunktioiden ($W = 4474,5$; $p < 0,001$) esiintymistäajuudessa. Alistuskonjunktioiden esiintymistäajuudessa ei nuorten aineistossa ole Kruskal-Wallis testin perusteella tilastollisesti merkitseviä eroja taitotasojen välillä ($Khiin\ neliö = 2,86$; $df = 3$; $p = 0,414$).



KUVIO 2. Tekstikohtaiset normalisoidut esiintymistäajuudet (esiintymiä / 100 sanaa).

Koska rinnastuskonjunktioita voidaan käyttää erilaisten syntaktisten elementtien yhdistämiseen, rinnastuskonjunktioiden esiintymiä tarkasteltiin myös jakamalla ne kolmeen ryhmään (kuvio 3): virkkeen ensimmäisen (tai ainoan) lauseen aloittaviin, virkkeen sisällä kahta lausetta yhdistäviin sekä muualla kuin lause- tai virkerajalla esiintyviin (vrt. Vyatkina, 2013). Yleisimpiä ovat virkkeen sisällä lauseita yhdistävät rinnastuskonjunktiot, joita on aikuisten aineistossa taitotasoilla A1–B2 noin 50 %, S2-nuorten aineistossa noin 60 % ja vertailuaineistossa 66 % kaikista rinnastuskonjunktioista. Aikuisten aineiston kaksi ylintä taitotasoa poikkeavat muusta aineistosta. Niillä rinnastuskonjunk-

tioista noin 40 % yhdistää saman virkkeen lauseita ja noin puolet (C1-tasolla 52 % ja C2-tasolla 51 %) rinnastuskonjunktioista sijoittuu muualle kuin lause- tai virkerajalle. Virkkeenalkuisten rinnastuskonjunktioiden osuus vaihtelee aikuisten aineistossa A1-tason 15 prosentista C2-tason 6 prosenttiin, S2-nuorten aineistossa 14 ja 25 prosentin välillä. S1-nuorten teksteissä rinnastuskonjunktioista 12 % aloittaa virkkeen, mikä on hyvin lähellä suomenkielisen asiaproosan aineistosta aiemmin raportoituja tuloksia, joiden mukaan 13 % rinnastuskonjunktioilla alkavista päälauseista sijaitsi virkkeen alussa (Hakulinen, Karlsson & Vilku, 1980).



KUVIO 3. Rinnastuskonjunktioiden normalisoidut esiintymistaajuudet esiintymispaikan mukaan ryhmiteltyinä.

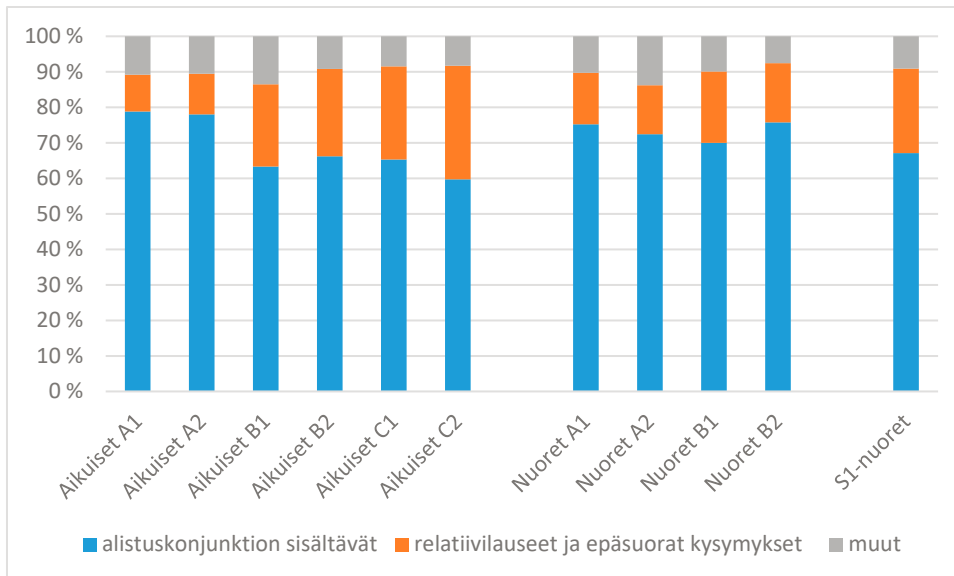
Kaikki sivulauseet eivät välttämättä sisällä alistuskonjunktioita. Siksi alistuskonjunktioiden esiintymiä tarkasteltiin myös suhteessa sivulauseiden määrään. Tarkasteltua varten aineiston sivulauseet luokiteltiin kolmeen ryhmään: 1) alistuskonjunktioita sisältävät sivulauseet, 2) relatiivilauseet ja epäsuorat kysymykset sekä 3) muut sivulauseet (vrt. Vilkuna, 2003, s. 68). Sekä rinnastuskonjunktioita että alistuskonjunktioita sisältävät sivulauseet (esimerkki 1) ja esimerkin 2 kaltaiset sivulauseet luokiteltiin alistuskonjunktioita sisältäviksi. Muihin sivulauseisiin laskettiin sivulauseet, joissa lauseraja tai lauseen alisteisuus on tulkinnanvarainen, sekä edelliselle sivulauseelle rinnasteiset sivulauseet, joiden pintarakenteesta alistuskonjunktio on jätetty pois (esimerkki 3). Tällaisissa sivulauseissa ellipsi ei välttämättä rajoitu alistuskonjunktioon (esimerkki 4).

Esimerkki 1. *Toivon että otatte minuun yhteyttä pikaisesti ja että saan uuden kahvinkeittimen tilalle.* (F-704 aikuiset B2)

Esimerkki 2. *Halusin kysyä, että milloin meillä on äidinkielen koe ja mitkä kappaleet pitää lukea?* (F-041 nuoret B1)

Esimerkki 3. *Huomasin heti, että paketti oli huonossa kunnossa ja levyyn kansi oli rikki.* (F-653 aikuiset B1)

Esimerkki 4. *Maa maailma olisi paljon kauniimpi paikka, jos kaikki kävelisivät tai ottaisivat bussia.* (F-689 aikuiset B2)



KUVIO 4. Alistuskonjunktioita sisältävien sivulauseiden, relatiivilauseiden ja epäsuorien kysymyslauseiden sekä muiden sivulauseiden osuus kaikista sivulauseista.

