

# **Ilmisubjektin esiintyvyys kirjoitetussa oppijansuomessa**

**Maisterintutkielma**

**Mikael Törö**

**Suomen kieli**

**Kielten laitos, Jyväskylän yliopisto**

**2017**



# JYVÄSKYLÄN YLIOPISTO

<b>Tiedekunta – Faculty</b> Humanistis-yhteiskuntatieteellinen tiedekunta	<b>Laitos – Department</b> Kielten laitos
<b>Tekijä – Author</b> Mikael Törö	
<b>Työn nimi – Title</b> Ilmisubjektin esiintyvyys kirjoitetussa oppijansuomessa	
<b>Oppiaine – Subject</b> Suomen kieli	<b>Työn laji – Level</b> Pro gradu -tutkielma
<b>Aika – Month and year</b> Maaliskuu 2017	<b>Sivumäärä – Number of pages</b> 58 + liitteet
<b>Tiivistelmä – Abstract</b> <p>Tämä maisterintutkielma selvittää ilmisubjektin esiintyvyyttä kirjoitetussa oppijansuomessa. Tutkimus on aihealueeltaan suomi toisena kielenä -tutkimusta. Ilmisubjektin käyttöä kirjoitetussa oppijansuomessa ei ole vastaavanlaaisesti tutkittu. Tutkimus pyrkii selvittämään, miten ilmisubjekti esiintyy yksikön 1. persoonassa Eurooppalaisen viitekehyksen (EVK) eri taitotasolla. Lisäksi tutkitaan sitä, esiintyykö ilmisubjekti poikkeavasti eri semanttisissa verbiluokissa. Myös selvitetään, vaikuttaako verbin esiintymisfrekvenssi siihen, kuinka usein verbi on ilmisubjektillinen.</p> <p>Ilmisubjektilla tarkoitetaan persoonamuotoisen verbin yhteyteen ilmipantua pronominisubjektia. Tutkimuksessa on eroteltu aineiston yksikön 1. persoonan verbit seuraaviin semanttisiin luokkiin: tila, teko ja muutos, liike, puhe, psykologinen tila ja prosessi, perpektio, abstraktit ja kieltö.</p> <p>Toinen kieli on kielenoppijan ensikielestä poikkeava kieli, jota hän oppii opittavan kielen käyttöympäristössä. Oppijankieli on kielenoppijan toisen kielen tuotosta, jossa on useita kehitysvaiheita. Oppijankielen tuotosta voidaan mitata EVK:n taitotasolla.</p> <p>Tutkimusaineistona toimii Jyväskylän yliopiston Topling-hankkeen suomenkielinen aineisto. Aineistoa on kerätty suomalaisissa ala- ja yläkouluissa sekä lukioissa suomea toisena kielenä opiskelevilta oppilailta. Aineiston tekstit on arvosteltu EVK:n taitotasolle A1–C2.</p> <p>Tutkimusta on tehty aineistopohjaisena korpustutkimuksena. Tutkimuksessa ilmenee, että ilmisubjektin esiintyvyys vähenee tasaisesti, kun taitotaso nousee. Eri semanttisten verbiluokkien välillä on myös eroavaisuuksia: ilmisubjektia esiintyy eniten teko- ja muutosverbeissä sekä liikeverbeissä, vähiten tilaverbeissä ja abstrakteissa verbeissä. Verbin esiintymisfrekvenssi vaikuttaa osassa semanttisia verbiluokkia niin, että frekventeimmät verbit ovat vähemmän ilmisubjektillisiä kuin vähemmän esiintyvät. Kaikki frekventeimmät verbit eivät kuitenkaan ole selkeästi heikosti ilmisubjektillisiä.</p> <p>Lisätutkimusta on mahdollista tehdä esimerkiksi siitä, miten eri tekstilajeissa ilmisubjektia käytetään. Myös S2-oppimateriaalien vaikutusta ilmisubjektin käyttöön olisi hedelmällistä tarkastella. Lisäksi mielenkiintoista olisi myös selvittää, poikkeako ilmisubjektin käyttö natiivien ja suomenoppijoiden kirjoituksissa toisistaan eri kehitysvaiheissa. Erityisen kiinnostavaa olisi tietää, käyttävätkö natiivit 2.–4. luokkalaiset juurikaan vähemmän ilmisubjektia kuin suomenoppijat.</p>	
<b>Asiasanat – Keywords</b> suomi toisena kielenä, persoonapronomini, subjekti, semantiikka, korpustutkimus, Topling	
<b>Säilytyspaikka – Depository</b> Jyväskylän yliopisto, kielen laitos (Fennicum)	
<b>Muita tietoja – Additional information</b>	



# SISÄLLYS

1. JOHDANTO .....	1
2. TOISEN KIELEN OPPIMINEN .....	5
2.1 Toisen kielen oppiminen.....	5
2.2 Oppijankieli .....	7
2.3 Suomi toisena kielenä .....	9
2.4 Kielen kehitys ja Eurooppalainen viitekehys.....	10
3. ILMISUBJEKTI JA VERBIEN SEMANTTINEN LUOKITTELU .....	13
3.1 Persoonapronominit ja persoonamuotoiset verbit.....	13
3.2 Ilmisubjekti .....	14
3.2.1 Ilmisubjekti ja subjektityypit .....	14
3.2.2 Subjektin ilmaisu maailman kielissä.....	16
3.3 Verbiin semanttinen luokittelu .....	17
3.3.1 Konkreettiset verbit.....	19
3.3.2 Mentaaliset verbit .....	22
3.3.3 Abstraktit verbit .....	23
3.3.4 Kieltoverbit .....	24
4. AINEISTO JA METODIT.....	25
4.1 Tutkimusaineisto.....	25
4.2 Metodit.....	26
5. TULOKSET JA ANALYYSI.....	30
5.1 Johdanto tuloksiin ja analyysiin.....	30
5.2 Ilmisubjektin esiintyvyys eri taitotasolla .....	31
5.2.1 Ilmisubjektin esiintyvyys taitotasolla A1 .....	32
5.2.2 Ilmisubjektin esiintyvyys taitotasolla A2 .....	33
5.2.3 Ilmisubjektin esiintyvyys taitotasolla B1 .....	34
5.2.4 Ilmisubjektin esiintyvyys taitotasolla B2.....	35
5.2.5 Ilmisubjektin esiintyvyys taitotasolla C1 ja C2 .....	35

5.2.6 Ilmisubjektin esiintyvyys yhteensä kaikilla taitotasoilla .....	36
5.3 Ilmisubjektin esiintyvyys neljässä semanttisessa pääverbiluokassa .....	37
5.4 Ilmisubjektin esiintyvyys kahdeksassa semanttisessa verbiluokassa .....	38
5.4.1 Ilmisubjektin esiintyvyys tilaverbeissä .....	38
5.4.2. Ilmisubjektin esiintyvyys teko- ja muutosverbeissä .....	39
5.4.3. Ilmisubjektin esiintyvyys liikeverbeissä .....	40
5.4.4 Ilmisubjektin esiintyvyys puheaktiverbeissä .....	42
5.4.5 Ilmisubjektin esiintyvyys psykologista tilaa ja prosessia ilmaisevissa verbeissä .....	42
5.4.6 Ilmisubjektin esiintyvyys perkeptioverbeissä .....	44
5.4.7 Ilmisubjektin esiintyvyys abstrakteissa verbeissä .....	45
5.4.8 Ilmisubjektin esiintyvyys kieltoverbeissä .....	46
5.5 Ilmisubjektin esiintyvyys aineiston frekventeimpien verbien kanssa .....	47
6. PÄÄTÄNTÖ .....	50
6.1 Tutkimustulosten pohdintaa .....	50
6.2 Tutkimuksen kriittistä tarkastelua ja jatkotutkimusehdotuksia .....	52
LÄHTEET .....	56
LIITTEET	

# TAULUKOT

Taulukko 1. Kirjoittaminen EVK:n yleisillä taitotasolla (EVK 2003: 96).	11
Taulukko 2. Persoonapäätteet eri moduksissa (VISK § 107).	14
Taulukko 3. Maailman kielten luokittelu pronominisubjektin käytön mukaan (Dryer: 2011).	16
Taulukko 4. Olen-verbin kaikki hyväksytyt muodot taitotasolla A1.	30
Taulukko 5. Ilmisubjektin esiintyvyys eri taitotasolla.	31
Taulukko 6. Ilmisubjektin esiintyvyys taitotasolla A1.	32
Taulukko 7. Ilmisubjektin esiintyvyys taitotasolla A2.	33
Taulukko 8. Ilmisubjektin esiintyvyys taitotasolla B1.	34
Taulukko 9. Ilmisubjektin esiintyvyys taitotasolla B2.	35
Taulukko 10. Ilmisubjektin esiintyvyys taitotasolla C1 ja C2.	36
Taulukko 11. Ilmisubjektin esiintyvyys yhteensä kaikilla taitotasolla.	36
Taulukko 12. Ilmisubjektin esiintyvyys semanttisissa yläluokissa ja kieltoverbeissä.	37
Taulukko 13. Ilmisubjektin esiintyvyys tilaverbeissä.	38
Taulukko 14. Ilmisubjektin esiintyvyys teko- ja muutosverbeissä.	39
Taulukko 15. Ilmisubjektin esiintyvyys liikeverbeissä.	41
Taulukko 16. Ilmisubjektin esiintyvyys puheaktiverbeissä.	42
Taulukko 17. Ilmisubjektin esiintyvyys psykologista tilaa ja prosessia ilmaisevissa verbeissä.	43
Taulukko 18. Ilmisubjektin esiintyvyys perkeptioverbeissä.	45
Taulukko 19. Ilmisubjektin esiintyvyys abstrakteissa verbeissä.	46
Taulukko 20. Ilmisubjektin esiintyvyys kieltoverbeissä.	47
Taulukko 21. Ilmisubjektin esiintyvyys aineiston frekventeimpien yksikön 1. persoonan verbien kanssa.	48

# 1. JOHDANTO

Elias Lönnrot, yksi suomen kielen kehittämisen tunnetuimmista pioneereista, on erityisesti *muukalaisuuksien poistamiseen* suomen kielestä tähdänneellä toiminnallaan vaikuttanut siihen, että kirjoitetussa suomessa persoonapronominin välttäminen on tullut aiheelliseksi (Paunonen 2001: 154). Kirjeessään Volmari Kilpiselle hän toteaa seuraavaa:

Valitettava asia on, että rahvas yleisesti on tottunut ruotsintapaiseen lukukieleen – – . Mutta ihmeteltävää se ei ole, sillä kansan selvinkin kirja, raamattu, on sillä tavoin suomennettu, että siitä kyllä joka lehdeltä, paitsi muuta, saisi kymmenittäin persoonallisia sianimukoita (= persoonapronomineja) pois pyyhkäistä; ja mihin kahtena vuosisatana on totuttu, sitä vuodessa eli kymmenessä ei heitetä. (Anttila 1985: 449.)

Lönnrotilta lähtöisin oleva periaate 1. ja 2. persoonan pronomien karsimisesta omaksuttiin suomen kielen kielioppeihin. Muun muassa E. N. Setälän *Suomen kielen lauseopeissa* tätä periaatetta on korostettu, kuten myös esimerkiksi E. A. Saarimaan kieli- ja tyylioppaassa *Huonoa ja hyvää suomea: oikeakielisyysohjeita* sekä useassa muussa tunnetussa suomen kielen kieliopissa. (Paunonen 2001: 154–156.) Matti Sadeniemen tarkistamassa Setälän Suomen kielen lauseopin 16. painoksessa (1973: 86) kerrotaan 1. ja 2. persoonan pronomien käytöstä seuraavaa:

Subjektina oleva 1:sen tai 2:sen persoonan pronomini pannaan yleensä ilmi ainoastaan, kun sillä on paino. Esim. Minä en sitä tehnyt vaan sinä. Sitä vastoin: Jos tulet minua tervehtimään, niin olen kiitollinen.

Pronominisubjektin käytön yleisyys on perin poikkeavaa puhutussa ja kirjoitetussa suomessa (Lappalainen 2006: 37). Tutkimukseni päähypoteesina on, että ilmipannun pronominisubjektin eli ilmisubjektin käyttö vähenee persoonamuotoisten verbien yhteydessä kirjoitetussa oppijansuomessa sen mukaisesti, kuin korkeamman taitotason tekstejä tarkastellaan. Perustelen tätä oletusta sillä, että alemmilla taitotasoilla kirjoitetaan enemmän suomen puhekielen tapaisesti, kun taas korkeammilla taitotasoilla kirjoitetun suomen konventiot ovat paremmin hallussa, minkä myötä kielenoppijan kirjoituksessa on havaittavissa vähemmän puhekielisyksiä. Suoranaisesti ilmisubjektin poisjättöön ei kuitenkaan ohjata, eikä sitä voida millään tavalla pitää virheellisenä, ja kuten myös edellä toin esiin, joissakin yhteyksissä ilmisubjektin käyttö on myös perusteltua. Kuitenkin esimerkiksi oppikirjat ja muut oppimateriaalit voivat vaikuttaa vahvasti persoonapronominin ilmipanoon (ks. Dal Pozzo 2014).



Kirjoitetussa kielessä perustelluksi pronominisubjektin käyttö on nähty ainoastaan hyvin korosteisissa yhteyksissä. Suomen kielen kielenhuoltajien pioneereilta lähtöisin oleva ohje on sisäistetty hyvin, mikä on nähtävissä tutkimuksissa, joissa on vertailtu pronominisubjektin yleisyyttä kirjoitettua kieltä ja arkipuhetta sisältävissä korpuksissa. Vaikka esimerkiksi kaulokirjalliset tekstit voidaan sijoittaa puhutun ja kirjoitetun kielen välille, on ero silti selkeä: keskusteluissa ja haastatteluissa pronominisubjektia käytetään kiistatta useammin kuin su-pisuomea ja käännössuomea sisältävissä korpuksissa. On kuitenkin edelleen huomattava, ettei kyseinen ero ole laadullinen vaan määrällinen, koska sekä puhutuissa että kirjoitetuissa korpuksissa esiintyy tapauksia, joissa subjekti ilmenee ainoastaan predikaattiverbillisissä (*juoksen*) tai ilmisubjektillisissä rakenteissa (*minä juoksen*). (Lappalainen 2006: 37–38.) Yksikön 1. persoonan ilmisubjektin kanssa toimivia verbimuotoja voi normaalissa puhekielessä olla jopa 80–90 prosenttia kaikista yksikön 1. persoonan verbimuodoista (ks. Paunonen 2001: 158; Lappalainen 2004: 81; Duvallon 2006: 203).

Variaationtutkimuksessa ilmisubjektin käyttö on nähty myös puhekielisyyden ilmaisijana (Lappalainen 2004: 60–63). Puhutussa suomessa on kuitenkin useita konteksteja, joissa merkitys muuttuu, mikäli subjektipronomini jätetään pois ja persoona indikoidaan ainoastaan verbin persoonapäätteellä. Ei voida täten puhua saman referenssin toistamisesta vaan valinnoista, joihin sisältyy omia merkityksiään. (Helasvuo ja Laitinen 2006: 179–183.) On siis todettava, että persoonapronominien käyttämättä jättäminen ei sovi kaikkiin tilanteisiin. Lappalaisen (2004: 80) mukaan persoonapronomini on joskus laitettava esiin, kun kyseessä olevan persoonan merkitystä halutaan korostaa. Jokaista pronominesiintymää kirjoitetussa suomessa ei voida mutkattomasti pitää puhekielisyytenä – kääntäen pronominitoman muodon käyttäminen ei aina ole merkki yleiskielisyydestä. Puhekielessä rinnasteisissa yhdyslauseissa jälkimmäistä subjektipronominia ei yleensä käytetä, jos se viittaa samaan tarkoitteeseen kuin edeltävä subjekti kuten esimerkissä (1).

1) Minä menen eka kauppaan ja sitten lähden lenkille.

Toisaalta esimerkin (2) kaltaisissa tilanteissa, joissa vastataan kysymykseen, perin yleistä on vastata toistamalla kysymyksessä ollut finiittiverbi.

2) – Ai menet?  
– Menen.

Näin ollen voidaan nähdä, ettei pronominin käyttö rajoitu vain painollisiin tilanteisiin (Duvallon 2006: 203). Erityyppisten tilanteiden analysointi tuo esiin varsin suuria poikkeamia pronominisubjektin esiintymismäärissä. Pronominisubjektin esiintyvyyttä ei voida täten täysin ongelmattomasti pitää puhekielisyyden ilmaisijana. Joissakin puhetilanteissa pronominisubjekti jää herkemmin pois kuin joissakin toisissa, mutta itse pronominisubjektin poisjättöä ja siihen liittyviä tekijöitä on tutkittu melko vähän. (Duvallon 2006: 204.) Pronominisubjektin esiintyvyyttä suomen vanhoissa murteissa on kuitenkin äskettäin tutkinut Väänänen (2016).

Mielenkiintoista näkökulmaa ilmisubjektin käyttöön tuo myös käännöstieteiden tutkimukset, joissa on vertailtu yksikön 1. persoonan käyttöä kaunokirjallisissa teksteissä. Vertailussa on ollut suomeksi käännettyjä tekstejä alun perin suomeksi kirjoitettuihin teksteihin ja käännösten lähdeteksteihin. Suomentajalle tuottaa jatkuvaa päänvaivaa se, mitä suomen persoonanilmaisukeinoista hän käyttää kääntäessään esimerkiksi englannin tai saksan persoonapronomineja (kieliä, joissa persoonapronomini on aina ilmipantu). Tutkimuksesta käy ilmi selkeä ero 1. persoonan persoonapronominin käytössä vertailtaessa suomenkielisiä käännöstekstejä ja alkuperäistekstejä. Eroja mittaavien frekvenssilistojen kärjessä esiintyvät persoonapronomit. Erityisen suuri ero käännössuomen ja supisuomen tekstien välillä on 1. ja 2. persoonan käytössä, joita käännössuomessa on selkeästi enemmän. (Mauranen & Tiittula 2005: 35–36.) Käännöstieteen tutkimusten yhteys tutkimukseni kanssa voidaan nähdä sitä kautta, miten kääntäjä muokkaa suomen kielensä käyttöä lähtötekstinsä mukaan. Myös suomenoppijat kirjoittavat suomeksi enemmän tai vähemmän oman ensikielensä kautta, jolloin persoonapronominin ilmisubjektillinen tai ilmisubjektiton käyttö voi määräytyä osittain sen mukaan, miten sitä käytetään suomenoppijan omassa äidinkieleessä.

Koska suomenkielissä kirjoitetuissa teksteissä tavanomaisempaa on persoonan ilmaisu pelkällä verbin persoonamuodolla, aiheuttaa ilmipantu *minä* sen, että lukija alkaa mahdollisesti etsiä sille tulkintaa. *Minä*-pronominia käytetään usein jonkin yleisen periaatteen esille tuomisessa, kuten esimerkissä (3).

3) Minä en tupakoi.

Tällaisissa tilanteissa, joissa puhuja haluaa tuoda selkeästi esiin jonkin periaatteensa tai ominaisuutensa, erityisesti kielteisen, *minä*-pronominia esiintyy tiuhaan. *Minä*-pronomini esiintyy usein näin ollen dialogeissa. (Mauranen & Tiittula 2005: 48–49, 67.) Tässä tutkimuksessa selvitän myös sitä, poikkeako kieltoverbin ilmisubjektillisuus muista verbeistä. Kielto voidaan usein nähdä melko painokkaana ilmauksena, joten edellä sanotun mukaan voidaan olet-

taa kieltoverbin esiintyvän mahdollisesti useammin ilmisubjektillisena kuin muiden verbien. Eri semanttisten verbiluokkien ilmisubjektillisuudesta kirjoitetussa oppijansuomesta minulla ei ole ollut alun perin pohjaa hypoteeseille. Vanhoissa suomen murteissa (aineistona puhetta) ilmisubjektillisimpina esiintyvät psykologista tilaa ja prosessia, puhetta ja tilaa ilmaisevat verbit, kun taas liikeverbit ja muutosverbit (transitiiviset) ovat herkimmin ilmisubjektittomia (Väänänen 2016: 224). Vanhat puhutut suomen murteet kuitenkin poikkeavat lähtökohtaisesti melko paljon kirjoitetusta oppijansuomesta. Huomioin tutkimuksessani myös verbien esiintymisfrekvenssiä. Frekventeimmin esiintyvien verbien oletan esiintyvän vähemmän ilmisubjektillisina kuin vähemmän frekventtien, koska usein esiintyvät ovat luonnollisesti enemmän kielenoppijan käytössä ja tämän myötä niiden käyttö ilmisubjektitta voi tuntua luontevammalta. Toisaalta on kuitenkin edelleen todettava, että jokaisella verbillä on omankaltainen luonteensa, joten toisten verbien kanssa ilmisubjekti sopinee useammin yhteen kuin toisten.

Tutkimuskysymykseni ovat seuraavat:

1. Kuinka paljon yksikön 1. persoonan ilmisubjektia käytetään Eurooppalaisen viitekehyksen eri taitotasoilla? Väheneekö ilmisubjektillisuus taitotason noustessa?
2. Miten yksikön 1. persoonan ilmisubjektin käyttö poikkeaa eri semanttisten verbiluokkien välillä?
3. Vaikuttaako yksikön 1. persoonan verbin frekventtiys ilmisubjektin esiintymistajuuteen?

Tämä maisterintutkielma on suomi toisena kielenä -tutkimusta, joka on laajemmassa mittakaavassa osa toisen kielen oppimisen tutkimusta. Tutkimusaineistona toimii Topling-hankkeen suomenkielinen aineisto (Topling – Toisen kielen oppimisen polut). Topling-aineistosta on viime vuosina tehty useita pro gradu -tutkielmia. Näistä esimerkkeinä mainittakoon Takkinen & Viinikainen (2015), jotka ovat tutkineet suomen kielen oppijoiden kertomustekstien rakennetta ja Salo (2015), joka taasen on tutkinut yläkouluikäisten suomen kielen oppijoiden kohteliaisuuden ilmaisuja sähköpostiviesteissä.

## 2. TOISEN KIELEN OPPIMINEN

### 2.1 Toisen kielen oppiminen

Toisen kielen oppiminen (*second language acquisition*) on kielitieteellistä tutkimusta, joka tutkii ihmisten kapasiteettia oppia kieliä, jotka poikkeavat heidän ensikielestään tai -kielistään (Ortega 2009: 1–2; toisen kielen oppimisen ja kaksikielisyyden eroista ks. Ortega 2009: 3–5). *Lähtökielellä* (L1) tarkoitetaan kielenoppijan äidinkieltä. Siitä voidaan myös käyttää nimityksiä *ensikieli* ja *lähdekieli*; opittavasta kielestä taasen käytetään nimitystä *kohdekieli* (L2) (Nissilä 2011: 38.) Toisen kielen oppimisen tutkimuksen tavoitteena on tutkia laaja-alaisesti erilaisia monimutkaisia vaikutteita ja ilmiöitä, jotka vaikuttavat erilaisiin lopputuloksiin opittaessa toista kieltä erilaisissa ympäristöissä ja konteksteissa (Ortega 2009: 2). Yleensä on tapana erottaa *toisen kielen oppiminen* (tapahtuu ympäristössä, jossa yhteisö puhuu oppijan kohdekieltä) *vieraan kielen oppimisesta* (tapahtuu ympäristössä, jossa yhteisö ei juuri puhu oppijan kohdekieltä) (VanPatten & Benati 2010: 1).

