



**TURUN
YLIOPISTO**

Selkokieli ja morfologia: Vartalonvaihdosten vaikutus S2-oppijoiden lukusujuvuuteen

Logopedian
pro gradu -tutkielma

Laatija:
Anniina Laurila

04.03.2024
Turku

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

Oppiaine: Logopedia

Tekijä: Anniina Laurila

Otsikko: Selkokieli ja morfologia: Vartalonvaihdojen vaikutus S2-oppiloiden lukusujuvuuteen

Ohjaajat: Rosa Salmela, Minna Lehtonen, Raymond Bertram

Sivumäärä: 34 sivua, 1 liitesivu

Päivämäärä: 04.03.2024

Selkokieli on suomen kielen muoto, joka on mukautettu sisällöltään, sanastoltaan ja rakenteeltaan yleiskieltä yksinkertaisemmaksi. Maahanmuuton lisääntyessä Suomessa jatkuvasti on tärkeää tehdä suomenkielistä tietoa helpommin saavutettavaksi, esimerkiksi kehittämällä selkokieltä ymmärrettävämmäksi. Suomen kielen morfologia on rikas, mikä aiheuttaa usein haasteita suomen kieltä toisena kielenä (S2) opiskeleville. Kieliopillisia funktioita ilmaistaan taivuttamalla sanoja eli yhdistämällä morfeemeja, kielen pienimpiä merkityksellisiä yksiköitä, toisiinsa. Lisäksi sanavartalossa tapahtuu usein morfofonologisia muutoksia, jotka muuttavat sanan ilmeä (esim. vesi: vettä, matto: maton). Näistä mahdollisista haasteista huolimatta selkokielen laatimisessa ei juurikaan ohjeisteta morfologian huomioimista. Tämän tutkimuksen tavoitteena oli selvittää katseenseuranta hyödyntäen, vaikuttavatko taivutuksen yhteydessä tapahtuvat vartalonvaihdokset lukusujuvuuteen S2-oppiloiden.

Kokeessa S2-oppilaat lukivat tietokoneen näytöltä lyhyitä selkokielistä tekstejä samalla, kun kamera kuvasi heidän silmänliikkeitään. Teksteissä oli joko paljon vartalonvaihdoksia sisältäviä sanoja, tai sanat oli korvattu semanttisesti samankaltaisilla ja esiintymistajuukseltaan verrattavissa olevilla sanoilla, joissa vartalonvaihdoksia ei esiintynyt. Lukusujuvuutta mitattiin näihin kohdesanoihin ja niitä seuraaviin sanoihin kohdistuneiden fiksaatioiden kestolla, regressioilla eli katseen taaksepäin suuntautumisilla, sekä tekstin kokonaislukuaajalla. Varianssianalyysissä havaittiin, että vartalonvaihdosten esiintyvyydellä ei ollut päävaikutusta lukusujuvuuteen koko aineistossa, mutta kun osallistujat jaettiin kahteen ryhmään luetun ymmärtämisellä arvioitun kielitaidon mukaan, havaittiin vartalonvaihdosten esiintyvyydellä ja kielitaidolla yhdysvaikutus ensimmäisen lukukerran keston ja fiksaatioiden kokonaiskeston kohdesanoja seuraavissa sanoissa. Vartalonvaihdoksen sisältävät sanat hidastivat näissä mittareissa kielitaidoltaan heikompien lukusujuvuutta, mutta kielitaidoltaan parempien lukusujuvuuteen vartalonvaihdokset eivät vaikuttaneet.

Tutkimuksessa saatu tieto on osittain samansuuntaista kuin aiemmissa morfologiatutkimuksissa, ja tukee maltillisesti hypoteesia, että suomen kielessä esiintyvät sanavartaloiden morfofonologiset muutokset aiheuttavat haasteita etenkin aloittelevimmille S2-oppiloiden. Tässä tutkimuksessa vartalonvaihdosten prosessointia tarkasteltiin ensimmäistä kertaa selkokielessä, mikä antaa viitteitä sille, että vartalonvaihdoksiin voisi olla hyödyllistä kiinnittää huomiota selkokieltä tuottaessa.

Avainsanat: morfologinen prosessointi, morfofonologiset muutokset, vartalonvaihtelu, allomorfi, toisen kielen oppijat, suomi toisena kielenä, S2, katseenseuranta, lukusujuvuus, lukeminen, selkokieli

Sisällysluettelo

1	Johdanto	4
1.1	Suomen kielen morfologia	5
1.2	Morfologinen prosessointi	6
1.2.1	Morfologinen prosessointi S1-puhujilla	7
1.2.2	Morfologinen prosessointi S2-oppijoilla	8
1.3	Selkokieli	10
1.4	Lukeminen ja silmänliikkeet	11
1.5	Tutkimuskysymykset	13
2	Menetelmät	14
2.1	Koehenkilöt	14
2.2	Tutkimusmateriaali	15
2.3	Kokeen asetelma ja muuttujat	16
2.4	Kokeen kulku	18
2.5	Tilastolliset analyysit	18
2.6	Eettisyys	19
3	Tulokset	20
4	Pohdinta	24
4.1	Vartalonvaihdojen vaikutus lukusujuvuuteen	24
4.2	Tutkimuksen arviointi	26
4.3	Lopuksi	28
	Lähteet	30
	Liitteet	
	Liite 1. Esimerkki koemateriaalista	

1 Johdanto

Maahanmuuttajien määrä kasvaa Suomessa jatkuvasti, ja vuonna 2022 Suomen väestöstä oli ulkomaalaistaustaisia noin 9.1 prosenttia (Tilastokeskus, 2024). On tärkeää tukea tehokkaasti suomen kielen opetusta toisena kielenä (S2-opetus), koska suomen kielen oppiminen on maahanmuuttajille lähes välttämätöntä yhteiskuntaan integroitumiseksi. Opetuksen lisäksi suomenkielistä tietoa voidaan tehdä helpommin saavutettavaksi, ja yksi siihen tarkoitettu väline on selkokieli. Selkokeskuksen määritelmän mukaan selkokieli on suomen kielen muoto, joka on mukautettu sisällöltään, sanastoltaan ja rakenteeltaan yleiskieltä luettavammaksi ja ymmärrettävämmäksi. Se on suunnattu apukieleksi ihmisille, joilla on vaikeuksia lukea tai ymmärtää yleiskieltä (Selkokeskus, 2021a). Vieraskieliset ovat jo yksi suurimmista selkokieltä tarvitsevista ryhmistä Suomessa (Juusola, 2019). Suomenkielisen tiedon saavutettavuutta voidaan parantaa kehittämällä selkokieltä yhä yksinkertaisemmaksi ja ymmärrettävämmäksi.

Yksi kieliopillisen järjestelmän osa-alueista on morfologia eli muoto-oppi. Morfologialla tarkoitetaan sanojen sisäisen rakenteen osajärjestelmää, jossa sanat koostuvat morfeemeista, kielen pienimmistä muotoaineksista, joilla on merkitys (Tieteen termipankki, 2022). Morfeemeja ovat sanavartalot, johtimet, tunnuksat, taivutuspäätteet, possessiivisuffiksit ja liitepartikkelit (Finn Lectura, 2001). Esimerkiksi sana 'kirja-sto-i-ssa-mme-kin' sisältää yhden morfeemin kustakin edellä mainitusta kategoriasta samassa järjestyksessä. Suomen kielen oppimisessa yksi merkittävä haaste on suomen rikas morfologia, jossa keskeistä on sanojen taivuttaminen ja taivutuksiin liittyvät sanavartaloiden vaihdokset, jotka voivat muuttaa huomattavasti sanan ilmeä (esim. tauko: tauolla, vesi: vettä, matto: maton) (Niilo Mäki Instituutti, 2023). Koska suomen kieli on agglutinoiva eli sanat muodostetaan yhdistämällä morfeemeja toisiinsa, yksittäiselle sanalle on olemassa valtava määrä taivutusmuotoja. Morfologinen segmentaatio eli merkityksellisten yksiköiden erottaminen on oleellinen prosessi kielen ymmärtämisessä, ja aiempien tutkimusten mukaan taivutetut suomen kielen nominit tunnistetaan pääasiassa dekomposition kautta eli pilkkomalla ne sanavartaloon ja päätteeseen (Bertram ym., 1999; Laine ym., 1999; Niemi ym., 1994; Vainio ym., 2014). Toisaalta hyvin yleisille sanamuodoille voi kehittyä kokosanaedustumat mentaaliseen leksikkoon, muistissa olevaan sanavarastoon, jolloin visuaalinen sanantunnistus tapahtuu myös useampimorfeemiselle sanalle yhtenä kokonaisuutena (esim. Lehtonen & Laine, 2003; Soveri ym., 2007).

Suomen morfologisen erityislaatuisuuden ja monimutkaisuuden voi olettaa tuottavan S2-oppijoille haasteita, ja tästä on hieman aiempaa empiiristä näyttöä. Sanantunnistuskokeet ja silmänliikkeitä lukemisen aikana mittaavat tutkimukset viittaavat siihen, että morfologinen monimutkaisuus ja vartalonvaihdokset vaikeuttavat ja hidastavat sanojen tunnistusta ja lukemista S2-oppijoilla verrattuna äidinkielenään suomea puhuviin (Salmela ym., lähetetty julkaistavaksi; Vainio ym., 2014).

Edellä mainituista syistä morfologiaan voisi olla hyödyllistä kiinnittää erityisesti huomiota selkokieltä tuotettaessa. Tästä huolimatta morfologian osalta selkokielen laatimiseen on hyvin vähän ohjeita. Ohjeistuksessa todetaan ainoastaan, että tavallisia sijamuotoja on suosittava ja harvemmin käytettyjä taas vältettävä (Selkokeskus, 2021b). Muuten morfologista monimuotoisuutta, kuten morfeemien määrää tai vartalonvaihdoksia, ei mainita. Ohjeiden vähäisyyteen saattaa vaikuttaa saatavilla olevan tutkimustiedon puute, johon tässä tutkimuksessa pyritään omalta osaltaan vastaamaan. Tämän tutkimuksen tavoitteena oli selvittää, vaikuttavatko suomen kielessä yleiset taivutuksiin liittyvät nominien sanavartaloiden morfofonologiset muutokset lukusujuvuuteen S2-oppijoilla. Lukemista tutkitaan katseenseurannalla, jolla saadaan tietoa silmän liikkeistä eli sakkadeista, katseen pysähdyksistä eli fiksaatioista ja lukemiseen kuluvan ajan kestosta suoraan lukemisprosessin aikana. Lisätiedon saaminen aiheesta on vuosi vuodelta merkityksellisempää maahanmuuton ja selkokielen tarpeen kasvaessa.

1.1 Suomen kielen morfologia

Morfologia on kielen osajärjestelmä, jossa keskiössä on sanojen muotorakenne. Siihen kuuluvat sananmuodostus ja sanojen taivutus. Kielen osajärjestelmiä ei kuitenkaan voida täysin erottaa toisistaan, sillä morfologia kytkeytyy niin semantiikkaan eli merkitysoppiin, syntaksiin eli lauseoppiin kuin fonologiaan eli äänneoppiinkin. Morfofonologia tarkoittaa morfeemien äännerakenteen kuvausta, esimerkiksi sitä, miten morfeemit saavat erilaisen ilmeilyn eri ympäristöissä. Ilmiö on nimeltään morfofonologinen vaihtelu. Saman morfeemin variantteja kutsutaan allomorfeiksi (esim. koti, kodi-; -lla, -llä).

Typologisesti suomen kieli on agglutinoiva eli sanavartalon perään liitetään sidonnoisia morfeemeja (esim. kirja-ssa-si). Näin ollen taivutusparadigma eli kaikkien taivutusmuotojen sarja on laaja, ja yleensä kullakin nominilla on periaatteessa yli 2000 mahdollista taivutettua ja liitteitä sisältävää sananmuotoa (Karlsson, 1983). Esimerkiksi englannin kielessä vastaavia merkityksiä ilmaistaan useista erillisistä sanoista koostuvilla lausekkeilla (esim. *in my book*).

Suomen kielen taivuttaminen ei kuitenkaan ole niin suoraviivaista, että morfeemien yhdisteleminen riittäisi sellaisenaan. Suomen kieli on myös fuusioiva eli morfeemin, erityisesti sanavartalon, sisällä voi tapahtua morfofonologisia muutoksia, kun siihen yhdistetään morfeemeja. Esimerkiksi morfeemi 'talo' säilyy aina samassa muodossa, mutta morfeemilla 'käsi' on neljä allomorfia (käsi, käte-, käde-, kät-). Sanat 'talo' ja 'käsi' kuuluvat eri taivutustyyppiin, koska ne taipuvat eri tavalla. Monimutkaisuudesta kertoo se, että Kielitoimiston sanakirjan (2022) mukaan erilaisia nominien taivutustyypppejä on peräti 49, mutta niiden määrä vaihtelee hieman eri lähteissä sen mukaan, katsotaanko harvinaiset omiksi tyypeikseen vai poikkeuksiksi sääntöihin.