Suurin osa aineiston sivulauseista sisältää alistuskonjunktion: aikuisten aineistossa 60–79 %, S2-nuorten aineistossa 70–76 % ja S1-nuorten aineistossa 67 % (kuvio 4). Relatiivilauseiden ja epäsuorien kysymysten osuus kasvaa aikuisten aineistossa A1-tason 10 prosentista C2-tason 32 prosenttiin, nuorten aineistossa A-tason noin 14 prosentista B1-tason 20 prosenttiin. S1-nuorten vertailuaineistossa relatiivilauseiden ja epäsuorien kysymysten osuus on 24 %. Muiden sivulauseiden osuus vaihtelee S2-aineistossa 8 ja 14 prosentin välillä ja on S1-vertailuaineistossa 9 % kaikista sivulauseista. Tämän aineiston perusteella alistuskonjunktiot voivat toimia sivulauseiden likiarvona alimmilla taitotasoilta, mutta jo niillä esiintyy myös muita kuin alistuskonjunktion sisältäviä sivulauseita.

Kun konjunktioiden valikoimaa tarkastellaan ryhämätasolla, havaitaan, että koko aineistossa yleisimmät konjunktiot *ja*, *koska*, *että*, *mutta*, *kun*, *jos* ja *tai* ovat käytössä jo taitotasolla A1. Muita heti taitotasolla A1 ilmaantuvia konjunktioita ovat *kuin* ja aikuisten aineistossa *eli*, nuorten aineistossa lisäksi *vai*, *eikä*, *vaikka*, *sillä* ja *ellei*. Osa näistä

konjunktioista tosin on taitotasolla A1 harvinaisia ja niistä on vain yksittäisiä esiintymiä. Aikuisten aineistossa eri konjunktioiden määrä kasvaa selvimmin taitotasojen A2 ja B1 välillä, jolloin se yli kaksinkertaistuu 11 konjunktiosta 25 eri konjunktion. Osa tästä muutoksesta selittävät vain kerran esiintyvät konjunktiot, joita on taitotasolla B1 kahdeksan ja taitotasolla B2 seitsemän, kun niitä tasoilla A2 ja C1 on vain kaksi molemmilla. Nuorten aineistossa valikoiman laajeneminen on tasaisempaa ja vain kerran esiintyvien konjunktioiden määrä vaihtelee neljän (taitotasot A1 ja B2) ja viiden (taitotasot A2 ja B1) välillä. Valikoiman laajeneminen ei ryhämätasolla ole selkeästi kumulatiivista, sillä kaikki jollakin taitotasolla esiintyvät konjunktiot eivät välttämättä esiinny seuraavalla tai muilla ylemmillä taitotasolla (taulukot 3 ja 4). Käytettyjen konjunktioiden valikoiman laajuutta tarkasteltaessa on otettava huomioon, ettei konjunktioiden käyttö jakaudu tasaisesti tekstien välillä. Mikään konjunktioista ei esiinny kaikissa teksteissä millään taitotasolla eikä S1-aineistossa, ei edes aineistossa kaikkein yleisin konjunktio *ja*.

TAULUKKO 3. Konjunktiot yleisyysjärjestyksessä ja useammin kuin kerran esiintyvien konjunktioiden suhteelliset frekvenssit (f%) aikuisten aineistossa taitotasoin.

Aikuiset A1	Aikuiset A2	Aikuiset B1	Aikuiset B2	Aikuiset C1	Aikuiset C2
<i>ja</i> (43) <i>koska</i> (20) <i>että</i> (10) <i>mutta</i> (10) <i>kun</i> (8) <i>jos</i> (5) <i>kuin</i> (2) <i>tai</i> (1) <i>eli</i> (1)	<i>ja</i> (42) <i>että</i> (16) <i>koska</i> (15) <i>mutta</i> (10) <i>jos</i> (6) <i>kun</i> (5) <i>tai</i> (3) <i>kuin</i> (2) <i>vai</i> (0) <i>eikä</i> , <i>eli</i>	<i>ja</i> (41) <i>että</i> (16) <i>mutta</i> (11) <i>jos</i> (6) <i>koska</i> (6) <i>kun</i> (6) <i>tai</i> (6) <i>kuin</i> (3) <i>eli</i> , <i>vaikka</i> (1) <i>eikä</i> (1) <i>sekä</i> (1) <i>sekä – että</i> (0) <i>vai</i> (0) <i>niin kuin</i> (0) <i>ettei</i> , <i>vaan</i> (0) <i>elikä</i> , <i>joko – tai</i> , <i>jotta</i> , <i>kunhan</i> , <i>kunnes</i> , <i>muttei</i> , <i>niin – kuin</i> , <i>sillä</i>	<i>ja</i> (40) <i>että</i> (15) <i>mutta</i> (10) <i>jos</i> (7) <i>kun</i> (6) <i>koska</i> (4) <i>tai</i> (4) <i>kuin</i> (3) <i>eikä</i> (2) <i>sekä</i> (1) <i>eli</i> (1) <i>sekä – että</i> , <i>vaikka</i> (1) <i>jotta</i> , <i>vaan</i> (1) <i>ettei</i> (1) <i>sillä</i> , <i>vai</i> (0) <i>elikä</i> , <i>joko – tai</i> , <i>kunnes</i> , <i>mikäli</i> , <i>muttei</i> , <i>niin kuin</i> , <i>niin – kuin</i>	<i>ja</i> (42) <i>että</i> (18) <i>jos</i> (7) <i>mutta</i> (6) <i>kun</i> (5) <i>tai</i> (3) <i>kuin</i> (3) <i>koska</i> (3) <i>eikä</i> , <i>vaikka</i> (2) <i>ettei</i> (2) <i>sekä</i> (1) <i>sillä</i> (1) <i>vai</i> (1) <i>eli</i> , <i>ennen kuin</i> , <i>jotta</i> , <i>vaan</i> (1) <i>sekä – että</i> (1) <i>ellei</i> , <i>niin kuin</i> (0) <i>joko – tai</i> , <i>muttei</i> (0) <i>kunhan</i> , <i>vaikkapa</i>	<i>ja</i> (45) <i>että</i> (17) <i>mutta</i> (6) <i>jos</i> (6) <i>kun</i> (4) <i>koska</i> (3) <i>tai</i> (3) <i>kuin</i> (3) <i>eikä</i> (2) <i>sekä</i> (2) <i>vaan</i> (1) <i>eli</i> (1) <i>ettei</i> , <i>vai</i> (1) <i>sillä</i> (1) <i>jotta</i> (1) <i>ennen kuin</i> , <i>vaikka</i> (1) <i>sekä – että</i> (1) <i>kunhan</i> (0) <i>kun taas</i> , <i>kunnes</i> , <i>mikäli</i>