Sunin (2008: 29–30) mukaan kuitenkin englanninkielisessä, toisin kuin suomenkielisessä (ks. 2.3), tutkimuksessa perinteisesti *second language* (toinen kieli) toimii kielenoppimisen perustyyppinä ja yläkäsitteenä suhteessa *foreign languageen* (vieraaseen kieleen). Perusteena tähän on se seikka, että englanninkielisissä yhteisöissä maahanmuutto ja maahanmuuttajien kielenoppiminen on jo pitkään ollut keskeistä – puolestaan Suomessa vieraiden kielten opetus ja oppiminen ovat olleet paljon pitempään huomion kohteena kuin maahanmuuttajien kielikyvykset. VanPatten ja Benati (2010: 1–2) korostavat kuitenkin, että tällainen erottelu on vain sosiologisesta näkökulmasta perusteltua. Toisaalta tutkittaessa oppijoita ja oppimismekanismeja, joita toisen kielen ja vieraan kielen oppiminen koskettaa, ei voida nähdä eroa oppimiskontekstien välillä, koska ihmisäivot toimivat kummassakin tilanteessa samalla tavalla. Ympäristö kuitenkin ilman muuta vaikuttaa siihen, kuinka hyvin kieltä opitaan, mutta se ei vaikuta siihen prosessiin, miten ihmisäivot omaksuvat, oppivat ja prosessoivat toista tai vierasta kieltä. Toisaalta Jantusen (2015: 109–110, 125–126) mukaan taasen oppimiskontekstilla (eli opitaanko kieltä toisena vai vieraana kielenä) on erityisesti merkitystä opittavan kielen pragmaattisten taitojen kehitykseen. Martinin (2003: 76) mukaan puolestaan toisen ja vieraan kielen erottamiseen on oppimisen ja erityisesti opettamisen kannalta perusteita. Toisaalta Nissilä (2011: 38) toteaa, että täydellistä jakoa toisen ja vieraan kielen välille ei ole mahdollista tehdä, vaan kyseessä on tietynlainen jatkumo, jonka toisessa päässä sijaitsee vieraan kielen

oppiminen ja toisessa toisen kielen omaksuminen. Asiassa on nähtävissä jonkinlaisia näkemyseroja, jotka toki ovat perusteltavissa erilaisilla näkökulmilla.

Toisen kielen oppimisen tutkimus sai alkunsa 1960-luvun lopulla, jolloin se toimi poikkitieteellisenä tieteenhaarana, joka sai suurimmat vaikutteensa kielen opettamisen, kielitieteen, ensikielen omaksumisen ja psykologian teorioista. 1980- ja 1990-luvulla alan tutkimus laajeni huomattavasti ja vuosituhannen loppuun mennessä sitä on alettu pitää itsenäisenä tieteenalana. Toisen kielen oppimisen tutkimus on tästä eteenpäin edelleen kehittynyt ja kasvanut jatkuvasti. Toisen kielen oppiminen nähdään kuitenkin yleisesti soveltavan kielitieteen yhtenä haarana. Soveltava kielitiede toimii erityisesti kielen ja yhteiskunnan sekä opetuksen ja kognition risteyskohdissa. (Ortega 2009: 2–3, 7.) Toisen kielen oppimisen tutkimus keskittyy ennen kaikkea oppimiseen ja oppijoihin opettajien ja opettamisen sijaan. Kuitenkin kielten opettamisen ollessa yksi tutkimuksen peruslähtökohdista toisen kielen oppimisen tutkimuksen tavoitteena on monesti tuottaa tietoa siitä, mitkä opetustavat ja -menetelmät antavat parhaan oppimistuloksen. Toisen kielen opettajien koulutuksen olisikin kuljettava käsi kädessä toisen kielen oppimisen tutkimustulosten kanssa, jotta opetus ja oppiminen antaisivat mahdollisimman hyviä lopputuloksia. (VanPatten & Benati 2010: 1, 6–7.)

Toisen kielen oppimisessa halutaan joka tapauksessa tutkia ohjatun oppimisen (*instructed learning*) lisäksi myös luonnollista (*naturalistic learning*) oppimista. Luonnollisella oppimisella tarkoitetaan kaikkea epämuodollista oppimista, jota tapahtuu esimerkiksi monikulttuurisilla asuinalueilla, kouluissa ja työpaikoilla ilman, että oppiminen olisi jotenkin ohjattua. Ohjattu oppiminen kattaa taasen kaiken oppimisen, jota tapahtuu esimerkiksi kouluissa, yliopistoissa ja yksityisopetuksessa. Monesti toista kieltä opitaan sekä ohjatusti että luonnollisesti. (Ortega 2009: 6.) Sajavaaran (1999: 75) mukaan luonnollisessa ympäristössä kieltä opittaessa kuitenkin tapahtuu paljon ohjausta ja opetusta, kun taas toisaalta luokkatilanteessa on myös paljon edellytyksiä luonnolliseen kielen omaksumiseen. Suomen kielen tutkimuksessa *omaksumista* ja *oppimista* on silloin tällöin käytetty samassa merkityksessä ihmisen kielen kehitykselle. Kuitenkin *omaksuminen* voidaan rajata tarkoittamaan luonnollisessa kielenkäyttöympäristössä tapahtuvaa kielenoppimista. *Oppimisen* puolestaan voidaan katsoa vaativan ainakin jonkintasoista muodollista ohjausta. Toki *oppimista* käytetään myös yleisterminä kaikenlaiselle kielen hallinnan edistykselle. Tämän tutkimuksen aineistosta (ks. 4.1) voidaan sikäli ennen kaikkea tutkia ohjattua oppimista, kuten esimerkiksi erilaisten tekstilajikonventioiden hallintaa ja niihin sopivaa kielenkäyttöä eri tehtävänannoissa (*mielipidekirjoitus, kertomus, viesti verkkokauppaan, viesti opettajalle*). Toisaalta *viesti ystävälle* on sellainen teksti ja tehtävänanto, johon voisi olettaa luonnollisesta kielenkäyttöympäristöstä tulevan enemmän

vaikutteita kuin ohjatun oppimisen kielenkäyttöympäristöistä – varsinkin siinä vaiheessa, kun opittavaa kieltä hallitaan jo jonkin verran.

## 2.2 Oppijankieli

Oppijankieli on kielenoppijoiden tuottamaa suullista tai kirjoitettua kieltä. Oppijankieli toimii toisen kielen oppimisen ensisijaisena tutkimusmateriaalina. On huomattava, ettei oppijankieli missään nimessä ole yhtenäinen ilmiö, vaan siinä on jatkuvasti suurta vaihtelua. Tämän myötä se eri tavoin kerättyinä materiaalina tuo esiin monia mielenkiintoisia näkökulmia siihen, miten kielenoppijat oppivat kieltä. (Ellis & Barkhuizen 2005: 4–5.) Oppijankieli muodostuu kielenoppijan lähtökielen, opittavan kielen ja oppijan kunkin oppijankielen vaiheen vaikutuksesta. On myös huomionarvoista, että oppijan osaamilla muilla kielillä on vaikutusta oppijankieleen. (Nissilä 2011: 39.)

Oppijankieli nähdään erilliseksi järjestelmäksi lähtö- ja kohdekielen välillä. Oppijankieli on jatkuvassa murroksessa, kun uusia sääntöjä omaksutaan ja vanhoja kehitetään. Täten se voidaan nähdä systemaattisena ja luovana jaksona kielenoppimisessa. Oppijankielen ominaisuuksiin kuuluvat oppijan kielimuodossa esiintyvät poikkeamat verrattuna kohdekieleen. Näitä poikkeamia ei näin ollen nähdä virheinä. Oppijankielen tutkimus painottaa oppimisprosessia, jota tarkastelemalla on havaittu todisteita erilaisille oppijoille yhteisistä oppimisvaiheista. Ennen kaikkea kiinnitetään huomiota oppijankielen muuttuvuuteen – oppijankieleessä tapahtuu jatkuvaa siirtoliikettä, kun kohdekielestä tulevat vaikutteet mukautuvat jo olemassa olevaan oppijankieleen. Jatkuvan muuttuvuuden vuoksi oppijankielen ennustettavuus on varsin vaikeaa, mutta tästä huolimatta on mahdollista löytää kollektiivisia piirteitä eri oppijoiden välillä. (Nissilä 2011: 39–40.)

Oppijankieleen ottaa osaa yhtäaikaaisesti useita eri kieliä. Siihen vaikuttavat lähtökieli, kohdekieli, parhaillaan menossa oleva oppijankielen vaihe sekä oppijan taitamat muut kielet. Toisen kielen oppijalla on hallussaan lähtökielensä systeemi, jonka lisäksi hän vastaanottaa kohdekielen syötöstä, jonka pohjalta hän tuottaa jatkuvasti uutta ja entisestä poikkeavaa kielisysteemiä. Oppijankieleen vaikuttaa kiistatta se, jos oppijan lähtö- ja kohdekieli ovat likeisiä sukukieliä. Kun oppija huomaa selkeää samankaltaisuutta lähtökielen ja kohdekielen välillä, on hänen lähtökohtaisesti helppoa assosoida opittavan kielen muotoja ja rakenteita jo osamiinsa lähtökielen piirteisiin ja sitä kautta sisällyttää ne joutuisasti laajenevaan oppijankieleensä. (Nissilä 2011: 40.) Myös suomen kielen persoonan ilmaisu on toki tällainen kielen

osa-alue, joka lähtökohtaisesti on helpompaa niille, joiden kielessä persoonaa ilmaistaan samanlaiseen tapaan kuin suomessa.

Ellis & Barkhuizen (2005: 5–6) jakavat lingvistisen tiedon implisiittiseen (*prosessuaaliseen*) ja eksplisiittiseen (*deklaratiiviseen*). Implisiittinen kielitietous on ennen kaikkea sellaista kielitietoa, jonka hallitsemme ensikielestämme. Tätä kielenkäyttäjää pystyy yleensä käyttämään ikään kuin tiedostamattaan eli implisiittinen kielitietous on automaattistunutta kielenkäyttöä spontaaneissa kielentuottamistilanteissa. Se voidaan jakaa kahteen osaan: kaavamaisiin fraasilohkoihin (*”en tiedä”*, *”mitä kuuluu?”*) ja tiedostamattomaan tietoon sekä vähäpätöisistä että merkittävistä abstrakteista rakenteista ja säännöistä. Eksplisiittinen kielitietous taas on tiedostettua. Siihenkin kuuluu äsken mainittu vähäpätöisten ja merkittävien abstraktien rakenteiden ja sääntöjen hallinta, mutta mentaalisesti tämä tieto prosessoidaan täysin eri tavalla. Lisäksi eksplisiittiseen tietoon kuuluu metakieli, jonka avulla analysoitu tieto voidaan ilmaista.

Toisen kielen oppimisen tutkimuksen päätavoitteena on ennen kaikkea tehdä selkoa oppijoiden implisiittisestä kielitietoudesta. Implisiittiseen kielitietouteen on kuitenkin varsin haastavaa päästä suoraan käsiksi, koska noin vain ei ole mahdollista katsoa oppijan aivoihin tai tietoisuuteen, millaisessa järjestyksessä kieli kenenkin aivoissa on representoitu tai minkälaista tietoa käytetään hyväksi silloin, kun suoritetaan jotakin tiettyä kielitehtävää. Jossakin määrin katsominen kuitenkin on mahdollista toiminnallisen magneettikuvauksen avulla. Toiminnallisessa magneettikuvauksessa pystytään tunnistamaan ne aivojen osiot, jotka aktivoituvat suoritettaessa kielitehtävää. Kuitenkaan tälläkään tekniikalla ei vielä voida erotella sitä, mitkä eri aivojen osiot suorittavat mitäkin kielellistä prosessia. Tämän vuoksi toisen kielen oppimisen tutkimuksen päämateriaalina ovat edelleen erilaiset kielenoppijoiden suoritukset ja tuotokset. (Ellis & Barkhuizen 2005: 6.) Kirjoitetusta oppijankielestä voidaan havainnoida kirjoitetun kielen normien noudattamista, tekstilajikonventioiden osaamista, puhekielisyyden määrää sekä runsaamminkin tyylikonventioiden hallintaa (Jantunen 2008: 79). Tämä tutkielma sijoittuu näistä kahden viimeiseksi mainitun välimaastoon, mutta kallistuu enemmän tyylikonventioiden hallintaan, koska ilmisubjektin käyttämistä ei voida suoranaisesti sanoa puhekielisyydeksi, vaikka sitä puhutussa suomessa käytetäänkin selkeästi enemmän kuin kirjoitetussa (ks. Lappalainen 2006: 37–38).

## 2.3 Suomi toisena kielenä

Globaalistumisen myötä toisen kielen oppiminen koskettaa eri ihmisiä monenlaisista syistä. Näitä ovat muun muassa matkustelu, työllisyys, maahanmuutto, sodat, monikulttuuriset avioliitot sekä monet muut näkökulmasta riippuen iloiset tai surulliset elämäntapahtumat. (Ortega 2009: 6.) 1990-luvulta lähtien Suomesta on tullut maastamuuton sijaan enenevässä määrin maahanmuuton kohde. Tämän myötä myös tilaus maahanmuuttajien suomen kielen opetusta kohtaan on luonnollisesti kasvanut melkoisesti. (Martin 2011: 1.) Ennen 1990-lukua toisen kielen käsitettä ei suomen kielessä ollut käytössä lainkaan. Sittemmin käyttöön on omaksuttu lähinnä ruotsin kielestä tuleva periaate siitä, etteivät toinen ja vieras kieli ole toisiinsa nähden ylä- ja alakäsitteitä, vaan niillä tarkoitetaan ensikielen jälkeen opittavia tai opittuja kieliä kieliympäristön toimiessa erottavana tekijänä. (Suni 2008: 29.)

Suomi toisena ja vieraana kielenä on soveltavaa kielentutkimusta ja sen tavoitteena on sekä alan opetuksen kehittäminen että kielen oppimista koskevan tieteellisen tiedon kartuttaminen. Se antaa myös pohjaa yhteiskunnalliselle päätöksenteolle, koska sen tutkimus keskittyy valtaosin vähemmistöryhmien kielenkäyttöön ja vähemmistöryhmille järjestettävään kielen opetukseen. (Martin 1999: 157.) Yhteiskunnallisesti merkittävää on erityisesti se, että maahanmuuttajien suomen kielen opetukselle asetetaan työllistymiseen, jatkokoulutusvalmiuksiin ja oppimistaitoihin liittyviä vaatimuksia. Näin ollen suomi toisena kielenä -opetuksen pedagogiset haasteet ovat monimuotoisia, mikä myös tarkoittaa sitä, että sen tulee poiketa pedagogiikaltaan vieraiden kielten opetuksesta. On varsin haasteellista saada tämänhetkiset yleisesti tunnustetut kielenoppimisen teoriat ja opetuksen käytäntö vielä entistä paremmin kohtaamaan. (Nissilä 2011: 65.) Kuitenkin suomi toisena kielenä -alalla opetus ja tutkimus kulkevat sikäli käsi kädessä, että useimmat alan tutkijat myös opettavat tai ovat opettaneet suomea toisena kielenä (Martin 1999: 164). Toisen kielen oppimisen tutkimuksesta suurin osa on tehty sellaisissa kieliympäristöissä, joissa opiskeltava kohdekieli on indoeurooppalainen kieli, monissa tapauksissa englanti. Englannin tai jonkin muun indoeurooppalaisen kielen oppimiseen liittyvät teoriat eivät monestikaan päde, kun tarkastellaan suomen oppimista toisena kielenä. Näin ollen S2-tutkimus voidaan nähdä eräänlaisena toisen kielen oppimisen universaalien koelaboratoriona. (Martin 2007: 64.) Myös oppijankielen korpustutkimuksissa oppijanenglannilla on edelleen hallitseva asema, mutta viime vuosina on kerätty kokoon sekä myös laajennettu useita oppijansuomea sisältäviä korpuksia (Jantunen & Pirkola 2015: 88).



Monesti suomea sanotaan vaikeaksi kieleksi. Suomen vaikeus – kuten minkä tahansa muunkin kielen – on kuitenkin perin suhteellista. Suomenoppijan lähtökieli ja -kulttuuri vaikuttavat ratkaisevasti asiaan. Vironkielisille ja muille suomen lähisukukielten puhujille suomi on rakenteeltaan ja myös sanastoltaan tutuntuntuinen kieli, kun taas tullessa kulttuurisesti tai maantieteellisesti kauempaa kielen outous lisääntyy. Helpoimmin kohdekielestä opitaan yleensä sellaiset asiat, jotka lähtökielessä voidaan ilmaista monella eri tavalla, mutta kohdekielessä vain yhdellä; sekä luonnollisesti ne seikat, jotka kohdekielessä eivät vaadi huomiota lainkaan. Suomessa yleisimpiin lähtökieliin verrattuna tällaisia ovat kieliopillisen suvun ja artikkelien puuttuminen, sanapainon sijoittuminen poikkeuksesta ensitavulle ja konsonanttivalikoiman vähäisyys. Sanat voivat alkuun näyttäytyä kummallisina ja pitkinä, mutta oppijan kielitaidon edistyessä suomen sanaston läpinäkyvyys, eli yhdyssanojen ja johdosten runsas osuus, keventää uusien sanojen oivaltamista. Toisaalta alkuun vaikeinta on morfologisten muotojen suuri määrä ja taivutusmuotojen variaatio. Rakenteellisten seikkojen lisäksi suomenoppijalle tähdellistä on suomen puhekielen ja kirjoitetun kielen suuri eroavaisuus – puhumattakaan alueellisista ja sosiaalisista eroista. (Martin 1999: 168–169.) Suomenoppijalla on kuitenkin aina muassaan kosolti kielellisiä ilmaisuvaroja, joita vahvistavat kontaktit Suomen kieliin ja kulttuureihin – oleellisinta on joka tapauksessa mahdollisuus olla alituisen interaktiossa suomalaisen kieliympäristön kanssa suomeksi. Suomenoppijat, jotka ovat päivittäin vähintään koulun kautta kosketuksissa suomen kieleen, tuottavat varsin nopeasti idiomaattisia ja sujuvia suomenkielisiä tekstejä. (Seilonen 2013:192.)

## 2.4 Kielen kehitys ja Eurooppalainen viitekehys

Kielen kehityksen tutkiminen voi olla olemukseltaan ymmärtävää, kuvailevaa ja/tai selittävää. Tutkimus on ajan myötä laajentunut äänne-, sana-, tai lausetasolta diskurssien, tekstien ja puheen tasolle, joten oppijankielen variaatiota pyritään tarkastelemaan kielenkäytön ja vuorovaikutuksen kontekstissa. Kielen kehitystä tutkittaessa on huomio suunnattava joihinkin tiettyihin ilmiöihin, koska käsitteenä *kehitys* on niin laaja, ettei sitä voida luonnehtia aineistossa erottavana, mitattavana ominaisuutena itsessään. (Martin & Alanen 2011: 38.)

Eurooppalaisen viitekehysten tarkoituksena on tarjota Euroopan maille yhteinen lähtökohta kielten opinto-ohjelmien, opetussuunnitelmien perusteiden, tutkintojen, oppimateriaalien ynnä muiden valmistukseen. Se luonnehtii laajasti, miten oppijoiden olisi opittava kieltään käyttämään, jotta viestintä olisi sujuvaa; sekä mitä tietoja ja taitoja heidän olisi kohennettava

viestinnän tehostamiseksi. Ennen kaikkea viitekehys määrittelee kielitaidon eri tasot, joiden avulla oppijoiden kielen kehittymistä pystytään mittaamaan oppimisen eri vaiheissa. (EVK 2003: 19.) Tarnasen (2002: 107) mukaan Eurooppalainen viitekehys perustuu vahvasti siihen näkemykseen, että kielen käyttö ja oppiminen ovat toimintaorientoituneita, joten kielenkäyttö edellyttää tiettyä kommunikatiivista kompetenssia sekä strategioiden käyttöä, jotta kielenkäyttäjän on mahdollista toimia jonkin kielenkäyttötilanteen edellyttämällä tavalla ja siten saavuttamaan tavoitteensa. Tässä tutkimuksessa Eurooppalaisella viitekehyksellä on myös tärkeä osa, koska Topling-aineiston kaikki tekstit on arvioitu EVK:n taitotasojen mukaan. Alla olevasta taulukosta 1 on nähtävissä, miten kirjoittamista yleisesti luokitellaan eri taitotasolla.

Taulukko 1. Kirjoittaminen EVK:n yleisillä taitotasolla (EVK 2003: 96).

A1	Pystyy kirjoittamaan yksinkertaisia, irrallisia ilmauksia ja lauseita.
A2	Pystyy yhdistämään yksinkertaisia ilmauksia ja lauseita tavanomaisilla sidesanoilla, kuten sanoilla 'ja', 'mutta', 'koska'.
B1	Pystyy kirjoittamaan yksinkertaisia, yhtenäisiä tekstejä tavallisista itseään kiinnostavista aiheista yhdistämällä lyhyempiä, irrallisia, yksinkertaisia ilmauksia yhtenäiseksi tuotokseksi.
B2	Pystyy kirjoittamaan selkeää, yksityiskohtaista tekstiä erilaisista itseään kiinnostavista aiheista. Pystyy tekemään yhteenvetoja sekä arvioimaan useista lähteistä poimimia tietoja ja perusteluja.
C1	Pystyy laatimaan selkeitä, hyvin jäsenneltyjä tekstejä monipolvisistakin aiheista ja korostamaan tärkeitä asioita. Pystyy jossain määrin tukemaan näkemyksiään lisäseikoilla, perusteluilla ja osuvilla esimerkeillä sekä päättämään kirjoituksensa sopivaan lopetukseen.
C2	Pystyy laatimaan selkeitä, hyvin jäsentyneitä tekstejä monipolvisista aiheista ja tuottamaan asianmukaisen ja tehokkaan johdonmukaisen rakenteen, joka auttaa lukijaa löytämään tekstistä keskeiset seikat.

Kuten on asian laita taulukossa 1, ei Eurooppalaisessa viitekehyksessä kuitenkaan ole mitään mainintaa persoonapronominien käytöstä ja sen minkäänlaisesta yhteydestä kielenkäytön kehitykseen, koska EVK:n lähtökohtana on ennen kaikkea arvioida kielen osaamista funktionaalista näkökulmasta. Ennen kaikkea on myös huomattava, että EVK on yleiseurooppalainen kielen oppimisen arviointimalli. Täten sen ei mitenkään ole edes mahdollista antaa tietoa siitä, mitä persoonapronominien tai muiden yksittäisten kieliopillisten osa-alueiden hallinta kertoo kielen osaamisen tasosta, koska Euroopan kielet poikkeavat todella paljon toisistaan

(erityisesti indoeurooppalaiset kielet vs. suomalais-ugrilaiset kielet; ilmisubjektin käytöstä eri maailman kielissä ks. 3.2.2). Tämä tutkimus pyrkiikin suomen kielen osalta osoittamaan, että taitotasolla ja ilmisubjektin käytöllä (tai käyttämättä jättämisellä) on selkeä yhteys tutkailtaessa suuria tekstimassoja, mutta samalla ymmärtäen, ettei yksittäisen suomenoppijan kirjoitetun suomen osaamista voida tekstin ilmisubjektillisuuden mukaan taitotasollisesti määrittää. Suomi toisena kielenä -tutkimusta, jossa eri taitotasojen välillä tapahtuvaa kehitystä jollakin kielen osa-alueella tutkaillaan, tehdään paljon.

