



## Visualisointi ja haastattelut kirjoitussujuvuuden tutkimuksen tukena

- Anne-Maj Åberg, Turun yliopisto
- Mari Mäkilä, Turun yliopisto
- Maarit Mutta, Turun yliopisto

Kirjoittajien yhteystiedot:

Anne-Maj Åberg, [anne-maj.berg@utu.fi](mailto:anne-maj.berg@utu.fi)

Mari Mäkilä, [mari.makila@utu.fi](mailto:mari.makila@utu.fi)

Maarit Mutta, [maamut@utu.fi](mailto:maamut@utu.fi)

Tässä tutkimuksessa tarkastelemme kirjoitussujuvuutta kolmella eri kielellä kirjoitettaessa. Tavoitteena on kartoittaa, miten laskennallista dataa, visualisointia sekä jälkikäteishaastatteluista ja taustakyselyistä saatua tietoa yhdistämällä voidaan tutkia kirjoitusprosesseja tarkemmin. Aineisto koostuu kuuden suomenkielisen yliopisto-opiskelijan tallennetuista kirjoitusprosesseista L1-suomeksi, L2-ruotsiksi ja L2-englanniksi. Kirjoitusprosessit tallennettiin GenoGraphiX-Log-näppäilyntallennusohjelmalla, jonka avulla on mahdollista visualisoida kirjoitusprosesseja. Tulokset viittaavat siihen, että kirjoitussujuvuudessa on eroja taukokäyttämisen ja purskeiden tuottamisen lisäksi myös muilla kirjoitusprosessiin liittyvillä osa-alueilla. Visualisointi on hyödyllinen työkalu näiden erojen tarkastelussa. Kirjoittajan oma näkökulma kirjoitusprosessiin voidaan tuoda esille jälkikäteishaastattelujen ja taustakyselyjen avulla. Tulosten perusteella ensikielellä kirjoittaminen koetaan helpommaksi kuin vieraalla kielellä kirjoittaminen, mutta esimerkiksi aiheen rajaaminen saattaa olla haasteellista. Etenkin vieraiden kielten osalta kirjoitusprosessin sujuvuuteen vaikuttavat mm. tehtävänanto ja kielten viimeaikainen käyttö.

**Asiasanat:** jälkikäteishaastattelu, kirjoitusprosessi, näppäilyntallennus, sujuvuus, visualisointi

## 1 JOHDANTO JA TUTKIMUKSEN TAVOITE

Sujuvuustutkimus on perinteisesti keskittynyt suulliseen tuotokseen erityisesti vieraiden kielten osalta, mutta hiljattain on alettu kiinnittää huomiota myös kirjoitussujuvuuteen. Kirjoitussujuvuudessa on aiemmin tarkasteltu esimerkiksi lopullisen tuotoksen sujuvuutta, mutta uudet aineistonkeruumenetelmät näppäilyntalennusohjelmineen mahdollistavat myös prosessin sujuvuuden tutkimuksen. Kirjoitusprosessin tutkimuksessa tavoitteena on usein saada esiin lopputuotoksen taustalla vaikuttava kognitiivinen prosessi tarkastelemalla kirjoitusprosessin aikana tapahtuneita toimintoja. (Leblay, Usoof & Caporossi, 2022; Lindgren & Sullivan, 2019; Lintunen, Mutta & Peltonen, 2020; Wengelin & Johansson, 2023.) Tässä tutkimuksessa havainnollistamme kirjoitusprosessin sujuvuutta visualisointien avulla ja tarkastelemme sujuvuuteen vaikuttavia tekijöitä kirjoittajan näkökulmasta.

Hyödynnämme tutkimuksessa kirjoitusprosesseista tallennettua laskennallista dataa, siihen perustuvaa visualisointia sekä jälkikäteishaastattelujen ja taustakyselyjen avulla saatua tietoa. Tavoitteenamme on kartoittaa, miten näitä eri tutkimusmetodeja yhdistämällä voidaan tutkia kirjoitussujuvuutta ja tuoda esiin kirjoittajan subjektiivinen näkemys kirjoittamisesta. Tutkimusasetelma ei kuitenkaan mahdollista sujuvuusmittareiden ja kognitiivisen prosessoinnin, esimerkiksi sanahaun, välisen yhteyden tarkempaa tarkastelua (ks. myös alaluku 2.2).

Tutkimuskysymyksemme ovat:

1. Millaisia eroja kirjoitussujuvuudessa esiintyy L1-suomeksi, L2-englanniksi ja L2-ruotsiksi kirjoitettaessa?
2. Miten visualisoinnit havainnollistavat eroja kirjoitusprosesseissa?
3. Miten kirjoittajat kommentoivat kirjoitusprosessejaan tutkimuksessa esiin nousseiden erojen osalta?

Kirjoitussujuvuudella tarkoitamme tekstin vaivatonta etenemistä ilman pitkiä taukoja ja korjauksia (Segalowitz, 2010). Sujuvuutta voidaan mitata esimerkiksi taukojen osuuden, pituuden ja

sijainnin sekä yhtenäisten kirjoitusjaksojen eli purskeiden pituuden avulla. Tauoiksi lasketaan vähintään kahden sekunnin mittaiset (= kognitiiviset) tauot, kun taas purskeet ovat kahden tauon, korjauksen tai muun toiminnon välissä tuotettua kieliaineista (Chenoweth & Hayes, 2001). Sujuvasti etenevässä kirjoitusprosessissa taukoja on tavallisesti vähän ja ne ovat lyhyitä, kun taas tuotetut purskeet ovat pitkiä (Cislaru & Olive, 2018; Spellman Miller, 2006).

Aiemmat tutkimustulokset kirjoitussujuvuuden ja kielitaidon välisestä suhteesta ovat olleet osittain ristiriitaisia. Tutkimuksissa on lisäksi peräänkuulutettu monipuolisempien aineistojen ja metodien käyttöä (Palviainen, Kalaja & Mäntylä, 2012; ks. myös alaluku 2.2). Varsinkin tutkimuksia, joissa verrataan yksilön tuotaman ensikielen ja vieraiden kielten välisiä eroja kirjoitussujuvuudessa, on Suomessa tehty vähän (Lintunen ym., 2020). Tässä tutkimuksessa mukana on samojen kirjoittajien tuottamat tekstit L1-suomeksi, L2<sup>1</sup>-englanniksi ja L2-ruotsiksi. Aineisto pohjautuu kuuden suomenkielisen yliopisto-opiskelijan kolmen kielen näppäilyntalennusohjelmalla nauhoitettuihin kirjoitusprosesseihin.

Tutkimuksessa käytettävä näppäilyntalennusohjelma GenoGraphiX-Log (GGXLog) tallentaa laskennallista dataa kirjoitusprosesseista ja sisältää erilaisia työkaluja syntyvän tekstin visualisointia varten (Bécotte-Boutin, Caporossi, Leblay & Hertz, 2019; Leblay ym., 2022). Tässä tutkimuksessa viittaamme visualisoinneilla työkalun tuottamiin tauko- ja purskekuvioidiin (ks. alaluku 3.2). Yhdistämme tutkimuksessa visualisointeja laskennalliseen dataan. Lisäksi hyödynnämme haastatteluja ja taustakyselyjä kirjoittajan näkökulman avaamiseksi. Näin ollen käyttämämme menetelmät vastaavat aiemman tutkimuskirjallisuuden osoittamiin puutteisiin kirjoitussujuvuuden tutkimusmenetelmissä ja tuovat uuden näkökulman kirjoitusprosessien tutkimukseen.

Tutkimus on osa hanketta *Monikielisten kirjoittajien kirjoitusprosessit: sanastokeskittymien ja sujuvuuden visualisointi graafiteorian avulla* (KISUVI)<sup>2</sup>. Tämä tutkimus on tärkeä ensiaskel,

1 L2 viittaa tässä tutkimuksessa sekä toiseen että vieraaseen kieleen, emmekä tee näiden välille eroa.

2 KISUVI-hanke on Koneen Säätiön rahoittama nelivuotinen tutkimushanke, joka käynnistyi Turun yliopiston kieli- ja käännöstieteiden laitoksella maaliskuussa 2022. Hankkeen keskiössä on kirjoittamisen sujuvuuden tutkimus ensi- ja vieraalla kielellä (ks. Mutta ym., 2022).

joka kartoittaa käytettyjen menetelmien yhdistämistä monipuoliseen aineistoon ja pohjustaa siten hankkeen tulevaa tutkimusta, jossa näitä menetelmiä sovelletaan laajemman aineiston analyysiin.

## 2 KIRJOITUSSUJUUVUUS

Sujuvuus yhdistetään usein suulliseen kielitaitoon, ja arkikielessä sujuvuudella voidaan tarkoittaa yksilön kokonaisvaltaista kielitaitoa. Kirjoitussujuvuutta tutkittaessa voidaan tehdä ero esimerkiksi tuotoksen sujuvuuden, taustaprosessien eli kognitiivisen sujuvuuden ja havaitun sujuvuuden välille. Havaittu sujuvuus perustuu lukijan arvioon lopullisen tuotoksen sujuvuudesta. (Chenoweth & Hayes, 2001; Lintunen ym., 2020; Segalowitz, 2010.) Sujuvuutta onkin mahdollista tarkastella sekä subjektiivisesti tutkimalla havaittua sujuvuutta että objektiivisesti tutkimalla esimerkiksi taukojen ja purskeiden pituutta kirjoitusprosessissa. Tässä artikkelissa tarkoitamme sujuvuudella vaivatonta tekstin tuottamista, jota voidaan tarkastella tallennetussa kirjoitusprosessissa taukojen, korjauksien ja tuotetun kieliaineksen kautta (Chenoweth & Hayes, 2001; Segalowitz, 2010). Myös havaittu sujuvuus on luonnollisesti osa kirjoittajien lopullisten tuotosten arviointia (ks. alaluku 3.1).

### 2.1 Tauot ja purskeet sujuvuusmittareina

Tekstiä tuotetaan tyypillisesti useamman merkkijonon kirjoitusrupeamina eli purskeina, ja näitä purskeita rytmittävät erimittaiset tauot, korjaukset tai muut toiminnot (Baaijen & Galbraith, 2018; Conijn, Dux Speltz & Chukharev-Hudilainen, 2021; Hayes & Chenoweth, 2007). Kirjoitussujuvuutta tarkasteltaessa keskiössä ovat siten kirjoittajan taukokäyttäytyminen sekä purskeiden tuottaminen (Chenoweth & Hayes, 2001; Olive, 2014; Spelman Miller, 2000).

Tauon pituudeksi määritellään usein vähintään kaksi sekuntia, jotta kyseessä olisi ajatuksellinen, kognitiivinen tauko (Chenoweth & Hayes, 2001; Cislaru & Olive, 2018; Révész, Michel & Lee, 2019). Tällä rajataan pois suuri osa teknisistä tauoista, esimerkiksi oikean kirjaimen etsiminen näppäimistöltä, vaikkakin näin saatetaan rajata pois myös osa typografisista ja kielellisistä koodausprosesseista. Myös tässä tutkimuksessa tauon rajaksi on asetettu kaksi sekuntia GGXLog-ohjelman tuottamiin analyyseihin. Tarkastelemme taukokäyttäytymistä taukojen pituuden, prosenttiosuuden ja

sijainnin avulla (ks. myös alaluku 3.2). Aiemmassa tutkimuksessa eripituiset tauot ja niiden sijainti on usein yhdistetty eritasoisin kognitiivisiin prosesseihin, kuten esimerkiksi sanahakuun ja suunnitteluun kirjoitusprosessin aikana (Lindström & Zetterholm, 2019; Olive, 2002). Tässä tutkimuksessa taukokäyttäytymisen analyysin on tarkoitus antaa yleiskuva kirjoitussujuvuudesta ja kirjoitusprosessin etenemisestä eri kielissä.

Tuotoksesta voidaan tunnistaa purskeita, joita voidaan rajata tutkimuksissa eri tavoin. P-purske (engl. *pause-burst*) alkaa ja päättyy kognitiiviseen taukoon, kun taas R-purske (engl. *revision-burst*) päättyy muokkaukseen. Lisäksi tutkimuksissa on käytetty I-purskeita (engl. *contextual insertion*), jotka rajautuvat tekstin lisäämiseen aiemmin tuotetun tekstin sisälle. (Baaijen & Galbraith, 2018; Chenoweth & Hayes, 2001; Conijn ym., 2021; Hayes, 2009; Hayes & Chenoweth, 2007.) GGXLog katkaisee tuotoksen kognitiiviseen taukoon, muokkaukseen tai muuhun toimintoon, esimerkiksi hiiren painallukseen. Näin ollen kaikki edellämainitut toiminnot rajaavat tuotoksen omiksi purskeikseen. Kutsumme näitä V-purskeiksi (engl. *visualisation-burst*). (ks. Mutta & Laine, Åberg & Mäkilä, tulossa.) Analyysissa tarkastelemme purskeiden pituutta ja tuotostiheyttä sekä tehokkuutta purskeen aikana. Näiden mittareiden avulla on mahdollista muodostaa käsitys sekä yksittäisten purskeiden tuottamisesta että niiden jakautumisesta kirjoitusprosesseissa (ks. alaluku 3.2). Siten ne kuvaavat tekstin rakentumista kirjoitusprosessin aikana.