TAULUKKO 4. Konjunktiot yleisyysjärjestyksessä ja useammin kuin kerran esiintyvien konjunktioiden suhteelliset frekvenssit (f%) S2-nuorten aineistossa taitotasoinen ja S1-nuorten aineistossa.

S2-nuoret A1	S2-nuoret A2	S2-nuoret B1	S2-nuoret B2	S1-nuoret
<i>ja</i> (31) <i>koska</i> (21) <i>jos</i> (12) <i>kun</i> (8) <i>mutta</i> (8) <i>että</i> (7) <i>tai</i> (7) <i>kuin</i> , <i>vai</i> , <i>vaikka</i> (1) <i>eikä</i> (1) <i>eli</i> , <i>ellei</i> , <i>sillä</i>	<i>ja</i> (31) <i>koska</i> (17) <i>että</i> (14) <i>jos</i> (10) <i>mutta</i> (8) <i>kun</i> (7) <i>tai</i> (7) <i>ettei</i> (1) <i>eikä</i> , <i>kuin</i> (1) <i>eli</i> , <i>sillä</i> , <i>vaikka</i> (1) <i>jotta</i> , <i>kunhan</i> , <i>niin kuin</i> , <i>sekä – että</i> , <i>vai</i>	<i>ja</i> (32) <i>että</i> (17) <i>koska</i> (11) <i>jos</i> (10) <i>mutta</i> (10) <i>tai</i> (7) <i>kun</i> (7) <i>eikä</i> (1) <i>eli</i> (1) <i>ettei</i> (1) <i>kuin</i> , <i>vaikka</i> (1) <i>jotta</i> (0) <i>vaan</i> (0) <i>kunhan</i> , <i>vai</i> (0) <i>ennen kuin</i> , <i>joko – tai</i> , <i>sekä</i> , <i>sekä – että</i> , <i>sillä</i>	<i>ja</i> (34) <i>että</i> (19) <i>kun</i> (13) <i>mutta</i> (11) <i>jos</i> , <i>koska</i> , <i>tai</i> (6) <i>eikä</i> (3) <i>eli</i> , <i>kuin</i> , <i>vaikka</i>	<i>ja</i> (31) <i>että</i> (14) <i>jos</i> (12) <i>mutta</i> (10) <i>kun</i> (9) <i>tai</i> (7) <i>koska</i> (6) <i>eikä</i> (3) <i>sillä</i> (2) <i>kuin</i> (1) <i>ettei</i> (1) <i>eli</i> , <i>vaikka</i> (1) <i>vaan</i> (1) <i>sekä</i> (0) <i>vai</i> (0) <i>jotta</i> , <i>niin kuin</i> (0) <i>joko – tai</i> (0) <i>ellei</i> , <i>ennen kuin</i> , <i>kunhan</i> (0) <i>mikäli</i> (0) <i>etteikö</i> , <i>joko – taikka</i> , <i>jottei</i> , <i>kunnes</i> , <i>niin – kuin</i> , <i>sekä – että</i>

S2-aineiston kaikilla taitotasoilla ja S1-nuorten vertailuaineistossa yleisin konjunktio on rinnastuskonjunktio *ja*. Sen esiintymät kattavat aikuisten aineistossa 40–45 %, S2-nuorten aineistossa 31–34 % ja S1-nuorten vertailuaineistossa 31 % kaikista konjunktioiden esiintymistä. Aineiston toiseksi yleisin konjunktio on aikuisten aineiston taitotasolla A1 ja nuorten aineiston taitotasolla A1 ja A2 alistuskonjunktio *koska*, jonka osuus on taitotasolla A1 noin 20 % kaikista konjunktioista. Seuraavilla taitotasoilla toiseksi yleisimmäksi

konjunktioksi nousee *että*, joka taitotason A1 jälkeen kattaa aikuisten aineistossa 15–18 % ja nuorten aineistossa 14–19 % kaikkien konjunktioiden esiintymistä. Se on toiseksi yleisin konjunktio myös S1-vertailuaineistossa, jossa sen osuus kaikista konjunktioista on 14 %. Toiseksi yleisin rinnastuskonjunktio on kaikilla taitotasoilla ja myös S1-aineistossa *mutta*, joka kattaa 6–11 % kaikista konjunktioiden esiintymistä ja on aikuisten aineiston eri taitotasoilla kolmanneksi tai neljänneksi yleisin konjunktio, nuorten aineistossa nel-

jänneksi tai viidenneksi yleisin.

Suurin osa aineiston konjunktiosta on yksiosaisia. Pari- ja liittokonjunktiot ovat harvinaisia niin S2- kuin S1-aineistossa. Yleisin on *sekä – että*, joka ilmaantuu aikuisten teksteihin taitotasolla B1 mutta esiintyy S2-nuorten teksteissä vain kahdesti ja S1-vertailuaineistossa vain kerran. Aikuisten aineistossa yleisin liittokonjunktio on vain C-tasolla esiintyvä *ennen kuin*, S1-aineistossa puolestaan *niin kuin*, joka esiintyy joitakin kertoja aikuisten aineiston kolmella ylimmällä taitotasolla. Kumpikin esiintyy S2-nuorten aineistossa vain kerran. Myös osa yksiosaisista konjunktiosta on aineistossa harvinaisia. Tällainen on esimerkiksi *kunnes*, joka esiintyy S2-aineistossa vain aikuisten teksteissä muutama kerran ja S1-vertailuaineistossa kerran. S1-vertailuaineistossa on kolme vain kerran esiintyvää konjunktiota, jotka eivät esiinny S2-aineistossa lainkaan, S2-aikuisten aineistosta puolestaan löytyy B- ja C-tasolta neljä konjunktiota, joita ei S2- ja S1-nuorten aineistossa ole käytetty lainkaan. Näiden esiintymistaajuus jää kuitenkin pieneksi, ja osa esiintyy koko aineistossa vain kerran. Aikuisten ja nuorten aineisto eroavat toisistaan myös konjunktioiden puhekielisten kirjoitusasujen osalta. Puhekielisiä kirjoitusasuja esiintyy lähes yksinomaan nuorten aineistossa⁵, ja ne ovat yleisempiä S1-nuorten kuin S2-nuorten teksteissä.