### 3. ILMISUBJEKTI JA VERBIEN SEMANTTINEN LUOKITTELU

#### 3.1 Persoonapronomit ja persoonamuotoiset verbit

Pronomit ovat nomineja, mutta niiltä uupuu muiden nominien kategorioiva ja luonnehtiva merkitys. Pronomineilla käsitellään puheenalaisia tarkoitteita ilman, että niitä luonnehditaan leksikaalisesti. (VISK § 713.) Yksikön ensimmäistä persoonaa ilmaisee *minä* ja monikkoa *me*, yksikön toista persoonaa *sinä* ja monikkoa *te*, yksikön kolmatta persoonaa *hän* (*se*) ja monikkoa *he* (*ne*). Persoonapronomit ovat itsenäisinä nominilausekkeina käyttäytyviä pronomineja, jotka viittaavat puhetilanteen osanottajiin (1. ja 2. persoona) ja puheena oleviin henkilöihin (3. persoona). 1. ja 2. persoonan pronominit, joita voidaan myös nimittää puheaktipronomineiksi, viittaavat puhetilanteen kahteen ehdottomaan osanottajaan, puhujaan ja kuulijaan. Puheaktipronomien tarkoite on tilanneriippuvainen, koska puhujan ja kuulijan roolit vaihtuvat puheenvuoron vaihtuessa. Näin ollen puheaktipronomit ovat deiktisiä. Kolmannen persoonan pronominit puolestaan ovat anaforis-deiktisiä, koska niiden kohde on kolmas osapuoli, joka ei ole puhuja eikä kuulija, joten niiden tarkoite ei vaihdu puhujan vaihtuessa. (VISK § 716.)

Persoonamuodossa esiintyvä verbi on finiittinen, mikä tarkoittaa sitä, että sen taivutus mukautuu nominatiivissa olevan subjektin persoonan mukaan. Finiittinen verbi itsessään sisältää tiedon subjektin kieliopillisesta persoonasta; lauseen subjektin ollessa yksikön tai monikon 1. tai 2. persoonan pronomini tulee finiittiverbin sisältää vastaavan persoonan tunnus. (VISK § 107, § 1268.) Subjektin ja predikaatin persoonakongruenssi käsittää verbin liittomuodon ja verbiketjun ensimmäisen jäsenen. Jokaista liittomuotoa, verbiliittoa ja verbiketjua kohden on ainoastaan yksi persoonamuoto. Tämä on nähtävissä esimerkissä (4).

4) Minä *ostin* leipää. | Minä *olen* ostanut leipää. | Minä *en* muistanut ostaa leipää. | Minä *olisin* kyllä voinut olla ostamatta leipää. | Minä *onnistuin* ostamaan leipää. (VISK § 1268.)

Taulukko 2. Persoonapäätteet eri moduksissa (VISK § 107).

Persoonaa	Indikatiivi, konditionaali, potentiaali	Pääte	Imperatiivi	Pääte
yks. 1.	kerro-n, kertoisi-n, kertone-n	-n	–	
yks. 2.	hyppää-t, hyppäisi-t, hypänne-t	-t	kerro <sup>x</sup> , hyppää <sup>x</sup>	–
yks. 3.	kerto-o, syö, söisi, syöne-e	-(V)	kertoko-on, hypätkö-ön	-On
mon. 1.	kerro-mme, söisi-mme, antane-mme	-mme	kertokaa-mme	-mme
mon. 2.	kerro-tte, söisi-tte, antane-tte	-tte	kertokaa, hypätkää	–
mon. 3.	kerto-vat, söisi-vät, hypänne-vät	-vAt	kertoko-ot	-Ot
passiivi	kerrota-an, syötäisi-in, annettane-en	-Vn	kerrottako-on	-On

Kuten taulukosta 2 näemme, finiittisissä taivutusmuodoissa persoonaa osoittaa verbin persoonapäite, joka myöntömuodoissa yhdistyy verbin vartaloon (indikatiivin preesensissä) tai vaihtoehtoisesti moduksen tai tempuksen tunnukseen. Näiden vastineina toimivat kielteiset liittomuodot, joissa persoonapäite liittyy kieltoverbiin (*en pääse* vs. *pääsen*). (VISK § 107.)

Yksikön 1. persoonan kannalta olennaista on, ettei sillä ole lainkaan imperatiivia. Indikatiivissa persoonapäätteet yhtyvät 1. ja 2. persoonassa astevaihtelun verbeillä heikkoon vartaloon (*kerro-n*); konditionaalissa taas on aina vahva-asteinen vartalo (*kerto-isi-n*). Potentiaalissa vaikuttava tekijä on suora ja käänteinen astevaihtelu. (VISK § 107.) Suoralla astevaihtelulla tarkoitetaan sitä, että sanan perusmuoto on vahva- ja taivutusvartalo heikkoasteinen, kuten on laita 1. ja 2. persoonan indikatiivissa – käänteisessä astevaihtelussa verbin perusmuoto on heikko- ja taivutusmuoto vahva-asteinen. (VISK § 42.) Potentiaali liittyy suoran astevaihtelun verbeissä vahva-asteiseen vartaloon (*kerto-ne-n*), kun taas käänteisen astevaihtelun verbeissä potentiaali liittyy heikkoasteiseen vartaloon (*hypä-nne-n*) (VISK § 107).

## 3.2 Ilmisubjekti

### 3.2.1 Ilmisubjekti ja subjektityypit

Keskityn tässä tutkielmassa yksikön 1. persoonan pronominin ilmipantuun (*minä* *menen*) ja ilmipanemattomaan (*menen*) esiintyvyyteen kirjoitetussa oppijansuomessa. Iso suomen kieloppi käyttää eksplisiittisestä eli ilmipannusta persoonapronominista nimitystä *ilmisubjekti* (VISK § 914). Subjekteja esiintyy kolmessa eri muodossa: *perussubjektina*, *e-subjektina* eli *eksistentiaalibjektina* sekä *genetiivisubjektina*. Aineistoni subjekteista huomioin ainoastaan

esimerkin (5) kaltaiset muodot, joissa 1. persoonan perussubjekti on mahdollinen, koska se on ainoa subjektimuoto, jonka kanssa finiittiverbi kongruoi persoonan ja luvun suhteen.

5) **Minä** tiedän kaiken.

Perussubjekti esiintyy siis aina nominatiivimuotoisena nominilausekkeena, muut subjektit käyttäytyvät eri tavoin. E-subjekti ei nominatiivimuotoisenakaan aiheuta verbin kongruenssia ja genetiivisubjekti puolestaan on nesessiivirakenteen ja infiniittisen rakenteen subjekti. Persoonamuotoinen genetiivisubjekti on esiintyessään lauserakenteen kannalta välttämätön kuten esimerkissä (6). (VISK § 910.)

6) **Minun** täytyy päteä.

Vaihtoehtona kuitenkin on myös esimerkin (7) *nollapersoona*, mutta se on ensimmäisen persoonan tapauksessa selkeä puhekielisyys ja verbi säilyy edelleen infiniittisenä, joten kyseessä olevaa persoonaa ei voida suoraan päätellä ilman kontekstia (VISK § 1365).

7) Täytyy päteä.

Ilmisubjekti esiintyy valinnaisena ensimmäisen ja toisen persoonan verbimuodon kanssa. Niiden yhteyteen sopii kullekin vain tietty subjekti: *minä*, *sinä*, *me* tai *te*. Kirjoitetulle suomen kielelle 1. ja 2. persoonan pronominisubjektin poisjätto on luonteenomaista kuten esimerkissä (8).

8) Olen todella tyytyväinen, että pääset paikalle.

Kolmannen persoonan verbin yhteyteen subjektipronomini kuitenkin yleensä merkitään ja niiden määrä onkin rajaton. Tätä varianssia kuvaa esimerkki (9). (VISK § 914.)

9) Mihin **Mikko** oikein lähti? **Hän** häviää aina ensimmäisenä paikalta, kuten teki aikoinaan myös hänen isänsä **Seppo**.

3. persoonan subjektipronominin poisjätto on laadultaan 1. ja 2. persoonasta poikkeavaa – pronominia vaille olevassa finiittilauseessa on joko esimerkin (10) nolla-anafora tai esimerkin (11) nollasubjekti (VISK § 1269).

10) Simo istui tovin, **meni** sitten menojaan.

11) Jos **osaa** luistaa hommasta kuin hommasta, **on** silloin ainakin jossakin lahjakas.

Puhutussa suomessa ilmisubjektit ovat puolestaan yleisempiä ja niiden esiintyvyys on eri persoonien välillä yhdenmukaisempaa (VISK § 914).

### 3.2.2 Subjektin ilmaisu maailman kielissä

Karkeasti suurin osa maailman kielistä voidaan jakaa nollasubjektikieliksi (*pro-drop-kieli*) ja ei-nollasubjektikieliksi (*ei-pro-drop-kieli*) (Dal Pozzo 2014: 200). Maailman kielistä ei-nollasubjektikieliset (esimerkiksi englanti, ruotsi ja saksa), joissa persoonan ilmaisuun tarvitaan aina persoonapronomini, ovat vähemmistössä. 75 % maapallon kielistä (tutkimuksessa tiedot yhteensä 711 maailman kielestä) on nollasubjektikieliä. (Posio 2012: 8.)

Taulukko 3. Maailman kielten luokittelu pronominisubjektin käytön mukaan (Dryer: 2011).

Luokittelu (ps = pronominisubjekti)	Lukumäärä	%
Ps:t ilmaistu subjektin asemassa olevalla (yleensä pakollisella) pronominilla	82	11,5 %
Ps:t ilmaistu verbiin liittyvällä affiksilla	437	61,5 %
Ps:t ilmaistu erilaisilla ja erilaisiin lauseenjäseneihin liittyvillä partikkeleilla	32	4,5 %
Ps:t esiintyvät syntaktisesti muista subjekteista poikkeavasti	67	9,4 %
Ps:t esiintyvät subjektin asemassa, mutta usein poisjätettyjä, vaikkakaan eivät millään affiksilla ilmaistuja	61	8,6 %
Ps:t toimivat enemmän kuin yhdellä edellä mainitulla tavalla, mutta mikään tapa ei ole hallitseva	32	4,5 %
Yhteensä	711	100 %

Kuitenkaan Dryerin (2011) luokittelu, jota voi tarkastella taulukosta 3, ei ota huomioon pronominisubjektien käytön frekvenssiä kielissä, joista se kieliopillisesti jää normaaliksi pois, mutta sitä kuitenkin käytetään (esimerkiksi espanjassa ja italiassa). Luokittelu onkin ennen kaikkea kieliopillinen kuin esiintymien frekvenssiin nojaava. (Posio 2012: 8–9.) Suomen kieli kuuluu viimeiseksi mainittuun, melkoisen harvinaiseen yhdistelmäluokkaan (esimerkiksi mayakieli jakaltekin kanssa), jossa 1. ja 2. persoona ilmaistaan yleensä verbin persoonapäätteellä, mutta 3. persoona taas pronominilla (Dryer 2011).

Dal Pozzon (2014: 200–201) Firenzen yliopistossa italiankielisten (nollasubjektikieli) suomen kielen opiskelijoiden teksteistä tekemä tutkimus osoittaa, että suomen kieli todennäköisesti näyttäytyy monille (erityisesti ulkomailla suomea opiskeleville) einollasubjektikielenä. Suomen kielessä ilmisubjektittomat lauseet sekä vapaamuotoinen sanajärjestys ovat yleisiä ominaisuuksia. Ulkomailla suomea opiskeleva kohtaa kuitenkin usein kielenkäyttötilanteissa ilmisubjekteja ja SV(O)-perussanajärjestystä etenkin, koska ne toistuvat jatkuvasti oppikirjojen teksteissä. Vaikka tutkielmani keskittyykin vain Suomessa suomea toisena kielenä opiskeleviin, voidaan perustellusti olettaa, että Suomessakin erityisesti alempien tasojen opetus ja oppimateriaalit saattavat ohjata suomenoppijoita samanlaiseen ajatteluun. Dal Pozzon (mp.) mukaan subjektin käyttö ei kytkeydy ainoastaan syntaksiin vaan myös diskurssiin ja pragmatiikkaan. Tämän vuoksi generatiivisen kieliopin teoriataustassa subjektin käyttöön liittyvä tutkimusalueen sijainti on syntaksin ja diskurssin raja-alueella (*syntax-discourse interface*). Vielä jopa varsin edistyneiden toisen ja vieraan kielen puhujien välikielessä tämän alueen on huomattu olevan problemaattinen.

### 3.3 Verbien semanttinen luokittelu

Luokittelemalla aineiston verbejä muutenkin kuin niiden taitotason ja esiintymisfrekvenssin mukaan on tähän tutkimukseen saatu paljon lisätietoa siitä, miten erityyppiset verbit suhtautuvat ilmisubjektiin. Täten aineiston verbit on jaoteltu semanttisesti toisistaan poikkeaviin verbi-  
luokkiin. Iso suomen kielioppi jakaa verbit luonnehtimiensa tilanteiden mukaan konkreettisiin, mentaalisiin ja abstrakteihin. Merkitykseltään konkreettiset verbit osoittavat konkreettisesti huomattavia tiloja, toimia ja tekoja, esimerkiksi olemista, liikkumista, rakentamista ja rikkomista (*olen, juoksen, rakennan, hävitän*). Näin ollen kyseiset verbit saavat täydennykseen nominaalisia lausekkeita, koska kuvattujen tilanteiden osanottajat ovat normaalisti konkreettisia olioita, joko eläviä tai elollisia. Mentaaliset verbit vuorostaan luonnehtivat mielen sisäisiä tiloja, tekoja ja toimia (*näen, tykkään, yritän*). Ne vaativatkin elollisen, omaa tajuntaansa käyttävän osallistujan, joten yleensä niillä on elollistarkoitteinen täydennys, joka on normaalisti subjekti. Mentaaliset verbit jakaantuvat tunnetta, aistimusta ja tiedontilaa indikoiviin. Abstraktien verbien tehtävänä on puolestaan asiaintilojen suhteuttaminen keskenään enemmän kuin niiden kuvaaminen. Abstrakteja verbejä ovat muun muassa modaaliset verbit, kuten *joutua* ja *voida* sekä asioiden välisiä suhteita ilmaisevat, kuten *aiheutua* ja *johtua*. Myös abstraktit verbit esiintyvät infinitiivien ja lauseiden kanssa. Täysin selvää rajaa konkreettisten,



mentaalisten ja abstraktien välillä ei kuitenkaan ole, koska verbit voivat metaforisesti laventaa käyttöalaansa konkreettisesta abstraktiin päin. (VISK § 445.) Semanttisten luokkien lisäksi erotan tässä tutkimuksessa kieltoverbit omaksi luokakseen, joten olen jaotellut aineiston verbit neljään eri pääluokkaan. Vaikka kieltoverbi taipuukin samalla tavalla persoonassa ja luvussa kuin muutkin verbit, se ei kuitenkaan itsessään edusta mitään edellä mainittua semanttista luokkaa, vaan se toimii liittomuodon apuverbinä, jolloin semanttinen luokka olisi nähtävissä pääverbistä, joka on kiello- tai partisiippimuodossa (ks. VISK § 108). Toisaalta samalla tavoin apuverbinä toimivat myös abstraktit verbit, mutta olen halunnut erottaa kieltoverbit omaksi luokakseen, jotta näkisin, vaikuttaako kiello poikkeavalla tavalla ilmisubjektin käyttöön.

Tämän tutkimuksen verbien semanttista luokittelua tarkasteltaessa on otettava huomioon tutkimuksen lähtökohtaiset rajoitteet. Ensimmäiseksi se, että aineiston verbiesiintymistä on huomioitu vain yksikön 1. persoonan verbit ja toiseksi se, että ainoastaan sellaiset verbit on huomioitu, joiden kanssa subjekti esiintyy perussubjektina eli valinnan alaisena (ks. 3.2.1). Näin ollen esimerkiksi *minua janottaa* tai *minun täytyy* eivät ole huomioitavia esiintymiä. Yllä olevasta Ison suomen kieliopin verbien semanttisesta jaottelusta huomattavasti yksityiskohtaisempi on Pajusen (2001) Argumenttirakenne: asiaintilojen luokitus ja verbien käyttäytyminen suomen kielessä. Olen käyttänyt Pajusen ansiokasta tutkimusta verbien semanttisen luokittelun pääasiallisena pohjana, mutta kaikkia hänen käyttämiään luokkia en tässä työssä edellä mainituista syistä käytä, koska osa niistä jää tarpeettomiksi esiintymien puutteen vuoksi. Olenkin päättänyt yhdistelemään joitakin luokkia isompiin kategorioihin. Kuvatessani näitä semanttisia luokkia käyn niitä ainoastaan läpi yksikön 1. persoonan näkökulmasta, vaikkakin esimerkkiverbit ovat infinitiivissä. Pajusen luokittelussa koen ongelmallisimmaksi luokitteluksi puheaktiverbien sijoittamisen mentaaliin, vaikka esimerkiksi puhumisesta ja kirjoittamisesta seuraa selkeästi konkreettinen tuotos – ääntä ja tekstiä. On toki samaan hengenvetoon todettava, että kaikki kognitiiviset toiminnotkin voidaan joillakin aivotoimintaa seuraavilla laitteilla havaita jonkinlaisena värähtelynä. Lähtökohtani on kuitenkin ennen kaikkea siinä, että konkreettisia verbejä ovat kaikki sellaiset, jotka suoraan ilmaisevat jotakin, jonka joku toinen kuin verbin tekijä voi normaaleissa olosuhteissa havaita, tai verbin kuvaama toiminta vaikuttaa suoraan johonkin toiseen. Toisaalta kaikki sellaiset verbit, jotka lähtökohtaisesti jäävät tekijän – tai tässä tapauksessa paremminkin kokijan – omaan sisäiseen havaintomaailmaan, ovat mentaalisia. Näin ollen en kaikissa kohdin täysin nojaa Pajusen verbien argumenttirakenteeseen, vaan pidän konkreettisen verbin pääpiirteinä sen (tekijästä ulkopuolista) havaittavuutta. Esittelen seuraavassa tutkimuksessani käyttämäni semanttisen luokittelun luokka

luokalta siten, että jokaisen luokan lopussa on lista esimerkiverbejä, jotka esiintyvät Pajusen teoksessa tai tämän tutkimuksen aineistossa.

### 3.3.1 Konkreettiset verbit

#### 3.3.1.1 Tila

Muuttumatonta tilassa olemista indikoivat verbit jakaantuvat kahteen pääluokkaan – verbaalisesti ja nominaalisesti koodattuihin. Ensimmäiset ovat leksikaalisia verbejä ja jälkimmäiset edelleen ovat toki syntaksisesta näkökulmasta verbejä mutta semanttisesti miltei tyhjiä (*olla*-verbin sisältävät rakenteet). (Pajunen 2001: 96.) Tila-luokan leksikaaliset verbit ovat joko asentoverbejä (*seistä*) tai tilassa oloaikkaa tai tilassa pysymistä kuvaavia verbejä. Tila-luokkaan kuuluvat myös asumis- ja elämisverbit (*asua, elää*) sekä aspektuaaliset tilaverbit (*odottaa, jäädä, lomailla*). (Mts. 113, 117–118.) Tilaverbeihin liitän tämän tutkimuksen luokittelussa myös fysiologista tilaa ilmaisevat verbit, jotka pääasiassa kuvailevat jonkinlaisia vähemmän miellyttäviä tiloja tai tuntemuksia, tai sairautta ja kipua (*palelen, sairastan*) (mts. 275–277).

**Tilaverbejä:** *olla, nukkua, valvoa, istua, seisoa, riippua, maata, asua, elää, jäädä, pysyä, viipyä, odottaa, lomailla, palella, sairastaa, potea, kuulua (johonkin), noudattaa, pärjätä, viettää, toimia (jossakin roolissa).*

#### 3.3.1.2 Teko ja muutos

Pajunen (2001: 120) erottaa teon, tapahtuman ja muutoksen tarkemmassa luokittelussa omiksi alaryhmikseen, mutta lähtökohtaisesti toteaa kaikkien olevan aspektityypeiltään prosesseja. Tässä tutkimuksessa olen päätenyt pitämään tämän ryhmän yhdessä ja poistamaan luokasta tapahtumaverbit, koska luonnollisesti ne eivät esiinny yksikön 1. persoonassa – niillä ei ole tietoista aiheuttajaa (*sataa, tuulee ~ ?sadan, ?tuulen*) (ks. mts. 168). Prosessien takana voidaan nähdä aina olevan kausaatiota, jonka tyyppi vaihtelee volitionaalisesta fysikaaliseen ja vaikuttavaan. Suomen kielessä kausaatio on mahdollista koodata suoraan verbilekseemiin, mutta sen ilmaisu on mahdollista myös johdinteitse tai analyttisesti. Kausaatiota kuvaavilla

teko- ja muutosverbeillä koodataan kontrolloituja tapahtumaketjuja. Tekoverbit kuvaavat esimerkiksi jotakin aktiviteettia (*laulaa, tapella*), jonkin nauttimista (*juoda, syödä*), huoltamista (*pukea*), käsittelemistä (siivota) tai valmistamista (*rakentaa*). (Mts. 120, 147, 159, 162–164, 166–168.) Ihmisen suorittamat muutosverbit ovat sellaisia, jotka aiheuttavat selkeän muutoksen teon kohteen olomuodossa tai -tilassa tietyssä ajassa tai saman tien. Muutosverbeihin kuuluu ennen kaikkea verbejä, jotka indikoivat jonkin esineen tai muun vastaavan avaruudellisten ulottuvuuksien tai materiaalistien ominaisuuksien muokkaamista. Muutosverbeihin kuuluu rikkomis-, hävittämisen-, tuhoamis-, tappamis- ja myrkyttämisenverbejä. (Mts. 155.) Teon ja muutoksen ryhmään liitän myös fysiologiset tilan muutokset (*laihtua, loukkaantua*) ja toiminnot (*hengittää, aivastaa*), fysiologisperustaiset reaktiot (*nauraa, itkeä*) (mts. 275–282) sekä puheaktiverbeistä irrottamani opettamis- ja kirjoittamisverbit (*opettaa, kirjoittaa, piirtää*), koska näen ne ensisijaisesti teko- ja tapahtumaverbeinä. Pajusen (2001: 299, 339) mukaan opettamis- ja kirjoittamisverbit ovat kognitiivista toimintaa edellyttäviä teko- ja tapahtumaverbejä ja sikäli kompleksisia, että niillä on sekä konkreettisten että mentaalisten verbien käyttäytymisominaisuuksia. Nämä ovat nähdäkseni kuitenkin sellaisia verbejä, jotka lähtökohtaisesti suuntaavat toimintaa johonkin tekijästä ulkoiseen ja aiheuttavat tässä (heti tai ajan kuluessa) näkyvän muutoksen (*opettaa* – joku oppii; *kirjoittaa* – joku voi lukea tekstin tai ainakin paperi tai tekstiedosto muuttaa ulkonäköään).

**Teko- ja muutosverbejä:** *tehdä, juoda, syödä, auttaa, herätä, nukahtaa, laulaa, tapella, pelata, soittaa, esittää, jatkaa, roikkua, pukea, kokeilla, peseytyä, siivota, imuroida, leikata, rakentaa, korjata, muuttua, muuttaa, täyttää, rikkoa, syntyä, kuolla, hukkua, hävitä, hävittää, poistaa, häiritä, kalastaa, tuhota, tappaa, myrkyttää, tulla (joksikin), laihtua, loukkaantua, parantua, satuttaa, menettää, hengittää, aivastaa, oksentaa, sairastua, hoitaa, yskiä, nauraa, itkeä, hymyillä, punastua, opettaa, opiskella, riidellä, valmistautua, hukata, kadottaa, löytää, käyttää, tavata, kohdata, uhata, nopeuttaa, aloittaa, yhdistää, järjestää, tarjota, pelastaa, pyörtyä, riehua, valita, tuhlata, jäätyä, tarkentaa.*

### 3.3.1.3 Liike

Liikettä pidetään monesti ihmisen käsitteistä perustavimpana. On kuitenkin huomattava, että liikkeellä on sekä teon että tapahtuman piirteitä. (Pajunen 2001: 52.) Liike on mahdollista jakaa lokomotionaaliseen ja orientaation muutosta olettavaan. Lokomotionaalisella tarkoite-

taan pääasiallisesti maantieteellisen paikan muuttumista eli olion A liikkumista B:stä C:n kautta D:hen – tosin paikallaan liikkuminenkin voidaan katsoa lokomotionaalisen liikkeen alatyypiksi. Orientaation muutos taas kuvaa ruumiin tai ruumiinosan asennon muuttumista. Liike saattaa olla joko kontrolloitua (ihmisen omaa tai aikaansaamaa liikettä: *kääntyä*) tai kontrolloimatonta (*horjahtaa*). Lokomotionaalisessa liikkeessä liikkuja on itse oman liikkeen aiheuttaja (*kävellä*), kun taas dynaamisessa liikkeessä liikkuja ei itse kontrolloi liikkumistaan, vaan kyseessä on fyysikaalinen kausaatio (*pudota*) – kausatiivisessa liikkeessä puolestaan ilmaistaan liikuttamisen kohteen liikkumista (*kuljettaa*). Lisäksi liike kytkeytyy esimerkiksi kontaktiin (*koskettaa*, *lyödä*). Huomattavaa on, että liikeverbeihin kuuluvat myös antamisverbit, jotka ilmaisevat jonkin asian vapaaehtoista tai pakollista antamista tai lahjoittamista (luovuttaa, antaa) sekä vastaanottamista (*ottaa*, *saada*). Myös omistussuhteen vaihtumisen muutosta jotakin korvausta vastaan ilmaisevat verbit (*ostaa*, *vaihtaa*) ovat liikeverbejä. (Mts. 185–272.)