### 2.2 Kirjoitussujuvuus ensi- ja vieraalla kielellä

Kirjoitettaessa teksti rakentuu kirjoitusprosessin eri vaiheiden dynaamisessa vuorovaikutuksessa. Näitä vaiheita ovat suunnitteleminen, tuottaminen sekä tarkastaminen, ja niiden taustalla vaikuttavat monet eri tekijät kuten työmuisti ja prosessointinopeus. (Cislaru & Olive, 2018; Hayes & Flower, 1980; ks. myös Roca de Larios, Coyle & Nicolás-Conesa, 2016.) Kirjoitusprosessin aikana kirjoittaja suunnittelee (eksplisiittisesti tai implisiittisesti), tuottaa ja jossain määrin myös tarkistaa ja/tai muokkaa tekstiään. Nämä vaiheet sekoittuvat keskenään etenkin tietokoneella kirjoitettaessa. (Roca de Larios ym., 2016.)

Esimerkiksi taukojen pituus heijastelee kirjoittajan kognitiivista toimintaa kirjoitusprosessin

eri vaiheiden aikana. Lyhyiden taukojen aikana kirjoittaja saattaa pohtia oikeinkirjoitusta, kun taas pitkien taukojen aikana kirjoittaja usein suunnittelee tekstiään. Näin ollen pitkät tauot ovat tyypillisiä kirjoitusprosessin alkupäässä. Lisäksi pitkiä taukoja esiintyy usein kirjoitusprosessin lopussa, kun kirjoittaja lukee ja tarkastaa tekstiään. (Lindström & Zetterholm, 2019, s. 70; Olive, 2002, s. 134–144.) Tässä tutkimuksessa eri vaiheiden ja kognitiivisen prosessoinnin välisen suhteen tarkempi tarkastelu rajataan analyysin ulkopuolelle, sillä tämä vaatisi vielä lisäksi esimerkiksi log-tiedostojen, *think aloud* -menetelmän ja silmänliikeanalyysien yhdistämistä (Révész ym., 2019).

Kirjoitussujuvuuteen vaikuttavia ulkoisia tekijöitä ovat mm. tuotettavan tekstin aihe, tehtävänäntoon ja ympäristöön liittyvät tekijät, kun taas sisäisiä tekijöitä ovat mm. kirjoittajan diskurssikompetenssi, motivaatio ja affektiiviset tekijät (ks. esim. Dörnyei, 2005; Dörnyei & Ushioda, 2009; Lindström & Zetterholm, 2019; Makonnen-Craig, 2011; Segalowitz, 2016; Van Waes & Schellens, 2003). Lisäksi kirjoittamiseen vaikuttaa kirjoittajan kirjoitusprofiili. Kirjoitusprofiililla tarkoitetaan tapaa, jolla kirjoittaja suunnittelee, tuottaa ja muokkaa tekstiään. Kirjoitusprofiili voi vaihdella tekstistä toiseen tai jopa saman tekstin sisällä. (Mutta, 2020; Van Waes & Schellens, 2003.) Tässä tutkimuksessa kirjoitusprofiilin vaikutuksen tarkastelu on osa taustakyselyjen ja haastattelujen analysointia.

Vieraalla kielellä kirjoitettaessa kirjoitussujuvuuteen vaikuttaa myös kirjoittajan kielitaito (Chenoweth & Hayes, 2001, 2003; Mutta, 2020; Palviainen ym., 2012; Schoonen, Snellings, Stevenson & van Gelderen, 2009; Spelman Miller, 2006). Vieraalla kielellä kirjoittaminen on usein kognitiivisesti kuormittavampaa kuin ensikielellä kirjoittaminen, ja koska (kielenoppijan) kognitiiviset resurssit ovat rajalliset, kognitiivinen kuormittavuus voi näkyä mm. kirjoitussujuvuudessa (Jacques, Alves, Fadaei & Barbosa, 2020). Kuormittavuutta lisää kielenoppijan rajalliset tiedot kohdekielestä. Esimerkiksi puutteellinen sanavarasto tai kohdekielen kieliopin soveltaminen tuotokseen saattavat hidastaa tuotosprosessia (Chenoweth & Hayes, 2001; Ellis & Yuan, 2004; Gunnarsson-Largy, 2012; Mutta, 2020; Schoonen ym., 2009). Kielitaidon, lopullisen tuotoksen laadun ja kirjoitussujuvuuden suhde ei kuiten-

kaan ole yksiselitteinen, koska tauot ja muokkaukset ovat usein tarpeellisia kielitaidon tasosta riippumatta, jotta lopullinen tuotos olisi laadukas. Tutkimuksessamme kielitaidon vaikutuksen tutkiminen rajataan pääosin ensikielen ja vieraiden kielten väliin eroihin.

Aiemman tutkimuksen mukaan erot kirjoitussujuvuudessa ensikielellä ja vieraalla kielellä kirjoitettaessa näkyvät etenkin taukokäyttäytymisessä ja purskeiden pituudessa. Ensikielellä kirjoitettaessa tuotosprosessi sisältää usein vähemmän taukoja, ja ne voivat olla lyhyempiä kuin vieraalla kielellä kirjoitettaessa. Tauot eivät kuitenkaan aina ole lyhyempiä ensikielellä kirjoitettaessa, koska esimerkiksi suunnittelu on kohdekielestä riippumatonta. Lisäksi lyhyet tauot ovat vieraalla kielellä kirjoitettaessa hyvin tavallisia esimerkiksi oikeinkirjoitukseen liittyen. (Lindström & Zetterholm, 2019; Mutta, 2017; Spelman Miller, 2000; Spelman Miller ym., 2008.)

Ero kirjoitussujuvuudessa ensikielellä ja vieraalla kielellä kirjoitettaessa näkyy myös taukojen sijainneissa. Koska vieraalla kielellä kirjoitettaessa kirjoittaja joutuu keskeyttämään tuotosprosessin useammin esimerkiksi pysähtyessään miettimään sanan oikeinkirjoitusta, sanansisäisiä taukoja voi esiintyä enemmän kuin ensikielellä kirjoitettaessa. Ensikielellä kirjoitettaessa tauot sijoittuvat useimmin suurempien kokonaisuuksien (lause, virke, kappale) rajoille, ja sujuvassa kirjoitusprosessissa ei todennäköisesti esiinny pitkiä sanansisäisiä taukoja. (Mutta, 2020; Spelman Miller, 2006.)

Ensikielellä kirjoitettaessa tuotetaan tavallisesti pidempiä purskeita kuin vieraalla kielellä. Kielitaidon karttuessa kirjoittaja pystyy yleensä tuottamaan pidempiä purskeita myös vieraalla kielellä pysähtymättä pohtimaan esimerkiksi sanan kirjoitusasua. (Chenoweth & Hayes, 2001, 2003; Gunnarsson-Largy, 2012; Hayes, 2009; Hayes & Chenoweth, 2007; Mutta, 2020; Palviainen ym., 2012.)

### 3 AINEISTO JA ANALYYSIMENETELMÄT

#### 3.1 Aineiston kuvaus ja keruu

Aineisto on kerätty vuonna 2022 suomalaisen yliopiston pohjoismaisten kielten syventävien opintojen kursseilla osana laajempaa KISU-VI-hankkeen tutkimusaineistoa. Aineisto koostuu kuuden suomenkielisen yliopisto-opiskelijan taus-

takyselyistä, tallennetuista kirjoitusprosesseista, näiden pohjalta tuotetuista visualisoinneista sekä kirjoitusprosessiin perustuvista haastatteluista. Nämä kuusi opiskelijaa valittiin, koska heillä oli kirjoitettuna kaikki kolme tutkittavaa kieltä, ja he olivat osallistuneet haastatteluun. Tutkimukseen osallistuneet opiskelijat ovat antaneet GDPR-tietosuojasetuksen mukaiset tutkimusluvut. Aineisto on pseudonymisoitu, ja jokaiselle koehenkilölle on annettu henkilökohtainen koodi.

Opiskelijoiden kirjoitusprosessit on tallennettu GGXLog-näppäilyntallennusohjelmalla (Leblay ym., 2022). Erilaiset näppäilyntallennusohjelmat tallentavat kirjoitusprosessin ja keräävät lokitiedostoon tiedot esimerkiksi jokaisen näppäinpainalluksen ajankohdasta ja esiintymisestä kuvaten näin kognitiivista toimintaa kirjoitusprosessin aikana (Spelman Miller ym., 2008). GGXLog sisältää lisäksi erilaisia työkaluja syntyvän tekstin analyysia ja visualisointia varten (Caporossi & Leblay, 2015; Leblay ym., 2022). Tässä tutkimuksessa ohjelman avulla kirjoitusprosesseista tuotettiin laskennallisen datan lisäksi purske- ja taukokuviot.

Opiskelijat asensivat GGXLog-ohjelman omille tietokoneilleen ja kirjoittivat yhteensä kolme argumentoivaa tekstiä: suomeksi (L1), englanniksi (L2) ja ruotsiksi (L2). Tekstit kirjoitettiin tässä järjestyksessä. Suomenkielinen ja englanninkielinen teksti kirjoitettiin oppitunnilla ja ruotsinkielinen teksti etätapaamisessa. Kirjoittamisessa ei saanut käyttää apukeinoja (esim. sanakirjaa). Jokaisen tekstin kirjoittamiseen oli varattu aikaa 30 minuuttia. Suomeksi opiskelijat kirjoittivat töiden ja opiskelun yhdistämisestä: *Mitä mieltä olet töiden tekemisestä opiskelun ohessa? Mitä hyviä ja huonoja puolia siinä on?* Englanninkielisen tekstin aiheena oli vaihtoon lähteminen ja siihen liittyvät hyvät ja huonot puolet: *What is your opinion on student exchange? What are the advantages? What are the disadvantages?* Ruotsiksi opiskelijat kirjoittivat ajokortin tarpeellisuudesta ja auton omistamisen hyödyistä ja haitoista: *Tycker du att det är viktigt att ha körkort? Vilka för- och nackdelar finns det med att ha bil?* Kirjoitustehtävien aihepiirit käsittelivät opiskelijoiden arkea, jotta heidän oli mahdollisimman helppo kirjoittaa aiheesta argumentoiva teksti. Kirjoi-

telmat kirjoitettiin GGXLog-ohjelmassa tekstieditori-ikkunassa. Ohjelma tallensi samalla kaikki näppäimistön lyönnit, hiiren liikkeet, tauot ja niiden kestot sekä poistot ja lisäykset sekä muut kirjoitusprosesseihin liittyvät tapahtumat.

Kaksi KISUVI-hankkeen tutkijaa arvioi lopulliset L2-tuotokset eurooppalaisen viitekehysten<sup>3</sup> taitotasojen A1–C2 mukaan DIALUKI-kriteerien perusteella (CEFR, 2003). Yhtenäisen arvioinnin saavuttamiseksi ristiriitatilanteessa konsultoitiin kolmatta arvioijaa. Informanttien lopullisten tuotosten taitotasoarviot on koottu taulukkoon 1.

TAULUKKO 1

Lopullisten L2-tuotosten taitotasoarviot.

Informantti	Tuotoksen taitotaso L2-ruotsi	Tuotoksen taitotaso L2-englanti
ID2002	B2	B1
ID2004	B2	A2
ID2022	C1	B1
ID2024	C1	B1
ID2027	C1	B2
ID2031	B2	B1

Lopulliset tuotokset sijoittuvat CEFR-tasolle A2–C1, suurin osa tuotoksista B-tasolle. Ainoastaan informantti ID2004:n L2-englannin tuotos on arvioitu tasolle A2. Kaikkien kirjoittajien lopullisen tuotoksen taitotaso oli korkeampi ruotsin kielellä kirjoitettaessa (B2–C1) kuin englannin kielellä kirjoitettaessa (A2–B2).

Tutkimuksen yhteydessä opiskelijat vastasivat myös taustakyselyyn, jolla kartoitettiin perustietojen lisäksi kielellistä taustaa, kielitaitoa ja kirjoittamistapoja. Kyselyllä kerättyjä tietoja käytetään analyysin tukena. Opiskelijat olivat iältään 20–25-vuotiaita (keskiarvo 23,3).

Jälkikäteishaastattelujen (engl. *retrospective interview*) avulla voidaan täydentää näppäilyntallennuksen tuottamaa tietoa kirjoitussujuvuudesta. Jälkikäteishaastattelussa kirjoittajat kommentoivat kirjoitusprosessejaan kirjoittamisen jälkeen.

3 Euroopan neuvoston kehittämä Eurooppalainen viitekehys (engl. *Common European Framework of Reference for Languages*, CEFR) määrittelee ja kuvaa kielitaidon taitotasoa asteikolla A1–C2 (CEFR 2003).

Tällöin erilaiset tukikeinot voivat auttaa kirjoittajaa kommentoimaan oman tekstin rakentumista ja muistamaan esimerkiksi taukojen aikana tapahtuvaa prosessointia (engl. *stimulated recall*). (Bowles, 2018.) Tässä tutkimuksessa kirjoittaja ja haastattelija seurasivat yhdessä kirjoitusprosessia tallenteelta haastattelun aikana.

Haastattelut toteutettiin ja tallennettiin Zoom-ohjelman avulla. Haastattelun aikana haastattelija jakoi ruudulle GGXLog-ohjelmalla tallennetun kirjoitusprosessin. Tallenne voitiin pysäyttää keskustelun ajaksi tai sen nopeutta säätää tilanteen mukaan. Haastattelu oli puolistrukturoitu (engl. *semistructured interview*). Kirjoittaja sai vapaasti kommentoida tuotosprosessiaan, mutta haastattelua varten oli myös valmisteltu kirjoittamisen teemoihin liittyviä kysymyksiä. Kysymykset pohjautuvat aiemmasta tutkimuksesta nousseisiin teemoihin, kuten tekstin suunnittelu ja aloittaminen, kirjoitusstrategiat, tauot, korjaukset ja kirjoittamisnopeus (Chenoweth & Hayes, 2001; Cisarlu & Olive, 2018). Tallennetut haastattelut litteroitiin.