Ison suomen kieliopin (2004) mukaan morfofonologinen vaihtelu voi kohdistua sekä vokaaleihin että konsonantteihin. Tyypillisintä on astevaihtelu, jota on sekä kvantitatiivista että kvalitatiivista. Ensimmäisessä kaksoisklusiilit /pp/, /tt/ ja /kk/ yksinkertaistuvat (esim. takki: takissa) tai käänteisesti yksinäisklusiilit /p/, /t/ ja /k/ kaksinkertaistuvat (esim. vaate: vaateen) silloin, kun niitä edeltää soinnillinen äänne. Jälkimmäisessä astevaihtelun tyyppissä yksinäisklusiili /p/, /t/ tai /k/ vaihtelee toisen konsonantin kanssa tai katoaa (esim. kampa – kamman, vika – vian). Kielitoimiston sanakirja (2022) jakaa astevaihtelun 13 tyyppiin. Myös muunlaista vaihtelua esiintyy sekä vokaaleissa että konsonanteissa. Esimerkiksi /i/ ja /e/ tai /m/ ja /n/ vaihtelevat nomineissa (pieni: pienenä, lumi: lunta).

Etenkin englanninkielisessä kirjallisuudessa käytetään morfologian yhteydessä termiä läpinäkyvyys (*transparency*). Sanat, joissa morfofonologisia muutoksia ei esiinny taivutettaessa, ovat läpinäkyviä (*transparent*) ja sanat, joissa näitä muutoksia tapahtuu, ovat osittain läpinäkyviä (*semitransparent*). Viittaan tässä tutkimuksessa sanalla 'läpinäkyvyys' morfofonologisten muutosten eli vartalonvaihdojen esiintyvyyteen muuttujana, mikä tarkoittaa, että sanoissa joko esiintyy vartalonvaihdoksia tai ei.

1.2 Morfologinen prosessointi

Kompleksisten ja monimorfeemisten sanojen morfologista prosessointia on tutkittu jo pitkään erityisesti äidinkielessä. Sen sijaan morfologinen prosessointi toisessa tai vieraassa kielessä on saanut tutkimuksissa tähän päivään mennessä vähemmän huomiota – erityisesti suomen kaltaisissa taivutussysteemeiltään monimutkaisissa kielissä. Alan tutkijat eivät ole yksimielisiä minkään teorian ylivoimaisesta paikkansapitävyydestä.

Psykolingvistisen morfologiatutkimuksen alkua ajoilta 1970-luvulta lähtien sanojen tunnistukselle on esitetty kahta kilpailevaa perusajatusta. Monet morfologisen prosessoinnin teoriat (esim. Murrell & Morton, 1974; Taft & Forster, 1975) nojaavat mekanismiltaan sanojen dekompositioon, jossa sanat pilkkotaan niistä koostuviin morfeemeihin (esim. talo-i-ssa-mme) ennen sanan vartalon semanttista muistista hakua. Sanavartalo jää jäljelle, kun taivutetusta sanasta erotetaan päätteet ja mahdolliset tunnukset (esim. nomineissa monikon tunnus t, i tai j), ja sen jälkeen se yhdistetään merkitykseen. Dekompositio perustuu siis opittuihin kielen morfologisiin operaatioihin (Karlsson, 1983), ja sen etu piilee siinä, että morfologinen systemaattisuus tekee etenkin sanojen varastoinnista muistille vähemmän kuormittavaa. Vastakkaisten teorioiden (esim. Butterworth, 1975; Manelis & Tharp, 1977) mukaan monimorfeemisetkin sanat ovat tallentuneet kokosanaedustumiksi mentaaliseen leksikkoon. Nämä teoriat mahdollistavat morfologisen prosessoinnin tehokkuuden, kun sanat voidaan tunnistaa hyvinkin nopeasti ilman erillistä pilkkomista, mutta toisaalta sanamuotojen suuri määrä on muistille kuormittavaa.

Käytännössä morfologinen prosessointi ei todennäköisesti ole näin mustavalkoista, vaan siihen voi vaikuttaa mm. kieli ja sen rakenne, sanan frekvenssi ja säännönmukaisuus sekä kielitaidon taso. Morfologiselle prosessoinnille on ehdotettu kahta rinnakkaista prosessointireittiä (*dual-route model of morphology*): säännöllisesti taipuvat sanat (esim. *walk: walked*) tunnistetaan dekomposition kautta ja epäsäännölliset sanamuodot (esim. *run: ran*) kokosanaedustuman avulla (esim. Pinker, 1991). Tämä sopii ainakin englannin kieleen. On myös oletettu, että mitä yleisempi sanamuoto on kyseessä, sitä todennäköisemmin sille muodostuu kokosanaedustuma (esim. Alegre & Gordon, 1999; Lehtonen & Laine, 2003; Taft, 1979; Taft & Ardasinski, 2006). Seuraavissa alaluvuissa käsitellän kielitaidon tason vaikutusta morfologiseen prosessointiin, mikä valottaa sitä, miksi morfologiaa voisi olla perusteltua ottaa enemmän huomioon selkokielen ohjeissa.

1.2.1 Morfologinen prosessointi S1-puhujilla

Suomea äidinkielenään puhuvilla tehdyissä tutkimuksissa on havaittu, että morfologisesti kompleksiset, kuten taivutetut sanat, tunnistetaan hitaammin ja huonommalla tarkkuudella kuin yksimorfeemiset sanat, jotka ovat verrattavissa pituudeltaan ja frekvenssiltään (Bertram ym., 1999; Laine ym., 1999; Niemi ym., 1994; Vainio ym., 2014). Hitaampi prosessointinopeus viittaa dekompositioon morfologisessa prosessoinnissa, sillä sanojen pilkkominen ja lopulta osien merkitysten integroiminen vaatii aikaa. Kuitenkaan

korkeafrekvenssisten taivutettujen sanojen kohdalla vastaavaa prosessoinnin hidastumista ei ole tutkimuksissa havaittu (Lehtonen & Laine, 2003; Soveri ym., 2007). Tämä viittaisi siihen, että hyvin yleisille sanoille on muodostunut vahva kokosanaedustuma muistiin, mikä nopeuttaa sanantunnistusta. Perinteisesti morfologista prosessointia on tutkittu leksikaalisen päätöksenteon tehtävillä (*lexical decision task*), joissa sanat esitetään yksittäin, ilman kontekstia. Hyönä ja kumppanit (2002) eivät kuitenkaan havainneet prosessointinopeuden hidastumista taivutettujen sanojen kohdalla verrattuna yksimorfeemisiin sanoihin, kun sanat esitettiin lauseissa ja prosessointia tutkittiin käyttäen silmänliikekameraa luettaessa. Konteksti saattaa siis tarjota tukea sanantunnistuksessa.

Allomorfeja koskevissa tutkimuksissa on havaittu, että suomea äidinkielenään puhuvat tunnistavat yhtä nopeasti taivutetut sanat, joissa esiintyy morfofonologisia muutoksia sanavartalossa, kuin sanat, joissa näitä muutoksia ei esiinny (Järvikivi & Niemi, 2002; Vainio ym., 2014). Näin ollen natiivipuhujilla olisi eri allomorfeille omat edustumansa mentaalisessa leksikossa. Suomen kielelle onkin kehitetty SAID-malli (*Stem Allomorph/Inflectional Decomposition*; (Niemi ym., 1994), jonka oletuksena on, että sanantunnistusta varten mentaalisessa leksikossa on morfeemiedustumat erikseen taivutusvartaloille ja -päätteille. Sanavartalot ovat edustettuina allomorfeina (esim. kukka ja kuka-), minkä ansiosta välttyään monimutkaisilta morfofonologisilta operaatioilta. Malli pätee nimenomaan taivutettuihin sanoihin, ei johdettuihin, ja poikkeuksena ovat hyvin yleiset sanamuodot, joille voi olla muodostunut kokosanaedustumia.

1.2.2 Morfologinen prosessointi S2-oppijoilla

Morfologisen prosessoinnin tutkimus S2-oppijoilla on vasta alussa, mutta sitä olisi tärkeää hyödyntää myös selkokieltä kehitettäessä, koska morfologian osalta selkokielen laatimiseen on vain vähän ohjeita. Vainio ja kumppanit (2014) tutkivat venäjää ja kiinaa äidinkielenään puhuvien aikuisten taivutettujen sanojen (substantiivien) tunnistusta heidän toisessa kielessään eli suomen kielessä. Venäjä on morfologisilta piirteiltään lähempänä suomea kuin kiina, koska venäjän kielessä ilmaistaan kieliopillisia merkityksiä sanoja taivuttamalla ja sanavartalon läpinäkyvyys vaihtelee, kun taas kiinan kielessä sanoja ei taivuteta. Venäjän kielessä esimerkiksi on kuusi sijamuotoa. Kielten typologinen ero näytti vaikuttavan morfologiseen prosessointiin, sillä venäjänkieliset tunnistivat morfologisesti kompleksisemmat suomen sanat hitaammin verrattuna yksimorfeemisiin sanoihin, kun taas kiinankielisillä tällaista eroa ei havaittu. Venäjänkieliset siis todennäköisesti prosessoivat taivutetut sanat dekomposition

kautta, kun taas kiinankieliset tunnistivat sanat kokosanaedustuman avulla. Lisäksi tutkittiin, vaikuttavatko sanavartalossa esiintyvät morfofonologiset muutokset prosessointinopeuteen. Sanat, joissa oli vartalonvaihdos, tunnistettiin venäjää ja kiinaa puhuvien keskuudessa hitaammin kuin sanat, joissa sitä ei esiintynyt, ja kiinaa puhuvilla ero oli suurempi. Lisäksi niissä tehtiin enemmän virheitä. Vartalonvaihdokset siis vaikuttivat S2-oppijoilla sanojen prosessointinopeuteen hidastavasti ja enemmän äidinkielen ollessa typologisesti kauempana suomea.

Myös Salmelan ja kumppanien (lähetetty julkaistavaksi) tutkimuksessa havaittiin, että vartalonvaihdokset aiheuttavat haasteita S2-oppijoille. Tutkimuksessa verrattiin yksimorfeemisten sanojen, läpinäkyvästi taipuneiden sanojen ja astevaihtelun alaisesti taipuneiden sanojen prosessointia leksikaalisen päätöksenteon tehtävässä ja katseenseurannalla. Leksikaalisen päätöksenteon tehtävässä vartalonvaihdoksen sisältävät sanat prosessoitiin hitaimmin, ja niissä oli eniten epätarkkuutta. Myös S1-puhujilla tavattiin reaktioajoissa vastaava havainto, mutta pienempi. Heillä tarkkuudessa ei esiintynyt eroja, vaan kattoefekti. Lauseen kontekstissa katseenseurannassa S2-oppijoilla vartalonvaihdoksiin liittyvä prosessoinnin hidastuminen säilyi, kun S1-puhujilla se katosi. Läpinäkyvästi taipuneiden ja vartalonvaihdoksellisten sanojen välillä oli tilastollisesti merkitsevä ero fiksaatioiden kokonaiskestossa. Tutkijoiden mukaan morfologisen prosessoinnin haasteet näyttävät liittyvän enemmän suomen kielen fusionaalisiin kuin agglutinoiviin piirteisiin.

Samansuuntaisia tuloksia on saatu myös muita toisia kieliä koskevissa tutkimuksissa, joissa morfologisen läpinäkyvyyden on todettu vaikuttavan toisen kielen prosessointiin niin, että läpinäkyvien sanojen prosessointi helpottuu (esim. Basnight-Brown ym., 2007; DeKeyser, 2005). Efekti voi johtua siitä, että harvinaisemmat allomorfit eivät ole yhtä tuttuja toisen kielen oppijoille kuin yleisimmät, vaikka tutkimuksissa kohdesanojen lemma- ja pintafrekvenssit onkin pyritty asettamaan yhtä suuriksi (Vainio, 2014). Toisaalta efekti voi johtua äidinkielen siirtovaikutuksesta. Jos äidinkielessä ei ole allomorfeja, tuottavat ne todennäköisemmin hankaluuksia, koska niihin ei ole tottunut. Kolmas vaihtoehto on, että toisen kielen oppijat nojaavat enemmän sanamuotojen samankaltaisuuteen kuin äidinkielen puhujat. Muoto vaikuttaa toisen kielen prosessointiin, sillä ilman vartalonvaihdosta taivutetut sanat muistuttavat enemmän yksinkertaisia sanamuotoja eli perusmuotoja, kuin sanat, joissa tapahtuu vartalonvaihdos.