**Liikeverbejä:** *mennä / tulla (johonkin / jostakin), lähteä, liikkua, kävellä, käydä, juosta, hiiptä, kulkea, lentää, heilua, uida, hiihtää, luistella, pyöräillä, ajaa, tanssia, surffata, matkustaa, lähestyä, poistua, saapua, hypätä, kiivetä, horjahtaa, kompastua, kaatua, istuutua, täristää, hätkähtää, kääntyä, pudota, tönäistä, lyödä, kampata, eksyä, upota, ampua, vetää, karata, laskea (mäkeä), koittaa (jotakin), pitää (jotakin käsissään), lähettää, laittaa, pistää, käynnistää, saada, antaa, panna, asettaa, maksaa, ostaa, vaihtaa, lainata, tarjota, pidättää, kaapata, palauttaa, hakea, viedä, koota, puristaa, näyttää, jättää, johdattaa, shoppailla, torjua (laukaus), puskea, vilkaista, kurkkia, näyttäytyä, seurata, paketoita, driftailla, piiloutua, räpäyttää, avata, sulkea.*

#### 3.3.1.4 Puheakti

Kuten jo mainitsin ylläluvussa 3.3, koen puheaktiverbien luokittelun mentaaliseen kategoriaan siinä mielessä ongelmallisena, että monet puheaktiverbit voidaan myös siinä mielessä nähdä (tai paremminkin kuulla) konkreettisina, koska puheesta ja huudosta koituu ääntä sekä käskyt ja lupaukset suunnataan useimmiten jollekin muulle kuin tekijäsubjektille itselleen – toisin sanoen ne antavat mahdollisuuden siihen, että toinen kokijäsubjekti ne suoraan havaitsee. Huomionarvoista on myös juuri tämä seikka, että puheaktiverbien subjektina toimii tekijä- eikä kokijäsubjekti (ks. Pajunen 2001: 340) kuten Pajusen luokittelun muissa mentaaliver-

beissä. Vaikka olenkin sijoittanut puheaktiverbit konkreettisten verbien luokkaan, osan Pajusen luokittelun puheaktiverbien alaluokista olen joutunut irrottamaan kokonaan toiseen luokkaan. Esimerkiksi lukemisverbit olen sijoittanut perkeptioon ja oppimisverbit psykologiseen tilaan ja prosessiin, koska niissä toimintaa ei voi täydellisesti havaita muut kuin itse toimija. Toisaalta opettamisverbit sijoitan konkreettisiin teko- ja muutosverbeihin, koska opettaminen voidaan myös nähdä konkreettisena tekemisenä. Pajusen (2001: 339) mukaan puheaktiverbit sanallistavat kielellistä kanssakäymistä. Ne on mahdollista jakaa yleensä kommunikaatiota ja toisaalta puhumista ja pääasiallista puheaktia leksikalisoiviin verbeihin (*keskustella, kysyä, puhua, sanoa*) ja verbaalisiin teko- ja muutosverbeihin (*kutsua, käskeä, pyytää*).

**Puheaktiverbejä:** *puhua, sanoa, kertoa, laulaa, huutaa, kiljua, keskustella nimittää, kiittää, myöntää, kiistää, kieltää, käskeä, pyytää, luvata, suostua, tilata, uhata, kysyä, vastata, haukkua (herjata), pahoitella, hyvästellä, vinkua, tarkoittaa, vaatia, suositella, mainita.*

### 3.3.2 Mentaaliset verbit

#### 3.3.2.1 Psykologinen tila ja prosessi

Psykologisen tilan ja prosessin verbit ilmaisevat pääasiassa jonkin kokijasubjektin emootioita tai kognitiivista toimintaa. Emootiot leksikalisoivat tunnetiloja (*rakastaa*) ja -tapahtumia (*rakastua*), kun taas kognitiiviset toiminnot propositionaalisia näkökantoja (*luulla*) sekä tiedollisia tiloja ja prosesseja (*ajatella, harkita*). Molemmat ryhmittyvät kahden argumentin, kokijan ja kohteen, kanssa. Kohde näyttäytyy emootioverbien kohdalla fysiologisten ja perpektiivisten verbien ärsykettä muistuttavana; vuorostaan kognitioverbien yhteydessä kohde on paremmin ajattelutoiminnan tulos. Emootioverbeillä ilmaistaan usein jonkin tilan alkamista, siinä olemista tai sen loppumista, kun puolestaan kognitioverbeillä indikoidaan tiloja ja aktiviteetteja. (Pajunen 2001: 309.)

Emootiot nähdään fysikalistisen näkökulman mukaan ruumiillisina, joten ne näin ollen toimivat reaktioina todellisiin tuntemuksiin. Ihmisen neurofysiologiaan pohjaavia perustunteita on kuusi: viha, pelko, hämmästyminen tai yllättyneisyys, suru, ilo tai inho. Ne korreloivat muun muassa kasvonilmeisiin, mutta fysikalistinen näkökulma ei ole kuitenkaan kulttuurista riippumaton. Kognitiivisesta näkökulmasta emootiot ovat olemukseltaan mentaalisia, jolloin esimerkiksi viha on reaktio ajatukseen, että on kokenut vääryyttä. Kognitioverbit taas ovat

kielellinen keino tuoda esiin asiointiloihin suuntautuvia propositionaalisia asenteita ja tiedon laatua. Näitä ovat esimerkiksi luulo, uskomus, tieto ja arvelu. (Pajunen 2001: 309–313.)

**Psykologista tilaa ja prosessia ilmaisevat verbit:** *suuttua, vihata, iloita, rakastaa, rakastua, ihastua, ajatella, ihmetellä, harkita, harmittaa, tuntea, hävetä, tykätä, pitää (jostakin), tietää, muistaa, uskoa, toivoa, luulla, kuvitella, epäillä, ymmärtää, käsittää, aikoa, yrittää, meinata, jaksaa, onnistua, osata, pystyä, haluta, oppia, opetella, koittaa (tehdä jotakin), selvitä (jostakin henkisesti), tunnistaa, stressata, tutustua, päättää, kaivata, kannattaa (olla jonkun puolella), tarvita (jotakin).*

### 3.3.2.2 Perkeptio

Perkeptioverbit jakaantuvat näkö-, kuulo-, maku-, haju-, ja tuntoaistin perusteella viiteen eri alaluokkaan. Aistityypit poikkeavat sanastollisesti siinä mielessä, että etäaisteilla (näkö ja kuulo) on paljon enemmän sanastoa kuin lähiaisteilla (maku, haju ja tunto) – kyseessä on myös spesifisyysero, sillä etäaistit ovat käytännössä pitemmälle erikoistuneet. Perkeptioverbit osoittavat joko aistimiskykyä (*kuulen ja näen erinomaisesti*), havaintoa itseään (*kuulla / näh-dä jotain*) tai toimintaa (*katsoa / kuunnella jotain*). (Pajunen 2001: 319–321.)

**Perkeptioverbejä:** *nähdä, katsoa, tuijottaa, vilkaista, lukea, kuulla, kuunnella, tuntea, haistaa, maistaa.*

### 3.3.3 Abstraktit verbit

Pajunen mukaan (2001: 47) modaalisten verbien luokka on syntaktisesti varsin yksinkertaisesti rajattava, koska pääverbi saa infinitiivimääritteen, jota määrittää subjektina tai objektina toimiva argumentti. Pajunen ei kuitenkaan juuri tarkemmin käsittele tutkimuksessaan modaalisia tai abstrakteja verbejä. Suomen kielioppiperinteessä modaaliverbien luokka on tutkijasta riippuen ollut erilaajuinen (VISK § 1563). Esimerkiksi Flintin (1980) mukaan kaikki dynaamiset voimisverbit ovat modaalisia, kun taas Kangasniemi (1992) rajaa modaaliset verbit paljon tiukemmin. Pajunen (2001: 47) muistuttaakin, että monet modaaliverbit sisältyvät tai ovat sisältyneet mentaaliverbien luokkaan, mutta niiden tarkoitusta on vaikea määrittellä millään yhdellä yhteisellä nimittäjällä. Tässä tutkimuksessa noudatan suhteellisen tiukkaa rajausta

abstraktien ja modaalisten verbien suhteen. Modaalisten verbien tunnusmerkkinä pidän niiden abstraktiutta – tarkoittaen sitä, että niihin on suoranaisesti vaikea liittää jotakin yksittäistä tunnetta, kuten on laita psykologista tilaa tai prosessia kuvaavien verbien suhteen; modaaliset verbit ovat epätarkempia ja näin ollen niiden nimittäminen abstrakteiksi on perusteltua.

Yleensä modaaliverbiin liittyy infiniittisenä ei-modaalinen verbi. Mahdollisia ovat kuitenkin myös kahden modaaliverbin yhdistelmät, joista jälkimmäinen edustuu A-infinitiivinä: *pitää voida, saattaa joutua*. (VISK § 1564.) Tutkimusaineistosta löytyy myös esiintymiä, joissa infiniittisen verbin paikka jää tyhjäksi kuten esimerkissä (12).

12) Sen jälkeen minä joutuin (–) röntken kuvaukeen.

Aineistosta huomioitavien modaalisten verbien määrää rajoittaa se seikka, ettei genetiivisubjektin vaativia verbejä (*minun pitää, minun täytyy*) voida tässä tutkimuksessa ottaa lukuun (ks. 3.2.1).

**Abstrakteja verbejä:** *voida, joutua, saada (tehdä jotakin), taitaa, pystyä, päästä, tulla (olemaan, pääsemään jne.), ehtiä, keretä, alkaa, ruveta*

### 3.3.4 Kieltoverbit

Vaikka kieltoverbit itsessään eivät sisällä kiellon lisäksi muuta semanttista informaatiota (kieltoverbiin liittyvää pääverbiä ei tässä tutkimuksessa huomioida), on ne tässä tutkimuksessa otettu omaksi erilliseksi verbiluokakseen, koska kieltoja esiintyy aineistossa paljon. Kieltoverbi *ei* on epätäydellisesti taipuva apuverbi, jonka taivutus mukautuu perussubjektiin. Kieltoverbillä on persoonamuodot aktiivin indikatiivin preesensissä (*en, et, ei, emme, ette, eivät*) ja imperatiivissa (*älä, älkää, älkäämme, älköön, älkööt*), mutta sillä ei itsessään ole konditionaalia ja potentiaalia, eikä se taivu aikamuotojen mukaan. Kieltoverbin ainoa aikamuoto on preesens; modukset ja tempukset, joiden mukaan kieltoverbi ei taivu, ilmaistuvat pääverbin vartalossa (*en juone, en juonut, en olisi saanut juoda*). Kieltoauseen predikaatissa on apuverbinä toimiva kieltosana ja pääverbi kieltomuodossa. (Savijärvi 1977: 11–13; VISK § 1615, 1623.) Suomenoppijoiden negaatioilmauksen frekvenssiä ja oikeinkirjoitusta on tutkinut muun muassa Halttunen (2014). Tässä tutkielmassa keskitytään kuitenkin vain kieltoverbiin liittyvään ilmisubjektin käyttöön ja verrataan sitä edellä lueteltuihin semanttisiin verbiluokkiin.

## 4. AINEISTO JA METODIT

### 4.1 Tutkimusaineisto

Tässä tutkimuksessa olen käyttänyt Topling-hankkeen koko suomi toisena kielenä -korpusta. Topling-hanke oli Jyväskylän yliopiston kielten laitoksen ja Soveltavan kielentutkimuksen keskuksen tutkimushanke vuosina 2010–2013. Sen rahoittajina toimivat Suomen Akatemia ja Jyväskylän yliopisto. Topling-korpusta on kerätty kolme eri kierrosta suomalaisissa kouluissa vuosina 2010–2012. Tavoitteena on ollut, että jokainen kirjoittaja on kirjoittanut kolmena eri lukuvuotena samat tehtävät, joten korpus antaa myös mahdollisuuden diakroniseen pitkittäis-tutkimukseen. Topling keskittyy koululaisten kirjoitustaitoon ja siihen, miten kirjoitustaito suomalaisessa koulutusjärjestelmässä kehittyy. (Topling – Toisen kielen oppimisen polut; Jantunen & Pirkola 2015: 90.)

Aineiston keräyskierrokset on suoritettu seuraavalla tavalla: alakoulun 2.–4. luokka, alakoulun 4.–6. luokka, yläkoulun 7.–9. luokka ja lukion 1.–3. vuosikurssi. Kaikki tekstit on arvioitu Eurooppalaisen viitekehyksen taitotasolle A1:stä C2:een, mutta hyvin pieni osa teksteistä on myös arvioitu alle alimman taitotason A1 (näitä tekstejä en ole ottanut tutkimuksessani huomioon) ja toisaalta varsin harva teksti ylittää tasolle C1 ja C2. Tekstit on jaoteltu keräyskierroksittain, luokka-asteittain, taitotasoittain ja tehtävittäin. Kaikista teksteistä on kolmen eri asiantuntijan arvio. Topling-hankkeessa tekstien arviointi on eritelty Facets-ohjelman tilastollisten menetelmien avulla. Facetsin avulla pystytään esimerkiksi huomioimaan poikkeamat siinä, kuinka tiukasti kukin arvioija on arvioinut suoritukset. Topling-korpukselta on Facets-analyysin avulla jätetty pois ne tekstit, joiden arvioinnit eroavat huomattavasti toisistaan. Tehtäviä on seitsemän erilaista: viesti ystävälle, viesti opettajalle, viesti verkkokauppaan, mielipide, kertomus, viesti ystävälle ulkomaille ja viesti ystävälle ulkomailta. Näistä kaksi viimeistä tehtävänantoa on ollut tehtävänä vain 2.–4. luokkalaisille. Kaikki tehtävänannot ovat tutkijoiden määrittelemiä. Kaikki tekstit on kirjoitettu käsin ja myöhemmin ne on litteroitu sähköiseen muotoon (mahdollisine virheineen ja tavuviivoineen; jos litteroija ei ole saanut jostakin sanasta selvää on sekin merkitty erikseen – annotoituun muotoon aineistoa ei ole saatettu). Topling-aineiston suomi toisena kielenä -korpuksessa on kaiken kaikkiaan 2548 tekstiä ja 101 000 sanetta. (Topling – Toisen kielen oppimisen polut; Jantunen & Pirkola 2015: 97–98.) Tutkielman liitteessä 1 on nähtävillä esimerkki Topling-tehtävänannosta ja liitteessä 2 on kaksi esimerkkitekstiä.



## 4.2 Metodit

Ojutkankaan (2008: 277–278) mukaan aineistopohjainen tutkimus on perusteltua ennen kaikkea siksi, että sen avulla tutkija pääsee käsiksi sellaiseen tietoon, johon hän ei pelkän yksilöllisen kielitajunsa avulla voisi päästä. Tutkija voi harvoin nojata pelkästään omaan intuitioonsa selvittäessään kielitieteellisiä kysymyksiä. Intuitiota tutkimuksen tekemisessä kuitenkin ehdottomasti tarvitaan erityisesti silloin, kun arvioidaan, mitä kannattaa tutkia, ja kun tulkitaan aineistoa. Poikkeavat tavat käyttää toisistaan poikkeavia aineistoja ovat metodisia valintoja, joista koostuu jatkumo. Aineistot voivat myös toimia inspiraation herättäjinä niiden paljastamassa monipuolisia asioita kielenkäyttäjien tyyleistä luonnehtia ja kielentää heitä ympäröivää maailmaa. Toisistaan poikkeavia aineistoja voidaan koota toisenlaisia tutkimuskysymyksiä varten, mutta aina on mahdollista, että aineistot voivat antaa vastauksia sellaisiin kysymyksiin, joita tutkija ei ole etukäteen aineistolle asettanut.

Jantusen (2009: 101–102) mukaan kieliaineistoja apunaan käyttävät tutkimukset on mahdollista jaotella aineistoesimerkein tuettuihin (*corpus-illustrated*), aineistopohjaisiin (*corpus-based*) ja aineistovetoisiin (*corpus-driven*) analyyseihin. Näistä *aineistoesimerkein tuettu* tutkimus antaa kaikkein vähiten merkitystä itse aineistolle. Tällaisessa tutkimuksessa aineisto palvelee tutkijaa lähinnä materiaalipankkina, josta tutkija ammentaa teoriaan osuvia esimerkkejä luonnehdittavasta ilmiöstä. Useimmissa tapauksissa tämä tarkoittaa sellaisia esiintymiä, jotka vahvistavat tutkijan kieli-intuitiota ja teoriaa. *Aineistopohjainen* lähestymistapa poikkeaa edeltävästä selkeästi, koska todelliseen kielenkäyttöön pohjautuva materiaali on tällaisessa tutkimuksessa paljon merkittävämmässä asemassa. Aineistosta ei ainoastaan poimita esimerkkejä, vaan tutkimusta tehdään ennen kaikkea aineistoa tarkasti analysoiden. Lähtökohtana useimmiten on tutkijan muodostama havainto kielestä, tutkijan intuitio tai edeltäviin tutkimuksiin pohjaava teoria tai hypoteesi. Kaikkein suurimman arvon aineistolle antaa kuitenkin *aineistovetoinen* metodi, jonka pääajatuksena on havainnoida tutkimusmateriaalia yhtenäisenä kokonaisuutena, minkä myötä aineisto saa suuren painoarvon määriteltäessä tutkimushypoteesia, teorian muodostamisessa sekä ilmiöiden luokittelussa. Tällä metodilla ilmiön toistuvuudella ja frekvenssillä on tärkeä rooli – niiden voidaankin nähdä olevan kielellisten luokittelujen lähtökohtana.

Aineistoja käytettäessä ensisijaista on niiden variaatio. Aitoa puhuttua tai kirjoitettua kielenkäyttöä sisältävä aineisto ilmentää sekä laajaa että vaihtelevaa kuvaa tutkittavasta ilmiöstä. Näin ollen murrekorpuksiakin voidaan pitää eräänlaisena internetin esiasteena. Suuret

kokoelmat normittamatonta kieltä eri alueilta ohjaavat kielentutkijan oman yksilöllisen – eli aina jossakin määrin rajallisen – kielitaitonsa ulkopuolelle. Samaten internetissä tekstien säännönmukaisuus varioi laidasta toiseen. Aineistoista tutkija voi myös hakea legimitointia intuitiivisille ennakkokäsityksilleen ja kääntäen myös tietoa, joka kumoaa ennakkokäsitykset. (Ojutkangas 278–280.) Tämä tutkimus on lähtenyt liikkeelle ennen kaikkea siitä hypoteesista, että ilmisubjektin käyttö vähenee siirryttäessä alemmilta EVK:n taitotasoilta ylemmille. Kie- len rakenneosat ovat usealla tunnetulla ja usealla toistaiseksi tutkimattomallakin tavalla toisiinsa ja laajempaan käyttökontekstiinsa liitoksissa. Kontekstissa tapahtuvat asiat ovat yksittäisen kielentutkijan kieliosaamisen ulkopuolella. Toisaalta formalistisen näkökannan mukaan kielioppi pitää sisällään paljon enemmän kuin aineistoista voi havainnoida. (Mts. 282–283.)

Näkemykseni mukaan tässä tutkimuksessa on eniten aineistopohjaiseen tutkimukseen viittaavia piirteitä. Aineistopohjaisuuteen viittaa ennen kaikkea se, että tutkimuskysymykseni perustuvat oletukseen siitä, miten ilmisubjektia käytetään eri taitotasolla. Aineistopohjaisessa tutkimuksessa kieltä eritellään ennalta tiedettyjen luokittelujen avulla aineiston suurimman merkityksen ollessa sen kyvyssä näyttää edeltävien teorioiden puutteet sekä potentiaalissa tuoda esiin kvantitatiivista tietoa (Jantunen 2009: 103). Tutkimukseni tavoitteena on ennen kaikkea luoda jonkinlainen näkemys siitä, miten ilmisubjektin käyttö muuttuu siirryttäessä alemmilta taitotasoilta ylemmille, ja miten ilmisubjektia käytetään erilaisissa semanttisissa verbiluokissa. Tutkimukseni nojaakin melko laajasta aineistosta saatuun kvantitatiiviseen tietoon. Aineistovetoisessa metodissa hypoteesin ja tutkimuskysymysten hahmotteleminen saatetaan peräti jättää aineiston motivoimaksi (mp.), mikä voi olla peräti eräänlaista tiedonlouhintaa, jolloin aineistosta pyritään löytämään sille ominaisia toistuvia piirteitä ilman, että niitä olisi ennalta määritetty – monesti pyritään myös kontrastiivisuuteen, jolloin eri korpuksia verrataan toisiinsa (Ivaska 2014: 164). Jantusen (2009: 103–104) mukaan aineistovetoinen analyysi vaatiikin tuekseen ”korpus- ja tilastollisten menetelmien hallintaa”. Tutkimukseni on siis ennen kaikkea aineistopohjainen, koska olen lähtenyt siitä hypoteesista liikkeelle, että ilmisubjektien määrä vähenee mentäessä alemmalta taitotasolta ylemmälle. Eri semanttisten verbiluokkien ilmisubjektillisuutta olen päätenyt tarkastelemaan sitä kautta, että esiintymistä löytyisi ilmisubjektillisuuden suhteen poikkeamia eri verbeissä kokonaismäärällisesti ja eri taitotasojen sisällä.

Topling-aineistoa olen käyttänyt ensin perin manuaalisesti, koska olen käynyt Microsoft Office Word -tekstitiedostot läpi teksti kerrallaan alleviivaten jokaisen ensimmäisessä persoonassa olevan verbiesiintymän, jolle ilmisubjekti on mahdollinen. Ilmisubjektittomat muodot olen merkinnyt violetilla värillä ja ilmisubjektilliset vihreällä. Merkintätöön jälkeen olen

laskenut kaikki esiintymät taitotasoin ja tehtävittäin sekä luonut laskelmista taulukoita. Huomioimatta olen jättänyt tekstien otsikot, koska monesti otsikko oli sama kuin tekstin ensimmäinen lause, mutta vielä useammin koko otsikkoa ei ollut. Vasta myöhemmin olen tarkastellut aineistoa AntConc-korpusohjelmiston (Anthony 2016) avulla tehdessäni taulukot aineiston frekventeimmistä verbeistä. AntConcin työkaluista tähän tehtävään olen käyttänyt Word List -työkalua frekventeimpien verbien löytämiseen. Word List -työkalua olen ensin käyttänyt frekvenssijärjestyksessä (*Sort by Freq*) ja sitten vielä aakkosjärjestyksessä (*Sort by Word*) huomioidakseni eri aikamuodot ja virheellisesti kirjoitetut muodot. Jokaisen riittävän frekventin verbin kohdalla olen tarkastanut verbien esiintymät Concordance-työkalun avulla. Joissakin tapauksissa olen myös tarkastellut esiintymää Concordancen kautta File View -työkalulla, jolla olen vielä tarkastellut jotakin yksittäistä tekstiä, josta jokin esiintymä on peräisin. Koko prosessi olisi varmasti ollut nopeampi, jos olisin heti vain syöttänyt aineiston korpusohjelmaan, mutta tällöin en varmasti tuntisi varsin monipuolista tutkimusaineistoani niin hyvin kuin nyt, ja ajattelen myös – vaikka AntConcin Word Listin avulla virheellisten muotojen huomioiminen ei lopulta ole vaikeaa – etten olisi osannut huomioida virheellisesti kirjoitettuja muotoja niin hyvin kuin ne nyt olen huomionut. Vaikka monet tekstit ovatkin olleet samankaltaisia, on joukosta löytynyt varsin viihdyttävää luettavaa.