Aineistossa yleisimmät konjunktiot ilmaantuvat S2-oppijoiden teksteihin heti taitotasolla A1, tosin joidenkin konjunktioiden kohdalla kyse on yksittäisistä esiintymistä. Kerran tai kaksi esiintyviä konjunktiota ei voi pitää taitotasolle tyypillisinä, mutta ne

havainnollistavat, ettei harvinaisen tai ylemmille taitotasolle tyypillisen konjunktion käytön ja kielitaidon tason välillä välttämättä ole yhteyttä. Tähän samaan viittaa se, ettei käytävissä olevien konjunktioiden varannon laajeneminen näy ryhmätasolla systemaattisena teksteissä käytettyjen konjunktioiden valikoiman laajentumisena eivätkä kaikki jollakin taitotasolla käytetyt konjunktiot välttämättä esiinny seuraavalla tai muilla ylemmillä taitotasolla. Esimerkiksi aikuisten aineistossa harvinainen ja vain taitotasolla C1 muutama kerran esiintyvä *ellei* esiintyy nuorten aineistossa vain kerran, taitotasolla A1.

5 POHDINTA

Tämän tutkimuksen aineistossa S2-kirjoittajat käyttävät sekä rinnastus- että alistuskonjunktiota kaikilla taitotasolla. Konjunktioiden ilmaantuminen S2-teksteihin tukee aiempia tuloksia konjunktioiden käytöstä jo kielitaidon alkuvaiheessa (Määttä, 2012; oppijansaksassa Vyatkinä, 2012, 2013). Tulokset eivät tältä osin tue EVK:n taitotasokuvauksista välittyvää näkemystä, että syntaktisten kytkösten merkitseminen konjunktiolla alkaisi yleistyä vasta taitotasolla A2, sillä tämän tutkimuksen aineistossa erot konjunktioiden normalisoidussa esiintymistaajuudessa eivät näytä olevan tilastollisesti merkitseviä taitotasojen A1 ja A2 välillä. Toisaalta tulokset myös osittain tukevat EVK:n taitotasokuvauksia, sillä täysin konjunktiottomien tekstien osuus kaikista teksteistä on suurin juuri taitotasolla A1 ja taitotasolla A2 niiden osuus on aikuisten aineistossa puolet ja nuorten aineistossa kolme neljäsosaa pienempi kuin taitotasolla A1. Jo taitotasolla A1 yli 80 % teksteistä kuitenkin sisältää vähintään yhden konjunktion.

Konjunktiota käytetään normalisoitu-

⁵ Puhekielistä kirjoitusasuista vain *et* esiintyy yhden kerran aikuisten aineistossa (taitotasolla B1).

jen esiintymistaajuuksien perusteella eniten suhteessa kokonaissanamäärään aikuisten aineistossa taitotasolla B1 ja B2, nuorten aineistossa taitotasolla A2 ja B1. Normalisoiduissa esiintymistaajuuksissa on tilastollisesti merkitseviä eroja taitotasojen välillä aikuisten aineistossa lähinnä tasojen A ja B välillä, nuorten aineistossa taitotasojen A1 ja B1 välillä. Vaikka tasojen A ja C välillä ei ole tilastollisesti merkitseviä eroja konjunktioiden normalisoiduissa esiintymistaajuuksissa, tulosta ei voi tulkita niin, että konjunktioiden käyttö tai syntaktinen kompleksisuus olisivat samanlaista aineiston alimmilla ja ylimmillä taitotasolla. Valittu tarkastelutapa kuvaa konjunktioiden suhteellista yleisyyttä eikä ota huomioon laadullisia eroja konjunktioiden käytössä tai syntaktisessa kompleksisuudessa. Lisäksi lauseiden keskimääräinen sanamäärä kasvaa aineistossa taitotason noustessa (Myläri, 2020a), millä voi olla vaikutusta normalisoiuihin esiintymistaajuuksiin.

Tilastollisissa analyysissa näkyviin taitotasojen väliin eroihin vaikuttaa myös taitotasojen sisäinen varianssi, joka kertoo siitä, että samalle taitotasolle sijoitetut tekstit eroavat toisistaan siinä, miten paljon niissä on käytetty konjunktioita. Sekä aikuisten että nuorten aineistossa varianssi on suurinta taitotasolla A1, minkä jälkeen se vähenee. Taitotasojen sisäistä varianssia ja taitotasojen välisten tilastollisesti merkitsevien erojen vähyyttä voi osin selittää se, että aineiston eri taitotasolla eri tehtävätyyppien osuus teksteistä vaihtelee, sillä tehtävätyyppien tai tekstilajien välillä voi olla eroja syntaktisessa kompleksisuudessa (esim. Michel, 2017; Pallotti, 2009). Lisäksi Cefling-aineistossa on havaittu tehtävätyyppien välisiä eroja joidenkin kielenpiirteiden kehityksessä taitotasolta toiselle (Kajander,

2013; Reiman, 2014; Seilonen, 2013). Taitotasojen sisäiseen varianssiin ja taitotasojen välisten tilastollisesti merkitsevien erojen vähyyteen voivat vaikuttaa myös eri taustatekijöistä johtuvat yksilölliset erot, joiden vaikutus konjunktioiden käyttöön ja sen avulla tarkasteltavaan syntaktiseen kompleksisuuteen on rajattu tämän tutkimuksen ulkopuolelle.