Morfologista prosessointia ei ole juurikaan tutkittu selkokielessä, eikä myöskään S2-oppijoiden lukiessa pidempää tekstiä. Dahlgren (2020) on kuitenkin tutkinut laadullisesti selkokielen oppaan työstämisprosessin aikana sitä, mitkä tekijät S2-alkeisoppijoiden ($n = 13$) subjektiivisen arvion mukaan ovat haastavia ymmärtää. Morfologisista rakenteista vaikeita olivat sijapäätteet postpositioiden sijaan (esim. avaamiseen – avaamista varten) ja possessiivisuffiksit (esim. passisi). Kuten tässä luvussa on esitelty, morfofonologisten muutosten aiheuttama prosessoinnin hitaus näyttää säilyvän S2-oppijoilla lauseiden lukemisen yhteydessä. Tämä antaa viitteitä siitä, että myös selkokielistä tekstiä prosessoitaessa suomen oppijoiden lukusujuvuus ja luetun ymmärtäminen voisivat parantua, kun sanat ovat morfologisesti mahdollisimman yksinkertaisia, esimerkiksi eivät sisällä vartalonvaihdoja.

1.3 Selkokieli

Selkokielen kehittäminen on keskittynyt erilaisten ohjeistusten laatimiseen. Selkokielen kehitystyö on painottunut kolmannelle sektorille, jossa Selkokeskus on ollut aikaansaavin taho (Leskelä & Uotila, 2020). Suomen selkokielen ohjeisiin on haettu myös mallia kansainvälisesti yleisluontoisista ohjeista (kehitysvammajärjestö Inclusion Europe ja kirjastoalan kattojärjestö IFLA), mutta ongelmana niissä on ollut se, että ne jättävät ulkopuolelle suomen kielen erityispiirteet, kuten juuri monimutkaisen morfologian (Leskelä & Uotila, 2020). Suomen selkokielen kriteerit ovat olleet melko karkeita, ja ne ovat liittyneet pikemminkin tekstikokonaisuuteen kuin yksittäisiin kielenpiirteisiin, mutta koko ajan lähestytään tarkempia kielellisiä normeja (Leskelä & Uotila, 2020). Yleisesti selkokieltä on tutkittu suhteellisen vähän ja kokeellisin menetelmin vielä vähemmän, vaikka sitä on kehitetty jo lähes 40 vuotta. Vasta 2000-luvulla selkokieli on alkanut saada tutkimuksista vankempaa lingvististä tukea (Leskelä & Uotila, 2020), kun myös selkokielisten aineistojen tarve on kasvanut.

Selkokeskus on laatinut yleisohjeita selkokirjoittajille selkokielen tuottamiseen tekstin kokonaisuudesta, selkokielen sanastosta ja helpoista kielen rakenteista sekä täsmäohjeita eri tekstilajien kirjoittajille. Rakenteisiin liittyvissä ohjeissa mainitaan morfologian osalta tavallisten sijamuotojen suosiminen ja harvemmin käytettyjen välttäminen, kun taas syntaksin osalta ohjeita on runsaammin (Selkokeskus, 2021b). Selkokeskus on tehnyt yhteistyötä myös Kotimaisten kielten keskuksen kanssa, ja he ovat kehittäneet selkokielen mittarin (2018, uudistettu 2022) tekstin selkokielisyyden arviointiin. Mittarin kriteereissä kielen rakenteiden osalta mainitaan lisäksi nominien perusmuotojen käyttäminen, jos se on lauseyhteydessä

mahdollista ja luontevaa, sekä useita erilaisia elementtejä (johtimia, taivutuspäätteitä ja liitteitä) sisältävien sanojen välttäminen. Selkokielen mittari ole tarkoitettu ohjeistukseksi, jonka avulla voisi tuottaa selkokielistä tekstiä.

Maahanmuuttajat ovat ensisijaisesti kielenoppijoita, ja eivät siis lähtökohtaisesti tarvitse samanlaista kieltä kuin esimerkiksi kehitysvammaiset tai muistisairaat, jotka ovat myös selkokielen kohderyhmiä. S2-oppijat hyötyvät rakenteellisesti mahdollisimman yksinkertaisesta kielestä, jonka tulkitseminen onnistuu alkeistason kielitaidolla. (Dahlgren, 2020). Siksi heille suunnatussa selkokielessä juuri morfologia ja syntaksi ovat oletettavasti tärkeässä asemassa. Selkokeskuksen asiantuntijoiden mukaan selkokielisestä aineistosta hyötyvät eniten A1–B1-taitotason kielenoppijat ja selkokielen tarve vähenee koko ajan kielitaidon kasvaessa (Juusola, 2019).

1.4 Lukeminen ja silmänliikkeet

Katseenseuranta on mainio tapa tutkia lukemista, koska sen avulla saadaan tarkkaa ajallista ja paikallista tietoa kirjoitetun kielen prosessoinnista. Lukeminen näkyy silmänliikekameran kuvaamina katseen pysähtymisinä ja siirtymisinä. Katseen pysähdyksiä kiinnekohtiin kutsutaan fiksaatioiksi, ja silmän liikkeitä fiksaatioiden välillä kutsutaan sakkadeiksi. Lisäksi on tavallista, että tekstissä palataan taaksepäin lukemaan uudelleen, ja näitä taaksepäin suuntaavia sakkadeita kutsutaan regressioiksi. Bertram ja Hyönä (2023) argumentoivat, että silmänliikekameran käyttö on hyödyllistä morfologiatutkimuksessa. Tällöin tutkittavat sanat esiintyvät luonnollisessa ympäristössä, kun ne luetaan lauseisiin sijoitettuina, minkä lisäksi katseenseuranta mahdollistaa useiden riippuvien muuttujien samanaikaisen tutkimisen. Vainio (2006) taustoittaa, että katseenseuranta lukemisessa perustuu ensinnäkin siihen, että prosessointiin vaikuttaa prosessoitavan rakenteen, sanan tai tekstin, kompleksisuus tai vaikeus eli hankalia kohtia prosessoidaan kauemmin. Lisäksi ns. silmä-mieli-hypoteesin (*eye-mind hypothesis*; Just & Carpenter, 1980) mukaan prosessoimme pääosin sitä, mitä katsomme. Tästä on kuitenkin poikkeuksena ns. valumisefekti (*spillover effect*; Rayner & Duffy, 1986), jossa haastava rakenne, esim. vieras sana, vaikuttaa myös sitä seuraavan sanan prosessointiin hidastavasti.

Lukemisen yksi mitattava ominaisuus on lukusujuvuus. Lukusujuvuuden määritelmä on hieman vaihteleva, mutta yleisesti se tarkoittaa teknisen lukemisen vaivattomuutta. Tekninen lukeminen eli dekodaus tarkoittaa visuaalista sanantunnistusta, jossa kirjaimista muodostetaan

sanoja. Sujuva lukija etenee tekstissä nopeasti, ja sanojen tunnistus on lähes automaattista. Lukusujuvuus on itse asiassa taitojen kokonaisuus, ja sen määritelmä sisältää usein kolme osataittoa: lukemisen tarkkuuden (muoto ja merkitys), lukemisen nopeuden ja ääneen luettaessa tekstin ilmeikkään lukemisen eli prosodian (esim. Kuhn & Stahl, 2003). Vaikka usein lukemisen tarkkuus ja lukusujuvuus erotetaan lukutaidon erillisiksi osa-alueiksi, on hyvä ottaa tarkkuus huomioon lukusujuvuuden määritelmässä, koska pelkkä nopeuden painottaminen johtaa helposti vain kovaan vauhtiin, jolloin tekstin merkitys ja ymmärtäminen kärsivät (Niilo Mäki Instituutti, 2023). Sujuva lukeminen on edellytys luetun ymmärtämiselle, ja joissakin lukusujuvuuden määritelmissä ymmärrys on osa sujuvuutta (esim. Harris & Hodges, 1995). Koska tässä tutkimuksessa on kyse silmänliikkeistä mielessä lukemisen aikana, sujuvuutta tarkasteltiin fiksaatioiden kestoina, regressioiden määrinä ja tekstin kokonaislukuajana, jotka heijastavat pääosin lukunopeutta. Koe oli kuitenkin suunniteltu niin, että koehenkilöiden pitäisi myös lukea mahdollisimman tarkasti ymmärtääkseen tekstejä.

Lukusujuvuuden tutkimus Suomessa on perinteisesti keskittynyt dysleksiatutkimukseen yksikielisillä lapsilla. S2-oppijoiden lukusujuvuutta koskevissa tutkimuksissa ei ole tiettävästi tarkasteltu tekstin lukusujuvuuteen vaikuttavia tekijöitä eikä hyödynnetty katseenseurantaa. Muualla tehdyissä toista kieltä koskevissa tutkimuksissa on tarkasteltu erityisesti kielitaidon vaikutusta mutta myös tekstin vaikeustason vaikutusta silmänliikkeisiin lukemisen aikana. Kun lukijat prosessoivat vaikeaa tekstiä, fiksaatiot ovat pidempiä, sakkadit lyhyempiä, sanojen yli hyppimistä on vähemmän ja regressioita enemmän (Castelhano & Rayner, 2008). Samoin luettaessa haastavia sanoja fiksaatioita on enemmän, ja ne ovat pidempikestoisia kuin luettaessa helppoja sanoja (Just & Carpenter, 1980).

Rayner ja kumppanit (2006) osoittivat, että lukijoiden subjektiiviset arviot tekstin vaikeudesta korreloivat heidän silmänliikkeidensä kanssa siten, että fiksaatioita oli enemmän ja ne olivat pidempiä vaikeampaa tekstiä lukiessa. Rets ja Rogaten (2021) tutkivat englantia toisena kielenä oppivien aikuisten silmänliikkeitä luettaessa alkuperäisiä yleiskielisiä ja yksinkertaistettuja selkokieliä tekstejä. Yksinkertaistetuissa teksteissä sanan ensimmäisen lukukerran fiksaatiot (*first pass fixations*) olivat pidempiä ja toisen lukukerran fiksaatiot (*second pass fixations*) lyhyempiä kuin alkuperäisissä teksteissä. Tutkijoiden mukaan tekstin yksinkertaistaminen mahdollisti tekstin tehokkaamman alustavan prosessoinnin, ja siten oli vähemmän tarvetta uudelleen lukemiselle. Tekstin luettavuutta on tutkittu englannin kielessä perinteisesti yksinkertaisten kielenpiirteiden, kuten sanojen ja lauseiden pituuden näkökulmasta (Nahatame, 2023).

1.5 Tutkimuskysymykset

Suomen morfofonologinen monimutkaisuus voi aiheuttaa haasteita toisen kielen oppijoille, ja siksi se voisi olla syytä huomioida selkokieltä tuotettaessa. Tämän tutkimuksen tavoitteena oli selvittää, voiko S2-oppijoiden lukemista helpottaa välttämällä erityisesti vartalonvaihdoksia eli morfofonologisia muutoksia sisältäviä sanoja, tarkemmin nomineja (kuten ilta: illan, lämmin: lämpimän), eli paraneeko tällöin lukusujuvuus. Lisäksi oltiin kiinnostuneita siitä, vaikuttaako vartalonvaihdosten esiintyvyys eri tavalla kielitaidoltaan eri tasolla oleviin S2-oppijoihin. Tutkimuskysymykset olivat:

1. *Vaikuttaako tekstissä esiintyvien vartalonvaihdosten määrä S2-oppijoiden lukusujuvuuteen?*
2. *Vaikuttaako kielitaidon taso siihen, millainen vaikutus vartalonvaihdoksilla on lukusujuvuuteen?*

Kokeessa S2-oppijat lukivat lyhyitä selkokielisiä tarinoita tietokoneen näytöltä samalla kun silmänliikekamera seurasi heidän katsettaan tekstissä. Puolet teksteistä sisälsi paljon vartalonvaihdoksia sisältäviä tutkimuksen kiinnostuksen kohteena olevia sanoja ja puolessa teksteistä nämä kohdesanat oli korvattu sanoilla, joissa vartalonvaihdoksia ei esiintynyt. Hypoteesina oli, että lukusujuvuus on parempi, kun vartalonvaihdosten määrää on rajoitettu. Morfofonologiset muutokset ja suffiksien lisääminen voivat muuttaa melko radikaalistikin sanan ulkoasua, mikä voi vaikuttaa sanantunnistukseen ja prosessointinopeuteen. Aiempaa tutkimusta vartalonvaihdosten merkityksestä S2-oppijoilla on vähän, mutta sen mukaan vartalonvaihdosten esiintyvyys vaikuttaa sanantunnistuksen nopeuteen ja virheiden määrään (Salmela ym., lähetetty julkaistavaksi; Vainio ym., 2014). Toisena hypoteesina oli, että vartalonvaihdosten rajoittaminen helpottaa lukemista enemmän aloittelevilla oppijoilla, joiden kielitaito on heikompi, verrattuna hieman edistyneempiin oppijoihin. Perusteena tälle on se, että kielitaidon kehittyessä voitaisiin nojata enemmän nopeampaan sanan tunnistukseen, jossa eri allomorfeille on muodostunut omat edustumansa mentaaliseen leksikkoon, kuten oletetaan tapahtuneen äidinkielen puhujilla (Järvikivi & Niemi, 2002; Niemi ym., 1994; Vainio ym., 2014). Aiemmissä tutkimuksissa ei ole tarkasteltu vartalonvaihdoksia sisältäviä sanoja pidemmän tekstin kontekstissa, vaan yksittäisinä tai lauseissa. Tutkimuksesta saatua tietoa voidaan hyödyntää luotaessa ohjeita sekä tuotettaessa selkokieltä, eli tehdä siitä mahdollisesti morfologian osalta helpommin ymmärrettävää.