Koska eri verbit esiintyvät tutkimusaineistossani varsin erilaisella frekvenssillä ja eri taitotasoilla on eri määrä tekstiä, eivät pelkät prosenttiluvut anna riittävästi tietoa esiintymien vertailuun. Siksi olen vertaillut esiintymismääriä analysoidessani prosenttilukujen poikkeavuuden merkitsevyyttä Z-testillä. Z-testi pohjautuu normaalijakaumaan, jonka todennäköisyysarvoja käytetään eri prosenttilukujen vertailuun. Laskettaessa otannan keskiarvojen poikkeavuutta oletetaan populaation keskiarvojen jakauman olevan otoksen keskiarvon ympärillä. Z-arvo ilmoittaa keskiarvojen poikkeavuuden tilastollisen merkitsevyyden. Z-arvo kahdesta otoksesta lasketaan seuraavalla kaavalla:

$$z = \frac{P_1 - P_2}{\sqrt{P(1-P) \cdot \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}$$

$P_1$  = tutkittavan asian suhteellinen osuus 1. otoksessa (desimaalimuodossa)

$P_2$  = tutkittavan asian suhteellinen osuus 2. otoksessa (desimaalimuodossa)

$n_1$  = ensimmäisen otoksen koko

$n_2$  = toisen otoksen koko

Z-arvojen merkitsevyytasoja taulukoista tarkasteltaessa ovat kriittiset Z-arvot seuraavat:

riskitasolla 5 % 1,96

riskitasolla 1 % 2,58

riskitasolla 0,1 % 3,29

Tässä tutkimuksessa erilaisia esiintymisfrekvenssejä vertaillessani on merkitsevyyden taso asetettu 1 %:iin. Ollakseen tilastollisesti *merkitsevä* on Z-testin ilmoittaman tuloksen mahdollista normaalijakauman 1 %:n marginaaliin. Z-arvo on siis merkitsevä ollessaan vähintään 2,58 tai alle 3,29. Z-arvon ollessa 5 %:n riskitasolla, eli vähintään 1,96 mutta alle 2,58, on arvo *melkein merkitsevä*. Jos Z-arvo on riskitasolla 0,1 %, eli 3,29 tai enemmän, on arvo *erittäin merkitsevä*. Saataessa merkitsevä arvo on siis vain 1 %:n mahdollisuus, että kyse on sattumasta, melkein merkitsevän kohdalla on 5 %:n ja erittäin merkitsevän kohdalla 0,1 % mahdollisuus siihen, että kyse on sattumasta. (Anttila 2005: 250; Stangroom 2017.)

## 5. TULOKSET JA ANALYYSI

### 5.1 Johdanto tuloksiin ja analyysiin

Tässä luvussa käydään läpi tutkimuksen tuloksia. Tavoitteena on esittää tuloksia ensin yleisesti taitotasoa ja verbien semanttisia luokkia eritellen ja lopulta analysoiden eri semanttisten luokkien frekventimpiä verbejä ja niiden suhtautumista ilmisubjektiin. Aineiston verbeistä on otettu huomioon kaikki yksikön ensimmäisessä persoonassa esiintyvät verbimuodot (kaikki mahdolliset tempukset ja modukset). Myös virheellisesti kirjoitetut muodot on huomioitu, mikäli kontekstista voidaan selkeästi päätellä, että kirjoittaja on tarkoittanut ilmaista verbin yksikön ensimmäisessä persoonassa. Huomioon on otettu myös verbit, jotka on kirjoitettu seuraavan tai edellisen sanan kanssa yhteen. Alla on esimerkkinä taulukko 4 oikein ja ”virheellisesti” kirjoitetuista *olen*-verbin eri muodoista aineistossa taitotasolta A1. Käytän tässä luvussa myös aineistoesimerkkejä, jotka olen lainannut aina täysin sellaisenaan. Tämän vuoksi esimerkiksi joissakin lainatuissa tekstipätkissä ei ole pistettä lainauksen lopussa, kun joissakin toisissa taas on. Joissakin teksteissä ei puolestaan ole isoja kirjaimia ja joissakin suurin osa sanoista alkaa isolla kirjaimella. Kaikkiin tässä luvussa käytettyihin esimerkkiteksteihin on pyritty lainaamaan kokonaisia virkkeitä, vaikka edellä mainituista syistä virkerajat ovat joissakin teksteissä olleet tulkinnanvaraisia.

Tämän luvun taulukoissa käytetään seuraavia lyhenteitä: *Ei-Is* = ilmisubjektiton esiintymä, *Is* = Ilmisubjektillinen esiintymä, *Yht* = yhteensä, *y. l.* = yksikön 1. persoonan esiintymä, *Te-mu* = teko- ja muutosverbit, *Pstpr* = psykologista tilaa ja prosessia kuvaavat verbit, *Perkep* = perkeptioverbit, *Abstr* = abstraktit verbit.

Taulukko 4. Olen-verbin kaikki hyväksytyt muodot taitotasolla A1.

verbi	Ei-Is	Is	Yht
olen	151	76	227
ole	6	10	16
olan	2	0	2
oleenpoika	1	0	1
olemajja	1	0	1
olenMatti	1	1	2
olensairas	2	0	2

verbi	Ei-Is	Is	Yht
olin	35	17	52
olinkipeä	0	1	1
olinsairas	1	0	1
ollen	1	1	2
ollin	0	1	1
oli n	1	0	1
oon	1	5	6
<b>olen (yhteensä)</b>	203	112	315
<b>% (yhteensä)</b>	64,4 %	35,6 %	100 %

Kuten edellä mainitsin ja yllä olevasta taulukosta 4 voidaan nähdä, tekstin lausekontekstia tutkailemalla myös oikeinkirjoituksesta poikkeavat muodot on voitu pääosin hyväksyä. Esimerkiksi virheellinen *ole*-muoto voitaisiin kontekstista irrotettuna liittää vain kieltoverbin apuverbiksi (*en ole sairas*), mutta yksikön ensimmäiseksi persoonaksi hyväksytyssä muodossa se esiintyy esimerkiksi näin: *Hei ole Maija ydekesän vuotias tyttö*. Analysoidessani seuraavissa alaluvuissa eri taitotasojen sekä eri semanttisten verbiluokkien ja verbien eroavaisuuksia kiinnitän erityisesti huomiota ilmisubjektin esiintyvyyden frekvenssiin, mutta sen lisäksi testaan paikoin esiintymien poikkeavuuden tilastollista merkitsevyyttä Z-testillä (ks. 4.2).

## 5.2 Ilmisubjektin esiintyvyys eri taitotasoilla

Tämän tutkimuksen keskeisiä lähtökohtia on ilmisubjektin esiintyvyyden tarkasteleminen eri taitotasojen välillä. Kuten alla olevasta taulukosta 5 on nähtävissä, on ilmisubjektien määrässä selkeästi vähenevä trendi kuljettaessa alimmasta taitotasosta ylimpään.

Taulukko 5. Ilmisubjektin esiintyvyys eri taitotasoilla.

Taitotaso	A1	A2	B1	B2	C1 & C2
<b>Is/y. 1.</b>	358/817	723/2358	344/1958	76/925	4/139
<b>%</b>	43,8 %	30,7 %	17,6 %	8,2 %	2,9 %

Taulukosta 5 näemme taitotasojen A1 ja A2 välillä ilmisubjektien määrän laskevan 43,8 %:sta (358/817) 30,7 %:iin (723/2358). Z-arvo näiden taitotasojen välillä on 6,84 eli erittäin merkitsevä. Siirryttäessä tasolta A2 tasolle B1 putoaa ilmisubjektien määrä 30,7 %:sta 17,6

:%iin (344/1958). Tällä siirtymällä Z-arvo on 9,93 eli vielä erittäin merkitsevämpi kuin edellinen siirtymä. Tasolta B1 tasolle B2 siirryttäessä ilmisubjektien määrä vähenee 17,6 %:sta 8,2 %:iin (76/925). Tämäkin siirtymä on edelleen erittäin merkitsevä Z-arvon ollessa 6,64. Mentäessä tasolta B2 tasolle C1 & C2 vähenee ilmisubjektien määrä 8,2 %:sta 2,9 %:iin (4/139). Näiden tasojen välinen Z-arvo on kuitenkin vain 2,23 eli melkein merkitsevä. Tämän taitotasovälin eroavaisuus jää siis selkeästi paljon pienemmäksi kuin edellisten. Suurin merkitsevyys on tasojen A2 ja B1 välillä. Tasojen C1 & C2 otanta on tutkimukseni aineistossa sen verran pieni, että kiinnitän siihen hiukan vähemmän huomiota kuin alempiin taitotasoisiin. Kuitenkin tälläkin otannalla voidaan sanoa, että ilmisubjektin käyttö tasoilla C1 & C2 on erittäin vähäistä.

### 5.2.1 Ilmisubjektin esiintyvyys taitotasolla A1

Alla olevasta taulukosta 6 voimme nähdä, että taitotasolla A1 yksikön 1. persoonan verbejä esiintyy kaiken kaikkiaan 817, joista 358 (43,8 %) on ilmisubjektillisia. Voidaankin sanoa, että taitotasolla A1 ilmisubjektia esiintyy varsin taajaan. Osa A1-taitotason semanttisista verbiluokista kuitenkin poikkeaa toisistaan varsin selkeästi ilmisubjektillisuudessa.

Taulukko 6. Ilmisubjektin esiintyvyys taitotasolla A1.

A1	Tila	Te-mu	Liike	Puhe	Pstpr	Perkep	Abstr	Kielto	Yht
Is/y. 1.	124/341	22/39	52/113	6/10	84/144	5/12	11/27	54/131	358/817
%	36,4 %	56,4 %	46,0 %	60,0 %	58,3 %	41,7 %	40,7 %	41,2 %	43,8 %

Taulukko 6 näyttää, että korkein esiintyvyyssprosentti on esimerkin (13) puheverbien kohdalla (60,0 %) ja matalin esimerkin (14) tilaverbeissä (36,4 %).

13) niin känykät pois **Minä ehdotin**

14) **asun** solkikatu

Puheverbejä on kuitenkin sen verran vähän, etteivät ne Z-testin mukaan poikkea tilastollisesti merkitsevästi tilaverbeistä ( $Z = 1,53$ ) tai kaikista A1 taitotason yksikön 1. persoonan verbeistä yhteensä ( $Z = 1,02$ ). Tilaverbien ja prosentuaalisesti toiseksi eniten ilmisubjektillisena esiintyvien (58,3 %) esimerkin (15) psykologista tilaa ja prosessia ilmaisevien verbien välillä on kuitenkin erittäin merkitsevä ero ( $Z = 4,47$ ).

15) **minä tykkään** Big brother pelistä koska se on kiva!

Tilaverbit poikkeavat kuitenkin kaikista taitotason A1 esiintymistä vain melkein merkitsevästi ( $Z = 2,35$ ), kun taas psykologisella tilalla ja prosessilla on kaikkiin verrattuna merkitsevä ero ( $Z = 3,22$ ).

### 5.2.2 Ilmisubjektin esiintyvyys taitotasolla A2

Taulukosta 7 voimme nähdä, että yksikön 1. persoonan verbejä esiintyy A2-taitotasolla 2358 kappaletta, joista 723 (30,7 %) on ilmisubjektillisia. Monien eri semanttisten verbiluokkien välillä ilmisubjektien esiintyvyys on selkeästi poikkeavaa.

Taulukko 7. Ilmisubjektin esiintyvyys taitotasolla A2.

A2	Tila	Te-mu	Liike	Puhe	Pstpr	Perkep	Abstr	Kielto	Yht
Is/y. 1.	190/799	67/128	154/405	26/69	114/336	21/61	31/121	120/439	723/2358
%	23,8 %	52,3 %	38,0 %	37,7 %	33,9 %	34,4 %	25,6 %	27,3 %	30,7 %

Taulukosta 7 voimme havaita, että esimerkin (16) teko- ja muutosverbit esiintyvät ilmisubjektillisina taitotasolla A2 prosentuaalisesti kaikkein taajimpaan (52,3 %) ja esimerkin (17) tilaverbit jälleen kaikkein vähiten (23,8 %).

- 16) **Minä syön** koulussa aina!  
 17) Moi **Oon** Maija ja **oon** tyttö.

Niiden välinen poikkeama on Z-testin mukaan erittäin merkitsevä ( $Z = 6,70$ ). Teko- ja muutosverbit poikkeavat myös erittäin merkitsevästi ( $Z = 5,13$ ) tason A2 esiintymien yhteenlasketusta määrästä. Nämä ovat suurimmat yhden taitotason sisäiset Z-arvopoikkeamat eri semanttisten verbiluokkien välillä mukaan lukien kaikki taitotasot, joten voidaan sanoa tason A2 sisältävän suurinta vaihtelua ilmisubjektien käytössä eri semanttisten luokkien välillä. Tilaverbien kohdalla poikkeama kaikista tason A2 esiintymistä on myös erittäin merkitsevä ( $Z = 3,71$ ). Merkitsevästi ilmisubjektillisempia ( $Z = 2,94$ ) yhteenlasketusta määrästä ovat esimerkiksi (18) liikeverbit.

- 18) kerran **minä kävelin** kadulla.



### 5.2.3 Ilmisubjektin esiintyvyys taitotasolla B1

Taulukosta 8 on nähtävissä, että yksikön 1. persoonan verbejä on B1-taitotasolla kaiken kaikkiaan 1958, joista 344 (17,6 %) esiintyy ilmisubjektillisena. Ilmisubjektin esiintyvyydessä ei ole enää nähtävissä niin suuria poikkeamia eri semanttisten verbiluokkien välillä kuin alemmilla taitotasoilla.

Taulukko 8. Ilmisubjektin esiintyvyys taitotasolla B1.

B1	Tila	Te-mu	Liike	Puhe	Pstpr	Perkep	Abstr	Kielto	Yht
Is/y. 1.	90/577	30/129	70/301	17/97	49/333	8/41	18/126	62/354	344/1958
%	15,6 %	23,3 %	23,3 %	17,5 %	14,7 %	19,5 %	14,3 %	17,5 %	17,6 %

Taulukko 8 kertoo, että prosentuaalisesti kaikkein eniten ilmisubjekteja esiintyy esimerkkien (19) liikeverbeissä (23,3 %) ja (20) teko- ja muutosverbeissä (23,3 %). Vähiten ilmisubjektillisena esiintyvät esimerkin (21) abstraktit verbit (14,3 %).

19) Mä näin audon tulossa mua päin ja **mä hyppäsin** ruohikkoon.

20) Aika pian syönnin jälkeen sisaruksillani sattui mahaan ja **minä sairastuin** vatsatautiin!

21) **Joudun** huolehtimaan pikkuveljestäni, koska vanhempani ovat iltavuorossa

Huomioin liikeverbit taajemmin ilmisubjektillisena kuin teko- ja muutosverbit niiden suuremman esiintymismäärän perusteella, mitä myös Z-testi tukee. Liikeverbien ja abstraktien verbien poikkeama on kuitenkin vain melkein merkitsevä ( $Z = 2,09$ ). Suurin poikkeama onkin liikeverbien ja esimerkin (22) tilaverbien välillä oleva merkitsevä ero ( $Z = 2,79$ ).

22) **Seisoin** bussin keski osassa missä ovet aukeavat.

Lähelle tällaista poikkeamaa päästään myös vertaamalla liikeverbejä esimerkin (23) psykologista tilaa ja prosessia ilmaiseviin verbeihin, joiden välillä on myös merkitsevä ero ( $Z = 2,75$ ).

23) **Pidän** enemmän vähän lyhyemmistä takkimalleista.

Kuitenkin kaikkien verbien yhteenlasketusta ilmisubjektien esiintymismäärästä vain liikeverbit poikkeavat melkein merkitsevästi ( $Z = 2,37$ ), joten voidaan sanoa, että taitotasolla B1 semanttiset luokat eivät poikkea suuresti toisistaan.

### 5.2.4 Ilmisubjektin esiintyvyys taitotasolla B2

Taulukko 9 kertoo, että yksikön 1. persoonan verbejä on 925 B2-taitotasolla, joista 76 (8,2 %) esiintyy ilmisubjektillisena. Prosentuaalisesti eri semanttisten verbiluokkien välillä ilmisubjektillisuudessa on suurempia eroja kuin taitotasolla B1, mutta Z-testin mukaan nämä poikkeamat ovat tilastollisesti vähemmän merkitseviä.

Taulukko 9. Ilmisubjektin esiintyvyys taitotasolla B2.

<b>B2</b>	<b>Tila</b>	<b>Te-mu</b>	<b>Liike</b>	<b>Puhe</b>	<b>Pstpr</b>	<b>Perkep</b>	<b>Abstr</b>	<b>Kielto</b>	<b>Yht</b>
<b>Is/y. 1.</b>	14/255	6/55	13/144	6/42	12/175	0/30	8/85	17/139	76/925
<b>%</b>	5,5 %	10,9 %	9,0 %	14,3 %	6,9 %	0 %	9,4 %	12,2 %	8,2 %

Taulukosta 9 voimme huomata, että taitotasolla B2 keskiarvoisesti yksikön 1. persoonan ilmisubjektia esiintyy vain 8,2 %:ssa kaikista yksikön 1. persoonan esiintymistä. Prosentuaalisesti eniten ilmisubjekteja esiintyy esimerkin (24) puheverbeissä (14,3 %), vähiten taas esimerkin (25) perkeptioverbeissä (0 %).

24) Minulla oli synttärjuhlat ja **minä kutsuin** minun viisi parasta kaveria.

25) Viiden minuutin jälkeen **nään** hissien ikkunasta että äitini juoksee portaita ylös Alas.

Puhe- ja perkeptioverbien välinen ero on Z-testin mukaan kuitenkin vain melkein merkitsevä ( $Z = 2,16$ ). Z-arvoltaan toisiinsa verrattuna eniten ääripäissä ovat esimerkin (26) kieltoverbit verrattuna esimerkin (27) tilaverbeihin, mutta niidenkin välinen ero on vain melkein merkitsevä ( $Z = 2,37$ ).

26) Anteeksi, mutta **mä en** valitettavasti pääse lauantaina keskustaan.

27) **Istun** melkein koko vapaa-aika koneella ja unohdan joskus läksyjen tekemisen.

Koko taitotason yhteenlasketusta ilmisubjektien keskiarvosta eniten poikkeavat perkeptioverbit ( $Z = 1,64$ ) ja kieltoverbit ( $Z = 1,56$ ) – nämä poikkeamat eivät kuitenkaan ole tilastollisesti merkitseviä.

### 5.2.5 Ilmisubjektin esiintyvyys taitotasolla C1 ja C2

Taulukko 10 kertoo, että taitotasolla C1 ja C2 esiintyy yhteensä 139 yksikön 1. persoonan verbiä, joista 4 (2,9 %) esiintyy ilmisubjektillisena. Ilmisubjektillisuudessa on eri semanttisten

verbiluokkien välillä prosentuaalisesti kohtuullisia eroja, mutta esiintymät ovat määrältään sen verran pieniä, ettei tilastollista merkitsevyyttä Z-testillä eri luokkien välille saada, joten kovinkaan hedelmällistä taitotasojen C1 ja C2 analysoiminen ei ole.

Taulukko 10. Ilmisubjektin esiintyvyys taitotasoilla C1 ja C2.

C1-C2	Tila	Te-mu	Liike	Puhe	Pstpr	Perkep	Abstr	Kielto	Yht
Is/y. 1.	2/32	1/11	0/20	0/7	1/27	0/13	0/9	0/20	4/139
%	6,3 %	9,1 %	0 %	0 %	3,7 %	0 %	0 %	0 %	2,9 %

Taulukosta 10 on nähtävissä, että prosentuaalisesti eniten ilmisubjekti esiintyy esimerkin (28) teko- ja muutosverbeissä. Ilmisubjektia ei esiinny lainkaan liike- (29), puhe- (30), perkeptio- (31) ja kieltoverbeissä (32) eikä abstrakteissa verbeissä (33).

28) Lopulta **minä** nukahdin.

29) **Käännysin** hetken päästä pois.

30) Vielä kerran **pyydän** hartaasti anteeksi, että en selvittänyt poissaoloani aikaisemmin.

31) Joskus **kuulen** rapinaa, joka ei tule oikein mistään.

32) **En** tiedä kannattaako uskoa niihin.

33) **Voinko** myös laittaa jotain kuvia työhöni?

### 5.2.6 Ilmisubjektin esiintyvyys yhteensä kaikilla taitotasoilla

Taulukko 11 näyttää, että yksikön 1. persoonan verbejä on kaiken kaikkiaan 6197, joista 1505 (24,3 %) esiintyy ilmisubjektillisina.

Taulukko 11. Ilmisubjektin esiintyvyys yhteensä kaikilla taitotasoilla.

Yht	Tila	Te-mu	Liike	Puhe	Pstpr	Perkep	Abstr	Kielto	Yht
Is/	420/	126/	289/	55/	260/	34/	68/	253/	1505/
y. 1.	2004	362	983	225	1015	157	368	1083	6197
%	21,0 %	34,8 %	29,4 %	24,4 %	25,6 %	21,7 %	18,5 %	23,4 %	24,3 %

Tarkasteltaessa taulukosta 11 ilmisubjektien yhteenlaskettua esiintyvyyttä eri semanttisissa luokissa voidaan huomata, että luokkien välillä on eroavaisuuksia. Prosentuaalisesti eniten ilmisubjektia käytetään teko- ja muutosverbien kanssa (34,8 %) ja vähiten abstraktien verbien kanssa (18,5 %). Näiden välinen ero Z-testauksella on erittäin merkitsevä ( $Z = 4,99$ ). Suurin ääripääpoikkeama on Z-testin mukaan kuitenkin teko- ja muutosverbien ja tilaverbien välillä oleva erittäin merkitsevä ero ( $Z = 5,76$ ). Teko- ja muutosverbit esiintyvät ilmisubjektin kera kaikkien verbien yhteenlasketussa ilmisubjektien määrästä erittäin merkitsevästi enem-

män ( $Z = 4,50$ ). Näin on asian laita myös verrattaessa liikeverbejä kaikkiin verbeihin ( $Z = 3,44$ ). Tilaverbien kanssa ilmisubjekti esiintyy merkitsevästi vähemmän kuin kaikkien verbien kanssa ( $Z = 3,06$ ). Verrattaessa abstraktien verbien ilmisubjektin esiintyvyyttä kaikkien verbien yhteenlaskettuun ilmisubjektin esiintyvyyteen jää  $Z$ -arvo kuitenkin niukasti vain melkein merkitsevään eroon ( $Z = 2,54$ ). Puhe-, psykologista tilaa ja prosessia ilmaisevat sekä perkeptio- ja kieltoverbit poikkeavat kaikkien verbien yhteenlasketusta melko vähän, koska vertailtaessa niitä kaikkiin verbeihin yhteensä, ei yhdenkään kohdalla saada  $Z$ -arvoa, joka olisi tilastollisesti merkitsevä.