Tutkimuksen aineistossa rinnastuskonjunktiot ovat alistuskonjunktioita yleisempiä kaikilla muilla taitotasolla kuin nuorten aineiston taitotasolla A1 ja A2. Alistuskonjunktioiden osuus kaikista konjunktioista ei ylemmillä taitotasolla ole suurempi kuin alemmilla. Tulos poikkeaa Vyatkinan (2012) tuloksista, joiden mukaan rinnastuskonjunktioiden ja alistuskonjunktioiden esiintymistaajuuden välillä vallitsi ryhmätasolla negatiivinen korrelaatio niin, että alistuskonjunktioiden määrän lisääntyessä rinnastuskonjunktioiden määrä väheni. Rinnastuskonjunktioiden käytön tarkempi tarkastelu osoitti, että aikuisten aineiston C-tasolla yli puolet rinnastuskonjunktioiden esiintymistä sijaitsee muualla kuin lauserajalla. Tämä tulos tukee näkemystä, että kielitaidon ylemmillä taitotasolla syntaktinen kompleksisuus voi olla erilaista kuin perus- tai keskitasolla ja että ylimmillä taitotasolla kompleksisuuden kehitys voi näkyä lauseiden yhdistämisen sijasta lausekkeiden yhdistämisessä (esim. Norris & Ortega, 2009).

Sivulauseiden tarkempi tarkastelu puolestaan osoitti, että alistuskonjunktion sisältävien sivulauseiden osuus kaikista sivulauseista on ylemmillä taitotasolla pienempi kuin alemmilla taitotasolla, mikä viittaa siihen, että kirjoittajien käyttämät alisteiset rakenteet monipuolistuvat kielitaidon ylemmillä taitotasolla. Jos syntaktista kompleksisuutta

tarkastellaan kielenpiirteiden monipuolisuutena (esim. Ortega, 2003), tulosta voidaan tulkita myös syntaktisen kompleksisuuden lisääntymisenä ryhmätasolla. Alistuskonjunktion sisältävien sivulauseiden vähenevä osuus kaikista sivulauseista on myös linjassaan kanssa, että relatiivilauseiden on todettu lisääntyvän kehityksellisesti niin toisessa kielessä (esim. Ortega, 2009) kuin ensikielessä (esim. Pajunen & Vainio, 2021b). Relatiivilauseiden ja epäsuorien kysymysten osuuden kasvu voi vaikuttaa myös siihen, että alistuskonjunktioiden normalisoitu esiintymistajuus on C-tasolla pienempi kuin B-tasolla, sillä sivulauseiden osuus kaikista lauseista ei tässä aineistossa kasva tilastollisesti merkittävästi näiden tasojen välillä (Mylläri, 2020a).

Yleisimpien konjunktioiden ilmaantuminen ja konjunktioiden suhteellisen suuri yhteenlaskettu esiintymistajuus jo taitotasolla A1 viittaavat siihen, että aikuiset ja yläkouluikäiset S2-kirjoittajat hallitsevat erilaisia syntaktisia yhdistämiskeinoja jo kielitaidon alimmilla taitotasolla. Tämä ei sinänsä ole yllättävää, sillä sekä rinnasteisten että alisteisten suhteiden ilmaisemisen voi olettaa olevan aikuisille ja myös yläkouluikäisille suomenoppijoille tuttua aiemmin opituista kielistä, jolloin kyse on enemmän olemassa olevan taidon siirtämisestä uuteen kieleen kuin uusista syntaktisten kytkösten luomisen keinoista. Tulokset tukevat näkemyksiä siitä, että toisen kielen taidon kehittymistä ei välttämättä ole perusteltua tarkastella sivulauseiden määrään tai T-yksikön pituuteen pohjautuvilla mittareilla, jotka on kehitetty ensikieltä oppivien lasten kielitaidon kehittymisen seuraamiseen (esim. Bardovi-Harlig, 1992). Samalla ne tukevat näkemyksiä (esim. Martin ym., 2010), että tämä pätee oppijansuomen kompleksisuuden mittaamiseen.

suuden mittaamiseen.

Viestintä tai kielitaito eivät itsessään edellytä kompleksisuutta tai monimutkaisten rakenteiden käyttöä (esim. De Clercq & Housen, 2017). Tämä näkyy myös konjunktioiden käytössä. Vaikka konjunktioiden monipuolinen käyttö tai jonkin alemmilla taitotasolla harvinaisen konjunktion esiintyminen tekstissä voi olla merkki edistyneestä kielitaidosta, minkään yksittäisen konjunktion puuttuminen tai konjunktioiden vähäinen määrä ei välttämättä tarkoita heikkoa kielitaitoa. Esimerkiksi seuraava epämuodollinen viesti sijoittuu taitotasolle C1, vaikka se ei ole syntaktisesti kompleksinen, jos kompleksisuutta tarkastellaan sivulauseiden tai konjunktioiden avulla:

Hyvät asukkaat! Kevät on tullut!

Pidetään pihan

SIIVOUSTALKOOT

lauantaina 25.4.2005 klo 11-14

Päälle mukavat vaatteet. Mukaan reipas ja iloinen mieli. Tarjolla limua ja makkaraa!

(F-1026 aikuiset C1)

Yksinkertaisempi syntaksi ei välttämättä tarkoita huonoa kielitaitoa, vaan voi olla merkki siitä, että kirjoittaja osaa sopeuttaa ilmaisuaan tilanteeseen (Kuiken & Vedder, 2019).