2 Menetelmät

2.1 Koehenkilöt

Tutkimukseen osallistui 34 aikuista koehenkilöä, jotka opiskelivat suomen kieltä toisena kielenä. Osallistujat rekrytoitiin Turun yliopiston ja Turun suomenkielisen työväenopiston suomen kielen kursseilta, joiden taso oli eurooppalaisen kielitaidon viitekehyksen mukaan A2–B2. Tutkittavilla oli normaali tai silmälasein korjattu näkö, ja heillä ei ollut lukihäiriötä. Analyysieihin otettiin lopulta mukaan 28 koehenkilöä (ks. kohta 2.5 Tilastolliset analyysit). Heistä 19 oli naisia ja 9 miehiä. Koehenkilöt olivat iältään 18–46-vuotiaita (*ka* 31.5 v). Heistä 80 %:lla oli korkeakoulututkinto, ja kaikilla oli vähintään peruskoulun jälkeinen ammattitutkinto. He olivat asuneet Suomessa puolesta vuodesta seitsemään vuoteen (*ka* 2.8 v). Osallistujien suomen kielitaidon itsearvio (asteikolla 1–10) vaihteli välillä 2–6 (*ka* 3.4). Osallistujat olivat 12 eri äidinkielen puhujia (ukraina, saksa, englanti, espanja, venäjä, turkki, viro, puola, hollanti, persia, arabia, kiina). Kolmen tutkimukseen mukaan otetun koehenkilön osalta puuttuivat tiedot iästä, koulutuksesta, Suomessa asumisen ajasta, kielitaidon itsearviosta ja äidinkielestä, joten ne eivät näy taustatiedoissa. Osallistujat jaettiin kielitaidon perusteella kahteen ryhmään (ks. kohta 2.3 Kokeen asetelma ja muuttajat), ja taulukossa 1 on esitetty taustatiedot paremmassa ja heikommassa kielitaitoryhmässä. Sukupuoli, ikä ja koulutustaso olivat hyvin samalla tavalla edustettuina molemmissa ryhmissä.

Taulukko 1

Taustatiedot kielitaitoryhmittäin

	Paremmat (<i>n</i> = 14)	Heikommat (<i>n</i> = 14)
Sukupuoli		
Nainen	<i>n</i> = 9	<i>n</i> = 10
Mies	<i>n</i> = 5	<i>n</i> = 4
Ikä (<i>ka</i>)	20–46 v (32 v 2 kk)	18–46 v (30 v 10 kk)
Koulutus		
Toinen aste	<i>n</i> = 2	<i>n</i> = 3
Korkeakoulu	<i>n</i> = 10	<i>n</i> = 10
Suomessa asuttu aika (<i>ka</i>)	1–7 v (3 v 2 kk)	6 kk–5 v (2 v 4 kk)
Kielitaidon itsearvio 1–10 (<i>ka</i>)	2–6 (3,75)	2–5 (3)

2.2 Tutkimusmateriaali

Graduntekijä laati käytettävän materiaalin. Lukemiskoetta varten laadittiin kahdeksan selkokielistä lyhyehköä, noin sadan sanan pituista tekstiparia, joissa oli joko paljon vartalonvaihdoksia sisältäviä sanoja (esim. ilta: illalla, kuppi: kupissa), tai joissa nämä sanat oli korvattu semanttisesti ja esiintymistaajuudeltaan samankaltaisilla sanoilla, joissa ei ollut vartalonvaihdoksia (esim. aamu: aamulla, pullo: pullossa). Ensin laadittiin lista kohdesanapareista, joista tutkimuksessa oltiin kiinnostuneita. Sanat valittiin mielivaltaisesti käyttäen apuna S2-oppikirjoja niin, että ne olivat melko tavallisia ja oletettavasti A2–B2-tasoisille suomen opiskelijoille tuttuja. Sanaparit koostuivat vartalonvaihdoksen sisältävästä ja läpinäkyvästä taivutetusta sanasta (esim. kadulle ja tielle). Sanat olivat nomineja, suurimmaksi osaksi substantiiveja ja adjektiiveja, ja sisälsivät myös muutaman numeraalin. Parit olivat samasta sanaluokasta ja semanttisesti samankaltaisia, tai niitä oli muuten mahdollista käyttää samassa kontekstissa. Lopuksi sanojen esiintymistaajuus tarkistettiin TurkuNLP:n laajasta suomen kielen internetkorpuksista (Luotolahti ym., 2015). Kohdesanoja vaihdettiin tarvittaessa niin, ettei minkään tekstiparin kohdesanojen yhteenlasketussa esiintymistaajuudessa tai pituudessa ilmennyt lopulta tilastollisesti merkitsevää eroa (kaikissa $p > .166$).

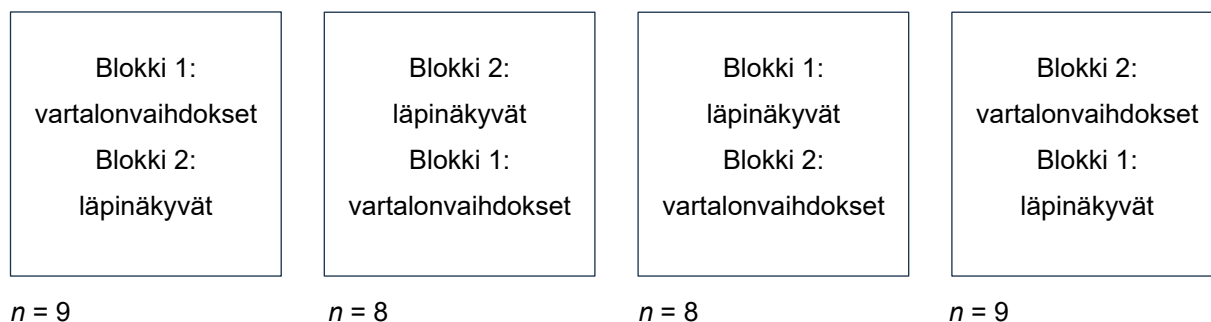
Seuraavaksi sanojen pohjalta luotiin kahdeksan pientä koherenttia kertomusta, eli yhteensä 16 eri tekstiä, joihin sanaparit sijoitettiin (ks. liite 1). Tekstiparit olivat muuten täysin identtisiä, mutta erosivat ainoastaan kohdesanojen osalta. Näin ollen myös juoni oli tekstiparin molemmissa teksteissä sama. Kohdesanoja taivutettiin ja johdettiin niin, että ne sopivat teksteihin. Sanaparit olivat kuitenkin pääosin keskenään samassa sijamuodossa, mutta 13 poikkeusta tehtiin, jotta sanat sopisivat luonnollisemmin kontekstiin (esim. ullakolta ja kellarista). Neljästä adjektiiviparista tuli tässä vaiheessa tapaa ilmaisevia adverbeja (esim. arka > arasti). Muuten kaikki sanat olivat edelleen nomineja. Sanoja taivutettiin 11 sijamuodossa (nominatiivi, partitiivi, genetiivi, inessiivi, elatiivi, illatiivi, adessiivi, ablatiivi, allatiivi, essiivi, translatiivi). Yksi adjektiivipari taipui vertailumuodossa (komparatiivi). Lisäksi kahdesta substantiiviparista tuli adjektiiveja johdettaessa (esim. tuuli: tuulinen, sade: sateinen). Kertomuksissa oli 19–23 kohdesanaa, ja yhteensä kaikissa teksteissä oli 167 kohdesanaparia (167 vartalonvaihdosta sisältävää sanaa ja 167 sanaa ilman vartalonvaihdosta). Kohdesanoissa tapahtuvista vartalonvaihdoksista 44 % oli kvalitatiivista astevaihtelua (esim. lupa: luvan), 33 % kvantitatiivista astevaihtelua (esim. pankki: pankissa) ja 23 % muunlaista vaihtelua (esim. hiiri: hiiren).

Kertomukset pyrittiin tekemään selkokielisiksi Selkokeskuksen kirjoittamisohjeiden mukaisesti ja ne olivat lauserakenteeltaan ja sanastoltaan melko yksinkertaisia, A2–B2-tason suomen opiskelijoille sopivia. Tekstin sopivaan haastavuustasoon otettiin mallia myös suomen kielen A2–B2-tasojen oppikirjoista. Teksteissä käytettiin enimmäkseen preesensaiamuotoa, joka on tutuin ja helpoin alkeisvaiheen toisen kielen oppijoille. Tekstien selkokielisyyttä arvioitiin selkokielen mittari 2.0:n (Selkokeskus, 2022) avulla soveltuvin osin. Mittari sisältää alakategoriat 'teksti kokonaisuutena', 'sanat', 'kielen rakenteet' sekä 'ulkoasu ja kuvitus'. Niissä on yhteensä 96 kriteeriä, joista arvioitiin 64 kohtaa asteikolla 1–3. Tekstit saivat arvosanaksi 2.9 eli ne ovat selkokieltä (raja-arvo 2.5). Mittarin ohjeissa mainitaan kuitenkin, että tulos ei täysin luotettavasti kerro tekstin selkokielisyydestä, jos mittarilla arvioidaan muita kuin informoivia tekstejä, ja tässä tutkimuksessa tekstit olivat lyhyitä tarinoita eli fiktiivisiä tekstejä. Lisäksi selkokielen mittari 2.0:n käyttö edellyttää selkokielen asiantuntemusta ja yleistä kielellistä osaamista.

Jokaista tekstiparia varten luotiin kaksi luetun ymmärtämistä testaavaa väittämää, joihin vastattiin 'oikein' tai 'väärin'. Toinen väittämä liittyi aina kertomuksen teemaan tai juoneen, ja toinen johonkin tekstin yksityiskohtaan. Oikein- ja väärinvastausten määrä oli tasapainotettu.

2.3 Kokeen asetelma ja muuttujat

Koeasetelma oli sekä koehenkilöiden sisäinen että välinen (*mixed-design*). Koehenkilöiden sisäinen riippumaton muuttuja oli vartalonvaihdojen esiintyvyys eli läpinäkyvyys, jolla oli kaksi tasoa: vartalonvaihdoksen sisältävät sanat ja läpinäkyvät sanat. Lukemiskokeessa oli neljä erilaista koetilannetta järjestysvaikutuksen kontrolloimiseksi, ja kukin tutkittava osallistui yhteen tilanteeseen (kuvio 1). Tekstiparit jaettiin kahteen blokkiin siten, että blokit olivat samankaltaisia kohdesanojen pituuden ja esiintymistaajuuden suhteen (kaikissa $p > .195$). Puolet tutkittavista luki ensimmäisen blokin vartalonvaihdoksia sisältävät tekstit sekä toisen blokin tekstit, joissa oli läpinäkyviä sanoja. Puolet luki päinvastaiset tekstit. Kaikki lukivat siten neljä tekstiä, joissa oli vartalonvaihdoksia, ja neljä, joissa niitä ei ollut. Lisäksi blokkien lukujärjestys oli tasapainotettu eli puolet tutkittavista luki ensin ensimmäisen blokin tekstit ja puolet luki ensin toisen blokin tekstit.

Kuvio 1*Eri koetilanteet*

Koehenkilöiden välinen riippumaton muuttuja oli kielitaidon taso, jonka perusteella tutkittavat jaettiin kahteen ryhmään: kielitaidoltaan parempiin ja heikompiin. Asetelma oli siis ryhmien välisen muuttujan osalta kvasikokeellinen. Kielitaidon tasoa arvioitiin luetun ymmärtämisellä, eli prosenttiosuudella, jonka tutkittavat vastasivat oikein tekstien lukemisen jälkeen esitettyihin väittämiin. Eniten oikein vastanneet 50 % sijoitettiin kielitaidoltaan parempaan ryhmään ja loput kielitaidoltaan heikompaan ryhmään. Kielitaidoltaan paremmassa ryhmässä vastattiin oikein keskimäärin 85 %:iin kysymyksistä (*kh* 7.4 %, vaihteluväli 75–100 %) ja heikommassa ryhmässä 59 %:iin kysymyksistä (*kh* 10.7 %, vaihteluväli 38–69 %). Luetun ymmärtäminen korreloi kohtalaisesti ja tilastollisesti merkitsevästi osallistujien kielitaidon itsearvion kanssa, $r = .461$, $p = .024$, ja kohtalaisesti, mutta ei tilastollisesti merkitsevästi Suomessa asuttujen vuosien määrän kanssa, $r = .314$, $p = .126$.