### 5.3 Ilmisubjektin esiintyvyys neljässä semanttisessa pääverbiluokassa

Verbien lokerointi isoihin semanttisiin verbiluokkiin ei lopulta tuo esille kovinkaan merkitsevästi poikkeavuutta eri luokkien välillä (luokittelusta ks. 3.3). Kuitenkin taulukosta 12 voidaan nähdä, että abstraktit verbit saavat, lukuun ottamatta taitotasoa B2, osakseen vähiten ilmisubjekteja. Taitotasolla A1 ja A2 mentaaliset verbit ovat eniten ilmisubjektillisia, mutta taitotasolla B1 ja C1 & C2 ilmisubjektillisimpia ovat konkreettiset verbit – taitotasolla B2 kieltoverbit ovat kaikkein ilmisubjektillisimpia.

Taulukko 12. Ilmisubjektin esiintyvyys semanttisissa yläluokissa ja kieltoverbeissä.

Taitotaso	Konkreettiset verbit	Mentaaliset verbit	Abstraktit verbit	Kieltoverbit
<b>A1 Is/y. 1.</b> %	204/421 48,5 %	90/154 58,4 %	11/27 40,7 %	54/131 41,2 %
<b>A2 Is/y. 1.</b> %	437/1401 31,2 %	135/397 34,0 %	31/121 25,6 %	120/439 27,3 %
<b>B1 Is/y. 1.</b> %	207/1104 18,8 %	57/374 15,2 %	18/126 14,3 %	62/354 17,5 %
<b>B2 Is/y. 1.</b> %	39/496 7,9 %	12/205 5,9 %	8/85 9,4 %	17/139 12,2 %
<b>C1&amp;C2 Is/y. 1.</b> %	3/70 4,3 %	1/40 2,5 %	0/9 0 %	0/20 0 %
<b>Yht Is/y. 1.</b> %	890/3492 25,5 %	295/1170 25,2 %	68/368 18,5 %	253/1083 23,4 %

$Z$ -testauksella isojen luokkien välillä ilmisubjektillisuuden kokonaismäärissä poikkeavat toisistaan eniten konkreettiset ja abstraktit verbit, joiden ilmisubjektillisuudessa on merkitsevä ero ( $Z = 2,96$ ). Myös mentaaliset verbit ovat merkitsevästi ( $Z = 2,65$ ) ilmisubjektilli-

sempia kuin abstraktit verbit. Kieltoverbit eivät kuitenkaan poikkea Z-testauksen mukaan abstrakteista verbeistä edes melkein merkitsevästi. Taitotasojen sisällä tilastollisesti merkitseviä poikkeamia ilmisubjektillisuudessa on melko vähän. Taitotasolla A1 mentaaliset verbit saavat ilmisubjektin merkitsevästi ( $Z = 2,90$ ) useammin kuin kieltoverbit. Taitotasolla A2 mentaaliset verbit ovat melkein merkitsevästi ( $Z = 2,09$ ) ilmisubjektillisempia kuin kieltoverbit. Taitotasolla B2 puolestaan kieltoverbit ovat melkein merkitsevästi ( $Z = 2,09$ ) ilmisubjektillisempia kuin mentaaliset verbit. Taitotasoilla B1 sekä taitotasoilla C1 & C2 ei Z-testauksella saada eri yläluokkien välille tilastollisesti merkitseviä eroja.

## 5.4 Ilmisubjektin esiintyvyys kahdeksassa semanttisessa verbiluokassa

### 5.4.1 Ilmisubjektin esiintyvyys tilaverbeissä

Tilaverbit ovat esiintymismäärältään tutkimukseni aineiston suurin semanttinen verbiluokka. Taulukosta 13 näemme ilmisubjektin esiintymisen yksikön 1. persoonan frekventeimpien tilaverbien kanssa. Luonnollisesti koko aineiston frekventein yksikön 1. persoonan verbi *olen* käsittää valtaosan kaikista tilaverbien esiintymistä ( $1792/2004 = 89,4\%$ ).

Taulukko 13. Ilmisubjektin esiintyvyys tilaverbeissä.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>olen</b>	112/315 35,6 %	157/662 23,7 %	72/538 13,4 %	14/249 5,6 %	0/28 0 %	355/1792 19,8 %
<b>asun</b>	8/13 61,5 %	8/22 36,4 %	2/15 13,3 %	0/4 0 %	0/0 0 %	18/54 33,3 %
<b>muut tilaverbit</b>	4/13 30,8 %	25/115 21,7 %	16/24 66,7 %	0/2 0 %	2/4 50,0 %	47/158 29,7 %
<b>tilaverbit yhteensä</b>	124/341 36,4 %	190/799 23,8 %	90/577 13,6 %	14/255 5,5 %	2/32 6,2 %	420/2004 21,0 %

Vertailtaessa taulukossa 13 esiintyviä frekventeimpiä tilaverbejä esimerkkien (34) *olen* ja (35) *asun* ilmisubjektillisuutta kaikkien taitotasojen yhteenlasketuissa määrissä saadaan näiden välille Z-testillä melkein merkitsevä poikkeama ( $Z = 2,44$ ). Näin on myös asian laita hiukan lievemmin verrattaessa *asun*-verbiä kaikkiin tilaverbeihin ( $Z = 2,19$ ). *Olen* poikkeaa esimerkkien (36) ja (37) muiden tilaverbien ryhmästä merkitsevästi ( $Z = 2,96$ ).

- 34) **Olen** tosi pahoillani että joudun peruun tämän näin.  
 35) **Asun** Suomessa ja minun tieni on Solkikatu.  
 36) Kun **Minä nukuin** sohvalla ja min tippuin sohvalta.  
 37) **minä** kerran **istuin** parvekkeen laidalla ja tipuin alas se ei sattunut.

Muiden ryhmä poikkeaa ilmisubjektillisuudessa myös merkitsevästi tilaverbien yhteenlasketusta ilmisubjektillisuudesta ( $Z = 2,58$ ). Muiden tilaverbien ryhmässä erittäin mielenkiintoinen tilastopoikkeama voidaan huomata taitotasolla B1, jossa peräti kaksi kolmesta (16/24) muusta tilaverbistä saa ilmisubjektin. Verrattuna muiden ryhmään taitotasolla B1 on *olen*-verbi erittäin merkitsevästi vähemmän ( $Z = 7,03$ ) ilmisubjektillinen kuten myös tilaverbit yhteensä ( $Z = 6,43$ ). Myös *asun* esiintyy B1-taitotasolla merkitsevästi vähemmän ilmisubjektillisena kuin muitten ryhmä ( $Z = 3,25$ ).

#### 5.4.2. Ilmisubjektin esiintyvyys teko- ja muutosverbeissä

Teko- ja muutosverbit ovat esiintymismäärältään tutkimukseni aineiston kuudenneksi suurin semanttinen verbiluokka. Kuten taulukosta 14 voidaan nähdä, teko- ja muutosverbeissä esiintyy aineistossa paljon variaatiota: *teen*- ja *kirjoitan*-verbien lisäksi mitään muuta yksittäistä temu-verbiä ei esiinny riittävästi, jotta niiden tarkasteleminen yksittäisenä olisi tilastollisessa mielessä perusteltua. *Teen*-verbi käsittää kaikista temu-verbeistä 12,2 % ja *kirjoitan* 9,9 %, joten peräti 77,9 % aineiston esiintymistä on muita temu-verbejä.

Taulukko 14. Ilmisubjektin esiintyvyys teko- ja muutosverbeissä.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>teen</b>	3/3 100 %	2/8 25,0 %	1/20 5,0 %	2/12 16,7 %	0/1 0 %	8/44 18,2 %
<b>kirjoitan</b>	1/2 50,0 %	4/10 40,0 %	2/16 12,5 %	0/8 0 %	0/0 0 %	7/36 19,4 %
<b>muut temu- muverbit</b>	18/34 52,9 %	61/110 55,5 %	27/93 29,0 %	4/35 11,4 %	1/10 10,0 %	111/282 39,4 %
<b>temuverbit yhteensä</b>	22/39 56,4 %	67/128 52,3 %	30/129 23,3 %	6/55 10,9 %	1/11 9,1 %	126/362 34,8 %

Taulukkoa 14 tarkasteltaessa esimerkkien (38) *teen* ja (39) *kirjoitan* näyttävät prosenttiluvuista päätellen olevan selkeästi vähemmän ilmisubjektillisiä kuin temu-verbit kokonaisuudessaan, mutta  $Z$ -testin mukaan vain *teen* ( $Z = 2,21$ ) kohdalla ero kaikkiin temu-verbeihin on

melkein merkitsevä, kun taas *kirjoitan* ( $Z = 1,86$ ) kohdalla ei aivan päästä tilastolliseen merkitsevyyteen.

38) Koska **teen** kokeen ja voisitko kertoa mistä tuli läksyksi.

39) **Kirjoitan** koska takki on minulle pieni ja en pidä tästä väristä.

*Teen* on kuitenkin merkitsevästi vähemmän ( $Z = 2,71$ ) ilmisubjektillinen kuin esimerkin (40) muiden temu-verbien ryhmä. Toisaalta *kirjoitan* poikkeaa ilmisubjektillisuudessa muiden ryhmästä vain melkein merkitsevästi ( $Z = 2,33$ ).

40) **Minä syön** koulussa aina!

Varsin varovaisesti voidaan väittää, että temu-verbien kohdalla ilmisubjektillisuus kasvaa sitä enemmän, mitä vähemmän frekventti verbi on kyseessä. Yhden taitotason sisällä yksittäisten verbien esiintymät ovat sen verran pieniä, ettei niitä ole syytä tarkastella tarkemmin. Mielenkiintoista on kuitenkin se, ettei temu-verbien ilmisubjektillisuus juurikaan vähene siirryttäessä taitotasolta A1 tasolle A2; muissa semanttisissa luokissa tässä siirtymävälissä ilmisubjektien määrä vähenee selkeästi enemmän, temu-verbeissä ei kokonaisuudessaan juuri lainkaan – muiden temu-verbien ryhmässä ilmisubjektien osuus tällä taitotasovälillä peräti hiukan nousee.

### 5.4.3. Ilmisubjektin esiintyvyys liikeverbeissä

Liikeverbit ovat tutkimukseni aineiston neljänneksi suurin semanttinen verbiluokka. Kuten taulukko 15 näyttää, aineistossa esiintyy useampia eri liikeverbejä, joiden esiintyvyys on kohtuullista (eli mahtuvat 24:n frekventeimmän yksikön 1. persoonan verbin joukkoon). Liikeverbien ilmisubjektillisuus vähenee taitotasoittain melko säännönmukaisesti. Myös liikeverbeissä esimerkkien (41) ja (42) muut verbit esiintyvät ilmisubjektillisempina kuin luokan frekventeimmät. Frekventeimmistä esimerkin (43) *tulen* esiintyy ilmisubjektillisena kokonaisuudessaan vähiten, esimerkin (44) *lähden* taasen eniten. *Tulin* ja *lähden* välillä ei kuitenkaan aivan ole  $Z$ -testin mukaan tilastollisesti merkitsevää eroa ( $Z = 1,92$ ). Verrattaessa *tulen*-verbin ilmisubjektillisuutta muiden liikeverbien ryhmään saadaan kuitenkin selkeästi erittäin merkitsevä poikkeama ( $Z = 4,15$ ). Muut ovat myös erittäin merkitsevästi ( $Z = 3,78$ ) ilmisubjektilli-

sempia kuin esimerkin (45) *saan* ja merkitsevästi ( $Z = 3,21$  ja  $2,57$ ) ilmisubjektillisempia kuin (46) *menen* ja (47) *ostan*.

- 41) Yhtäkkiä joku hyppäsi puskista eteeni ja **minä kaadu**in.  
 42) Sitten **minä tönäisin** Kaisaa ja työntekijä katsoi Kaisa tuimasti ja sitten lähdimme kaupasta nauraen.  
 43) Kun **tulin** takaisin niin silloinkin menin hissillä ylös.  
 44) **Minä lähden** kesällä tivoliin äidin kanssa ja isän kanssa.  
 45) Kerran kun **sain** lahjoja niin ne olivat hienoja.  
 46) Kun **menin** ekaa kertaa Särkänniemeen niin se oli hauskaa.  
 47) Se peli ei toimi jonka **ostin**.

Taulukko 15. Ilmisubjektin esiintyvyys liikeverbeissä.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>menen</b>	16/40 40,0 %	33/101 32,7 %	9/57 15,9 %	2/26 7,7 %	0/1 0 %	60/225 26,7 %
<b>saan (liike)</b>	4/13 30,8 %	19/69 27,5 %	14/53 26,4 %	0/25 0 %	0/3 0 %	37/163 22,7 %
<b>tulen (liike)</b>	5/15 33,3 %	11/48 22,9 %	5/41 12,20 %	0/9 0 %	0/1 0 %	21/114 18,4 %
<b>lähden</b>	2/2 100 %	10/26 38,5 %	2/15 13,3 %	2/7 28,6 %	0/0 0 %	16/50 32,0 %
<b>ostan</b>	2/5 40,0 %	2/10 20,0 %	4/14 28,6 %	1/15 6,7 %	0/1 0 %	9/45 20,0 %
<b>juoksen</b>	1/2 50,0 %	4/10 40,0 %	2/16 12,5 %	0/8 0 %	0/0 0 %	7/36 19,4 %
<b>muut liike- verbit</b>	22/36 61,1 %	75/141 53,2 %	34/105 32,4 %	8/54 14,8 %	0/14 0 %	139/350 39,7 %
<b>liikeverbit yhteensä</b>	52/113 46,0 %	154/405 38,0 %	70/301 23,3 %	13/144 9,0 %	0/20 0 %	289/983 29,4 %

Taulukon 15 muut liikeverbit poikkeavat ilmisubjektillisuudessaan erittäin merkitsevästi ( $Z = 3,55$ ) myös kaikkien liikeverbien ilmisubjektillisuudesta yhteensä. Frekventeimmistä liikeverbeistä ainoastaan *tulen* ( $Z = 2,46$ ) poikkeaa melkein merkitsevästi kaikkien liikeverbien ilmisubjektillisuudesta. Tarkasteltaessa frekventeimpiä liikeverbejä suhteessa muiden liikeverbien ryhmään, voidaan havaita ilmisubjektillisuudessa siinä määrin merkitseviä eroja, että on mahdollista sanoa liikeverbien ilmisubjektillisuuden kasvavan, mitä vähemmän frekventti liikeverbi on kyseessä.



#### 5.4.4 Ilmisubjektin esiintyvyys puheaktiverbeissä

Puheaktiverbit ovat tutkimusaineiston seitsemänneksi suurin eli toiseksi pienin semanttinen verbiluokka. Taulukosta 16 näemme puheaktiverbien esiintymät aineistossa.

Taulukko 16. Ilmisubjektin esiintyvyys puheaktiverbeissä.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>tilaan</b>	1/2 50,0 %	2/10 20,0 %	3/27 11,1 %	2/15 13,3 %	0/2 0 %	8/56 14,3 %
<b>sanon</b>	1/2 50,0 %	3/10 30,0 %	6/19 31,6 %	1/1 100 %	0/0 0 %	11/32 34,4 %
<b>muut puheaktiverbit</b>	4/6 66,7 %	21/49 42,9 %	8/51 15,7 %	3/26 11,5 %	0/5 0 %	36/137 26,3 %
<b>puheaktiverbit yhteensä</b>	6/10 60,0 %	26/69 37,7 %	17/97 17,5 %	6/42 14,3 %	0/7 0 %	55/225 24,4 %

Taulukossa 16 esiintyvistä frekventeimmistä puheaktiverbeistä esimerkin (48) *tilaan* esiintyy Z-testin mukaan vain melkein merkitsevästi ( $Z = 2,20$ ) vähemmän ilmisubjektillisena kuin esimerkin (49) *sanon*, vaikka prosentuaalista eroa on kaiken kaikkiaan 20,1 %. Aiemmin käsitellyistä semanttisista luokista poiketen esimerkin (50) puheaktiverbien muiden ryhmä ei käyttäydy ilmisubjektin suhteen frekventeistä puheaktiverbeistä merkitsevästi poiketen. Puheaktiverbit ovat ilmisubjektillisuuden suhteen melko yhtenäinen semanttinen luokka, mutta toki esiintymistaajuudeltaan myös melko pieni.

48) **Tilasin** teiltä viime kuussa valkoisen talvitakin.

49) **mä sanoin** okei pite maksaa laskun

50) Kaikki sanoivat, että **lauloin** liian hiljaa.

#### 5.4.5 Ilmisubjektin esiintyvyys psykologista tilaa ja prosessia ilmaisevissa verbeissä

Psykologista tilaa ja prosessia ilmaisevat verbit ovat tutkimusaineiston kolmanneksi suurin semanttinen verbiluokka. Taulukosta 17 voidaan nähdä, että psykologista tilaa ja prosessia ilmaisevissa verbeissä on varsin suurta hajontaa niin verbien esiintyvyyden kuin niiden ilmisubjektillisuuden suhteen.

Taulukko 17. Ilmisubjektin esiintyvyys psykologista tilaa ja prosessia ilmaisevissa verbeissä

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>haluan</b>	13/34 38,2 %	19/144 13,2 %	23/204 11,3 %	2/67 3,0 %	0/6 0 %	57/455 12,5 %
<b>tykkään</b>	29/47 61,7 %	24/56 42,9 %	2/7 28,6 %	1/1 100 %	0/0 0 %	56/111 50,5 %
<b>pidän (pstpr)</b>	33/44 75,0 %	32/46 69,6 %	0/1 0 %	0/3 0 %	0/1 0 %	65/95 68,4 %
<b>toivon</b>	0/4 0 %	2/10 20,0 %	1/20 5,0 %	0/21 0 %	0/3 0 %	3/58 5,2 %
<b>yritän</b>	0/1 0 %	5/8 62,5 %	5/18 27,8 %	0/4 0 %	0/0 0 %	10/31 32,3 %
<b>muut pstpr. verbit</b>	9/14 64,3 %	32/72 44,4 %	18/83 21,7 %	9/79 11,4 %	1/17 5,9 %	69/265 26,0 %
<b>kaikki pstpr. verbit</b>	84/144 58,3 %	114/336 33,9 %	49/333 14,7 %	12/175 6,9 %	1/27 3,7 %	260/1015 25,6 %

Prosentuaalisesti taulukon 17 ääripäistä löytyvät esimerkkien (51) *pidän* ja (52) *toivon*. *Pidän* saa ilmisubjektin kaikkein useimmin (peräti 68,4 %), kun taas *toivon* kaikkein harvimminkin (vain 5,2 %). Huomionarvoista on myös, että esimerkin (53) *tykkään*- ja *pidän*-verbiä esiintyy suurissa määrin vain tasoilla A1 ja A2. Tähän olennaisesti vaikuttavana tekijänä on se, että viesti ystävälle ulkomaille ja viesti ystävältä ulkomailta ovat olleet tehtävänantoina vain 2.–4. -luokkalaisilla. Viesti ystävälle ulkomailta -tehtävänannossa nimenomaan pyydetään kertomaan, mistä kirjoittaja pitää. Tehtävänantojen vaikutusta ilmisubjektillisuuteen ei tässä tutkimuksessa kuitenkaan käsitellä, vaikka niillä toki on oma vaikutuksensa aineistoon.

- 51) Hei **minä pidän** preikarista ja **minä pidän** sinisestä **minä pidän** hotokista ja minä tykkään paljon Lauuluja  
 52) Siispä **toivon** että vanhemmat päättävät asiasta.  
 53) **Minä tykkään** miniclip koska se on peli.

*Pidän* esiintyy erittäin merkitsevästi ilmisubjektillisempänä kuin *toivon*. Tässä tapauksessa kokonaismäärien vertailu ei kuitenkaan ole kokonaislukujen osalta kovinkaan perusteltua, koska taitotasolla A1 ja A2 *pidän* esiintyy 90 kertaa *toivon* esiintyessä vain 14 kertaa, kun taas taitotasolla B1 ja B2 *toivon* esiintyy 41 kertaa ja *pidän* vain neljä kertaa. Vertailukelpoinen taitotaso on A2, jossa *pidän* esiintyy 46 kertaa ja *toivon* 10 kertaa. Z-testi osoittaa, että tällä taitotasolla *pidän* esiintyy merkitsevästi enemmän ilmisubjektillisena ( $Z = 2,91$ ) kuin *toivon*, joten voidaan kuitenkin melko luotettavasti todistaa *pidän* olevan herkemmin ilmisubjektillinen kuin *toivon*.

Ilmisubjektillisuuden merkitsevyydessä vertailukelpoisin pari on kuitenkin esimerkin (54) *haluan* ja *pidän*. Kokonaisuudessaan *pidän* on erittäin merkitsevästi ilmisubjektillisempi kuin *haluan* ( $Z = 11,93$ ). Erittäin merkitsevä poikkeavuus niiden välisessä ilmisubjektillisuudessa on myös taitotasolla A2 ( $Z = 7,51$ ) ja merkitsevä taitotasolla A1 ( $Z = 3,27$ ).

54) Niin **haluan** että sen korjatan.

Myös *haluan* ja *tykkään* ovat ilmisubjektillisuudessaan samankaltaisesti erittäin merkitsevästi poikkeava pari kokonaisuudessaan ( $Z = 8,96$ ) sekä taitotasolla A2 ( $Z = 4,58$ ) – taitotasolla A1 *tykkään* on kuitenkin vain melkein merkitsevästi ( $Z = 2,09$ ) ilmisubjektillisempi kuin *haluan*. Poiketen monesta muusta semanttisesta luokasta esimerkin (55) muiden pstpr-verbien luokka käyttäytyy ilmisubjektin suhteen miltei samaan tapaan kuin pstpr-verbit kokonaisuudessaan.

55) **Uskon**, että enemmistö oppilaista on kanssani samaa mieltä siitä, että ylioppilaskirjoitukset pitäisi poistaa tai ainakin niiden vaikutustasoa pitäisi vähentää.

Pstpr-verbien suhteen ei voida missään nimessä sanoa, että vähemmän frekventit verbit saisivat ilmisubjektin useammin kuin frekventit verbit. Frekventeimmissä verbeissä on kuitenkin suurta poikkeamaa ja näyttääkin siltä, että *pidän* ja *tykkään* ovat poikkeuksellisen herkästi ilmisubjektillisiä verbejä, vaikka niiden esiintymät ovatkin pääosin alimmilta taitotasoilta. Vertailu muihin koko aineiston frekventeimpiin verbeihin taitotasoilla A1 ja A2 tämän myös osoittaa (ks. taulukko 21 luvusta 5.5).

#### 5.4.6 Ilmisubjektin esiintyvyys perkeptioverbeissä

Perkeptioverbit ovat tutkimusaineistossa kaikkein vähiten esiintyvä semanttinen verbiluokka. Taulukosta 18 voimme tarkastella perkeptioverbien esiintyvyyttä ja ilmisubjektillisuutta.

Taulukko 18. Ilmisubjektin esiintyvyys perkeptioverbeissä.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>näen</b>	2/6 33,3 %	8/32 25,0 %	5/20 25,0 %	0/6 0 %	0/7 0 %	15/71 21,1 %
<b>huomaan</b>	1/1 100 %	0/6 0 %	0/13 0 %	0/11 0 %	0/2 0 %	1/33 3,0 %
<b>muut perkeptioverbit</b>	2/5 40,0 %	13/23 56,5 %	3/8 37,5 %	0/13 0 %	0/4 0 %	18/53 34,0 %
<b>perkeptioverbit yhteensä</b>	5/12 41,7 %	21/61 34,4 %	8/41 19,5 %	0/30 0 %	0/13 0 %	34/157 21,7 %

Taulukon 18 mukaan voisi päätellä, että frekventeimmissä perkeptioverbeissä näyttäisi olevan selkeää poikkeamaa, kun tarkastelee prosenttilukuja. Z-testin mukaan esimerkkien (56) *näen* ja (57) *huomaan* välillä oleva ero ( $Z = 2,38$ ) jää kuitenkin melkein merkitseväksi. *Huomaan* esiintyy ilmisubjektillisena kuitenkin erittäin merkitsevästi vähemmän ( $Z = 3,36$ ) esimerkin (58) muiden perkeptioverbien ryhmään verrattaessa.