6 LOPUKSI

Pelkkä konjunktioiden tarkastelu antaa rajatun kuvan tekstien sisältämisestä syntaktisista kytköksistä, sillä se jättää tarkastelun ulkopuolelle esimerkiksi asyndeettiset rinnastukset ja muut kuin konjunktion sisältävät sivulauseet. Konjunktioiden esiintymillä voi kuitenkin olla merkitystä oppijankielen

syntaktisen kompleksisuuden mittaamisen näkökulmasta, sillä ne ovat tekstien pinta-rakenteessa näkyviä merkkejä kirjoittajan käyttämistä syntaktisista kytköksistä ja niiden poimiminen aineistosta on helpommin automatisoitavissa kuin esimerkiksi lauseiden tai T-yksiköiden laskeminen. Niiden avulla voidaan myös tehdä havaintoja syntaktisesta kompleksisuudesta kielellisten keinojen monipuolisuutena, kun tarkastellaan eri taitotasoille sijoittuvissa teksteissä esiintyvien konjunktioiden valikoimaa tai alistuskonjunktion sisältävien lauseiden osuutta kaikista sivulauseista. Konjunktioiden esiintymien tarkastelu voi näin täydentää muilla mittareilla saatua tietoa oppijankielen syntaktisesta kompleksisuudesta.

Konjunktioiden käyttöä ja sen kautta syntyvää kuvaa syntaktisesta kompleksisuudesta on tässä tutkimuksessa tarkasteltu suhteessa kielitaidon taitotasoon. Tavoitteena oli kartoittaa ilmiötä ryhmätasolla, ja tutkimuksen keskiössä olivat eri taitotasoilla mahdolliset

ja tyypilliset piirteet sekä konjunktioiden käytössä näkyvät erot taitotasojen välillä. Tällainen tarkastelu ei tavoita kielitaidon tai syntaktisen kompleksisuuden yksilöllistä kehitystä, joka voi olla erilaista kuin ryhmätason kehitys (esim. Larsen-Freeman, 2006). Pitkittäisaineiston avulla olisi mahdollista tarkastella myös yksilötason kehitystä ja verrata sitä nyt havaittuihin ryhmätason kehityskulkuihin.

Tässä artikkelissa konjunktioita ja niiden avulla luotuja rinnastus- ja alistussuhteita on tarkasteltu syntaksin alaan kuuluvana ilmiönä lause- ja virketasolla määrällisin menetelmin. Kokonaisten tekstien tarkastelu voisi tarjota uudenlaisen näkökulman syntaktiseen kompleksisuuteen (esim. Rimmer, 2009). Laadullisempi tutkimusote voisi myös tehdä näkyväksi sellaisia konjunktioiden käytön puolia, joita tämän tutkimuksen määrällinen tarkastelu ei tavoita, ja näin täydentää nyt saatua tietoa oppijansuomen syntaktisesta kompleksisuudesta.

LÄHTEET

- Alanen, R., Huhta, A. & Tarnanen, M. (2010). Designing and assessing L2 writing tasks across CEFR proficiency levels. Teoksessa I. Bartning, M. Martin & I. Vedder (toim.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*, (s. 21–56). Edisegno srl.
- Alisaari, J. (2016). *Songs and poems in the second language classroom. The hidden potential of singing for developing writing fluency*. Väitöskirja. Turun yliopisto.

- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26(2), 390–395.
- Benevento, C. & Storch, N. (2011). Investigating writing development in secondary school learners of French. *Assessing Writing*, 16(2), 97–110.
- Bernardini, P. & Granfeldt, J. (2019). On cross-linguistic variation and measures of linguistic complexity in learner texts: Italian, French and English. *International Journal of Applied Linguistics, Special issue*, 1–22.

- Biber, D., Gray, B. & Poonpon, K. (2011) Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5–35.
- Bulté, B. & Housen, A. (2012). Defining and operationalising L2 complexity. Teoksessa A. Housen, I. Vedder & F. Kuiken, (toim.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, (s. 21–46). John Benjamins Publishing Company.
- Cristofaro, S. (2014). Is there really a syntactic category of subordination? Teoksessa L. Visapää, J. Kalliokoski & H. Sorva (toim.), *Contexts of subordination. Cognitive, typological and discourse perspectives*, (s. 1–17). John Benjamins Publishing Company.
- Crossley, S. & McNamara, D. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66–79.
- De Clercq, B. & Housen, A. (2017). A cross-linguistic perspective on syntactic complexity in L2 development: Syntactic elaboration and diversity. *The Modern Language Journal*, 101(2), 315–334.
- Ellis, R. & Barkhuizen, G. (2005). *Analysing learner language*. Oxford University Press.
- EVK 2003 = Eurooppalainen viitekehys (2003). *Kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys*. WSOY.
- Grant, L. & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9(2), 123–145.
- Gyllstad, H., Granfeldt, J., Bernardini, P. & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR: The case of syntactic complexity in written L2 English, L3 French and L4 Italian. Teoksessa L. Roberts, I. Vedder & J.H. Hulstijn (toim.), *Eurosla Yearbook 14*, (s. 1–30). John Benjamins Publishing Company.
- Hakulinen, A., Karlsson, F. & Vilkkuna, M. (1980). *Suomen tekstilauseiden piirteitä: kvantitatiivinen tutkimus*. Helsingin yliopisto.
- Herlin, I., Visapää, L. & Kalliokoski, J. (2014). Introduction. Teoksessa L. Visapää, J. Kalliokoski & H. Sorva (toim.), *Contexts of subordination. Cognitive, typological and discourse perspectives*, (s. 1–17). John Benjamins Publishing Company.
- Honko, M. (2013). *Alakouluikäisen leksikaalinen tieto ja taito: Toisen sukupolven suomi ja S1-verrokkit*. Väitöskirja. Tampereen yliopisto.
- Housen, A., De Clercq, B., Kuiken, F. & Vedder, I. (2019). Multiple approaches to complexity in second language research. *Second Language Research*, 35(1), 3–21.
- Housen, A., Vedder, I. & Kuiken, F. (toim.) (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. John Benjamins Publishing Company.
- Huhta, A., Alanen, R., Tärnänen, M., Martin, M. & Hirvelä, T. (2014) Assessing learners' writing skills in a SLA study: Validating the rating process across tasks, scales and languages. *Language Testing*, 31(3), 307–328.
- Jagaiah, T., Olinghouse, N. G. & Kearns, D. M. (2020). Syntactic complexity measures: variation by genre, grade-level, students' writing abilities, and writing quality. *Reading and Writing*, 33, 2577–2638.
- Jantunen, J. & Pirkola, S. (2015). Oppijansuomen sähköiset tutkimusaineistot. Nykytilanne. *Virittäjä*, 119(1), 88–103.
- Kajander, M. (2013). *Suomen eksistentiaalilause toisen kielen oppimisen polulla*. Väitöskirja. Jyväskylän yliopisto.