Riippuvat muuttujat olivat lukusujuvuusmittareita, joita oli 11: ensimmäisen lukukerran kesto (*gaze duration*), selektiivisen regressiopulun kesto (*selective regression-path duration*), fiksaatioiden kokonaiskesto (*total fixation duration*), ja saapuvien ja lähtevien regressioiden määrä, erikseen kohdesanoissa ja niitä seuraavissa sanoissa, sekä tekstin kokonaislukuaika (*trial time*). Ensimmäisen lukukerran keston laskettiin fiksaatiot ennen katseen siirtymistä pois sanasta ensimmäistä kertaa, kun taas selektiivisen regressiopulun keston laskettiin kaikki sanaan kohdistuvat fiksaatiot ennen siirtymistä oikealle seuraavaan sanaan. Fiksaatioiden kokonaiskesto sisälsi kaikki sanaan kohdistuneet fiksaatiot eli myös uudelleenlukemiset. Saapuvien regressioiden määrään laskettiin sanaan suuntautuvat regressiot, kun taas lähtevien regressioiden määrään laskettiin sanasta poispäin suuntautuvat regressiot. Fiksaatioiden lyhyempi kesto ja regressioiden vähäisempi määrä kertoo paremmasta lukusujuvuudesta. Kohdesanoja seuraavat sanat otettiin mukaan tarkasteluun ns.

valumisefektin takia. Tekstin kokonaislukuaika oli koko tekstin lukemiseen eli myös muiden kuin kohdesanojen ja niitä seuraavien sanojen lukemiseen käytetty aika.

2.4 Kokeen kulku

Aineisto kerättiin Turun yliopistolla maaliskuun 2023 aikana. Koe tehtiin tutkimushuoneessa, jossa häiriötekijät oli pyritty minimoimaan. Tutkija oli huoneessa koko koetilanteen ajan. Ensin tutkittavat tekivät erillisellä tietokoneella Lexize-sanastotestin (Salmela ym., 2021), jonka alussa kerättiin taustatietoja (mm. ikä, äidinkieli). Teknisten ongelmien vuoksi sanastotestin tuloksia ei voitu lopulta hyödyntää kielitaidon arvioinnissa. Sanastotestissä ei ollut tutkimuksen kohdesanoja. Lukemiskoe tehtiin käyttäen SR Research Ltd. -merkkistä EYELINK Portable Duo -silmänliikekameraa (Versio 6.50), joka rekisteröi tutkittavien silmänliikkeitä samalla, kun he lukivat tekstejä Asus-merkkisen kannettavan tietokoneen näytöltä (17") mielessään. Lukuetaisyys oli 70 cm. Tutkittavat käyttivät leuka- ja otsatukea, jotta pää pysyi mahdollisimman paikallaan lukemisen aikana. Pöytä ja tuet säädettiin sopivan korkuisiksi ennen aloittamista. Tutkittava asetti myös sormensa valmiiksi tietokoneen näppäimille kysymyksiin vastaamista varten niin, että oikea etusormi oli vihreällä 'oikein'-näppäimellä (M) ja vasen etusormi punaisella 'väärin'-näppäimellä (Z). Lisäksi peukalolla oli mahdollista painaa välilyöntinäppäintä, jota painettiin seuraavan tekstin tai kysymyksen ilmestymiseksi näytölle. Tutkittavat lukivat ohjeet näytöltä. Kamera seurasi oikeaa silmää ja se kalibroitiin 9 pisteen kalibroinnilla ennen harjoitustekstiä, ennen ensimmäistä tekstiä ja kokeen puolivälissä ennen viidettä tekstiä. Kalibrointi toistettiin, kunnes vahvistusvirhe oli keskimäärin alle 0.50 visuaalista astetta ja jokainen yksittäinen virhe alle 0.80. Ensin luettiin harjoitusteksti, jonka jälkeen esitettiin kaksi tekstiä koskevaa väittämää. Tutkittavat vastasivat tietokoneen kahdella värillisellä näppäimellä, oliko väittämä oikein vai väärin. Sen jälkeen luettiin kahdeksan varsinaista tekstiä ja neljän tekstin lukemisen jälkeen oli mahdollista pitää lyhyt tauko. Jokaisen tekstin jälkeen esitettiin kaksi luetun ymmärtämistä testaavaa väittämää, joihin vastattiin jälleen tietokoneen värillisillä näppäimillä. Lopuksi tutkittaville kerrottiin tutkimuksesta ja näytettiin omaa silmänliikedataa kiinnostuksen mukaan. Koe kesti yhteensä noin tunnin.

2.5 Tilastolliset analyysit

Ennen analyysieja silmänliikeaineiston kalibrointia korjattiin tarvittaessa manuaalisesti siirtämällä fiksaatioita vertikaalisesti EYELINKIN Data Viewer -ohjelmassa. Alle 50 ms kestoiset

fiksaatiot poistettiin. Kuuden koehenkilön data poistettiin kokonaan sen heikon laadun takia, ja 12 koehenkilön datasta jouduttiin poistamaan 1–3 tekstin data. Nämä koehenkilöt olivat mahdollisesti liikkuneet kalibroinnin jälkeen. Näin ollen analyyseissa oli lopulta mukana 28 koehenkilöä ja 201 heidän lukemaansa tekstiä, joista 102 (50.7 %) oli vartalonvaihdoksellista. Vartalonvaihdosten ja kielitaitotason vaikutusta lukusujuvuuteen testattiin toistettujen mittausten varianssianalyysillä, joka oli muotoa 2×2 . Koehenkilöiden sisäinen riippumaton muuttuja oli läpinäkyvyys ja koehenkilöiden välinen riippumaton muuttuja kielitaito. Riippuvat muuttujat eli käytetyt lukusujuvuuden mittarit olivat erikseen kohdesanoissa ja niitä seuraavissa sanoissa ensimmäisen lukukerran kesto, selektiivisen regressiopulun kesto, fiksaatioiden kokonaiskesto, saapuvat regressiot ja lähtevät regressiot sekä tekstin kokonaisluku-aika. Ennen testaamista riippuvien muuttujien arvoille, lukuun ottamatta regressioita, tehtiin logaritmimuunnos. Muuttujien sfäärisyys testattiin Mauchlyn sfäärisyystestillä ja normaalisuus testattiin Kolmogorov-Smirnovin testillä ennen analyysejä. Sfäärisyyttä ei voitu olettaa, joten tuloksia analysoitiin Greenhouse-Geisser-korjattua testiä käyttäen. Varianssianalyysit ja jatkovertailut läpinäkyvyyden vaikutuksesta kielitaitoryhmittäin tehtiin parametrisillä testeillä, koska normaalisuusoletus oli voimassa. Monivertailua ei huomioitu parittaisissa testeissä, koska se vähentää tilastollista voimaa, mikä on haitaksi eksploraatiiviselle tutkimukselle. Tilastolliset analyysit suoritettiin IBM SPSS Statistics -ohjelmistolla (versio 27.0.1).

2.6 Eettisyys

Tutkimus on osa laajempaa tutkimusprojektia, jonka tavoitteena on kehittää uusia välineitä S2-oppijoiden morfologisten taitojen tueksi. Tätä osatutkimusta varten haettiin lupa Åbo Akademin eettiseltä toimikunnalta. Kokeeseen osallistuminen oli vapaaehtoista eikä siitä voinut koitua tutkittaville mitään haittaa. Tutkimus oli mahdollista keskeyttää tutkittavan pyynnöstä ilman seurauksia tai perusteluja, ja silmien väsyessä tietokoneen näytöltä luettaessa oli mahdollista pitää tauko. Osallistujat antoivat luvan vastausten tallentamiselle ja käsittelylle tutkimuksen alussa sähköisellä lomakkeella (Lexize-testin alussa). Koehenkilöillä oli oikeus saada tietonsa ja aineistonsa poistettua halutessaan. Kerätty data käsiteltiin luottamuksellisesti ja anonyymisti Turun yliopiston tietoturvakäytännön mukaisesti, eikä yksittäistä tutkittavaa voitu tunnistaa tuloksista. Suorat tunnistetiedot, kuten sähköpostiosoite, joka kerättiin palkitsemista varten, korvattiin analyysivaiheessa osallistujakohtaisella tunnistenumeroilla. Osallistujat palkittiin lahjakortilla.

3 Tulokset

Taulukossa 2 on esitetty lukusujuvuusmuuttujiin liittyvät keskiarvot ja -hajonnat kielitaitoryhmittäin ja kaikkien koehenkilöiden osalta.

Taulukko 2

Lukusujuvuusmuuttujien keskiarvot ja -hajonnat (suluissa) kielitaitoryhmittäin ja kaikkien koehenkilöiden osalta

	Vartalonvaihdokset			Läpinäkyvät		
	Heikommat	Paremmat	Kaikki	Heikommat	Paremmat	Kaikki
Ensimmäisen lukukerran kesto (ms)						
Kohdesana	728 (174)	727 (316)	728 (250)	750 (318)	732 (198)	741 (260)
Kohdesanan jälk.	592 (125)	514 (205)	553 (171)	546 (131)	536 (151)	541 (139)
Selektiivisen regressiopulun kesto (ms)						
Kohdesana	908 (223)	847 (305)	887 (268)	907 (266)	860 (290)	883 (279)
Kohdesanan jälk.	714 (196)	601 (190)	657 (201)	665 (146)	604 (139)	635 (146)
Fiksaatioiden kokonaiskesto (ms)						
Kohdesana	1304 (432)	1091 (346)	1197 (399)	1305 (491)	1081 (324)	1193 (424)
Kohdesanan jälk.	1001 (325)	749 (211)	875 (298)	889 (208)	751 (177)	820 (202)
Saapuvat regressiot (lkm.)						
Kohdesana	0.26 (0.15)	0.21 (0.09)	0.24 (0.13)	0.24 (0.16)	0.20 (0.07)	0.22 (0.12)
Kohdesanan jälk.	0.23 (0.13)	0.17 (0.07)	0.20 (0.11)	0.24 (0.14)	0.16 (0.06)	0.20 (0.11)
Lähtevät regressiot (lkm.)						
Kohdesana	0.27 (0.10)	0.22 (0.11)	0.24 (0.10)	0.25 (0.14)	0.21 (0.10)	0.23 (0.12)
Kohdesanan jälk.	0.23 (0.15)	0.20 (0.07)	0.22 (0.12)	0.20 (0.11)	0.18 (0.08)	0.19 (0.10)
Kokonaisluku aika (s)	99.7 (26.7)	80.4 (22.6)	90.0 (26.5)	97.1 (25.7)	80.2 (20.6)	88.7 (24.8)

Ensimmäisessä tutkimuskysymyksessä haluttiin selvittää, vaikuttavatko taivutuksessa esiintyvät sanavartaloiden morfofonologiset muutokset lukusujuvuuteen S2-oppijoilla. Siihen vastaamiseksi tarkasteltiin läpinäkyvyyden päävaikutusta lukusujuvuusmittareihin, joista keskityn ensin ensimmäisen lukukerran keston ja fiksaatioiden kokonaiskeston. Tuloksista havaittiin, että läpinäkyvyydellä ei ollut päävaikutusta ensimmäisen lukukerran keston kohdesanoissa, $F(1,26) = 1.258$, $p = .272$, $n_p^2 = 0.046$, eikä niitä seuraavissa sanoissa, $F(1,26) = 0.002$, $p = .969$, $n_p^2 = 0.000$. Myöskään fiksaatioiden kokonaiskeston kohdesanoissa, $F(1,26) = 0.599$, $p = .446$, $n_p^2 = 0.023$, eikä niitä seuraavissa sanoissa, $F(1,26) = 2.492$, $p =$

.127, $n_p^2 = 0.087$, kohdesanan läpinäkyvyydellä ei ollut päävaikutusta. Lisäksi efektikoot olivat hyvin pieniä.