56) Olin laivalla, laivalla **mä näin** joulupukin ja laiva kokki ja ville viiking.

57) Kuin katsoin takse ja **huomasin** että se on naamio.

58) Joka tapauksen **minä kuulin** että meidän luokalle oon jotain lopputyötä?

Taitotasojen sisällä esiintymät ovat melko pieniä, mutta taitotasoilla A2 ja B1 muut perkeptioverbit esiintyvät melkein merkitsevästi ilmisubjektillisempänä kuin *huomaan* (A2:  $Z = 2,48$  ja B1:  $Z = 2,38$ ). Perkeptioverbien kohdalla voidaan varovaisesti väittää frekventeimpien verbien saavan ilmisubjektin hiukan harvemmin kuin vähemmän frekventtien sillä poikkeuksella kuitenkin, että selkeästi frekventeimminkin esiintyvä *näen* esiintyy melkein merkitsevästi ilmisubjektillisempänä kuin toiseksi frekventein *huomaan*.

#### 5.4.7 Ilmisubjektin esiintyvyys abstrakteissa verbeissä

Abstraktit verbit ovat tutkimusaineiston viidenneksi suurin semanttinen verbiluokka. Taulukosta 19 näemme abstraktien verbien esiintyvyyden ja ilmisubjektillisuuden.

Taulukko 19. Ilmisubjektin esiintyvyys abstrakteissa verbeissä.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>voin</b>	8/20 40,0 %	22/96 22,9 %	6/73 8,2 %	5/32 15,6 %	0/4 0 %	41/225 18,2 %
<b>saan (abstr)</b>	0/1 0 %	0/6 0 %	3/21 14,3 %	0/6 0 %	0/1 0 %	3/35 8,6 %
<b>joudun</b>	0/0 0 %	0/6 0 %	2/10 20 %	2/16 12,5 %	0/1 0 %	4/33 12,1 %
<b>muut abstr. verbit</b>	3/6 50,0 %	9/13 69,2 %	7/22 31,9 %	1/31 3,2 %	0/3 0 %	20/75 26,7 %
<b>abstr. verbit yhteensä</b>	11/27 40,7 %	31/121 25,6 %	18/126 14,3 %	8/85 9,4 %	0/9 0 %	68/368 18,5 %

Taulukosta 19 voimme havaita, että kokonaismäärällisesti prosentuaalisesti eniten ilmisubjektillisuudessa poikkeavat esimerkkien (59) *saan* ja (60) muut abstraktit verbit. Ne eivät kuitenkaan eroa Z-testauksella toisistaan tilastollisesti merkitsevästi ( $Z = 1,67$ ). Koska Z-testauksella ryhmien välille ei saada kokonaisuudessaan ilmisubjektin käytön suhteen tilastollisesti merkitseviä eroja, voidaan sanoa, ettei abstraktien verbien ilmisubjektillisuudessa ole suuria eroja. Muut abstraktit verbit ovat kuitenkin erittäin merkitsevästi ilmisubjektillisempia kuin esimerkin (61) *voin* taitotasolla A2 ( $Z = 3,47$ ) ja B1 ( $Z = 2,82$ ), joten tässäkin semanttisessa luokassa on jossakin määrin viitteitä siihen suuntaan, että vähemmän frekventit verbit saavat herkemmin ilmisubjektin kuin frekventit.

59) Ope **sain** tietää että meillä on ollut koe oliko se koe helppo vai vaikea?

60) kivaa oli että **mä pääsin** egyptiin

61) Mutta **voin** tulla sun kaa ensi viikon tiistai klo 18:00 linnanmäkeen.

#### 5.4.8 Ilmisubjektin esiintyvyys kieltoverbeissä

Kieltoverbit ovat tutkimusaineiston toiseksi suurin semanttinen verbiluokka. Kuten taulukosta 20 voimme nähdä, kieltoverbeissä on vain yksi mahdollinen esiintymä *en*, joten luokan sisältä vertailua ei voida tehdä samaan tapaan kuin kaikissa muissa semanttisissa verbiluokissa. Kieltoverbien ilmisubjektillisuus vähenee varsin tasaisesti mentäessä alimmalta taitotasolta ylimpään.

Taulukko 20. Ilmisubjektin esiintyvyys kieltoverbeissä.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>en = kaikki kieltoverbit</b>	54/131 41,2 %	120/439 27,3 %	62/354 17,5 %	17/139 12,2 %	0/20 0 %	253/1083 23,4 %

Kieltoverbien semantiikka syventyisi toki tutkailtaessa niihin liittyviä pääverbejä ja sitä kautta kieltoverbin ilmisubjektillisuutta, mutta tämän tutkimuksen puitteissa kieltoverbin pääverbi on jätetty huomiotta. Kieltoverbin esiintymisestä tutkimusaineistossa on nähtävissä esimerkit (62), (63) ja (64).

62) **Minä en** voi tulla elokuvaan koska minä menen eväretkelle perheen kanssa ja sitten kauppaan.

63) Itseasiassa **en** välitä jos koulussa ei saisi käyttää puhelinta, koska mitä sillä tekee kun on oppitunti.

64) **En** ole koskaan aikaisemmin törmännyt tällaiseen.

Kieltoverbi on otettu tähän tutkimukseen semanttisten verbiluokkien kanssa tasavertaisena luokkana sen vuoksi, jotta voitaisiin nähdä poikkeako kieltoverbi ilmisubjektillisuudessa muista semanttisista verbiluokista. Tuloksista on kuitenkin nähtävissä, että kieltoverbit sijoittuvat ilmisubjektillisuudessaan varsin keskiverroksi luokaksi muihin semanttisiin verbiluokkiin nähden (ks. 5.2.6). Ainoastaan taitotasolla B2 se on selkeästi monia muita luokkia ilmisubjektillisempi (ks. 5.2.4).

## 5.5 Ilmisubjektin esiintyvyys aineiston frekventeimpien verbien kanssa

Taulukko 21 esittelee aineiston kymmenen kokonaismäärältään frekventeintä yksikön 1. persoonan verbiä ja niiden ilmisubjektillisuutta. Kymmenen frekventeimmän joukossa on kolme psykologista tilaa ja prosessia kuvaavaa verbiä (*haluan*, *tykkään* ja *pidän*), kolme liikeverbiä (*menen*, *saan* ja *tulen*), yksi tilaverbi (*olen*), kieltoverbi (*en*), abstrakti verbi (*voin*) ja perkeptioverbi (*näen*). Kuten taulukosta 21 voimme havainnoida, kokonaisuudessaan kaikkein ilmisubjektillisimmat eli *pidän* ja *tykkään* esiintyvät kuitenkin pääosin vain taitotasolla A1 ja A2 (ks. 5.4.5), joten kokonaismäärissä voidaan katsoa olevan pientä vääristymää tämän myötä, mutta tarkastelen ilmisubjektillisuuden ääripäitä myös taitotasoittain.

Taulukossa 21 esiintyvistä yksikön 1. persoonan verbeistä kokonaismääriltään ilmisubjektillisuuden ääripäät löytyvät psykologista tilaa ja prosessia ilmaisevista verbeistä *pidän* ja *haluan*, joiden välillä on erittäin merkitsevä ero ( $Z = 11,93$ ). Eri semanttisista luokista olevien

verbien suurin ilmisubjektillisuuden erittäin merkitsevä ( $Z = 9,48$ ) poikkeavuus löytyy *pidän* ja *voin* väliltä. Kuten jo aiemmin mainitsin, *pidän* esiintyy kohtuullisissa määrin vain kahdella alimmalla taitotasolla, joten on syytä vielä huomioida kohtuullisissa määrin kaikilla taitotasolla esiintyviä verbejä. Tällaisista suurin poikkeavuus kokonaismäärien ilmisubjektillisuudessa on *menen* ja *haluan* välillä oleva erittäin merkitsevä ( $Z = 4,60$ ) ero.

Taulukko 21. Ilmisubjektin esiintyvyys aineiston frekventeimpien yksikön 1. persoonan verbien kanssa.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>1. olen</b> (tila)	112/315 35,6 %	157/662 23,7 %	72/538 13,4 %	14/249 5,6 %	0/28 0 %	355/1792 19,8 %
<b>2. en</b> (kiel- to)	54/131 41,2 %	120/439 27,3 %	62/354 17,5 %	17/139 12,2 %	0/20 0 %	253/1083 23,4 %
<b>3. haluan</b> (pstpr)	13/34 38,2 %	19/144 13,2 %	23/204 11,3 %	2/67 3,0 %	0/6 0 %	57/455 12,5 %
<b>4. menen</b> (liike)	16/40 40,0 %	33/101 32,7 %	9/57 15,9 %	2/26 7,7 %	0/1 0 %	60/225 26,7 %
<b>5. voim</b> (abstr)	8/20 40,0 %	22/96 22,9 %	6/73 8,2 %	5/32 15,6 %	0/4 0 %	41/225 18,2 %
<b>6. saan</b> (liike)	4/13 30,8 %	19/69 27,5 %	14/53 26,4 %	0/25 0 %	0/3 0 %	37/163 22,7 %
<b>7. tulen</b> (liike)	5/15 33,3 %	11/48 22,9 %	5/41 12,20 %	0/9 0 %	0/1 0 %	21/114 18,4 %
<b>8. tykkään</b> (pstpr)	29/47 61,7 %	24/56 42,9 %	2/7 28,6 %	1/1 100 %	0/0 0 %	56/111 50,5 %
<b>9. pidän</b> (pstpr)	33/44 75,0 %	32/46 69,6 %	0/1 0 %	0/3 0%	0/1 0 %	65/95 68,4 %
<b>10. näen</b> (perk)	2/6 33,3 %	8/32 25,0 %	5/20 25,0 %	0/6 0 %	0/7 0 %	15/71 21,1 %

Taulukon 21 verbeistä taitotasolla A1 suurin tilastollinen poikkeavuus on esimerkkien (65) *pidän* ja (66) *olen* välillä. *Pidän* esiintyy ilmisubjektillisena erittäin merkitsevästi enemmän ( $Z = 4,99$ ) kuin *olen*.

65) **minä pidän** laulusta Annasesoida

66) **Olen** tyttö sen huomaa nimestä.

Taitotasolla A2 suurin poikkeavuus on samasta semanttisesta luokasta löytyvien esimerkkien (67) *pidän* ja (68) *haluan* välillä oleva erittäin merkitsevä ero ( $Z = 7,51$ ). Eri semanttisista luokista löytyvien verbien välillä oleva suurin ero on jälleen esimerkkien (69) *pidän* ja (70) *olen* välillä oleva erittäin merkitsevä ero ( $Z = 6,80$ ).

- 67) **Minä pidän** siitä Crash Bandicoot, koska se on seikkaileva, hauska, paljon tasoja ja kaikkea hauskaa.  
 68) **Haluan** kysyä koesta kaksi asia monta sivunen koe on ja onko vaikea.  
 69) hei ope **olen** lähdössä Sudaniin katsomaan mumoa ja omaa kotimaata  
 70) **Mä pidän** kesäkuusta koska voi olla paljon ulkona- ja nautia.

Taitotasolla B1 eniten ilmisubjektillisuudessaan poikkeavat verbit ovat esimerkkien (71) *saan* (liikeverbi) ja (72) *haluan*. *Saan* esiintyy merkitsevästi ( $Z = 2,80$ ) ilmisubjektillisempänä kuin *haluan*.

- 71) Mitä **mä sain** kokeesta?  
 72) **Haluaisin** palauttaa kännykän ja ottaa rajat takaisin jos mahdollista.

Taitotasolla B2 suurin poikkeavuus Z-testin mukaan on esimerkkien (73) *tykkään* ja (74) *haluan* välillä oleva erittäin merkitsevä ilmisubjektillisuuden ero ( $Z = 4,69$ ), mutta *tykkään* esiintyy tällä taitotasolla vain kerran. Esimerkin (75) *voin* esiintyy melkein merkitsevästi ( $Z = 2,29$ ) ilmisubjektillisempänä kuin *haluan*.

- 73) ja **mä tykkään** olla yksin kotona.  
 74) **Haluan** rahani takaisin, kiitos.  
 75) **Voinko minä** tehdä jotain korvaavia tehtäviä viikon aikana, jotta en jäisi jälkeen?

Jos ei oteta huomioon *tykkään* 1/1 ilmisubjektillisuutta taitotasolla B2, on *voin* ainoa frekventimmistä verbeistä, jonka ilmisubjektillisuus kohtuullisilla esiintymismäärillä nousee yksittäisellä taitotasovälillä (B1–B2) katsottaessa alemmalta taitotasolta ylempään. Taitotasolla C1 & C2 ei esiinny lainkaan ilmisubjekteja frekventimpien verbien kohdalla. Aineistosta vielä erikseen eri semanttisten luokkien analysoinnissa huomioidut frekventtiysjärjestyksessä sijoilla 11.–24. olevat yksikön 1. persoonan verbit ovat tarkasteltavissa taulukosta 22, joka löytyy liitteestä 3.



## 6. PÄÄTÄNTÖ

### 6.1 Tutkimustulosten pohdintaa

Tässä aineistopohjaisessa korpustutkimuksessa yksikön 1. persoonan ilmisubjektin esiintyvyydestä kirjoitetussa oppijansuomessa on saatu monipuolisia tuloksia. Ensimmäiseksi on pystytty toteamaan, että ilmisubjektin käyttö varsin selkeästi vähenee, kun EVK:n taitotasoylenee. Taitotasolla A1 ilmisubjektia käytetään melko paljon, mutta jo taitotasolla A2 ilmisubjektin käyttö on erittäin merkittävästi vähentynyt. Vähentyminen jatkuu tasaisesti tilastollisesti erittäin merkittävänä taitotasojen välillä, joskin taitotasolla B2 käytetään jo sen verran vähän ilmisubjektia, ettei sen käyttö siitä enää vähene tasoille C1 & C2 enää kuin melkein merkittävästi. Koska seikkaperäisempää tutkimusta kirjoitetun oppijansuomen ilmisubjektin käytöstä (tai yleensäkin suomen ilmisubjektin käytöstä kirjoitetun kielen kehitysvaiheissa) ei ole tehty, on jokseenkin vaikea suoralta kädeltä sanoa, mistä ilmisubjektin vähentyminen varmuudella johtuu. Voidaan kuitenkin katsoa, että ilmisubjektin käytön vähentyminen on osa kirjoitetun suomen kielen taidon kehittymistä. Kuten aiemminkin olen todennut, ilmisubjektin käyttö kirjoitetussa suomessa ei itsessään ole minkäänlainen virhe. Lopulta kuitenkin kirjoitettu kieli pyrkinee jonkinlaiseen taloudellisuuteen, jolloin sellaiset ilmaukset, joilla ei itsessään tuoda mitään lisämerkitystä tekstiin, vähenevät. Kirjoitus- ja lukutaito kulkevat myös käsi kädessä – lukutaidon kehittyessä suomenoppija lukee jatkuvasti enemmän suomenkielistä tekstiä, jossa ilmisubjektia tavataan käyttää säästeliäästi, minkä myötä hän saa koko ajan lisää mallia kirjoittamiseensa.

Ilmisubjektin käyttöön voi kuitenkin vaikuttaa myös se, miten persoonaa ilmaistaan suomenoppijan ensikielessä. Jos esimerkiksi suomenoppijan ensikielessä persoonaa ilmaistaan suomen kaltaisesti, voi olla todennäköisempää, että suomenoppija omaksuu suomen kielen persoonailmauksen valinnaisuuden helpommin kuin suomenoppija, jonka ensikielessä persoonapronominia käytetään aina. Toisena varmasti vaikuttavana tekijänä ovat oppikirjat ja muut oppimateriaalit, joissa ilmisubjektia käytetään (ks. Dal Pozzo 2014: 200–201). Suomen kielen oppikirjoissa autenttinen kielenkäyttö ei aina ole ohjenuorana siinä, miten tai mitä kielen ilmiötä kuvaillaan. Muoto voi mennä todenmukaisen kielenkäytön edelle, kun halutaan varmistaa, että rakenteen muodostaminen tulee oppijalle selväksi, tai ettei jonkin ilmiön monimerkisyisyys tuota hämmennystä. (Seilonen 2013: 193.) Tässä tutkimuksessa ei ole kuitenkaan ollut mahdollista ottaa huomioon, mitä oppikirjoja ja -materiaaleja Topling-aineiston

eri kirjoittajat ovat käyttäneet. Aineistoa ei ole myöskään jaoteltu oppijoiden ensikielen mukaan. Alimmilla EVK:n taitotasolla voi myös olla niin, että suomenoppijat eivät vielä hahmota selkeästi puhutun ja kirjoitetun suomen eroja. Tämän vuoksi suomenoppija pyrkinee kirjoittamaan puheenkaltaisesti, jolloin myös ilmisubjekteja tulee kirjoitettuun tekstiin melko paljon, koska ilmisubjekteja käytetään suomen puhekielessä kosolti.

Lopulta tutkimukseni tuloksissa on ollut mielenkiintoisinta huomata se, ettei ilmisubjektin frekvenssi ole riippuvainen ainoastaan taitotasosta, vaan että eri verbit saavat osakseen ilmisubjektin paikoin perin poikkeavin määrin. Kun verbejä tarkastellaan vain neljässä suuressa semanttisessa luokassa, ei eri luokkien välille tule muuten suuria eroja kuin siinä, että abstraktit verbit ovat kokonaismääriä katsottaessa merkitsevästi vähemmän ilmisubjektillisiä kuin konkreettiset, mentaaliset ja kieltoverbit. Jos ajatellaan frekventeintä abstraktia verbiä *voin*, joka on yleisesti melko taajaan käytetty ja merkittävä verbi, voidaan ajatella sen käyttöä opeteltavan siinä määrin, että kyseisen verbin käyttö harjaantuu nopeasti, jolloin myös ilmisubjekti alkaa käyttötottumuksen myötä piakkoin pudota sen seurasta pois. Näin voidaan arvella käyvän myös tutkimusaineiston frekventeimmän yksikön 1. persoonan verbin *olen* kanssa, joka on moniin muihin verrattuna varsin vähäisesti ilmisubjektinen jo tasolla A1. *Olen*-verbi tietysti on siinä määrin keskeinen verbi kielessä kuin kielessä, että sen käyttö opetetaan ja opitaan ensimmäisenä, minkä myötä sen käyttöön totutaan nopeasti.

Jaettaessa verbit kahdeksaan eri semanttiseen luokkaan alkaa suurempia eroja erityyppisten verbiryhmien välille kuitenkin selkeästi muodostua. Lopulta teko- ja muutosverbit sekä liikeverbit ovat selkeästi ilmisubjektillisimpia, kun taas tilaverbit ja abstraktit verbit edustuvat vähiten ilmisubjektillisinä. Esimerkiksi Väänäsen (2016: 229) vanhoista suomen murteista (aineistona puhetta) tehdyn tutkimuksen mukaan ilmisubjektillisimpina esiintyivät tila-, kognitio- (tässä tutkimuksessa psykologinen tila ja prosessi) ja puheverbit, kun taas liike- ja transitiiiviset muutosverbit (tässä tutkimuksessa osa teko- ja muutosverbejä) esiintyivät vähiten ilmisubjektillisimpina. Oma tutkimukseni kirjoitetusta oppijansuomesta antaa Väänäsen tutkimukseen verrattuna varsin poikkeavia tuloksia, joten voidaan sanoa, että kirjoitettu oppijansuomi käyttäytyy ilmisubjektin suhteen eri tavalla kuin puhuttu vanha murre-suomi. Ilmisubjektin käyttöön liittyvät seikat rakentunevat täten jokseenkin eri tavalla oppijansuomessa kuin natiivien puhutussa suomessa.

Yksi tutkimukseni lähtöhypoteeseista oli myös se, että kieltoverbit oletetun painokkuutensa vuoksi olisivat mahdollisesti muita semanttisia verbiluokkia ilmisubjektillisempi. Osoitautui kuitenkin, että kieltoverbit ovat lopulta ilmisubjektillisuudessaan varsin keskiverto

luokka muihin verrattuna. Tosin kaikki kiellot eivät välttämättä lopulta ole kovinkaan painokkaita (*en tiedä vs. minä en tykkää!*)

Teko- ja muutosverbien kohdalla on mielenkiintoista, että kyseisessä luokassa mikään yksittäinen verbi ei esiinny kovinkaan frekventtinä. Kohtuullisissa määrissä esiintyvät *teen* ja *kirjoitan* eivät kuitenkaan ole kovinkaan ilmisubjektillisiä, mutta muiden temu-verbien ryhmä on lopulta varsin ilmisubjektillinen. Tämä seikka viittaisi siihen, että vähemmän frekventit verbit esiintyisivät herkemmin ilmisubjektillisinä. Useissa semanttisissa verbiluokissa frekventeimmät verbit ovat selkeästi vähemmän ilmisubjektillisiä kuin muiden ryhmä. Merkitsevästi asia on näin tila- ja temu-verbien kohdalla. Tarkasteltaessa aineiston kymmenen frekventeimmän yksikön 1. persoonan verbin ilmisubjektillisuutta voidaan kuitenkin huomata, että frekventeimmätkin verbit saavat osakseen ilmisubjektin perin poikkeavissa määrin. Vaadittaisiin vielä tarkempia laskumalleja, jotta voitaisiin tyhjentävästi todeta frekventtiyden selkeästi vaikuttavan ilmisubjektillisuuteen, mutta muutamia poikkeuksia lukuun ottamatta tämä tutkimus antaa väittämän totuuspohjalle selkeitä viitteitä.

Kun tarkastellaan eri ryhmien sisällä frekventeimpien verbien variaatiota ilmisubjektillisuuden suhteen, voidaan huomata, että psykologista tilaa ja prosessia ilmaisevat verbit ovat kaikkein eniten sisäisesti poikkeava ryhmä. *Haluan* on merkitsevästi vähemmän ilmisubjektillinen kuin *tykkään-* ja *pidän-*verbit myös A1 ja A2 taitotasolla. Edelleen tehtävänannot voivat vaikuttaa ilmisubjektillisuuteen, koska eri tehtävänannot ohjaavat erilaisten verbien käyttöön, kuten voidaan nähdä verrattaessa esimerkiksi viesti ystävälle ulkomaille -tehtävää viesti verkkokauppaan -tehtävään. Ensimmäisessä kehoitetaan muun muassa kertomaan, mistä kirjoittaja pitää, kun taas toisessa pyydetään kirjoittamaan reklamaatiosähköposti jostakin viallisesta tuotteesta. Voidaankin ajatella, että ilmisubjektittomuuteen liittyy jonkinlaista virallisuutta, kun taas ilmisubjektillisuudessa voisi toisaalta olla erityistä oman itsensä esilletuontia. Joka tapauksessa tämä tutkimus jättää monia perustavanlaatuisia kysymyksiä avoimeksi, joihin laajemmissa oppijansuomen tutkimuksissa voidaan yksityiskohtaisemmin perehtyä.