- Khushik, G. & Huhta, A. (2020). Investigating syntactic complexity in EFL learners' writing across Common European framework of reference levels A1, A2, and B1. *Applied Linguistics*, 41(4), 506–532.
- Kuiken, F. & Vedder, I. (2019). Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. *International Journal of Applied Linguistics*, 29, 192–210.
- Lambert, C. & Kormos, J. (2014). Complexity, accuracy and fluency in task-based research: Toward more developmentally-based measures of second language acquisition. *Applied Linguistics*, 35, 607–614.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590–619.
- Martin, M. (2013). The complex simple – a problematic adjective in the CEFR writing scales. *Nordand*, 8(2), 63–85
- Martin, M., Mustonen, S., Reiman, N. & Seilonen, M. (2010). On becoming an independent user. Teoksessa I. Bartning, M. Martin & I. Vedder (toim.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*, (s. 57–80). Edisegno srl.
- Michel, M. (2013). The use of conjunctions in cognitively simple versus complex oral L2 tasks. *The Modern Language Journal*, 97(1), 178–195.
- Michel, M. (2017). Complexity, accuracy, and fluency in L2 production. Teoksessa S. Loewen & M. Sato (toim.), *The Routledge handbook of instructed second language acquisition*, (s. 50–68). Routledge.
- Mustonen, S., (2015). *Käytössä kehittyvä kieli. Paikat ja tilat suomi toisena kielenä -oppijoiden teksteissä*. Väitöskirja. Jyväskylän yliopisto.
- Mylläri, T. (2020a). Measuring syntactic complexity in learner Finnish. *Apples – Journal of Applied Language Studies* 14(2), 67–92.
- Mylläri, T. (2020b). Words, clauses, sentences, and T-units in learner language: Precise and objective units of measure? *Journal of the European Second Language Association*, 4(1), 13–23.
- Määttä, T. (2012). Oppikirjan sanaston vaikutuksesta ruotsinkielisten alkeistason suomenoppijoiden kirjallisiin tuotoksiin. *Läbivertailuja*, 22, 188–218.
- Norris, J. & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: the case of complexity. *Applied Linguistics*, 30(4), 555–578.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492–518.
- Ortega, L. (2009). *Understanding second language acquisition*. Hodder Education.
- Pajunen, A. & Vainio, S. (2021a). Kielen rakenteen hallinta peruskoululaisilla ja nuorilla aikuisilla. Teoksessa A. Pajunen & M. Honko (toim.), *Suomen kielen hallinta ja sen kehitys: peruskoululaiset ja nuoret aikuiset*. Suomalaisen Kirjallisuuden Seuran toimituksia, 1472, (s. 367–420). Suomalaisen Kirjallisuuden Seura.
- Pajunen, A. & Vainio, S. (2021b). Syntaktista rakennetta tiivistävät ja kompleksistavat tekijät peruskoululaisilla ja nuorilla aikuisilla. Teoksessa A. Pajunen & M. Honko (toim.), *Suomen kielen hallinta ja sen kehitys: peruskoululaiset ja nuoret aikuiset*. Suomalaisen Kirjallisuuden Seuran toimituksia, 1472, (s. 421–479). Suomalaisen Kirjallisuuden Seura.
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590–601.

- Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research*, 31(1), 117–134.
- Reiman, N. (2011). Two faces of complexity: Structural measures and diversity of constructions. *Nordand*, 6(2), 9–23.
- Reiman, N. (2014). Yläkoulun S2-oppilaiden transitiivi-ilmausten käyttö Eurooppalaisen viitekehysten taitotasolla. *Lähi-vördlusi. Lähi-vertailuja*, 24, 183–220.
- Rimmer, W. (2009). Can what counts in complexity be counted? *University of Reading: Language studies working papers*, 1, 25–34.
- Saukkonen, P., Haipus, M., Niemikorpi, A. & Sulka, H. (1979). *Suomen kielen taajuusanasto*. WSOY.
- Seilonen, M. (2013). *Epäsuora henkilöön viittaminen oppijansuomessa*. Väitöskirja. Jyväskylän yliopisto.
- Spoelman, M. & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31(4), 532–553.
- Taguchi, N., Crawford, W. & Wetzell, D. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly*, 47(2), 420–430.
- Tilma, C. (2014). *The dynamics of foreign versus second language development in Finnish writing*. Väitöskirja. Jyväskylän yliopisto.
- Vilkuna, M. (2003). *Suomen lauseopin perusteet*. Edita.
- VISK = Hakulinen, A., Vilkuna, M., Korhonen, R., Koivisto, V., Heinonen, T. & Alho, I. (2004). *Iso suomen kielioppi*. Suomalaisen Kirjallisuuden Seura. Verkkooversio, viitattu 15.2.2021. Saatavissa: <http://scripta.kotus.fi/visk>.
- Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *The Modern Language Journal*, 96(4), 576–598.
- Vyatkina, N. (2013). Specific syntactic complexity: Developmental profiling of individuals based on an annotated learner corpus. *The Modern Language Journal* 97 (S1), 11–30.
- Wolfe-Quintero, K., Inagaki, S. & Kim, H. (1998). *Second language development in writing: measures of fluency, accuracy, and complexity*. University of Hawai'i, Second Language Teaching and Curriculum Center.

**CONJUNCTIONS AND SYNTACTIC COMPLEXITY:
DEVELOPMENT OF THE USE OF CONJUNCTIONS IN WRITTEN LEARNER FINNISH**

Taina Mylläri, Department of Language and Communication Studies, University of Jyväskylä

Learner language development is often analysed by measuring complexity, accuracy and fluency. Complexity can be defined as the range and sophistication of the structures available to the learner, yet syntactic complexity is typically analysed using quantitative measures tapping subordination. This article focuses on the development of syntactic complexity in written learner Finnish across the CEFR proficiency levels by exploring changes in the use of conjunctions. The data are drawn from the corpus of the Jyväskylä University CEFLING project. The development of syntactic complexity in terms of coordination and subordination is explored by tracking changes in the use of coordinate and subordinate conjunctions. Syntactic complexity as the range of forms used by the learner is explored by investigating the emergence and frequency of individual conjunctions. The results question the role of subordination in the development of syntactic complexity in learner Finnish, as they show that both coordinate and subordinate conjunctions are already used by learners on the lowest CEFR levels. The results also indicate that development in the use of conjunctions may be different for adult learners and adolescent learners.