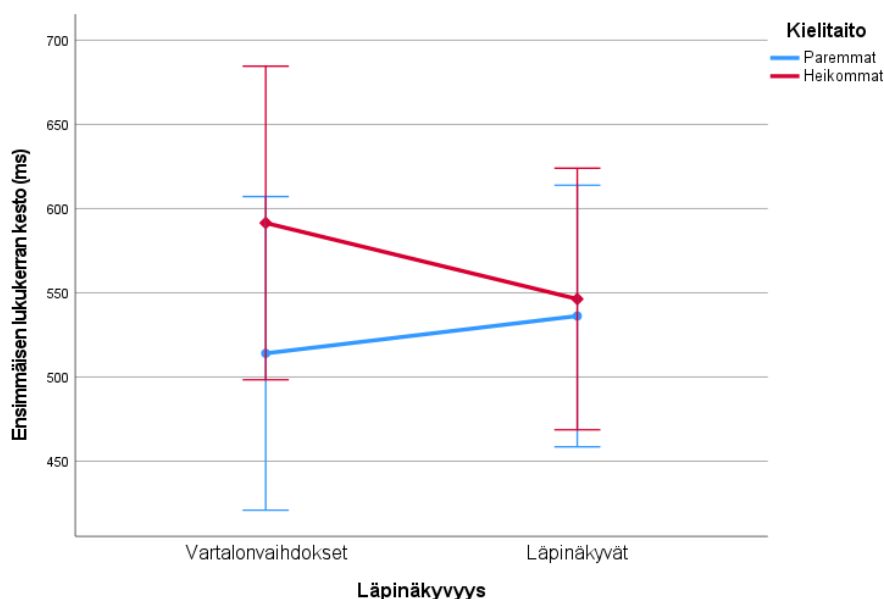
Kielitaidolla ei ollut päävaikutusta ensimmäisen lukukerran kestoon kohdesanoissa, $F(1,26) = 0.000$, $p = 1.000$, $n_p^2 = 0.000$, eikä niitä seuraavissa sanoissa, $F(1,26) = 0.556$, $p = .463$, $n_p^2 = 0.021$. Päävaikutusta ei ollut myöskään fiksaatioiden kokonaiskestoon kohdesanoissa, $F(1,26) = 1.897$, $p = .180$, $n_p^2 = 0.068$, mutta niitä seuraavissa sanoissa se löytyi, $F(1,26) = 5.089$, $p = .033$, $n_p^2 = 0.164$. Parempi kielitaito viittasi lyhyempään fiksaatioiden kokonaiskestoon melko suurella efektillä.

Toiseen tutkimuskysymykseen vastaamiseksi tarkasteltiin läpinäkyvyyden ja kielitaidon yhdysvaikutusta lukusujuvuuteen. Yhdysvaikutusta ei ollut ensimmäisen lukukerran kestoon kohdesanoissa, $F(1,26) = 0.008$, $p = .928$, $n_p^2 = 0.000$, mutta kohdesanoja seuraavissa sanoissa sellainen havaittiin, $F(1,26) = 4.465$, $p = .044$, $n_p^2 = 0.147$, ja efektikoko oli melko suuri.

Yhdysvaikutus on esitetty kuviossa 2. Jatkovertailujen perusteella läpinäkyvyys ei vaikuttanut tilastollisesti merkitsevästi ensimmäisen lukukerran kestoon kielitaidoltaan heikommassa, $t(13) = 1.378$, $p = .191$, $d = 0.368$, eikä paremmassa ryhmässä, $t(13) = -1.660$, $p = .121$, $d = -0.444$. Kielitaitoryhmien välillä ei ollut eroa ensimmäisen lukukerran kestossa vartalonvaihdojen sisältäviä kohdesanoja seuraavissa sanoissa, $t(26) = 1.338$, $p = .196$, $d = 0.506$, eikä läpinäkyviä kohdesanoja seuraavissa sanoissa, $t(26) = -0.036$, $p = .971$, $d = -0.014$.

Kuvio 2

Läpinäkyvyyden ja kielitaidon yhdysvaikutus ensimmäisen lukukerran kestoon kohdesanan jälkeisissä sanoissa

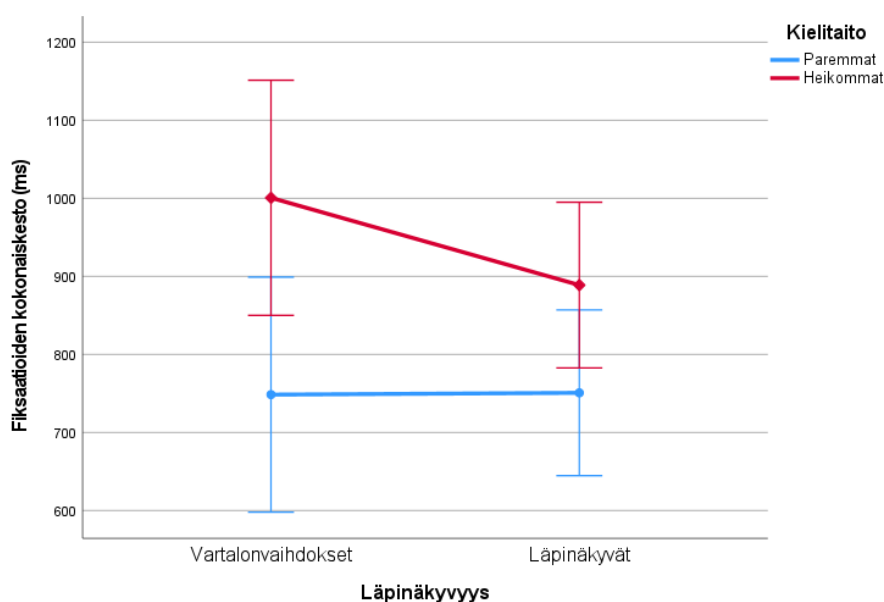


Virhepalkit kuvaavat 95 % luottamusvälejä

Läpinäkyvyyden ja kielitaidon yhdysvaikutusta ei ollut myöskään fiksaatioiden kokonaiskestoon kohdesanoissa, $F(1,26) = 0.004$, $p = .948$, $\eta_p^2 = 0.000$, mutta niitä seuraavissa sanoissa oli, $F(1,26) = 4.248$, $p = .049$, $\eta_p^2 = 0.140$, efektikoon ollessa melko suuri. Yhdysvaikutus esitetään kuviossa 3. Jatkovertailujen perusteella läpinäkyvyys vaikutti fiksaatioiden kokonaiskestoon kielitaidoltaan heikommassa ryhmässä niin, että läpinäkyviä kohdesanoja seuraavissa sanoissa kesto oli lyhyempi kuin vartalonvaihdoksen sisältäviä kohdesanoja seuraavissa sanoissa keskisuurella efektillä, $t(13) = 2.184$, $p = .048$, $d = 0.584$. Kielitaidoltaan paremmassa ryhmässä läpinäkyvyydellä ei ollut vaikutusta fiksaatioiden kokonaiskestoon, $t(13) = -0.436$, $p = .670$, $d = -0.117$. Ryhmien välillä oli ero fiksaatioiden kokonaiskestossa vartalonvaihdoksen sisältäviä kohdesanoja seuraavissa sanoissa suurella efektillä, $t(26) = 2.426$, $p = .023$, $d = 0.917$, eli heikommalla ryhmällä fiksaatioiden kokonaiskesto oli pidempi. Läpinäkyviä kohdesanoja seuraavissa sanoissa ryhmien välillä ei ollut eroa, $t(26) = 1.866$, $p = .073$, $d = 0.705$. Efektikoko oli kuitenkin melko suuri.

Kuvio 3

Läpinäkyvyyden ja kielitaidon yhdysvaikutus fiksaatioiden kokonaiskestoon kohdesanan jälkeisissä sanoissa



Virhepalkit kuvaavat 95 % luottamusvälejä

Analyyseissa ei havaittu tilastollisesti merkitseviä tuloksia, päävaikutuksia eikä yhdysvaikutuksia, lukusujuvuuden mittareista selektiivisen regressiopolon kestossa eikä saapuvissa ja lähtevissä regressioissa kohdesanoissa tai niitä seuraavissa sanoissa, kuten ei myöskään tekstin kokonaislukuajassa. Näiden muuttujien osalta varianssianalyysien tulokset ovat nähtävissä taulukossa 3.

Taulukko 3

Varianssianalyysien F- ja p-arvot selektiivisessä regressiopolkussa, saapuvissa ja lähtevissä regressioissa sekä tekstin kokonaislukuajassa

	Läpinäkyvyys		Kielitaito		Läpinäkyvyys x kielitaito	
	<i>F</i> (1,26)	<i>p</i>	<i>F</i> (1,26)	<i>p</i>	<i>F</i> (1,26)	<i>p</i>
Selektiivinen regressiopolku						
Kohdesana	0.000	.985	0.339	.566	0.017	.896
Kohdesanan jälk.	0.284	.599	0.055	.231	2.179	.111
Saapuvat regressiot						
Kohdesana	1.526	.228	0.931	.344	0.120	.732
Kohdesanan jälk.	0.003	.960	2.888	.101	0.080	.780
Lähtevät regressiot						
Kohdesana	0.955	.337	1.245	.275	0.097	.758
Kohdesanan jälk.	2.189	.151	0.379	.544	0.115	.737
Kokonaisluku aika	0.854	.364	3.796	.062	0.662	.423

4 Pohdinta

Tutkimuksen tavoitteena oli selvittää katseenseurantaa hyödyntäen, vaikuttavatko suomen kielessä yleiset, taivutettaessa esiintyvät sanavartaloiden morfofonologiset muutokset lukusujuvuuteen S2-oppijoilla. Morfofonologisia muutoksia eli vartalonvaihdoksia sisältäviä sanoja luettiin osana selkokieliä lyhyitä tarinoita, jotta tuloksia voitaisiin hyödyntää selkokieltä tuottaessa. Morfofonologisten muutosten esiintyvyydellä ei ollut koko aineistossa päävaikutusta lukusujuvuuden mihinkään mittariin: ensimmäisen lukukerran keston, selektiivisen regressiopolun keston, fiksaatioiden kokonaiskeston, saapuviin tai lähteisiin regressioihin eikä tekstin kokonaisluku-aikaan. Näin ollen tutkimus ei tue suoraan hypoteesia, että lukusujuvuutta voitaisiin parantaa rajoittamalla nomineissa esiintyviä morfofonologisia muutoksia ja suosimalla läpinäkyvästi taipuvia sanoja tekstissä. Kuitenkin, kun otettiin kielitaitotason vaikutus mukaan analyysiin, havaittiin sillä ja läpinäkyvyydellä olevan yhdysvaikutus ensimmäisen lukukerran keston ja fiksaatioiden kokonaiskeston, mutta vain kohdesanoja seuraavissa sanoissa. Läpinäkyvyys vaikutti fiksaatioiden kokonaiskeston kielitaidoltaan heikommassa ryhmässä niin, että läpinäkyviä kohdesanoja seuraavissa sanoissa kesto oli lyhyempi kuin vartalonvaihdoksen sisältäviä kohdesanoja seuraavissa sanoissa. Lisäksi heikommassa ryhmässä fiksaatioiden kokonaiskesto vartalonvaihdoksen sisältäviä kohdesanoja seuraavissa sanoissa oli pidempi kuin paremmassa ryhmässä. Yhdysvaikutuksesta voidaan päätellä, että kielitaitotasoltaan kaikista aloittelevimmat hyötyivät morfofonologisten muutosten eli vartalonvaihdosten rajoittamisesta, koska silloin heidän lukemisensa oli tutkitulta osin nopeampaa, mikä puoltaa vartalonvaihtelun huomioimista selkokielessä. Selkokieli onkin suunnattu pääasiassa A1–B1-tason kielenoppijoille, jonka jälkeen sen tarve vähenee. Paremmassa kielitaitoryhmässä läpinäkyvyydellä ei ollut vaikutusta lukusujuvuuteen.

4.1 Vartalonvaihdosten vaikutus lukusujuvuuteen

Mielenkiintoista on, että vartalonvaihdosten vaikutukset lukusujuvuuteen näkyivät nimenomaan vain kohdesanoja seuraavissa sanoissa eivätkä itse kohdesanoissa. Kohdesanoja seuraavat sanat otettiin mukaan kokeeseen, koska aiemmissa tutkimuksissa (Rayner & Duffy, 1986; Reichle ym., 2003; Reilly & Radach, 2006; suomen kielessä Vainio ym., 2011) on useasti osoitettu valumisefekti, joka tarkoittaa sitä, ettei sanan prosessointi lopu samalla hetkellä, kun sanan lukeminen loppuu, vaan jatkuu vielä seuraavan rakenteen kohdalla, mikä näkyy senkin prosessoinnin hidastumisena. Koska itse kohdesanojen lukemisessa ei

tapahtunut hidastumista, voidaan itse asiassa puhua viivästetystä efektistä (*delayed effect*, Rayner & Duffy, 1986; suomen kielessä Vainio ym., 2008), jossa prosessointivaikeus ilmenee vasta vaikean rakenteen jälkeisessä rakenteessa. Valumisefektiä ja viivästettyä efektiä on selitetty usein leksikaalisella haulla. Vaikean sanan oletetaan myöhästyttävän sen jälkeisen sanan kognitiivista prosessointia enemmän kuin se hidastaa sakkadia tähän sanaan, jolloin prosessointi tapahtuu enimmäkseen fiksaation ollessa jo seuraavassa sanassa. Toisin sanoen vaikea sana estää seuraavan sanan parafoveaalista esiprosessointia (Engbert ym., 2005; Reichle ym., 2003). Parafoveaalinen esiprosessointi tarkoittaa sanan alustavaa prosessointia jo siinä vaiheessa, kun katse ei ole vielä siirtynyt siihen, vaan se nähdään epätarkasti edellistä sanaa katsottaessa. On myös mahdollista, että haastavat sanat on vaikeampi integroida lauseeseen ja kontekstiin, mikä vaikuttaa seuraavan sanan prosessointiin hidastavasti (Rayner & Duffy, 1986).