### 3.2 Analyysimenetelmät

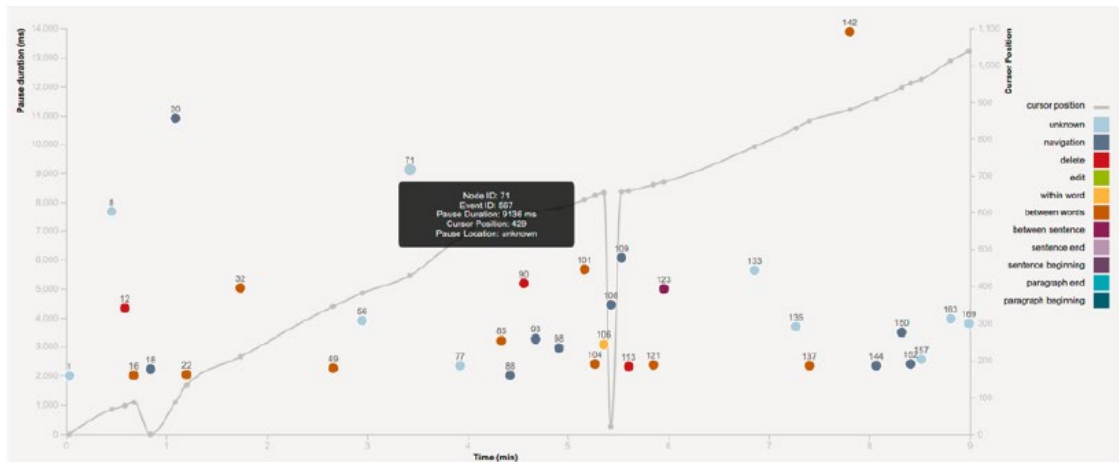
Kokosimme kirjoitusprosesseista deskriptiivistä dataa ja tuotimme tauko- ja purskeanalyysit GGXLog-ohjelman avulla (ks. Usoof ym., 2020). Tauko- ja purskeanalyysien keskiarvojen pohjalta valitsimme *extreme case sampling* -metodilla (ks. esim. Dörnyei, 2007, s. 128) tarkemman tarkas-

telun kohteeksi prosesseja, joissa sujuvuusmittarit osoittivat ryhmän ääripäitä tai koehenkilöitä, joiden kirjoitusprosesseissa oli selviä eroja eri kielillä kirjoitettaessa. Tämä rajausta mahdollistaa visualisointien tarkastelun sekä taustakyselyjen ja haastattelujen yhdistämisen analyysiin. Pienen otannan vuoksi aineistoa ei ole analysoitu tilastollisesti, ja näin ollen erot ääripäiden välillä eivät välttämättä ole tilastollisesti merkitseviä. Seuraavaksi vertasimme sujuvuuserojen havainnollistumista valituissa tuotoksissa visualisointien avulla. Lopuksi analysoimme kirjoittajien taustakyselyt ja puolistrukturoidut haastattelut kvalitatiivisen sisällönanalyysin avulla (Tuomi & Sarajarvi, 2018). Keskityimme kyselyissä ja haastatteluissa kohtiin, joissa kirjoittajat kommentoivat omaa kirjoitusproffiaa, kirjoitusprosessien sujuvuuteen vaikuttavia tekijöitä ja kieltenvälisiä eroja sujuvuudessa. Artikkelissa annamme esimerkkejä näistä sitaattien muodossa. Analyysissa tarkastelimme sekä yksilön tuottamien kielten välisiä eroja että yksilöiden välisiä eroja kielten sisällä.

Tutkimme taukojen osuutta tuotosprosessista, sillä eripituisissa tuotoksissa taukojen määrän vaihtelu ei aina osoita eroa prosessin sujuvuudessa. Analysoimme myös taukojen keskimääräistä pituutta, joka on yleisesti sujuvuustutkimuksissa käytetty mittari (ks. esim. Olive, 2014) sekä taukojen paikkaa, jonka oletetaan eroavan ensikielellä ja vieraalla kielellä kirjoitettaessa (ks. luku 2). Kuviossa 1 annetaan esimerkki GGXLog-ohjelman tuottamasta taukokuviosta.

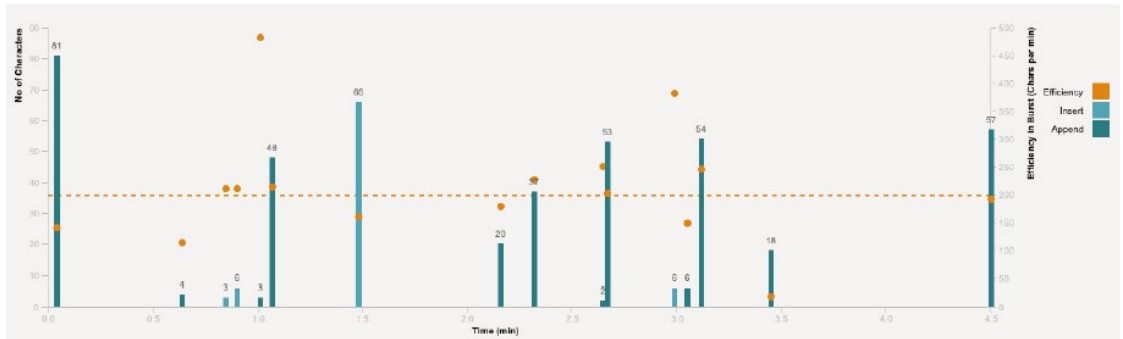
#### KUVIO 1

Esimerkki taukokuviosta (Usoof ym., 2020, s. 34).



## KUVIO 2

Esimerkki purskekuviosta (Usoof ym., 2020, s. 44).



Kaikki artikkelin värilliset kuviot ovat saatavilla QR-koodin takaa.<sup>4</sup>

Taukokuviolla voidaan havainnollistaa tuotosprosessin sisältämien taukojen pituutta ja sijaintia. Vasen y-akseli kertoo tauon pituuden, x-akselilla näkyy tuotosprosessin aika. Y-akselin asteikko mukautuu pisimmän tauon mukaan. Taukojen sijainti merkitään kuviossa värein ja värien selitykset on annettu kuviossa oikealla (ks. QR-koodi). Tauot, joita ei voi luokitella, kategorisoidaan tuntemattomiksi (engl. *unknown*). Esimerkiksi tauot hiiren liikkeiden välillä luokitellaan tuntemattomiksi tauoiksi (Usoof ym., 2020, s. 33). Analyysissa keskityimme kuviossa keltaisella merkittyihin sanansisäisiin taukoihin (engl. *within word*), punaruskealla merkittyihin taukoihin sanojen välissä (engl. *between words*) ja tummansinisellä merkittyihin navigointiin liittyviin taukoihin (engl. *navigation*). Lisäksi kuviosta voidaan nähdä kursorin paikka tauon aikana (oikea y-akseli) sekä kursorin sijainti tuotoksessa ja näin ollen tarkastella, miten kirjoittaja on navigoinut tekstissä. Tässä tutkimuksessa olemme muokanneet taukokuviota luettavuuden parantamiseksi.<sup>5</sup> Olemme vahvistaneet värikontrasteja, suurentaneet fonttikokoa ja jättäneet kuvioista pois taukojen numeroinnit, joita ei analysoida tässä tutkimuksessa. Taukokuviota vertailtaessa tulee huomioida ero erityisesti vasemman Y-akselin asteikossa, joka osoittaa tauon pituuden. Asteikkoja ei voi skaalata vastaamaan toisiaan,

koska usein erot tuotosprosessien pisimmissä tauoissa ovat liian suuret.

Purskeista tutkimme sekä purskeiden keskimääräistä pituutta, tuotostiheyttä että näppäilynopeutta purskeen aikana (= tehokkuus). Purskeiden keskimääräinen pituus lasketaan merkkimäärästä, koska morfologisesti erilaisten kielten (esim. suomi ja englantia/ruotsi) vertailu sanamäärien avulla on ongelmallista (ks. esim. Mutta & Laine, 2022). Tuotostiheys näyttää, kuinka paljon purskeita kirjoittaja tuottaa ja miten nämä jakautuvat tuotosprosessissa. Tehokkuus kertoo keskimääräisen näppäilynopeuden purskeen aikana. Kuviossa 2 annetaan esimerkki purskekuviosta.

Purskekuviot visualisoivat purskeiden pituutta, tuotostiheyttä sekä tehokkuutta purskeen aikana. Vasen y-akseli osoittaa purskeen pituuden ja x-akseli tuotosprosessin aikaa. Y-akselin asteikko on mukautettu pisimmän purskeen mukaan. Purskeiden tuotospaikka merkitään kuviossa eri värein (ks. QR-koodin takaa löytyvät värilliset kuviot). Tekstin perään tuotetut purskeet (engl. *append*) näkyvät kuviossa tummina pylväinä, kun taas tekstin sisälle tuotetut purskeet (engl. *insert*) ovat vaaleita. Lisäksi purskekuvioista käyvät ilmi näppäilynopeus purskeen aikana (oikea y-akseli ja kuvion oranssit pallot, engl. *fluency*) ja tämän leikattu keskiarvo (oranssi katkoviiva, engl. *trimmed mean*)<sup>6</sup>. Tässä tutkimuksessa olemme muokanneet myös purskekuviota luettavuuden

4 QR-koodi värillisiin kuvioihin

5 Kiitämme KISUVI-hankkeen projektitutkija Antti Suomelaa kaikesta teknisestä avusta ja kuvioiden muokkauksesta.

6 Purskeen aikaisen näppäilynopeuden 10 %:n leikattu keskiarvo, jolloin 10 % pienimmistä ja suurimmista arvoista on jätetty pois ennen keskiarvon laskemista (Grönroos, 2004: 41).



parantamiseksi. Olemme vahvistaneet värikont-rasteja, suurentaneet fonttikokoa ja annamme kuvioissa purskeen pituuden ainoastaan niissä tapauksissa, joissa purskeen pituus on yli kaksin-kertainen suhteessa keskimääräiseen purskeen pituuteen. Lisäksi keskenään vertailtavat purs-kekuviot on skaalattu saman asteikon mukaan vertailtavuuden parantamiseksi, koska mukau-tuminen purskeen pituuden mukaan vaikeuttaa vertailua purskeen keskimääräisessä pituudessa.

## 4 KIRJOITUSPROSESSIN SUJUVUUS KOLMEN ERI KIELEN KIRJOITUS- PROSESSEISSA

### 4.1 Kirjoitusprosessien deskriptiivinen analyysi

Taulukkoon 2 on koottu tiedot kirjoitusprosessin kokonaisajasta (h:min:sek) ja lopullisen tuotok-sen sanamäärästä L1-suomessa, L2-englannissa ja L2-ruotsissa. Taulukossa annetaan nämä tiedot yksilöittäin sekä näiden tietojen pienin arvo (min), suurin arvo (max), keskiarvo (ka) ja keski-hajonta (kh) ryhmätasolla.

Taulukosta 2 nähdään, että ruotsiksi kirjoitet-taessa kirjoitusprosessiin käytettiin keskimäärin eniten aikaa. Englanniksi kirjoitettaessa käytetty

aika oli lyhin. Yksilöiden kirjoitusajoissa oli suurta vaihtelua. Kirjoitusaika vaihteli 9 minuut-tista 40 sekunnista (informantti ID2022, englantii) 29 minuuttiin 50 sekuntiin (informantti ID2004, ruotsi).

Sanamääristä ei voi tehdä suoraan vertailuja, koska kirjoitusprosessiin käytetty aika vaihteli, ja koska kielet ovat morfologisesti erilaisia. Sana-määrää käytetään kuvaamaan aineiston laajuutta. Eniten sanoja lopputuotoksessa oli keskimäärin suomeksi, sitten ruotsiksi ja vähiten englanniksi. Myös lopullisen tuotoksen sanamäärissä oli suuria yksilöllisiä eroja. Lyhin lopputuotos sisälsi vain 90 sanaa (informantti ID2004, englantii) ja pisin loppu-tuotos 454 sanaa (informantti ID2031, englantii).

### 4.2 Kirjoitusprosessien tauot

Taulukossa 3 esitellään kognitiivisten taukojen lukumäärä, kokonaisaika (h:min:sek), keski-määräinen pituus millisekunneina ja osuus koko kirjoitusprosessista. Taukojen osuus kirjoitus-prosessista on laskettu jakamalla kognitiivisten taukojen kokonaisaika koko kirjoitusprosessiin käytetyllä ajalla (ks. Taulukko 2).

Kognitiiviset tauot ovat keskimäärin pisimpiä L1-suomessa ja lyhimpiä L2-englannissa. Vaihte-luväli kognitiivisten taukojen keskimääräisessä

#### TAULUKKO 2

Kirjoitusprosessin kokonaisaika ja lopullisen tuotoksen sanamäärä.