Keywords: conjunctions, learner Finnish, syntactic complexity

LIITE 1

TAULUKKO 1A. Konjunktioiden esiintymät (n) ja esiintymistaajuus tuhatta sanaa kohden (n/1000) aikuisten aineistossa.

	YKI A1		YKI A2		YKI B1		YKI B2		YKI C1		YKI C2		Yhteensä	
	n	n/1000	n	n/1000	n	n/1000	n	n/1000	n	n/1000	n	n/1000	n	n/1000
<i>ja</i>	172	37	186	35	421	43	330	40	402	37	380	38	1 891	39
<i>mutta</i>	39	8	42	8	113	11	81	10	60	6	52	5	387	8
<i>tai</i>	5	1	14	3	60	6	36	4	28	3	25	3	168	3
<i>eli</i>	3	1	1	0	10	1	9	1	6	1	11	1	40	1
<i>vai</i>			2	0	4	0	3	0	7	1	10	1	26	1
<i>eikä</i>			1	0	8	1	14	2	18	2	18	2	59	1
<i>sekä</i>					6	1	10	1	11	1	15	2	42	1
<i>sekä – että</i>					5	1	7	1	5	0	5	1	22	0
<i>vaan</i>					2	0	6	1	6	1	12	1	26	1
<i>sillä</i>					1	0	3	0	8	1	8	1	20	0
<i>muttei</i>					1	0	1	0	2	0			4	0
<i>joko – tai</i>					1	0	1	0	2	0			4	0
<i>elikä</i>					1	0	1	0					2	0
<i>niin – kuin</i>					1	0	1	0					2	0
Rinnastus- konjunktiot yhteensä	219	48	246	46	634	64	503	60	555	52	536	54	2 693	55
<i>koska</i>	79	17	64	12	62	6	37	4	25	2	27	3	294	6
<i>että</i>	41	9	71	13	162	16	129	15	174	16	145	15	722	16
<i>kun</i>	32	7	24	5	61	6	54	6	50	5	32	3	253	5
<i>jos</i>	21	5	25	5	63	6	62	7	66	6	48	5	285	6
<i>kuin</i>	8	2	8	2	26	3	28	3	27	3	21	2	118	2
<i>vaikka</i>					10	1	7	1	18	2	6	1	41	1
<i>niin kuin</i>					3	0	1	0	3	0			7	0
<i>ettei</i>					2	0	5	1	17	2	10	1	34	1
<i>jotta</i>					1	0	6	1	6	1	7	1	20	0
<i>kunnes</i>					1	0	1	0			1	0	3	0
<i>kunhan</i>					1	0			1	0	2	0	4	0
<i>mikäli</i>							1	0			1	0	2	0
<i>ennen kuin</i>									6	1	6	1	12	0
<i>ellei</i>									3	0			3	0
<i>vaikkapa</i>									1	0			1	0
<i>kun taas</i>											1	0	1	0
Alistus- konjunktiot yhteensä	181	39	192	36	392	40	331	40	397	37	307	31	1 800	37
Kaikki konjunktiot yhteensä	400	87	438	83	1 026	104	834	100	952	89	843	84	4 493	92

TAULUKKO 1B. Konjunktoiden esiintymät (n) ja esiintymistaajuus tuhatta sanaa kohden (n/1000) nuorten aineistossa.

	S2-nuoret A1		S2-nuoret A2		S2-nuoret B1		S2-nuoret B2		Yhteensä		S1-nuoret	
	n	n/1000	n	n/1000	n	n/1000	n	n/1000	n	n/1000	n	n/1000
<i>ja</i>	71	31	203	34	262	36	40	35	576	35	613	31
<i>mutta</i>	17	7	54	9	83	12	13	11	167	10	199	10
<i>tai</i>	15	6	46	8	61	8	7	6	129	8	131	7
<i>vai</i>	3	1	1	0	2	0			6	0	9	0
<i>eikä</i>	2	1	5	1	11	2	3	3	21	1	53	3
<i>eli</i>	1	0	4	1	9	1	1	1	15	1	18	1
<i>sillä</i>	1	0	4	1	1	0			6	0	38	2
<i>sekä – että</i>			1	0	1	0			2	0	1	0
<i>vaan</i>					3	0			3	0	12	1
<i>sekä</i>					1	0			1	0	10	1
<i>joko – tai</i>					1	0			1	0	5	0
<i>niin – kuin</i>											1	0
<i>joko – taikka</i>											1	0
Rinnastus- konjunktiot yhteensä	110	48	318	53	435	61	64	56	927	56	1 091	55
<i>koska</i>	48	21	114	19	89	12	7	6	258	16	122	6
<i>jos</i>	27	12	68	11	85	12	7	6	187	11	231	12
<i>kun</i>	18	8	47	8	58	8	15	13	138	8	186	9
<i>että</i>	16	7	95	16	137	19	23	20	271	16	287	14
<i>kuin</i>	3	1	5	1	5	1	1	1	14	1	25	1
<i>vaikka</i>	3	1	4	1	5	1	1	1	13	1	18	1
<i>ellei</i>	1	0							1	0	3	0
<i>ettei</i>			6	1	8	1			14	1	23	1
<i>jotta</i>			1	0	4	1			5	0	6	0
<i>kunhan</i>			1	0	2	0			3	0	3	0
<i>niin kuin</i>			1	0					1	0	6	0
<i>ennen kuin</i>					1	0			1	0	3	0
<i>mikäli</i>											2	0
<i>etteikö</i>											1	0
<i>jottei</i>											1	0
<i>kunnes</i>											1	0
Alistus- konjunktiot yhteensä	116	50	342	57	394	55	54	47	906	55	918	46
Kaikki konjunktiot yhteensä	226	98	660	111	829	115	118	104	1 833	110	2 009	101