Tulokset ovat vain osittain samansuuntaisia kuin aiemmissa tutkimuksissa (Salmela ym., lähetetty arvioitavaksi; Vainio ym., 2014), koska niissä vartalonvaihdoksen sisältävissä sanoissa on osoitettu prosessoinnin hidastumista, kun tässä tutkimuksessa hidastuminen näkyi vain kohdesanoja seuraavissa sanoissa. Koska tulokset eivät tue kaikkien kielenoppijoiden, eivätkä kaikkien mittareiden osalta hypoteesia, että vartalonvaihdoksia välttämällä voitaisiin parantaa lukusujuvuutta, on syytä pohtia miksi. Tutkimus on ensimmäinen laatuaan, sillä vartalonvaihdosten vaikutusta tarkasteltiin selkokielen pidemmän tekstin kontekstissa, kun aiemmin sitä on tarkasteltu yksittäisissä sanoissa tai lauseissa. Siksi tässä yhteydessä vaikutusmekanismi voi olla erilainen kuin aiemmissa tutkimuksissa. On mahdollista, että konteksti vaikutti sanojen tunnistamiseen eikä vartalonvaihdoksiset sanatkaan aiheuttaneet haasteita prosessoinnissa. Aiemmassa tutkimuksessa vain suomea äidinkielenään puhuvilla konteksti on poistanut vartalonvaihdosten aiheuttaman prosessoinnin hidastumisen, ja S2-oppijoilla efekti on kuitenkin säilynyt myös lauseen kontekstissa (Salmela ym., lähetetty julkaistavaksi). Toiseksi, etenkin S1-puhujien lukemista koskevissa tutkimuksissa on havaittu, että hyvin yleisille sanamuodoille muodostuu tunnistusta nopeuttavia kokosanaedustumia muistiin (Lehtonen & Laine, 2003; Soveri ym., 2007). Kun kiinnostuksen kohteiksi tässä tutkimuksessa valittiin suhteellisen helppoja sanoja ajatellen A2–B2-tason kielenoppijoita, ne voivat olla esiintymistaajuudeltaan niin yleisiä, että niiden allomorfeillekin saattaa olla jo muodostunut omat edustumansa mentaaliseen leksikkoon lukijoiden aloittelevasta tasosta huolimatta. Tällöin sanantunnistus voisi tapahtua nopeammin eikä fiksaatioiden kestossa näkyisi eroja läpinäkyvien ja vartalonvaihdoksen sisältävien sanojen välillä.

Kohdesanojen ensimmäisen lukukerran kestossa ei ollut tilastollisesti merkitsevää eroa vartalonvaihdoksellisissa ja läpinäkyvissä teksteissä. Vastaavasti Salmelan ja kumppanien (lähetetty julkaistavaksi) tutkimuksessa läpinäkyvien ja vartalonvaihdoksen sisältävien sanojen ensimmäisessä lukukerrassa ei S2-oppijoilla ollut eroa, kun ero fiksaatioiden kokonaiskestossa löytyi. Heidän tulkintansa tästä myöhemmästä prosessoinnista oli, että vartalonvaihdoksen sisältävien sanojen lukeminen ei yleensä onnistu täysin heti ensimmäisellä kerralla, vaan vaatii uudelleen lukemista ja regressioita. Retsin ja Rogatenin (2021) tutkimuksessa havaittiin, että yksinkertaistetuissa teksteissä ensimmäisen lukukerran fiksaatiot olivat pidempiä ja toisen lukukerran fiksaatiot lyhyempiä kuin alkuperäisissä yleiskielisissä teksteissä. Tutkijoiden mukaan tekstin yksinkertaistaminen mahdollisti tekstin tehokkaamman alustavan prosessoinnin. Tämä voisi kertoa siitä, että ensimmäisen lukukerran fiksaatiot eivät ole välttämättä sopivin mittari lukusujuvuudelle selkokielistä tekstiä luettaessa. Tosin Rets ja Rogaten eivät tutkineet morfologian vaikutusta, vaan tekstejä oli helpotettu sanaston, syntaksin ja kontekstin osalta.

4.2 Tutkimuksen arviointi

Tutkimuksen rajoituksena voidaan ehdottomasti pitää osallistujien vähäistä määrää, minkä takia tilastollisesti merkitseviä tuloksia on vaikeaa saada esiin ja yleistettävyyks on heikko. Kun kielitaidon perusteella vielä jaettiin osallistujat kahteen ryhmään, jäivät ryhmät erittäin pieniksi. Osallistujien äidinkieli vaihteli, mikä on hyvä otoksen edustavuuden kannalta. Näin yksittäinen kieli ei pääse vaikuttamaan tuloksiin. Esimerkiksi Vainion ja kumppanien (2014) tutkimuksessa todettiin, että venäjänkieliset tunnistivat nopeammin suomen vartalonvaihdoksen sisältäviä sanoja kuin kiinankieliset. Venäjä on morfologisilta piirteiltään lähempänä suomea kuin kiina, koska siinäkin tapahtuu morfofonologisia muutoksia. Toisaalta moninainen kielitausta estää päätelmien tekemisen siitä, että tietyn äidinkielen puhujille, esim. jollekin S2-opetusryhmälle, suomen kielen morfofonologinen vaihtelu aiheuttaisi haasteita. Tutkimuksen osallistujat eivät myöskään äidinkieltä puolesta edusta kaikkia suomen oppijoita tai maahanmuuttajia. Heistä myös huomattavan suurella osalla (80 %) oli korkeakoulututkinto.

Aiemmissa tutkimuksissa, joissa vartalonvaihdokset ovat vaikuttaneet sanantunnistukseen hidastavasti ja vaikeuttavasti, tutkittavien kielitaito on ollut eurooppalaisen kielitaidon viitekehyksen mukaan vastaava kuin tässä tutkimuksessa. Suomen kielitaitoa ei huomioitu sisäänottokriteereissä tarkemmin kuin siten, että osallistujat rekrytoitiin A2–B2-tasoisilta

kielikursseilta. Heidän taidoissaan ja suomen kielelle altistumisen määrässä on siis todennäköisesti ollut suurta vaihtelua. Tutkittavat tekivät kokeen alussa Lexize-sanastotestin (Salmela ym., 2021), joka perustuu oikeiden sanojen ja epäsanojen tunnistamiseen. Testillä oli tarkoitus arvioida tutkittavien yksilöllistä kielitaitoa, mutta teknisen virheen vuoksi sitä ei pystytty käyttämään luotettavasti. Siksi kielitaidon tasoa päätettiin mitata ainoastaan luetun ymmärtämistä testaavilla kysymyksillä. Luetun ymmärtäminen on siinä mielessä hyvä kielitaidon mittari, että se vaatii sanaston tuntemisen lisäksi tietoa mm. kielen morfologiasta ja syntaksista. Koska lyhyempi fiksaatioiden kokonaiskesto oli yhteydessä parempaan luetun ymmärtämiseen, ne todennäköisesti heijastavat juuri kielitaidon tasoa eivätkä esimerkiksi tarkkuutta. Luetun ymmärtäminen oli myös yhteydessä S2-oppijoiden arvioon omasta suomen taidostaan. Luetun ymmärtämistä testaavia kysymyksiä oli kuitenkin tutkimuksessa suhteellisen vähän, mikä voi heikentää reliabiliteettia. Mitä enemmän kysymyksiä on, sitä lähemmäs päästään vastaajan todellista ymmärtämisprosenttia.

Osallistujille ei annettu tarkempaa ohjetta esimerkiksi siihen, kuinka monta kertaa tekstit voi lukea. Osa saattoi lukea tekstit useampaan kertaan, mikä tietysti vähentää lukusujuvuutta tämän tutkimuksen mittareilla, koska uudelleen luettaessa myös fiksaatioita tulee enemmän ja lukuaika pitenee. Periaatteessa lukemisen keston voisi ajatella lisäävän ymmärtämistä, mutta näin ei kuitenkaan ollut, koska heikko ymmärtäminen ja fiksaatioiden pidempi kokonaiskesto olivat yhteydessä. Toisaalta useampaan kertaan lukemista voidaan pitää merkinä tekstin haastavuudesta. Tekstin kokonaislukuajassa ero paremman ja heikomman kielitaitoryhmän välillä oli lähellä merkitsevää eli vaikuttaa siltä, että kauemmin ja useampaan kertaan tekstejä lukeneet ovat myös ymmärtäneet niitä heikommin. Läpinäkyvyys ei kuitenkaan vaikuttanut kokonaisluku aikaan.

Jos pureudutaan vielä syvemmälle tutkimuksen kohdesanojen laatuun, voidaan todeta, että merkittävän suuri osa, kolmannes kaikista vartalonvaihdoiksi kohdesanoista sisälsi kvantitatiivista astevaihtelua (esim. kauppa: kaupassa). Äänteiden kvantiteetti, kuten kaksoiskonsonantit, joka on suomen kielen erityispiirre, tuottaa oppijoille haasteita (Brown, 2004). Suomen oppijat joutuvat opettelemaan ulkoa, missä sanoissa on kaksoiskonsonantti. He voivat miettiä esimerkiksi sitä, onko 'kauppa' vai 'kaupa' oikea suomen sana. Voidaan siis pohtia, vaikuttavatko tämännäköiset vartalonvaihdot morfologiseen prosessointiin sanantunnistuksessa, kun ne muuttavat sanan ilmeä vain äänten kvantiteetin osalta. Sen sijaan kvalitatiivisen astevaihtelun (esim. vika: vialla) voisi uskoa aiheuttavan suurempia pulmia, koska äänteissä tapahtuu korvautumisia, jotka muuttavat sanojen ulkoasua vielä

enemmän kuin saman äänteen yksinkertaistuminen tai kaksinkertaistuminen. Tämä olisi hyödyllistä ottaa huomioon tutkimuksissa jatkossa. Lisäksi tässä tutkimuksessa kohteena oli vain kaksimorfeemisia sanoja, mutta monimorfeemiset sanat ovat suomen kielessä yleisiä. Monimutkaisuuden kasvattaminen lisännee haasteita, mikä olisi myös hyvä huomioida jatkotutkimuksissa.

Aineisto kerättiin käyttäen korkeatasoisia tutkimuslaitteita ja silmänliikekamera kalibroitiin jokaisen tutkittavan kohdalla kolme kertaa eri kohdassa koetta. Silti laitteisto on herkkä esimerkiksi pään liikkumiselle ja erilaisille heijastuksille. Noin kolmasosalla tutkittavista oli ainakin yhden tekstin lukemisen aikana epätarkkuutta silmän liikkeiden kohdistuksissa. Poistettu data voi heikentää tutkimuksen validiteettia.

4.3 Lopuksi

Selkokieltä on tärkeää tutkia, jotta löydetään uusia keinoja tehdä suomenkielistä tietoa helpommin saavutettavaksi maahanmuuton lisääntyessä ja samalla S2-oppijoiden määrän kasvaessa. Etenkin morfologisten piirteiden osalta selkokielen tutkimus ja kehittäminen on tähän päivään mennessä ollut vähäistä. Tätä tutkimusta voidaan pitää pilottitutkimuksena vartalonvaihtelun vaikutuksesta S2-oppijoiden lukusujuvuuteen luettaessa selkokielistä tekstiä. Tutkimuksen päähavainto oli, että vartalonvaihdoiset vaikuttivat lukusujuvuuteen hidastavasti kohdesanojen jälkeisissä sanoissa heikommassa, mutta ei paremmassa, kielitaitoryhmässä. Vaikka tutkimus antoi vain melko vähän viitteitä siitä, että S2-oppijoiden lukusujuvuutta voitaisiin parantaa rajoittamalla vartalonvaihdoksia sisältäviä sanoja tekstissä, tukee se osittain aiempia tutkimustuloksia siitä, että suomen kielen morfologia ja erityisesti vartalonvaihdoiset ovat haastavia S2-oppijoille. Eniten hyötyä vartalonvaihdosten välttämisestä oli kaikkein aloittelevimmille oppijoille, joiden tekstin ymmärtämisen taso oli heikkoa. Selkokieli onkin suunnattu juuri alkeistason kielenoppijoille. Heillä morfologian huomioiminen selkokieltä tuottaessa saattaisi olla hyödyllistä, jotta tieto saavuttaisi heidät mahdollisimman hyvin. Vartalonvaihdoksen sisältävät sanat ovat kuitenkin kiistämätön osa suomen kieltä, ja kielitaidon kehittyessä niiden oppiminen on välttämätöntä kielen hallitsemiseksi. Tässä tutkimuksessa tutkittiin vartalonvaihdosten vaikutusta lukusujuvuuteen, mutta jatkossa olisi hyödyllistä tutkia niiden vaikutusta myös tarkemmin luetun ymmärtämiseen. Ymmärrettävyys on kuitenkin tärkein selkokielen ominaisuus. Koska tutkimus kertoi jonkin verran vartalonvaihtelun haastavuudesta S2-oppijoiden lukemisessa,

mutta oli kuitenkin eksploratiivinen, tarvitaan aiheesta lisää tutkimusta isommalla otoksella ja laajemmalla kielitaitoskaalalla.