Informantti	Kirjoitusprosessin kokonaisaika (h:min:sek)			Sanamäärä (lkm)		
	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi
ID2002	0:19:55	0:21:25	0:23:48	107	139	122
ID2004	0:25:44	0:21:54	0:29:50	221	90	202
ID2022	0:26:07	0:09:40	0:19:17	431	180	207
ID2024	0:19:42	0:14:20	0:26:39	158	152	256
ID2027	0:29:30	0:14:05	0:28:17	114	140	230
ID2031	0:26:28	0:21:49	0:25:09	430	454	197
min	0:19:42	0:09:40	0:19:17	107	90	122
max	0:29:30	0:21:54	0:29:50	431	454	256
ka	0:24:34	0:17:12	0:25:30	243,5	192,5	202,3
kh	0:03:55	0:05:13	0:03:44	150,4	131,4	45,1

min = pienin arvo, max = suurin arvo, ka = keskiarvo, kh = keskihajonta

## TAULUKKO 3

Kognitiiviset taudit kirjoitusprosesseissa.

Informantti	Kognitiivisten taukojen lukumäärä			Kognitiivisten taukojen kokonaisaika (h:min:sek)			Kognitiivisten taukojen keskimääräinen pituus (ms)			Kognitiivisten taukojen osuus (%)		
	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi
ID2002	40	92	98	0:13:56	0:11:49	0:13:08	20918,9	7714,4	8049,0	70,0	55,2	55,2
ID2004	87	77	107	0:13:13	0:16:19	0:20:54	9125,4	12726,8	11728,0	51,4	74,6	70,1
ID2022	91	36	79	0:07:51	0:04:06	0:12:19	5184,9	6853,5	9244,1	30,1	42,5	63,1
ID2024	36	55	67	0:12:43	0:08:25	0:19:36	21218,8	9190,2	17563,6	64,6	58,7	73,6
ID2027	107	49	123	0:16:59	0:07:23	0:17:23	9524,9	9051,1	8487,6	57,6	52,5	61,5
ID2031	98	78	104	0:10:48	0:09:02	0:17:39	6618,4	6958,6	10190,6	40,8	41,5	70,2
min	36	36	67	0:07:51	0:04:06	0:12:19	5184,9	6853,5	8049,0	30,1	41,5	55,2
max	107	92	123	0:16:59	0:16:19	0:20:54	21218,8	12726,8	17563,6	70,0	74,6	73,6
ka	76,5	64,5	96,3	0:12:35	0:09:31	0:16:50	12098,6	8749,1	10877,1	52,4	54,2	65,6
kh	30,6	21,2	20,2	0:03:04	0:04:10	0:03:27	7131,2	2189,9	3529,3	15,0	12,2	6,9

min = pienin arvo, max = suurin arvo, ka = keskiarvo, kh = keskihajonta

pituudessa on suuri. Yksilötasolla keskimääräinen tauon pituus on pisin L1-suomessa informantilla ID2024 (21218,8 ms), mutta myös lyhin L1-suomessa informantilla ID2022 (5184,9 ms). Informantilla ID2022 keskimääräinen tauon pituus on lyhin kaikissa tutkittavissa kielissä. Vaikka taudit ovat keskimääräisesti pisimpiä ensikielellä kirjoitettaessa, on kognitiivisten taukojen osuus kirjoitusprosessista kuitenkin pienin L1-suomessa ja suurin L2-ruotsissa. Myös taukojen osuudessa yksilöllinen vaihtelu on suurta. Suurin taukojen osuus on informantilla ID2004 L2-englannissa (74,6 %, taitotaso A2, ks. Taulukko 1) ja pienin informantilla ID2022 L1-suomessa (30,1 %).

Seuraavaksi tarkastelemme GGXLog-ohjelman tuottamaa analyysia taukojen sijainneista. Tässä tutkimuksessa keskitymme sanansisäisiin taukoihin, taukoihin sanojen välissä ja navigointiin liittyviin taukoihin.<sup>7</sup> Taulukkoon 4 on koottu tiedot näiden taukojen osalta yksilöittäin jokaisesta kielestä ja pienin arvo, suurin arvo, keskiarvo ja keskihajonta on ilmoitettu kielittäin ryhmätasolla. Taulukossa näkyy kunkin taुकотyyppin luku-

määrä ja osuus kaikista tauoista.

Sanansisäisten taukojen keskimääräinen osuus on suurin L2-ruotsissa ja pienin L2-englannissa. Sanansisäisten taukojen keskimääräistä osuutta L1-suomessa nostaa yksittäisen informantin ID2027 korkea sanansisäisten taukojen määrä (10 kpl). Etenkin ero L2-englannin ja L1-suomen välillä on näiden taukojen osalta suhteellisen pieni.

Sanojen välissä sijaitsevien taukojen keskimääräinen osuus on suurin L2-englannissa ja pienin L2-ruotsissa. Taulukosta käy myös ilmi, että tietyt taुकотyyppit vaikuttavat painottuvan tietyille yksilöille. Esimerkiksi informantilla ID2022 on verrattain paljon navigointitaukoja kaikissa kielissä. Informantilla ID2004 on runsaasti sanansisäisiä taukoja etenkin L2-englannissa ja L2-ruotsissa, myös navigointitaukoja on paljon etenkin L1-suomessa ja L2-englannissa.

Seuraavaksi tarkastelemme, miten taulukoissa 3 ja 4 esiin tulevat erot taुकотyyppien käytössä näkyvät visualisoinneissa valittujen informanttien kohdalla. Tarkasteluun valiokoitu-

<sup>7</sup> Koska taulukossa annetaan vain näiden tarkasteltavien taुकотyyppien lukumäärä ja osuus kaikista tauoista, prosentteista ei muodostu 100 %. Muut taुकотyyppit jäävät analyysin ulkopuolelle.

## TAULUKKO 4

Taukojen sijainnit kirjoitusprosesseissa.

Informantti	Sanansisäinen tauko %-osuus (lkm)			Tauko sanojen välissä %-osuus (lkm)			Navigointi %-osuus (lkm)		
	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi
ID2002	2,50 (1)	3,26 (3)	7,14 (7)	47,50 (19)	48,91 (45)	32,65 (32)	7,50 (3)	13,04 (12)	13,27 (13)
ID2004	5,75 (5)	11,69 (9)	16,82 (18)	27,59 (24)	45,45 (35)	29,91 (32)	24,14 (21)	9,09 (7)	28,04 (30)
ID2022	2,08 (2)	2,78 (1)	2,50 (2)	31,25 (30)	47,22 (17)	37,50 (30)	23,96 (23)	16,67 (6)	13,75 (11)
ID2024	5,56 (2)	1,82 (1)	2,99 (2)	44,44 (16)	32,73 (18)	52,24 (35)	5,56 (2)	23,64 (13)	2,99 (2)
ID2027	9,35 (10)	2,04 (1)	3,25 (4)	43,93 (47)	53,06 (26)	47,15 (58)	5,61 (6)	6,12 (3)	4,88 (6)
ID2031	2,04 (2)	1,28 (1)	2,88 (3)	46,9 (46)	44,87 (35)	35,58 (37)	9,18 (9)	6,41 (5)	21,15 (22)
min	2,04 (2)	1,28 (1)	2,50 (2)	27,59 (24)	32,73 (18)	29,91 (32)	5,56 (2)	6,12 (3)	2,99 (2)
max	9,35 (10)	11,69 (9)	16,82 (18)	47,50 (19)	53,06 (26)	52,24 (35)	24,14 (21)	23,64 (13)	28,04 (30)
ka	4,55 (3,67)	3,81 (2,67)	5,93 (6,00)	40,28 (30,33)	45,37 (29,33)	39,17 (37,33)	12,66 (10,67)	12,50 (7,67)	14,01 (14,00)
kh	3.39	3.20	6.16	13.40	10.97	10.42	9.14	3.98	10.37

min = pienin arvo, max = suurin arvo, ka = keskiarvo, kh = keskihajonta

vat informantit ID2022 ja ID2002. Informantilla ID2022 on pienin kognitiivisten taukojen osuus L1-suomessa (30,1 %). Taukoja on tuotosprosessissa paljon (91 kpl), mutta ne ovat verrattain hyvin lyhyitä (5184,9 ms). Informantilla ID2002 kognitiivisten taukojen osuus L1-suomessa on taas ryhmän suurin (70,0 %). Taukoja on kuitenkin vähän (40 kpl), mutta ne ovat keskimäärin hyvin pitkiä (20918,9 ms). Kuvio 3 havainnollistaa informantin ID2022 ja kuvio 4 informantin ID2002 taukokäyttämistä.<sup>8</sup> Värilliset kuvat löytyvät QR-koodin takaa (ks. alaluku 3.2).

Erot taukokäyttämässä informanttien ID2022 ja ID2002 välillä näkyvät kuvioissa hyvin. Lisäksi kuvioista käy ilmi taukojen sijainnit. Sanansisäisiä taukoja on molemmissa tuotoksissa ensikielelle tyypillisesti vähän. Informan-

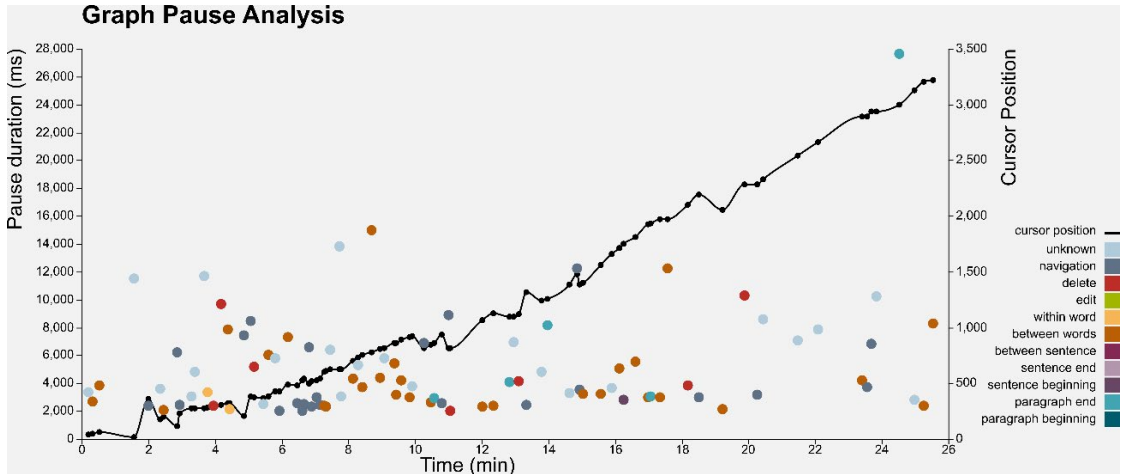
tilla ID2002 tällaisia taukoja on yksi (2,50 %) ja informantilla ID2022 kaksi (2,08 %). Suurin osa tauoista esiintyy sanojen välissä. Informantilla ID2022 on lisäksi navigointiin liittyviä taukoja huomattavasti enemmän kuin informantilla ID2002 (23,96 % vs. 7,50 %). Ero navigoinnissa näkyy selvästi myös taukokuvioissa kursorin paikkaa osoittavassa käyrässä. Tauot jakautuvat molemmissa tuotoksissa tasaisesti prosessin aikana.

Kvantitatiivisen analyysin perusteella informantin ID2022 kirjoittaminen L1-suomeksi oli ryhmän sujuvinta taukojen osalta. Informantin ID2022 mukaan kirjoittaminen on hänelle mielekästä, ja tekstiä syntyy yleensä helposti, mikä saattaa olla yksi tekijä kuviossa näkyvien lyhyiden taukojen taustalla. Lisäksi kuviossa näkyy,

<sup>8</sup> Kuvioita vertailtaessa on huomioitava ero asteikossa, joka määräytyy pisimmän tauon mukaan. Informantin ID2022 taukokuviossa asteikko on 28 000 millisekuntiin asti, kun taas informantin ID2002 taukokuvio jopa 120 000 millisekuntiin asti (1 s = 1000 ms).

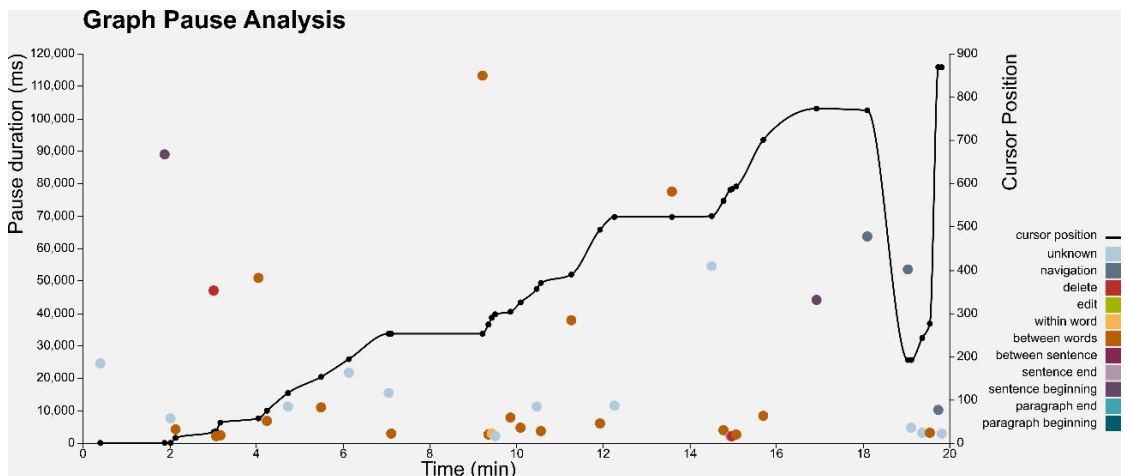
## KUVIO 3

Informantin ID2022 taukokuvio L1-suomessa (asteikko 28 000 ms).