## 6.2 Tutkimuksen kriittistä tarkastelua ja jatkotutkimusehdotuksia

Toisen kielen oppimisen tutkimuksessa merkittävintä on mielestäni se, että tutkimus pyrkii tarjoamaan kielen oppimiseen ja opettamiseen työkaluja ja suuntaviivoja. Millä tavoin toista kieltä opitaan tehokkaimmin ja millaisin tavoin opetus tehokkaaseen oppimiseen voi ohjata? Toisen kielen oppimiseen, kuten kaikenlaiseen oppimiseen, liittyy paljon sellaisia seikkoja,

joista tutkijat eivät ole yksimielisiä – voisi myös sanoa, että erimielisyys johtuu ennen kaikkea siitä, että oppimiseen liittyy paljon sellaisia prosesseja, joiden olemuksesta olemme epävarmoja. Yhtä kaikki tämä tutkimus ei suoranaisesti tuo mitään uutta ja mullistavaa toisen kielen oppimiseen ja opettamiseen. Tämä tutkimus osoittaa suuressa kuvassa ennen kaikkea, että ilmisubjektia käytetään sitä vähemmän, mitä korkeammalla EVK:n taitotasolla ollaan. Näin ollen tutkimuksen pää- ja lähtöhypoteesi on osoitettu toteen. Tutkimukseni voidaan nähdä antavan tietoa suomea toisena kielenä oppivien kielellisestä kehityksestä ja kirjoitetun suomen konventioiden omaksumisesta. Kuten aiemmin olen todennut, EVK ei juuri millään tavalla määritä sitä, mitä kieliopillisia osa-alueita milläkin taitotasolla hallitaan. Yksittäisen suomenoppijan kirjoitettua tekstiä tarkasteltaessa olisi kuitenkin mielestäni mahdollista kiinnittää huomiota ilmisubjektin liialliseen käyttöön, mikäli teksti olisi jo vähintään B1-tasoa. Ilmipanun persoonapronominin painollisuutta voi suomenoppijoille opettaa siinä vaiheessa, kun verbien persoonapäätteet ovat kunnolla hallinnassa. Alkuvaiheessa persoonapronominin käyttö persoonamuotoisen verbin kanssa opetuksessa ja oppimateriaaleissa lienee kuitenkin perusteltua, jotta persoonapronominien ja persoonapäätteiden yhteys hahmotetaan varmuudella.

Tutkimukseni on ollut prosessina senkaltainen, että pikkuhiljaa tutkimusaihetta ja tutkimuskysymyksiä on syvennetty kohti yksityiskohtaisempaa tietoa ilmisubjektin käytöstä. Ensin lähestyin pelkästään oppijansuomen ilmisubjektia kontekstistaan irrallaan. Koska pelkä ilmisubjektin määrällinen esiintyminen eri taitotasojen teksteissä ei lopulta kerro kovinkaan paljon, oli tutkimuksessa syvennyttävä erityisesti ilmisubjektin liittolaiseen eli persoonamuotoiseen verbiin. Verbien semanttinen luokittelu antaa tietoa siitä, miten ilmisubjektillisinä semanttisesti erityyppiset verbit esiintyvät. Edelleen yksittäisten frekventeimpien verbien vertailu antaa tietoa siitä, miten eri verbit poikkeavat ilmisubjektillisuudessa. Koska pronominisubjekti itsessään ei kerro itseään enempää, voitaisiin tämä tutkimus nähdä ennemminkin verbitutkimuksena kuin persoonapronomitutkimuksena. Ilmisubjektin ollessa kuitenkin tässä tutkimuksessa verbejä erittelevä tekijä on mielestäni perusteltua pitää tutkimustani ennen muuta ilmisubjektitutkimuksena. Tutkimusprosessinani haastavinta on ehkä ollut verbisemantiikkaan syventyminen. Lopulta en ole juuri avannut verbien semanttista luokittelua syntaksin näkökulmasta, so. semanttisesti samankaltaisilla verbeillä on samankaltainen argumenttirakenne eli tietty määrä tietynlaisia verbin merkityksen vaatimia osallistuja-argumentteja sekä kehysmääritteitä (ks. Pajunen 2001: 19). Argumenttirakennetta ja muutenkin tarkemmin lausekontekstia huomioimalla olisi erilaisten verbien ilmisubjektillisuuden analysoimiseen saattanut löytyä enemmän työkaluja – samalla tosin tutkimus olisi laajentunut melkoisesti. Hieman lisätyötä aiheuttivat myös sellaiset kaksimerkityksiset verbit, kuten *saan*

(sekä liike että abstrakti) ja *tulen* (sekä liike että abstrakti), joiden semanttinen verbiluokka oli aina selvitettävä tekemisen kohteen kautta. Päädyin myös joiltakin osin hajottamaan ja uudelleen muodostamaan Pajusen semanttisia verbiluokkia, koska sellaisenaan ne eivät olisi kovinkaan käytännöllisesti sopineet tutkimusaineistoni luokitteluun. Eniten päänvaivaa minulle aiheuttivat Pajusen luokittelun puheaktiverbit, jotka päädyin sijoittamaan Pajusen vastaisesti pääosin konkreettisiksi. Lopulta minulle ratkaisevaa konkreettisten ja mentaalisten verbien karkeassa erottelussa on se, onko verbin kuvaama teko muiden kuin tekijän havaittavissa. Konkreettisia tekoja on muidenkin kuin tekijän itse mahdollista havaita, kun taas mentaalisten tekojen tekijä ei oikeastaan ole tekijä, vaan omien mielten liikkeittensä kokija.

Tämän tutkimuksen luonne aineistopohjaisena korpustutkimuksena avautui minulle vähitellen. Toimiessani aluksi varsin manuaalisesti lukuisten tekstitiedostojen kimpussa en tähännyt kovinkaan yksityiskohtaiseen tietoon, mutta lopulta otettuani aineiston AntConc-korpusohjelmiston avulla käsittelyyn aloin vasta kunnolla sisäistämään, kuinka joutuisasti korpustyökaluilla saadaan yksityiskohtaista tietoa isoista tekstimassoista. Aineiston manuaalinen läpikäynti kuitenkin mahdollisti minulle sen, että olen pystynyt tunnistamaan aineiston erityispiirteitä (yhteen kirjoitettuja eri sanoja, välimerkittämiä tekstejä, tehtävänantojen aiheuttamia verbivalintoja).

Jatkotutkimusta ilmisubjektin käytöstä kirjoitetussa oppijansuomessa voitaisiin tehdä muun muassa siten, että vertailtaisiin samanlaisessa kehitysvaiheessa olevia natiiveja suomenkielisiä suomen kielen oppijoihin. Olisi mielenkiintoista tietää, käyttävätkö esimerkiksi 2.–4. luokkalaiset natiivit erityisen selkeästi vähemmän ilmisubjektia kuin suomen kielen oppijat. Voiko olla mahdollista, että ilmisubjektin käyttöön liittyisi myös jonkin kehitysvaiheen ”minäkeskeisyyttä”? Tällaisessa tutkimuksessa olisi ilman muuta tärkeää, että aineisto olisi kerätty samoilla tehtävänannoilla sekä natiiveilta että oppijoilta. Mielenkiintoista olisi myös kirjoitetun suomen lisäksi vertailla, poikkeako suomenoppijoiden puhekielen ilmisubjektillisuus natiivien puhekielen ilmisubjektillisuudesta. Yksi syventymisvaihtoehto olisi myös selvittää, kuinka paljon suomenoppijan lähtökieli vaikuttaa ilmisubjektin käyttöön. Käyttääkö ensikieleltään espanjalainen (nollasubjektikieli) suomenoppija selkeästi vähemmän ilmisubjektia kuin ensikieleltään saksalainen (ei-nollasubjektikieli)?

Perusteltua saattaisi olla myös tutkailla erilaisten tekstilajien ilmisubjektillisuutta. Topling-aineiston kohdalla se tarkoittaisi vertailua eri tehtävänantojen välillä, mikä olisi toki ollut mahdollista tässäkin tutkimuksessa, mutta päätin rajata tehtävänantojen vaikutuksen tutkimuskysymysten ulkopuolelle. Eri tehtävänantojen vertailu johtaisi kuitenkin näkemykseni mukaan todennäköisesti siihen, että tutkailtaisiin ennemminkin tehtävänannon vaikutusta ver-

bivalintoihin, joita mikäkin tehtävänanto missäkin määrin aiheuttaa. Samalla olisi ilman muuta mahdollista tutkailla eri verbien ilmisubjektillisuutta yhtä lailla kuin tässä tutkimuksessa, mutta nähdäkseni verbien semanttinen luokittelu ajaa saman asian.

Lopulta voidaan todeta, että tämän tutkimuksen tulosten myötä Elias Lönnrotin ei kuitenkaan tarvitse kääntyä haudassaan kirjoittaakseen Volmari Kilpiselle siitä, kuinka vielä pidempäänkin suomea toisena kielenä oppinut kirjoitetussa suomessaan viljelee turhan tiuhaan *persoonallisia sianimukoita*. Ilmipannut persoonapronominit eli ilmisubjektit vähenevät suomenoppijan kirjoitetusta suomesta tasaisen varmasti samalla, kun oppijan suomen osaaminen muutenkin edistyy.

## LÄHTEET

Anthony, Laurence 2016: *AntConc (4.3.3) [Computer Software]*. Tokyo: Waseda University. – <http://www.laurenceanthony.net/> 16.12.2016

Anttila, Aarne 1985: *Elias Lönnrot: elämä ja toiminta*. Helsinki: Suomalaisen kirjallisuuden seura.

Anttila, Pirkko 2005: *Ilmaisu, teos, tekeminen ja tutkiva toiminta*. Artefakta 16. Hamina: Akaatiimi Oy.

Dal Pozzo, Lena 2014: Havaintoja subjektin käytöstä S2-puhujien kielessä ja oppikirjoissa. – Yrjö Lauranto & Marjut Vehkanen (toim.): *Päättymätön projekti. Puhetta eri S2-foorumeilla* s. 199–210. Helsinki: Helsingin yliopiston suomen kielen, suomalais-ugrilaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos.

Dryer, Matthew S. 2011: Expression of Pronominal Subjects. M. S. Dryer & M. Haspelmath (eds.): *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library. – [wals.info/chapter/101](http://wals.info/chapter/101) 16.11.2014.

Duvallon, Outi 2006: Milloin pronominisubjekti jää pois puhutussa suomessa? – Anneli Pajunen ja Hannu Tommola (toim.): *XXXII Kielitieteen päivät Tampereella 19.–20.5.2005* s. 203–217. Tampere Studies in Language, Translation and Culture. Series B 2. Tampere: Tampere University. – <http://urn.fi/urn:isbn:951-44-6622-5>

Ellis, Rod & Barkhuizen, Gary 2005: *Analysing learner language*. Oxford: Oxford University.

EVK = *Eurooppalainen viitekehys. Kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys 2003*. Helsinki: WSOY.

Flint, Aili 1980: *Semantic structure in the Finnish lexicon: verbs of possibility and sufficiency*. Helsinki: Suomalaisen kirjallisuuden seura.

Halttunen, Sari 2014: *"Sori mä en voin tullan": negaatioilmaukset suomi toisena kielenä -oppijoiden teksteissä*. Maisterintutkielma. Jyväskylän yliopiston suomen kielen laitos.

Helasvuo, Marja-Leena & Laitinen, Lea 2006: Person in Finnish: Paradigmatic and syntagmatic relations in interaction. – Marja-Leena Helasvuo & Lyle Campbell (eds.): *Grammar from the human perspective. Case, space and person in Finnish* s. 173–208. Amsterdam: John Benjamins.

Ivaska, Ilmari 2014: Edistyneen oppijansuomen avainrakenteita. Korpusnäkökulma kahden kielimuodon tyypillisiin rakenteellisiin eroihin. – *Virittäjä* 118 (2) s. 161–191.

Jantunen, Jarmo Harri 2008: Haasteita oppijankielen korpusanalyysille: oppijankielen universaalit. Pille Eslon (toim.): *Õppijakeele analüüs: võimalused, probleemid, vajadused* s. 67–92. Tallina Ülikooli eesti filoloogia osakonna toimetised 10. Tallinn: Tallinna Ülikooli.

——— 2009: Ei pelkästään mielikuvituksen puutteen vuoksi – Kieliaineistojen systemaattinen käyttö kielentutkimuksessa. – *Virittäjä* 113 (1) s. 101–113.

——— 2015: Oppimiskontekstin vaikutus oppijanpragmatiikkaan: astemääritteet leksikaalisina nallekarhuina. *Lähivõrdlusi. Lähivertailuja*, 25, 105–136. Tallinn: Eesti Rakenduslingvistika Ühing 2015.

Jantunen, Jarmo Harri & Pirkola, Silja 2015: Oppijansuomen sähköiset tutkimusaineistot. Nykytilanne. – *Virittäjä* 119 (1) s. 88–103.

Kangasniemi, Heikki 1992: *Modal expressions in Finnish*. Studia Fennica, Linguistica. 2. Helsinki: Suomalaisen Kirjallisuuden seura.

Lappalainen, Hanna 2004: *Variaatio ja sen funktiot*. Helsinki: Suomalaisen kirjallisuuden seura.

——— 2006: Pronominisubjektin käytöstä ja poisjätöstä Kelan asiointikeskusteluissa. – Taru Nordlund, Tiina Onnikki-Rantajääskö & Toni Suutari (toim.): *Kohtauspaikkana kieli. Näkökulmia persoonaan, muutoksiin ja valintoihin*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Martin, Maisa 1999: Suomi toisena ja vieraana kielenä. – Kari Sajavaara & Arja Piirainen-Marsh (toim.): *Kielenoppimisen kysymyksiä* s. 157–178. Jyväskylä: Jyväskylän yliopisto.

——— 2003: Kieli on kuin lammikko – johdatusta toisen kielen oppimiseen. – Leena Nissilä & Heidi Vaarala & Maisa Martin (toim.): *Suolla suomea. Perustietoa maahanmuuttajien suomen kielen opettajille* s. 75–90. Helsinki: Äidinkielen opettajain liitto.

——— 2007: A square peg into a round hole? Fifteen years of research into Finnish as a second language. – *Nordand. Nordisk tidskrift for andrespråksforskning* s. 63–86.

——— 2011: Finnish and Estonian in Second-Language Acquisition Research. – Ain Haas & Peter B. Brown (eds.): *Proceedings of the XIVth, XVth, and XVIth Conferences of the Finno-Ugric Studies Association of Canada* s. 137–148. Rhode Island College: Finno-Ugric Studies Association of Canada (FUSAC).

Martin, Maisa & Alanen, Riikka 2011: Oppijankieli – kehitystä ja virheitä. – Paula Kalaja, Riikka Alanen & Hannele Dufva (toim.): *Kieltä tutkimassa: tutkielman laatijan opas* s. 33–48. Helsinki: Finn Lectura.

Mauranen, Anne & Tiittula, Liisa 2005: Minä käännessuomessa ja supisuomessa. – Anne Mauranen & Jarmo Harri Jantunen (toim.): *Käännessuomeksi. Tutkimuksia suomennosten kielestä* s. 35–70. Tampere: Tampere Studies in language.

Nissilä, Leena 2011: *Viron kielen vaikutus suomen kielen verbien ja niiden rektioiden oppimiseen*. Oulu: Oulun yliopisto. – <http://urn.fi/urn:isbn:9789514296161>

Ojutkangas, Krista 2008: Kielentutkijan mielikuvituksen puute eli mihin todelliseen kielenkäyttöön perustuvaa aineistoa tarvitaan. – *Virittäjä* 112 (2) s. 276–284.

Ortega, Lourdes 2009: *Understanding second language acquisition*. London: Hodder Education.



- Pajunen, Anneli 2001: *Argumenttirakenne: asiointilojen luokitus ja verbien käyttäytyminen suomen kielessä*. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Paunonen, Heikki 2001: Yksikön 1. persoonan pronominin käyttö Helsingin puhekielessä. – Marianne Blomqvist & Mirja Saari & Peter Slotte (red.): *Våra språk i tid och rum* s. 154–168. Helsingfors: Helsingfors universitet.
- Posio, Pekka 2012: *Pronominal subjects in Peninsular Spanish and European Portuguese. Semantics, pragmatics, and formulaic sequences*. Helsinki: University of Helsinki. – <http://urn.fi/URN:ISBN:978-952-10-83203-7>
- Salo, Aino 2015: *Kohteliaisuus S2-oppilaiden sähköposteissa*. Maisterintutkielma. Jyväskylän yliopiston suomen kielen laitos.
- Savijärvi, Ilkka 1977: *Itämerensuomalaisten kielten kieltoverbi*. Joensuun Korkeakoulun Julkaisuja. Sarja A No 8. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Seilonen, Marja 2013: *Epäsuora henkilöön viittaaminen oppijansuomessa*. Jyväskylä: Jyväskylän yliopisto. – <http://urn.fi/URN:ISBN:978-951-39-5025-5>
- Setälä, E. N. 1973: *Suomen kielen lauseoppi*. Helsinki: Kustannusosakeyhtiö Otava.
- Stangroom, Jeremy 2017: Social Science Statistics. – <http://www.socscistatistics.com/> 2.2.2017.
- Suni, Minna 2008: *Toista kieltä vuorovaikutuksessa. Kielellisten resurssien jakaminen toisen kielen omaksumisen alkuvaiheessa*. Jyväskylä: Jyväskylän yliopisto. – <http://urn.fi/URN:ISBN:978-951-39-3209-1>
- Takkinen, Emma & Viinikainen, Tiia 2015: *Suomi toisena kielenä -oppilaiden kertomustekstien rakentuminen*. Maisterintutkielma. Jyväskylän yliopiston suomen kielen laitos.
- Tarnanen, Mirja 2002: *Arvioija valokeilassa: suomi toisena kielenä -kirjoittamisen arviointia*. Jyväskylä: Jyväskylän yliopisto.
- Topling – Toisen kielen oppimisen polut. – <https://www.jyu.fi/hytk/fi/laitokset/kivi/tutkimus/hankkeet/paattyneet-tutkimushankkeet-kansio/topling> 2.2.2017
- VanPatten, Bill & Benati, Alessandro 2010: *Key terms in second language acquisition*. London: Continuum International Publishing Group.
- VISK = Hakulinen, Auli, Vilkuna, Maria, Korhonen, Riitta, Koivisto, Vesa, Heinonen, Tarja Riitta & Alho, Irja 2004: *Iso suomen kielioppi*. Helsinki: Suomalaisen Kirjallisuuden Seura. – <http://scripta.kotus.fi/visk>
- Väänänen, Milja 2016: *Subjektin ilmaiseminen yksikön ensimmäisessä persoonassa. Tutkimus suomen vanhoista murteista*. Turku: Turun yliopisto. – <http://urn.fi/URN:ISBN:978-951-29-6664-6>

**LIITE 1****Esimerkki tehtävölmakkeesta**

Nimi: \_\_\_\_\_

Koulu ja luokka: \_\_\_\_\_

Päivämäärä: \_\_\_\_\_

***Viesti ystävälle***

Olet luvannut kaverillesi, että menet illalla käymään hänen kotonaan. Sinulla on kuitenkin muuta tekemistä. Lähetä ystävälle sähköpostiviesti.

- Kerro, miksi et voi tulla.
- Ehdota uusi aika.

Kirjoita selvällä käsialalla **suomeksi**. Muista tervehtiä ja lopettaa viestisi sopivasti.

Lähtettäjä:

Vastaaottaja:

Aihe:

**LIITE 2****Aineistoesimerkkejä**

Aineiston litteroija on muuttanut teksteissä esiintyvät nimet, osoitteet ja puhelinnumerot.

**Topling, 2. kierros, 3. luokka, Viesti vstävältä ulkomailta, taso A1**

**ID: F-2065-AK3-2-AUT11-HKI29LL-T7B**

Lähettäjä: matti  
Vastaanottaja: kalle  
Aihe: vastaus

moi! minä taas tykkään hamsterista minun  
kissasta ja Big brother pelistä ja siskosat  
isästä aidistä sitten tykkään juoda  
vettä minä tykkään Big brother pelistä  
koska se on kiva!

T: Matti

**Topling, 2. kierros, 8. luokka, Viesti verkkokauppaan, taso B1**

**ID: F-2263-YK8-2- AUT11- TRE09EE-T3B**

Lähettäjä: Maija  
Vastaanottaja: Nettikauppa  
Aihe: Takki

Hei!  
Olen Maija Solki Solkilasta.  
Tilasin viime kuussa teiltä valkoisen  
talvitakin. Takki on minulle vähän  
iso ja se on minulle liian pitkä.  
Haluaisin palauttaa takki teille, jos se  
olisi mahdollista. Pidän enemmän  
vähän lyhyemmistä takkimalleista.  
Haluaisin tilata uuden, erilaisen  
talvitakin teiltä. Jos talvitakkia ei voi  
palauttaa, haluan sen vaihtaa pienmpä-  
än kokoon.  
040-1234567, Maija.  
KIITOS!

## LIITE 3

## Aineiston frekventeimmät yksikön 1. persoonan verbit, sijat 11.–24.

Taulukko 22. Aineiston frekventeimmät yksikön 1. persoonan verbit sijoilla 11.–24.

Verbi	A1 Is/y. 1. %	A2 Is/y. 1. %	B1 Is/y. 1. %	B2 Is/y. 1. %	C1 & C2 Is/y. 1. %	Yhteensä Is/y. 1. %
<b>11. toivon</b> (pstpr)	0/4 0 %	2/10 20,0 %	1/20 5,0 %	0/21 0 %	0/3 0 %	3/58 5,2 %
<b>12. tilaan</b> (puhe)	1/2 50,0 %	2/10 20,0 %	3/27 11,1 %	2/15 13,3 %	0/2 0 %	8/56 14,3 %
<b>13. asun</b> (tila)	8/13 61,5 %	8/22 26,4 %	2/15 13,3 %	0/4 0 %	0/0 0 %	18/54 33,3 %
<b>14. lähden</b> (liike)	2/2 100 %	10/26 38,5 %	2/15 13,3 %	2/7 28,6 %	0/0 0 %	16/50 32,0 %
<b>15. ostan</b> (liike)	2/5 40,0 %	2/10 20,0 %	4/14 28,6 %	1/15 6,7 %	0/1 0 %	9/45 20,0 %
<b>16. teen</b> (temu)	3/3 100 %	2/8 25,0 %	1/20 5,0 %	2/12 16,7 %	0/1 0 %	8/44 18,2 %
<b>17. käyn</b> (liike)	3/7 42,9 %	4/15 26,7 %	1/9 11,1 %	0/6 0 %	0/1 0 %	8/38 21,1 %
<b>18. kirjoitan</b> (temu)	1/2 50,0 %	4/10 40,0 %	2/16 12,5 %	0/8 0 %	0/0 0 %	7/36 19,4 %
<b>19. saan</b> (abstr)	0/1 0 %	0/6 0 %	3/21 14,3 %	0/6 0 %	0/1 0 %	3/35 8,6 %
<b>20. huomaa</b> (perk)	1/1 100 %	0/6 0 %	0/13 0 %	0/11 0 %	0/2 0 %	1/33 3,0 %
<b>21. joudun</b> (abstr)	0/0 0 %	0/6 0 %	2/10 20 %	2/16 12,5 %	0/1 0 %	4/33 12,1 %
<b>22. sanon</b> (puhe)	1/2 50,0 %	3/10 30,0 %	6/19 31,6 %	1/1 100 %	0/0 0 %	11/32 34,4 %
<b>23. yritän</b> (pstpr)	0/1 0 %	5/8 62,5 %	5/18 27,8 %	0/4 0 %	0/0 0 %	10/31 32,3 %
<b>24. juoksen</b> (liike)	3/3 100 %	4/10 40,0 %	0/8 0 %	1/3 33,3 %	0/0 0 %	8/24 33,3 %

Taulukosta 22 näemme loput tutkimuksessa erikseen huomioidut verbit eli sellaiset, joiden esiintymisfrekvenssin olen katsonut vielä jokseenkin merkittäväksi. Taitotasoin näiden verbien esiintymismäärä on kuitenkin sen verran pieni, etten katso aiheelliseksi vertailla niitä samalla tavalla kuin tein edellä kymmenen frekventeimmän verbin kohdalla. Kokonaismäärällisesti ilmisubjektillisuuden ääripäistä löytyvät *asun* ja *toivon*, joista *asun* esiintyy erittäin merkitsevästi ( $z = 3,82$ ) ilmisubjektillisempänä kuin *toivon*.