Lähteet

- Alegre, M. & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40(1), 41–61.
- Basnight-Brown, D. M., Chen, L., Hua, S., Kostić, A. & Feldman, L. B. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of memory and language*, 57(1), 65–80.
- Bertram, R. & Hyönä, J. (2023). The Importance of Eye Movement Research for Gaining Insight into Morphological Processing. Teoksessa D. Crepaldi (toim.), *Linguistic Morphology in the Mind and Brain* (s. 121-136). Routledge.
- Bertram, R., Laine, M. & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. *Journal of Psycholinguistic Research*, 28(3), 213–226.
- Brown, A. (2004). Kvantiteetin oppimisesta. Teoksessa B. Statszer & A. Brown (toim.), *Suomen kielen prosodian opettamisen ja oppimisen kysymyksiä* (s. 117–144). Kakkoskieli 5. Helsingin yliopiston suomen kielen laitos.
- Butterworth, B. (1983). Lexical representation. Teoksessa B. Butterworth (toim.), *Language production* (s. 257–294). Academic Press.
- Castelhano, M. & Rayner, K. (2008). Eye movements during reading, visual search, and scene perception: An overview. Teoksessa K. Rayner, D. Shen, X. Bai & G. Yan (toim.), *Cognitive and cultural influences on eye movements* (s. 3–33). Tianjin People's Publishing House.
- Dahlgren, S. (2020). Selkokieltä S2-oppijalle: Kaikille selviä hallinnollisia tekstejä. *Puhe ja Kieli*, 39(4), 325–344.
- Deans, P., O’Laughlin, L., Brubaker, B., Gay, N., & Krug, D. (2010). Use of eye movement tracking in the differential diagnosis of attention deficit hyperactivity disorder (ADHD) and reading disability. *Psychology*, 1(04), 238–246.
- DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language learning*, 55(1), 1–25.
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological review*, 112(4), 777–813.
- Finn Lectura (2001). Verkkokielioppi. Muoto-oppi eli morfologia.
<https://fl.finnlectura.fi/verkkosuomi/Morfologia/sivu2.htm>.

- Harris, T. & Hodges, R. (1995). *The literacy dictionary: the vocabulary of reading and writing*. International Reading Association.
- Hyönä, J., Vainio, S. & Laine, M. (2002). A morphological effect obtains for isolated words but not for words in sentence context. *European Journal of Cognitive Psychology*, 14 (4), 417–433.
- Just, M. & Carpenter, P. (1980). A theory of reading: from eye fixations to comprehension. *Psychological review*, 87(4), 329–354.
- Juusola, M. (2019). *Selkokielen tarvearvio 2019*. Selkokeskus. <https://selkokeskus.fi/wp-content/uploads/2021/05/Tarvearvio-2019.pdf>.
- Järvikivi, J. & Niemi, J. (2002). Form-based representation in the mental lexicon: Priming (with) bound stem allomorphs in Finnish. *Brain and Language*, 81(1–3), 412–423.
- Karlsson, F. (1983). Suomen kielen äänne- ja muotorakenne. Werner Söderström.
- Kielitoimiston sanakirja (2022). Kotimaisten kielten keskuksen verkkojulkaisuja 35. Verkkoersio. <https://www.kielitoimistonsanakirja.fi>.
- Kuhn, M. & Stahl, S. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95(1), 3–21.
- Laine, M., Vainio, S. & Hyönä, J. (1999). Lexical access routes to nouns in a morphologically rich language. *Journal of Memory and Language*, 40(1), 109–135.
- Lehtonen, M. (2006). Morfologinen prosessointi yksi- ja kaksikielisillä: neurokognitiivisia tutkimuksia suomen ja ruotsin kielessä. *Puhe ja kieli*, 26(1), 39–56.
- Lehtonen, M., Cunillera, T., Rodríguez-Fornells, A., Hultén, A., Tuomainen, J. and Laine, M. (2007). Recognition of morphologically complex words in Finnish: evidence from event-related potentials. *Brain Research*, 1148(1), 123–137.
- Lehtonen, M. & Laine, M. (2003). How word frequency affects morphological processing in monolinguals and bilinguals. *Bilingualism: Language and Cognition*, 6(3), 213–225.
- Leskelä, L. & Uotila, E. (2020). Selkokieli saavutettavan viestinnän välineenä. Teoksessa M. Hirvonen & T. Kinnunen (toim.), *Saavutettava viestintä: Yhteiskunnallista yhdenvertaisuutta edistämässä* (s. 227–248). Gaudeamus.
- Luotolahti, J., Kanerva, J., Laippala, V., Pyysalo, S., & Ginter, F. (2015). Towards universal web parsebanks. *Proceedings of the third international conference on Dependency Linguistics (Depling 2015)*, 211–220.
- Manelis, L. & Tharp, D. A. (1977). The processing of affixed words. *Memory & Cognition*, 5(6), 690–695.

- Murrell, G. A. & Morton, J. (1974). Word recognition and morphemic structure. *Journal of Experimental Psychology*, 102(6), 963–968.
- Nahatame, S. (2023). Predicting processing effort during L1 and L2 reading: The relationship between text linguistic features and eye movements. *Bilingualism: Language and Cognition*, 26(4), 724–737.
- Niemi, J., Laine, M. & Tuominen, J. (1994). Cognitive morphology in Finnish: Foundations of a new model. *Language and cognitive processes*, 9(3), 423–446.
- Niilo Mäki Instituutti (2023). LukiMat. Morfologia. <http://www.lukimat.fi/sanasto/lukimat-sanasto/morfologia>.
- Niilo Mäki Instituutti (2023). LukiMat. Sujuvuuden opettaminen. <http://www.lukimat.fi/lukeminen/tietopalvelu/lukemaan-opettaminen-1/sujuvuuden-opettaminen>.
- Olkkonen, S. (2017). *Second and foreign language fluency from a cognitive perspective: inefficiency and control of attention in lexical access*. Jyväskylä Studies in Humanities 314. Jyväskylän yliopisto.
- Pinker, S. (1991). Rules of language. *Science*, 253(5019), 530–535.
- Rayner, K., Chace, K., Slattery, T. & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific studies of reading*, 10(3), 241–255.
- Rayner, K., & Duffy, S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14(3), 191–201.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The EZ Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences*, 26(4), 445–476.
- Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, 7(1), 34–55.
- Rets, I. & Rogaten, J. (2021). To simplify or not? Facilitating English L2 users' comprehension and processing of open educational resources in English using text simplification. *Journal of Computer Assisted Learning* 37(3), 705–717.
- Salmela, R., Lehtonen, M., Garusi, S., & Bertram, R. (2021). Lexize: a test to quickly assess vocabulary knowledge in Finnish. *Scandinavian Journal of Psychology*, 62(6), 806–819.

- Salmela, R., Lehtonen, M., Vainio, S. & Bertram, R. (lähetytty julkaistavaksi). Challenges in Inflected Word Processing for L2 Speakers: The Role of Stem Allomorphy. *Studies in Second Language Acquisition*. Artikkelii vertaisarviossa.
- Selkokeskus (2021a). Selkokielen määritelmä. <https://selkokeskus.fi/selkokieli/selkokielen-maaritelma/>.
- Selkokeskus (2021b). Helpot kielen rakenteet. <https://selkokeskus.fi/selkokieli/nain-kirjoitat-selkokielta/helpot-kielen-rakenteet/>.
- Selkokeskus (2022). Selkokielen mittari 2.0. <https://selkokeskus.fi/wp-content/uploads/2022/04/Selkokielen-mittari-2.0.pdf>.
- Soveri, A., Lehtonen, M. & Laine, M. (2007). Word frequency and morphological processing in Finnish revisited. *The Mental Lexicon*, 2(3), 359–385.
- Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7(4), 263–272.
- Taft, M. & Ardasinski, S. (2006). Obligatory decomposition in reading prefixed words. *The Mental Lexicon*, 1(2), 183–199.
- Taft, M. & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 638–647.
- Tilastokeskus (2024). Maahanmuuttajataustaisten määrät ja osuudet alueittain. https://pxdata.stat.fi/PxWeb/pxweb/fi/Maahanmuuttajat_ja_kotoutuminen/Maahanmuuttajat_ja_kotoutuminen__Maahanmuuttajat_ja_kotoutuminen/maakoto_pxt_11vu.px/.
- Tieteen termipankki (2022). Kielitiede:morfologia. <https://tieteentermipankki.fi/wiki/Kielitiede:morfologia>
- Vainio, S. (2006). Lausetason kokeellisesta tutkimuksesta: tuloksia, haasteita ja tulevaisuuden näkymiä. *Puhe ja kieli*, 26(1), 3–21.
- Vainio, S., Bertram, R., Pajunen, A., & Hyönä, J. (2011). Processing modifier-head agreement in long Finnish words: Evidence from eye movements. *Acta Linguistica Hungarica*, 58(1), 134–156.
- Vainio, S., Hyönä, J., & Pajunen, A. (2008). Processing modifier-head agreement in reading: Evidence for a delayed effect of agreement. *Memory & cognition*, 36(2), 329–340.
- Vainio, S., Pajunen, A. & Hyönä J. (2014). L1 and L2 word recognition in Finnish: Examining L1 Effects on L2 Processing of Morphological Complexity and Morphophonological Transparency. *Studies in Second Language Acquisition*, 36(1), 133–62.

VISK = Auli Hakulinen, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta

Heinonen ja Irja Alho (2004): Iso suomen kielioppi. Suomalaisen Kirjallisuuden Seura.

Verkkoversio. <http://scripta.kotus.fi/visk> URN:ISBN:978-952-5446-35-7

Liitteet

Liite 1. Esimerkki koemateriaalista

JÄNNITTÄVÄ PÄIVÄ

Keskiviikko oli ollut tavallinen päivä, mutta torstaina **illalla** Emmi on **hädässä**. Hän on tulossa **kaupasta**, kun huomaa **kodin** edessä lasinsiruja. Samassa Emmi näkee **varkaan**, joka juoksee **kadulle** päin. Mies on mustassa **puvussa** ja hiukset ovat **peitossa tiukan lakin** alla. Emmi on **peloissaan**, mutta hän tapaa pihassa **kiltin** naapurin. Yhdessä he menevät **hitaasti** sisälle tarkistamaan **vahingot** joka **paikasta**. Mies on rikkonut ikkunan ja ottanut **pöydältä kullasta** tehdyn **palkinnon**. Muuta ei ole viety. Emmi ilmoittaa poliisille **murrosta**. **Soiton** jälkeen hän tajuaa, että syyllinen on jättänyt itsestään **jäljen**. Hän on unohtanut **laatikon** päälle **kolikon**. Emmi jää odottamaan poliisin saapumista.

JÄNNITTÄVÄ PÄIVÄ

Keskiviikko oli ollut tavallinen päivä, mutta torstaina **aamulla** Emmi on **pulassa**. Hän on tulossa **koulusta**, kun huomaa **talon** edessä lasinsiruja. Samassa Emmi näkee **rosvon**, joka juoksee **tielle** päin. Mies on mustassa **asussa** ja hiukset ovat **piilossa kapean pipon** alla. Emmi on **kauhuissaan**, mutta hän tapaa pihassa **mukavan** naapurin. Yhdessä he menevät **rohkeasti** sisälle tarkistamaan **vauriot** joka **tilasta**. Mies on rikkonut ikkunan ja ottanut **hyllyltä hopeasta** tehdyn **mitalin**. Muuta ei ole viety. Emmi ilmoittaa poliisille **ryöstöstä**. **Puhelun** jälkeen hän tajuaa, että syyllinen on jättänyt itsestään **vihjeitä**. Hän on unohtanut **lipaston** päälle **setelin**. Emmi jää odottamaan poliisin saapumista.