## KUVIO 4

Informantin ID2002 taukokuvio L1-suomessa (asteikko 120 000 ms).



ettei hän liiku tekstissä paljoakaan, vaan teksti rakentuu suhteellisen suoraviivaisesti. Taustakyselyn mukaan hän pyrkiikin heti kirjoittamaan melko valmista tekstiä:

*Kirjoittaminen on minulle mielekästä, ja tekstiä syntyy yleensä helposti. Jos kirjoitan esimerkiksi esseitä, pyrin heti ensimmäisellä kerralla kirjoittamaan aika valmista tekstiä, eli en kauheasti muokkaile teksti sen kirjoittamisen jälkeen.*

Haastattelussa hän kommentoi L1-kirjoitusprosessiaan seuraavasti:

*No se ne totta kai tekstin tuottaminen suomeksi kun se on äidinkieli, niin se on kaikista sujuvinta ja nopeinta ja helpointa.*

Myös tehtävänanto helpotti kirjoitusprosessia. L1-kielen aihe oli hänelle sopiva:

*Kun oli suomeksi se niin jotenkin siitä niin se oli se, että työskentelyä opintojen ohessa, niin se oli tavallaan itselle semmoinen mistä oli helppo tuottaa tekstiä.*

Informantin ID2002 taukoanalyysi osoittaa pitkiä taukoja L1-kirjoitusprosessissa. Informantti ID2002 kommentoikin taustakyselyssä, ettei ole kovin nopea kirjoittamaan valmista tekstiä. Haastattelun mukaan mahdollisia syitä pitkiin taukoihin voivat olla tilanteen aiheuttama jännitys ja aikaraja:

*Nyt kun ties et se nauhotetaan niin sitten halus*

*jotenkin miettii eka rauhassa ja sitten kirjoittaa vasta. --- Kun on kuitenkin sellai niinku testi ja sit siin on niinku aikaraja ja sit se on vähän jotenkin siihen sellai pieni jännitys sen takii.*

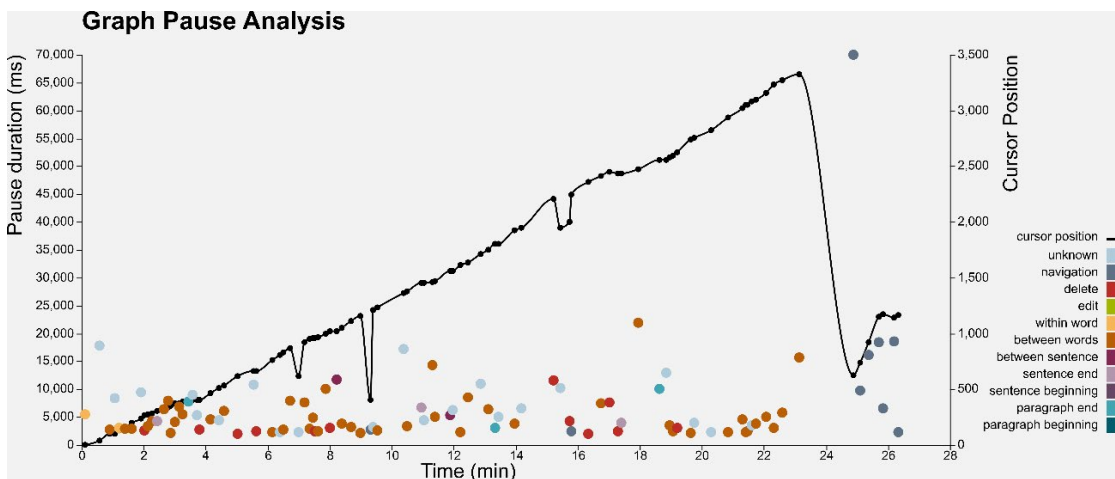
Yksilön tuottamien kielten välisiä eroja tarkastellaan informantin ID2031 tuotosprosesseissa. Tämä informantti valikoitui tarkasteluun, koska hänen taukokäyttäytymisensä eroaa selvästi tutkittavien kielten välillä. Kuviot 5, 6 ja 7 havainnollistavat näitä eroja.<sup>9</sup>

Informantilla ID2031 on L1-suomessa toiseksi eniten kognitiivisia taukoja (98 kpl), mutta tauko-

jen osuus prosessista on pieni (40,8 %), ja tauot ovat keskimäärin lyhimpiä (ka = 6618,4 ms). Kirjoitusprosessin loppuvaiheessa on kuitenkin yksi pidempi tauko, jolloin kirjoittaja palaa tekstin alkuun, kuten kursorin sijaintia osoittavasta käyrästä näkyy. L2-englannissa taukoja on vähiten (78 kpl), taukojen osuus on toiseksi pieni (41,5 %), ja tauot ovat myös toiseksi lyhimpiä (6958,6 ms). L2-ruotsissa kognitiivisia taukoja on eniten (104 kpl), ja ne ovat myös keskimääräisesti pisimpiä (10190,6 ms). Myös taukojen osuus prosessista (70,2 %) on suurin L2-ruotsissa. Nämä erot kielten

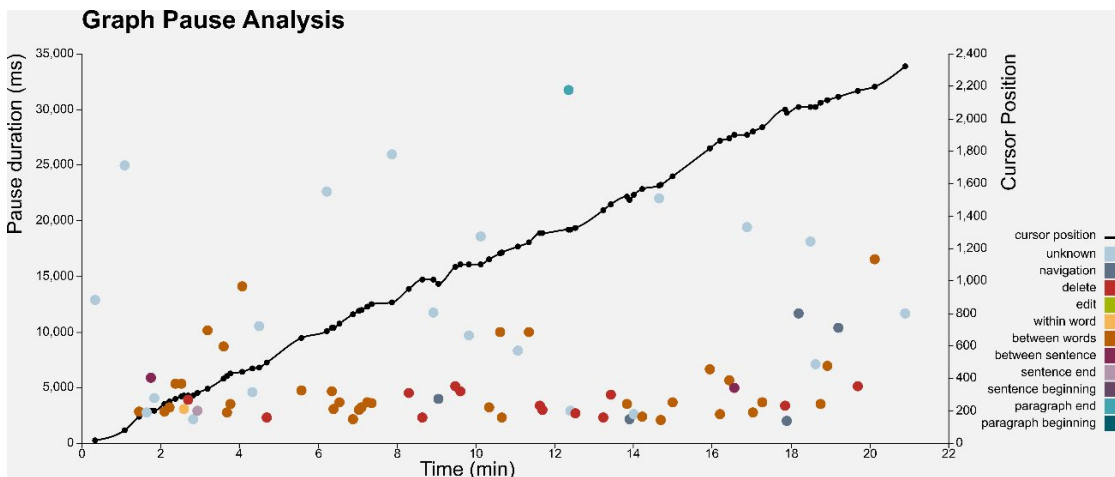
#### KUVIO 5

Informantin ID2031 taukokuvio L1-suomessa (asteikko 70 000 ms).



#### KUVIO 6

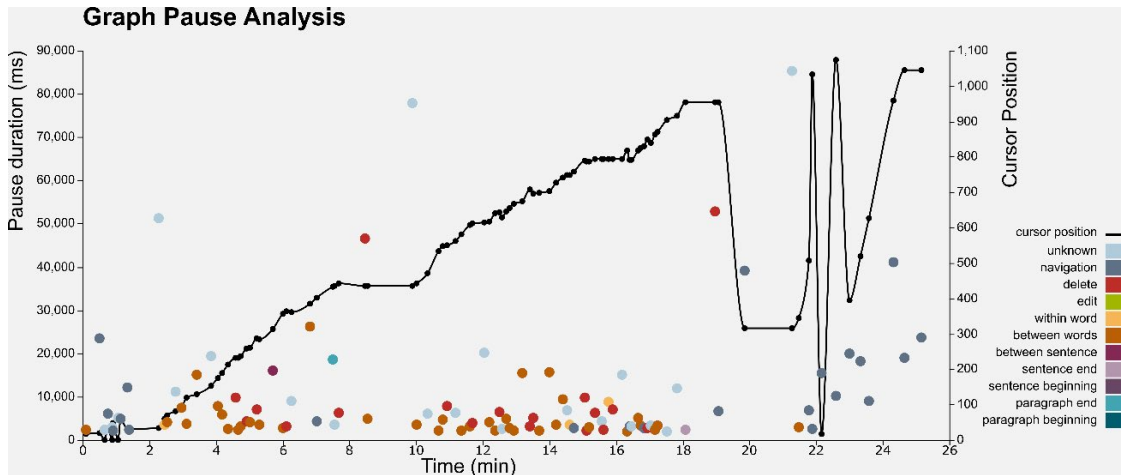
Informantin ID2031 taukokuvio L2-englannissa (asteikko 35 000 ms).



<sup>9</sup> Kuviaita vertailtaessa on huomioitava ero asteikossa, joka määräytyy pisimmän tauon mukaan. Informantin ID2031 L1-suomen taukokuviossa asteikko on 70 000 millisekuntia, L2-englannin 35 000 millisekuntia ja L2-ruotsin 90 000 millisekuntia (1 s = 1000 ms).

## KUVIO 7

Informantin ID2031 taukokuvio L2-ruotsissa (asteikko 90 000 ms).



välillä ovat havaittavissa kuvioita vertailtaessa, vaikka asteikon mukautuminen tauon pituuden mukaan saattaa vaikeuttaa tulkintaa.

Informantilla ID2031 on myös keskimäärin enemmän taukoja sanojen välissä L1-suomessa (46,90 %) ja L2-englannissa (44,87 %) kuin L2-ruotsissa (35,58 %). Sanansisäisiä taukoja on kaikissa kielissä vähän, L1-suomessa 2,04 %, L2-englannissa 1,28 % ja L2-ruotsissa 2,88 %. Erot taukojen sijainnissa ovat suurempia navigointitaukojen osalta. L1-suomessa navigointiin liittyviä taukoja on 9,18 %, ja kuten kuvioista 5 voidaan havaita, ne painottuvat kirjoitusprosessin loppuun. Tämä osoittaa, että kirjoittaja on kirjoitusprosessin lopussa palannut tekstissä taaksepäin. Kuvioista 6 ja 7 käy ilmi, että L2-englannin tuotos on hyvin suoraviivainen, ja navigointiin liittyviä taukoja on suhteellisen vähän (6,41 %), kun taas L2-ruotsin tuotoksessa informantti ID2031 navigoi paljon ja tähän liittyviä taukoja on jopa 21,15 %. Kuvioista 7 nähdään, että L2-ruotsin osalta navigointitaukoja on sekä prosessin alussa että lopussa. Erot navigoinnissa näkyvät selvästi myös kursorin paikkaa tekstissä osoittavasta käyrästä. Yleisesti ottaen taukoja esiintyy kaikissa kielissä suhteellisen tasaisesti koko prosessin ajan, mutta L2-ruotsissa suuri osa tauoista keskittyy prosessin keskivaiheille.

Informantti ID2031 kommentoi haastattelussa, että suomeksi on lähtökohtaisesti helppo kirjoittaa, mutta hänellä oli vaikeuksia aiheen rajaamisen kanssa:

*Suomi nyt on äidinkielen, niin siinä ei periaatteessa melkein ikinä pelota kauheasti. --- Mä olin silleen, että silloin piti kirjoittaa niinku työnteosta opiskeluitten ohessa, niin siinä mä olin vähän silleen niinku hämilleni niin että kun oli niin paljon ajatuksia niin että mitä mä nyt sit oikeasti haluan tuoda esiin siinä.*

L2-kielten välisiä eroja saattaa informantin mukaan selittää esimerkiksi se, että hän on viime aikoina käyttänyt enemmän englantia opinnoissaan:

*Ehkä en ole vaikka ollut sellainen, että apua että nyt täytyy kirjoittaa englanniksi tai tälleen, kun viime aikoina on sitten käyttänyt aika paljon silleen englantia. --- On aika silleen tuoreessa muistissa, jotkut niinku Academic English kurssit tai tällaiset, ja sitten mä oon nyt yhtä kurssikirjaa lukenut mikä on englanniksi, niin ehkä joo, englantia on ollut siellä aika pinnalla viime aikoina.*

Ruotsiksi kirjoittaessa prosessi ei ollut yhtä sujuva kvantitatiivisen analyysin perusteella, mikä näkyy myös visualisoinnissa (ks. kuvio 7). Informantin mukaan aihe oli vaikeampi. Lisäksi hän ei ole viime aikoina kirjoittanut paljon ruotsiksi, mikä voi vaikuttaa kirjoitussujuvuuteen:

*[Ruotsi pelottaa] ehkä sitten taas ihan vähän ja --- tässä tehtävässä just kun oli vähän silleen itelle ei niin läheinen aihe. --- Viime aikoina vähän ärsyttänyt, kun on tullut semmoisia vähän tyhmiä virheitä, siksi kun ei ole pitkään aikaan kirjoittanut ruotsiksi.*

### 4.3 Kirjoitusprosessien purskeet

Taulukossa 5 on koottuna tiedot kirjoittajien keskimääräisistä purskeen pituuksista (= tuotettu merkkimäärä sisältäen välilyönnit jaettuna purskeiden lukumäärällä), purskeiden maksimipituuksista ja tehokkuudesta purskeen aikana (= purskeen sisältämä merkkimäärä jaettuna purskeen kestolla)<sup>10</sup>. Taulukossa annetaan nämä tiedot yksilöittäin sekä näiden tietojen pienin arvo (min), suurin arvo (max), keskiarvo (ka) ja keskihajonta (kh) ryhmätasolla.

Ryhmätasolla purskeet ovat keskimäärin pidempiä L1-suomessa kuin L2-englannissa ja L2-ruotsissa. Ero L2-kielten välillä on suhteellisen pieni. Yksilölliset erot taas ovat suurempia L2-kielissä kuin L1-kielillä kirjoitettaessa. L2-englannissa lyhin purskeen keskimääräinen pituus on 7,87 merkkiä (informantti ID2004) ja pisin 17,69 (informantti ID2022). L2-ruotsissa lyhin keskimääräinen purskeen pituus on 9,28 (informantti ID2004) ja pisin keskimääräinen purskeen pituus 15,60 (informantti ID2024).

Taulukko osoittaa, että tuotettujen purskeiden pituus on melko yksilöllinen. Informantti ID2004

tuottaa lyhyitä purskeita etenkin L2-kielissä, kun taas informantilla ID2022 purskeet ovat ryhmätasolla tarkasteltuna pitkiä kaikissa kielissä. Informantti ID2022 tuotti aineiston pisimmän purskeen L1-kielillä (214 merkkiä). Tehokkuudessa erot ryhmätasolla ovat hyvin pieniä, mutta yksilötasolla vaihtelu on suurta. Tehokkuus on matalin informantin ID2027 L2-englannin tuotoksessa (168,36 merkkiä minuutissa purskeen aikana) ja korkein informantin ID2024 L2-englannin tuotoksessa (429,92 merkkiä minuutissa purskeen aikana). Taulukosta käy myös ilmi, että informantti ID2024 tuottaa purskeet tehokkaasti kaikissa kielissä, kun taas informantin ID2027 purskeen aikainen tuottaminen on hitaampaa.

Kuvioissa 8 ja 9 verrataan kahden informantin ID2004 ja ID2024 tuottamia purskeita L2-ruotsissa. Informantti ID2004 tuotti ryhmän lyhimät ja informantti ID2024 ryhmän pisimmät purskeet L2-ruotsiksi kirjoitettaessa. Värilliset kuviot löytyvät QR-koodin takaa (ks. alaluku 3.2).

Kuviot osoittavat selvästi eron informantin ID2004 ja informantin ID2024 tuotosprosessien välille purskeiden tuottamisen osalta sekä purs-

TAULUKKO 5

Purskeet kirjoitusprosessissa.

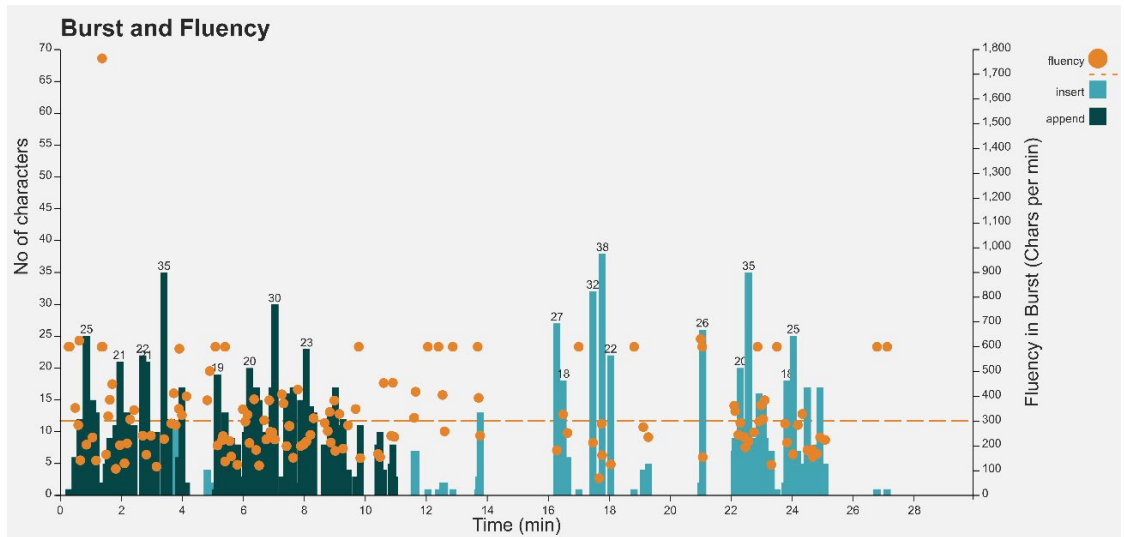
Informantti	Keskimääräinen purskeen pituus (merkkiä)			Pisimmän purskeen pituus (merkkiä)			Tehokkuus (merkkiä minuutissa purskeen aikana)		
	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi	L1-suomi	L2-englanti	L2-ruotsi
ID2002	13,67	9,50	9,41	59	39	48	323,94	381,65	350,61
ID2004	16,61	7,87	9,28	86	41	38	290,35	231,80	301,62
ID2022	25,03	17,69	15,06	214	73	81	419,46	374,07	367,51
ID2024	18,26	14,34	15,60	80	59	64	390,67	429,92	364,11
ID2027	11,67	14,98	10,76	49	95	59	237,86	168,36	229,44
ID2031	-	14,52	10,35	-	82	47	-	373,37	379,23
min	11,67	7,87	9,28				237,86	168,36	229,44
max	25,03	17,69	15,60				419,46	429,92	379,23
ka	17,05	12,88	12,02				332,46	326,53	332,09
kh	5,14	3,70	2,84				73,74	102,15	57,11

min = pienin arvo, max = suurin arvo, ka = keskiarvo, kh = keskihajonta

10 Informantin ID2031 L1-suomen kirjoitusprosessista ei saanut purskeanalyysia teknisen vian takia.

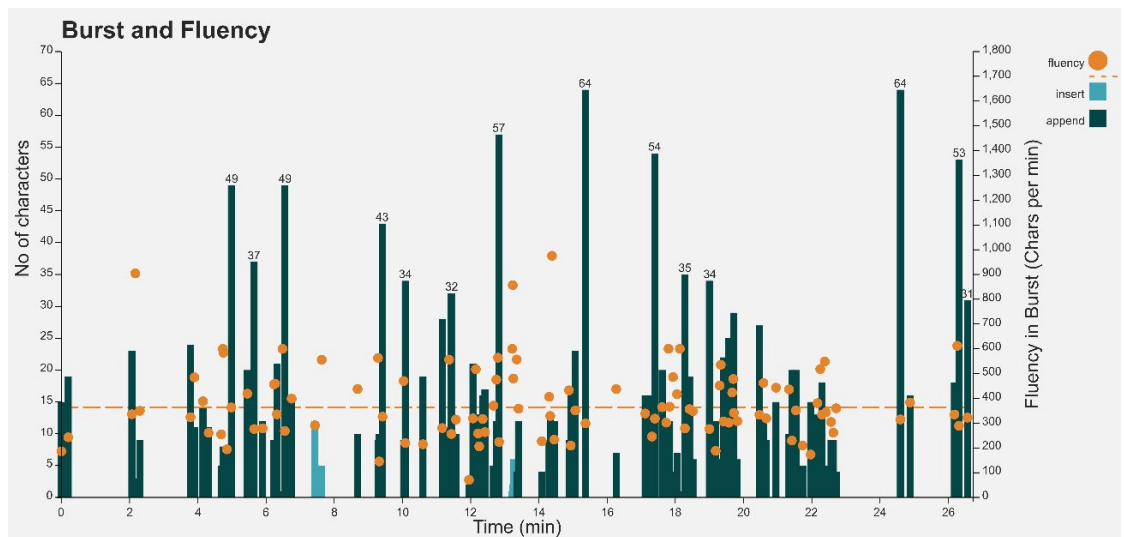
## KUVIO 8

Informantin ID2004 purskekuvio L2-ruotsissa (asteikko 70 merkkiä).



## KUVIO 9

Informantin ID2024 purskekuvio L2-ruotsissa (asteikko 70 merkkiä).



keiden pituuksissa että tuotostiheydessä. Asteikko on skaalattu molemmissa kuvioissa 70 merkkiin vertailtavuuden parantamiseksi. Informantti ID2004 on tuottanut tekstiään lyhyiden purskeiden (ka = 9,28) ryppäissä, kun taas informantti ID2024 tuottaa tekstiään prosessin aikana tasaisemmin pitkinä purskeina (ka = 15,6). Lisäksi kuvioista käy ilmi, että informantti ID2024 on myös keskimäärin tehokkaampi tekstiä tuottaessaan (ks. kuvioiden oranssit katkoviivat).

Taustakyselyssä informantti ID2004 kertoi, ettei kirjoittaminen ole hänelle helppoa eikä nopeaa. Tämä saattaa olla yksi tekijä lyhyiden purskeiden taustalla. Hän ei myöskään suunnittele tekstiään etukäteen:

*Kirjoittaminen ei ole minulle helppoa eikä nopeaa. Korjaan tekstejäni liian vähän. Aloitan kirjoittamisen lähes aina suunnittelematta, koska koen, että suunnitelmasta ei ole mitään hyötyä.*

Suunnittelu prosessin aikana näkyy myös kuvioista. Suunnittelutauot näkyvät purskeiden tuotostiheydessä, ja informantti ID2004 lisää tekstiä jo tuotetun tekstin sisään (vaaleat pylväät, ks. alaluku 3.2).

Informantin ID2024 purskekuvio osoittaa pitkien purskeiden tuottamista tasaisesti L2-ruotsissa. Tästä huolimatta myös hän tuottaa tekstiä tavallisesti hitaasti eikä suunnittele tekstiä etukäteen taustakyselyn mukaan:

*En oikein pidä kirjoittamisesta eikä se ole minulle luontaista. Minulla kestää kauan saada tekstiä aikaiseksi. En useimmiten suunnittele tekstejäni tarkasti etukäteen.*

Taustakyselyistä käy ilmi, että molemmilla informanteilla on samankaltainen näkemys omasta kirjoitusprofiilista. Kuvioita 8 ja 9 vertaamalla ero purskeiden tuottamisessa on kuitenkin selvä. Vaikka informantti ID2024 ei suunnittele tekstiään etukäteen, hän tuottaa pitkiä purskeita tasaisesti.

Kuvioissa 10, 11 ja 12 vertaamme informantin ID2002 tuottamia purskeita L1-suomessa, L2-englannissa ja L2-ruotsissa, koska laskennallinen datan mukaan informantin tuotosprosessit eroavat selvästi L1-kielellä ja L2-kielillä kirjoitettaessa.

Kuvioista käy ilmi, että informantin ID2002 tuotosprosessi L1-kielessä eroaa selvästi L2-kielistä. Asteikko on skaalattu kaikissa kuvioissa 70 merkkiin vertailtavuuden parantamiseksi. L1-kie-

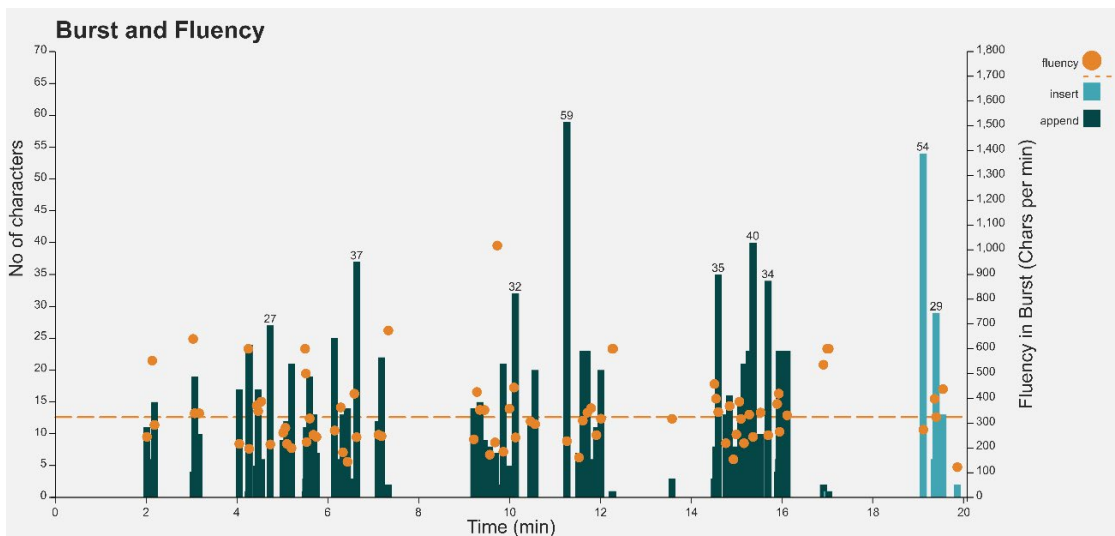
lessä purskeet ovat pidempiä (keskimääräinen purskeen pituus 13,67 merkkiä) kuin L2-kielissä (englannissa keskimääräinen purskeen pituus 9,50 ja ruotsissa 9,41 merkkiä). Tekstiä on tuotettu L1-prosessin alkuvaiheessa yksinomaan tekstin perään (tummat pylväät, ks. alaluku 3.2), vasta prosessin lopussa informantti ID2002 on tehnyt lisäyksiä tekstin sisälle (vaaleat pylväät). Teksti on rakentunut tuotosryppäinä, ja ryppäiden välillä on paljon taukoa. Tehokkuus on matalampi kuin L2-kielissä. Vaikka ero keskimääräisessä purskeen pituudessa L2-englannin ja L2-ruotsin välillä on pieni, voidaan kuvioita 11 ja 12 vertaamalla huomata selviä eroja tuotosprosesseissa. Englannin kielellä informantti ID2002 tuottaa purskeita enemmän ja tuotos on tiheämpää ja tehokkaampaa, kun taas ruotsinkielisessä tekstissä purskeiden tuottaminen on harvempaa eikä yhtä tehokasta.

Informantin ID2002 purskekuvioita vertailemalla voi nähdä, että purskeiden tuotostiheys on harvempaa L1-kielessä kuin L2-kielissä. Haastattelussa kävikin ilmi, että L2-kielissä kirjoittaminen tuntui informantista helpommalta, kun taas L1-kielellä hänellä oli vaikeuksia tuottaa tekstiä. Vaikeudet tuottamisessa johtuivat mahdollisesti jännittämisestä:

*No musta tuntuu että mulla meni se suomen kielen teksti kaikista huonoiten vaikka se on mun äidinkieli ja en sitten tiedä, että johtuuko se siitä, että se oli ensimmäinen --- jotenkin jännitti siinä --- ei*

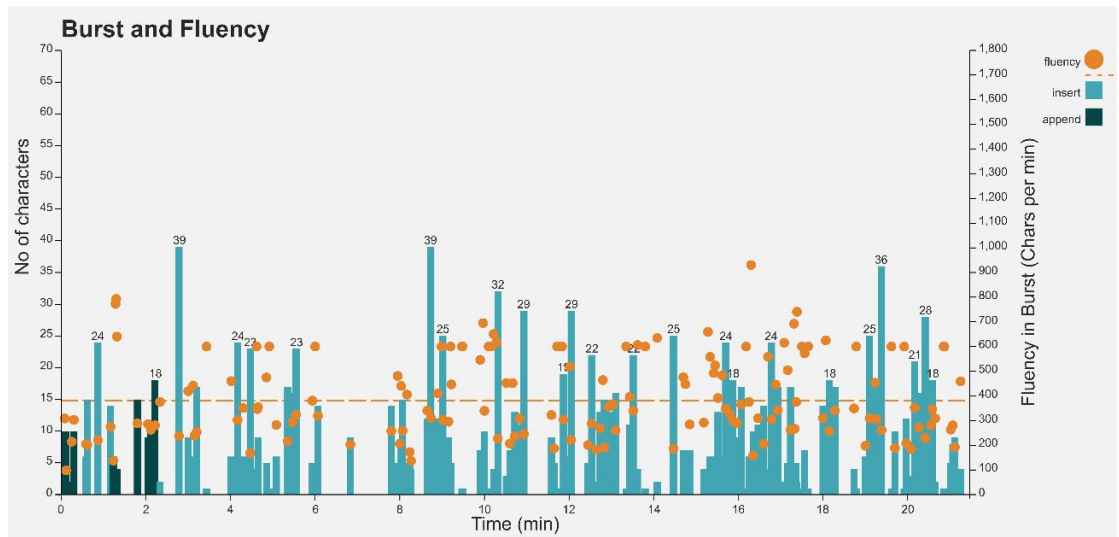
KUVIO 10

Informantin ID2002 purskekuvio L1-suomessa (asteikko 70 merkkiä).



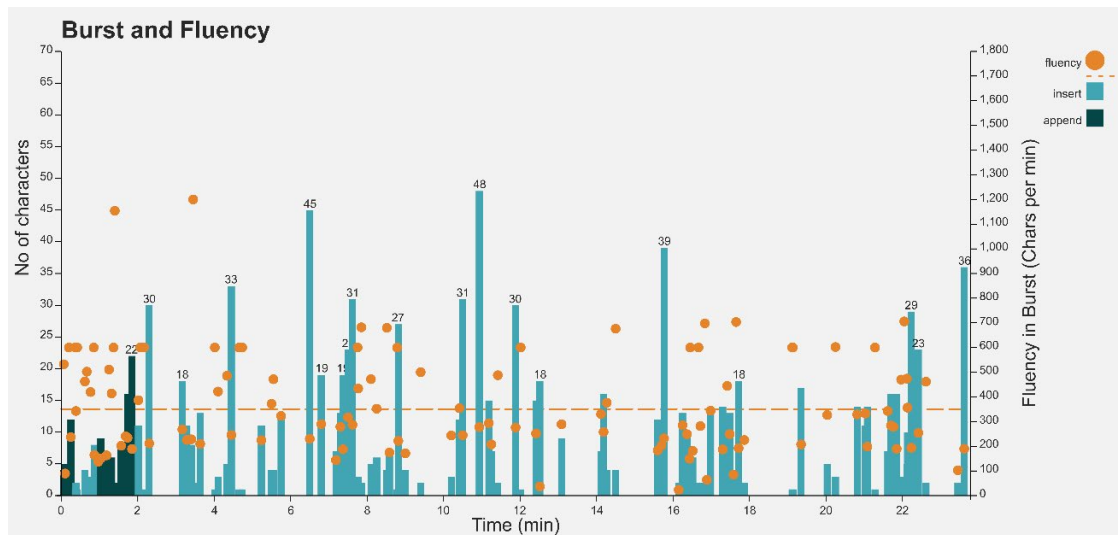
## KUVIO 11

Informantin ID2002 purskekuvio L2-englannissa (asteikko 70 merkkiä).



## KUVIO 12

Informantin ID2002 purskekuvio L2-ruotsissa (asteikko 70 merkkiä).



*tullut mulle tota ideotakaan silleen tai ajatus ei liikkunut samalla tavalla kuin kahessa muussa.*

## 5 JOHTOPÄÄTÖKSET JA POHDINTA

Tässä tutkimuksessa tavoitteenamme oli selvittää, miten kirjoitusprosessista näppäilyntallennusohjelmalla tallennettua laskennallista dataa, siihen perustuvaa visualisointia sekä taustakyselyjen ja jälkikäteishaastattelujen avulla saatavaa tietoa

yhdistämällä voi havaita ja havainnollistaa eroja kirjoitussujuvuudessa. Analysoimme samojen kirjoittajien taukokäyttätymistä ja purskeiden pituuksien sekä tiheyden ja tehokkuuden vaihtelua osana ensikielen ja vieraiden kielten kirjoitusprosesseja. Tutkimuksen keskiössä olivat sekä kielten että yksilöiden väliset erot. Tutkimuksen aineisto koostui GGXLog-näppäilyntallennusohjelmalla tallennetuista L1-suomen, L2-englannin

ja L2-ruotsin kirjoitusprosesseista. Näiden lisäksi analysoimme ohjelman tuottamia kuvioita sekä taustakyselyjä ja haastatteluja.

Ryhmätason tulokset osoittivat, että L1-kielessä kognitiiviset tauot ovat keskimääräisesti pidempiä kuin L2-kielissä. Kuten aiemmassakin tutkimuksessa on todettu, keskimääräinen tauonpituus voi olla ensikielellä kirjoitettaessa pidempi kuin vieraalla kielellä kirjoitettaessa (Lindström & Zetterholm, 2019). Kognitiivisten taukojen osuus kirjoitusprosessista on kuitenkin ensikielessä pienempi kuin vieraissa kielissä, eli L1-kirjoitusprosessin aikana aikaa käytetään enemmän tuottamiseen kuin taukoihin (ks. myös Spelman Miller, 2000).

Yksilötason analyysi osoitti selviä eroja taukojen pituudessa, osuudessa ja sijainnissa sekä kielten että yksilöiden välillä. Näitä eroja voivat selittää esimerkiksi erot kirjoitusprofiileissa, mutta myös välineen hallinnassa (Van Waes & Schellens, 2003; Van Waes, Leijten, Roeser, Olive & Grabowski, 2021). Erot eivät kuitenkaan ole kovin selviä esimerkiksi sanansisäisissä tauoissa, sillä sanansisäisiä taukoja on aineistossa pääosin vähän. Tähän saattaa vaikuttaa kirjoittajien suhteellisen korkea taitotaso L2-kielissä. Sanansisäisiä taukoja esiintyy usein enemmän alemmilla taitotasoilla (Mutta, 2020; Spelman Miller, 2000, 2006). Navigointitaukojen osalta analyysi osoitti mielenkiintoisia yksilöllisiä eroja tavoissa rakentaa tekstiä. On yksilöllistä, miten paljon kirjoittaja muokkaa tekstiään prosessin aikana. Muokkaaminen prosessin aikana saattaa näyttäytyä navigointitaukojen vähäisenä määränä, kun taas loppuvaiheen muokkaaminen tuottaa paljon navigointitaukoja.

Tuotosprosessin ja lopullisen tuotoksen välinen suhde ei varsinkaan kirjoituksessa ole yksiselitteinen. Toisin kuin puheessa, jossa viestintä tapahtuu usein reaaliajassa, kirjoitusprosessin aikana esiintyvät tauot eivät haittaa ymmärrystä tai vaikuta negatiivisesti havaittuun sujuvuuteen. Kirjoitettaessa on mahdollista suunnitella ja muokata välitettävää viestiä sujuvammaksi lukijan näkökulmasta ilman että viestin vastaanottaja havaitsee esimerkiksi prosessoinnin aiheuttamia taukoja. Tauot ja muokkaukset ovat usein kirjoituksessa tarpeellisia, jotta lopullinen tuotos on laadukas. Kirjoitussujuvuuden tutkimuksissa perinteiset nopeuteen ja taukoihin perustuvat mittarit saattavat siis osoittaa prosessin ja lopullisen tuotoksen välille erilaisen suhteen kuin suullisessa tuotok-

sessä. Prosessin ja tuotoksen välistä suhdetta on kuitenkin tärkeä tutkia, jotta voidaan paremmin ymmärtää prosessia tuotoksen takana. (Lintunen ym., 2020; Palviainen et al., 2012.)

Purskeiden osalta tuloksemme tukevat aiemman tutkimuksen tuloksia siitä, että kognitiivisesti vähemmän kuormittava tuotosprosessi L1-kielessä mahdollistaa pidempien purskeiden tuottamisen ja että kielitaito vaikuttaa purskeen pituuteen (esim. Chenoweth & Hayes, 2001, 2003; Gunnarsson-Largy, 2012; Hayes & Chenoweth, 2007; Mutta, 2020; Palviainen ym., 2012). Tuloksemme osoittavat, että purskeet ovat pisimpiä L1-suomessa. L2-kielten välillä purskeiden pituudessa on vain pieni ero purskeiden ollessa aavistuksen pidempiä englannissa kuin ruotsissa. Ryhmätasolla tähän saattaa vaikuttaa kirjoittajien suhteellisen korkea kielitaitotaso. Tulokset viittaavat toisaalta myös siihen, että tuotettujen purskeiden pituus on jossain määrin yksilöllinen. Tämä saattaa johtua yksilöiden välisistä eroista esimerkiksi välineen hallinnassa (Van Waes ym., 2021). Toisaalta yksilölliset erot ovat suurempia L2-kielissä kuin L1-kielellä kirjoitettaessa, mikä saattaa johtua erosta ensi- ja vieraalla kielellä kirjoittamisen välillä (Schoonen ym., 2009; Spelman Miller, 2006).

Visualisointeja vertaamalla voidaan tarkastella eroja kirjoitussujuvuudessa monipuolisemmin kuin esimerkiksi pelkästään laskennallista dataa tutkimalla. Taukojen osalta erot näkyvät visualisoinneissa selvimmin taukojen jakautumisessa tuotosprosessin aikana. Tällainen taukoanalyysi tuo uuden näkökulman kirjoittajien taukokäyttäytymisen vertailuun, sillä kuvio näyttää, missä tuotosprosessin kohdassa eripituisia taukoja esiintyy ja miten kirjoittaja navigoi tekstissä. Visualisointi syventää tuotosprosessin analyysia mahdollistamalla purskeiden tuostiheyden ja tehokkuuden tarkastelun purskeiden pituuden lisäksi. Visualisointi auttaa näin ollen ymmärtämään ja hahmottamaan kirjoittamista eri näkökulmista ja täydentää laskennallisen datan tulkintaa (vrt. Leblay ym., 2022).

Jälkikäteishaastattelut ja taustakyselyt täydensivät visualisointien ohella laskennallista dataa ja antoivat lisätietoa sujuvuuserojen taustalla vaikuttavista syistä kirjoittajan näkökulmasta. Jälkikäteishaastatteluisia ja taustakyselyissä nousi esiin sekä ulkoisia että sisäisiä prosessin sujuvuuteen vaikuttavia tekijöitä (Lindström &

Zetterholm, 2019; Lintunen ym., 2020; Makkonen-Craig, 2011; Segalowitz, 2016; Van Waes & Schellens, 2003). Keskeisimpiä teemoja olivat oma kielitaito, tehtävän aihe, kielen viimeaikainen käyttäminen, kielten väliset erot ja kielten mahdollinen sekoittuminen, tehtävän aiheuttama ahdistus tai jännitys, yleinen suhtautuminen kirjoittamiseen sekä oma kirjoitusprofiili, esimerkiksi tekstin suunnittelu ja korjaaminen. Nämä havainnot vahvistavat aiemmassa tutkimuksessa osoitettua tarvetta tutkimuksille, joissa samalta henkilöltä verrataan sujuvuusprofiilia useammassa kielessä (ks. Lintunen ym., 2020). Taustakyselyt ja haastattelut täydentävät kirjoitusprosessista saatua tietoa, vaikkakin tämä tieto pohjautuu koehenkilön subjektiiviseen tulkintaan omasta kirjoittamisesta.

Tässä tutkimuksessa käytetyt menetelmät täydentävät toisiaan hyvin. Tarkemman kognitiivisen prosessoinnin ja prosessin aikana tapahtuvien toimintojen välistä suhdetta voidaan jatkossa tarkastella yksityiskohtaisemmin lyhyemmässä tuotoksessa, keskittymällä prosessin tiettyyn osaan tai esimerkiksi silmänliiketutkimuksen avulla (ks. esim. Révész ym., 2019).

## 6 LOPUKSI

Tämän tutkimuksen puitteissa oli mahdollista analysoida rajattu määrä aineistoa. Tuloksia tulkittaessa on huomioitava, että kyseessä on vain tähän otantaan perustuvat havainnot, eikä

ääritapausten välille ole voitu osoittaa tilastollista eroa. Jatkossa laajennamme tutkimustamme niin, että myös tilastollisia analyysejä on tarkoituksenmukaista soveltaa osaksi hankkeen tulevaa tutkimusta. On myös huomioitava, että tutkimuksessa esitetyt tulokset kirjoitussujuvuudesta perustuvat valittuihin mittareihin. Tutkimuksessa ei ole kontrolloitu kirjoitusvälineen hallintaa, mutta jatkossa myös näppäilynopeus huomioidaan analyysissa. Visualisointien tulkinta ei myöskään ole täysin haasteetonta, sillä esimerkiksi asteikon mukautuminen pisimmän tauon mukaan vaikeuttaa taukojen pituuden vertailua. Näistä rajoituksista huolimatta tämä tutkimus toimii hyvänä lähtökohtana valittujen menetelmien yhdistämiselle kirjoitussujuvuuden tutkimuksessa.

Kirjoitussujuvuus on monitahoinen ilmiö, jonka tutkimiseen vaaditaan monipuoliset menetelmät (ks. myös Palviainen ym., 2012). Kirjoitusprosessin kokonaisvaltaisessa tarkastelussa visualisointi esimerkiksi tauko- ja purskekuvioiden avulla on hyödyllistä. Taustakyselyjen ja haastattelujen avulla saadun tiedon soveltaminen tutkimukseen ja opetukseen tukee kompleksisen ilmiön ymmärtämistä kirjoittajan näkökulmasta (Vieira, Parsons & Byrd, 2018; Wengelin & Johansson, 2023). Kirjoitusprosessin tarkastelu ja sanoittaminen voivat auttaa kirjoittajaa, tutkijoita ja opettajia hahmottamaan kirjoittamisen taustalla vaikuttavia tekijöitä entistä monipuolisemmin.

## LÄHTEET

- Baaijen, V. M. & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, 36(3), 199–223. <https://doi.org/10.1080/07370008.2018.1456431>
- Bécotte-Boutin, H.-S., Caporossi, G., Leblay, C. & Hertz, A. (2019). Writing and rewriting: Keystroke logging's colored numerical visualization. Teoksessa E. Lindgren & K. Sullivan (toim.), *Observing writing: Insights from keystroke logging and handwriting*, (s. 96–124). Brill Academic Publishers.
- Bowles, M. A. (2018). Introspective verbal reports: Think-alouds and stimulated recall. Teoksessa A. Phakiti, P. De Costa, L. Plonsky & S. Starfield (toim.), *The Palgrave handbook of applied linguistics research methodology*, (s. 339–357). Palgrave Macmillan. [https://doi.org/10.1057/978-1-137-59900-1\\_16](https://doi.org/10.1057/978-1-137-59900-1_16)
- Caporossi, G. & Leblay, C. (2015). A graph theory approach to online writing data visualization. Teoksessa G. Cisarlu (toim.), *Writing(s) at the crossroads: The process-product interface*, (s. 171–181). John Benjamin.

- CEFR (2003) = *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Euroopan neuvosto 2003. Eurooppalainen viitekehys: kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys. WSOY.
- Chenoweth, A. & Hayes, J. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, 18, 80–98.
- Chenoweth, A. & Hayes, J. (2003). The inner voice in writing. *Written Communication*, 20, 99–118.
- Cislaru, G. & Olive, T. (2018). *Le processus de textualisation: Analyse des unités linguistiques de performance écrite*. De Boeck Supérieur. <https://doi.org/10.3917/dbu.cisla.2018.01>
- Conijn, R., Dux Speltz, E. & Chukharev-Hudilainen, E. (2021). Automated extraction of revision events from keystroke data. *Reading and Writing*. <https://doi.org/10.1007/s11145-021-10222-w>
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Dörnyei, Z. & Ushioda, E. (2009). *Motivation, language identity and the L2 self*. Multilingual matters. <https://doi.org/10.21832/9781847691293>
- Ellis, R. & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26(1), 59–84.
- Grönroos, M. (2004). *Johdatus tilastotieteeseen. Kuvailu, mallit ja päättely*. Finn Lectura.
- Gunnarsson-Largy, C. (2012). The development of complexity, accuracy and fluency in the written production of L2 French. Teoksessa A. Housen, F. Kuiken & I. Vedder (toim.), *Dimensions of L2 performance and proficiency. Complexity, accuracy and fluency in SLA*, (s. 247–276). John Benjamins, <https://doi.org/10.1075/llt.32.11gun>
- Hayes, J. (2009). From idea to text. Teoksessa R. Beard, D. Myhill, M. Nystrand & J. Riley (toim.), *The SAGE handbook of writing development*, (s. 65–79). SAGE.
- Hayes, J. & Flower, L. S. (1980). Identifying the organization of writing processes. Teoksessa L.W. Gregg & E. R. Steinberg (toim.), *Cognitive processes in writing*, (s. 3–30).
- Hayes, J. & Chenoweth, A. (2007). Working memory in an editing task. *Written Communication*, 24, 283–294. <https://doi.org/10.1177/074108830730482>
- Jacques, T., Alves, R. A., Fadaei, S. & Barbosa, F. (2020). Realtime psychophysiological and writing correlates of expressive writing. *Experimental Psychology*, 67(4), 237–245. <https://doi.org/10.1027/1618-3169/a000495>
- Leblay, C., Usoof, H. & Caporossi, G. (2022). Kirjoittamisprosessin dynaaminen visualisointi. *Kieli, koulutus ja yhteiskunta*, 13(3).
- Lindgren, E. & Sullivan, K. (2019). *Observing writing: Insights from keystroke logging and handwriting*. Studies in Writing 38. Brill. <https://doi.org/10.1163/9789004392526>
- Lindström, E. & Zetterholm, E. (2019). Tangenttedslag och pauser i flerspråkiga elevers skrivprocess. *RASK*, 49, 67–86.
- Lintunen, P., Mutta, M. & Peltonen, P. (2020). Sujuvuustutkimuksen käännteitä: katsaus sujuvuustutkimuksen menetelmiin. Teoksessa S. Grasz, T. Keisanen, F. Oloff, M. Rauniomaa, I. Rautiainen & M. Siromaa (toim.), *Menetelmällisiä käännteitä soveltavassa kielentutkimuksessa – Methodological Turns in Applied Language Studies*. AFinLAN vuosikirja 2020. Suomen soveltavan kielitieteen yhdistyksen julkaisuja 78, (s. 174–197).
- Makkonen-Craig, H. (2011). Kirjoittajan kompetenssit ja äidinkielellä kirjoittaminen. Teoksessa A. Kauppinen, H. Lehti-Eklund, H. Makkonen-Craig & R. Juvonen (toim.), *Lukiolaisten äidinkieli. Suomen- ja ruotsinkielisten lukioiden opiskelijoiden teksti-maailmat ja kirjoitustaitojen arviointi*, (s. 62–88). Suomalaisen Kirjallisuuden Seura.
- Mutta, M. (2017). Pausal behavior in the writing processes of foreign and native language writers: The importance of defining the individual pause length. Teoksessa S. Plane, C. Bazerman, F. Rondelli, C. Donahue, A. N. Applebee, C. Boré, P. Carlino, M. Marquilló Larray, P. Rogers & D. Russell (toim.), *Research on writing: Multiple perspectives. International exchanges on the study of writing*, (s. 511–529). WAC Clearinghouse and CREM. <https://doi.org/10.37514/INT-B.2017.0919.2.28>
- Mutta, M. (2020). L2 fluency and writer profiles. Teoksessa P. Lintunen, M. Mutta & P. Peltonen (toim.), *Fluency in L2 learning and use*, (s. 63–80). Multilingual Matters.

- Mutta, M. & Laine, P. (2022). Tapaustutkimus monikielisten kirjoittajien kirjoitusprosesseista: Näppäilyntallennus apuna sanastokeskittymien tunnistamisessa. *Lähivõrdlusi. Lähi-vertailuja*, 32, 120–149.
- Mutta, M., Laine, P., Mäkilä, M., Åberg, A-M., Hintikka, A., Mäntylä, K., Toropainen, O. & Lahtinen, S. (2022). Kirjoittamisprosessin tallentaminen ja visualisointi kielenoppimisen apuna. *Kieli, koulutus ja yhteiskunta*, 13(6). <https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-marraskuu-2022/kirjoittamisprosessin-tallentaminen-ja-visualisointi-kielenoppimisen-apuna/>
- Mutta, M., Laine, P., Åberg, A-M. & Mäkilä, M. (tulossa). *Visualising multilingual writers bursts and profiles in the initial writing phase*.
- Olive, T. (2002). La gestion en temps réel de la production verbale : méthodes et données. Teoksessa M. Fayol (toim.), *Production du langage*, (s. 131-147). Hermès Science Publications-Lavoisier.
- Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, 6 (2), 173–194. <https://doi.org/10.17239/jowr-2014.06.02.4>
- Palviainen, Å., Kalaja, P. & Mäntylä, K. (2012). Development of L 2 writing fluency and proficiency. *AFinLA E Soveltavan Kielitieteen Tutkimuksia*, 4, 47–59.
- Révész, A., Michel, M. & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors. A mixed-methods study. *Studies in Second Language Acquisition*, 41(3), 605–631. doi:10.1017/S027226311900024X
- Roca de Larios, J., Coyle, Y. & Nicolás-Conesa, F. (2016). Focus on writers: Processes and strategies. Teoksessa R. M. Manchón & P. K. Matsuda (toim.), *Handbook of second and foreign language writing*, (s. 267–286). De Gruyter Mouton.
- Schoonen, R., Snellings, P., Stevenson, M. & van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. Teoksessa R. Manchón (toim.), *Writing in foreign language contexts: Learning, teaching, and research*, (s. 77–101). Multilingual Matters. <https://doi.org/10.21832/9781847691859-007>
- Segalowitz, N. (2010). *The cognitive bases of second language fluency*. Routledge.
- Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching*, 54(2), 79–95.
- Spelman Miller, K. (2000). Academic writers on-line: Investigating pausing in the production of text. *Language Teaching Research*, 4(2), 123–148.
- Spelman Miller, K. (2006). Pausing, productivity and the processing of topic in online writing. Teoksessa K. Sullivan & E. Lindgren (toim.), *Computer key-stroke logging and writing*, (s. 131–156). Studies in Writing 18. Brill. [https://doi.org/10.1163/9780080460932\\_009](https://doi.org/10.1163/9780080460932_009)
- Spelman Miller, K., Lindgren, E. & Sullivan, K. (2008). The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke logging. *TESOL Quarterly*, 42(3), 433–454.
- Tuomi, J. & Sarajärvi, A. (2018). *Laadullinen tutkimus ja sisällönanalyysi*. Tammi.
- Usoof, H., Leblay, C. & Caporossi, G. (2020). *GenoGraphiX-Log Version 2.0 User Guide*. Les Cahiers Du GERAD G-2020-68, 1–63.
- Van Waes, L. & Schellens, P. J. (2003). Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics*, 35, 829–853.
- Van Waes, L., Leijten, M., Roeser, J., Olive, T. & Grabowski, J. (2021). Measuring and assessing typing skills in writing research. *Journal of Writing Research*, 13(1), 107–153. <https://doi.org/10.17239/jowr-2021.13.01.04>
- Vieira, C., Parsons, P. & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers and Education*, 122, 119–135. <https://doi.org/10.1016/j.compedu.2018.03.018>
- Wengelin, Å. & Johansson, V. (2023). Investigating writing processes with keystroke logging. Teoksessa O. Kruse, C. Rapp, C. M. Anson, K. Benetos, E. Cotos, A. Devitt & A. Shibani (toim.), *Digital writing technologies in higher education*, (s. 405–420). Springer, Cham. [https://doi.org/10.1007/978-3-031-36033-6\\_25](https://doi.org/10.1007/978-3-031-36033-6_25)

# Visualisation and retrospective interviews in analysing writing fluency

- Anne-Maj Åberg, University of Turku
- Mari Mäkilä, University of Turku
- Maarit Mutta, University of Turku

This study examines writing fluency in three languages. The aim is to explore how writing processes can be studied in more detail by using visualisations, background forms and retrospective interviews in addition to quantitative data. The data are based on six university students' writing processes in L1 Finnish, L2 Swedish and L2 English and are gathered with the keystroke-logging programme GenoGraphiX-Log that contains tools for visualising writing processes. The results indicate that writing fluency differs not only regarding pauses and bursts but also other features, and that visualisation is a helpful tool in studying these differences. According to the background forms and interviews, writing in L1 is considered to be relatively easy, although narrowing down the subject, for instance, may cause problems. Regarding writing fluency in L2, the assignment and recent use of the languages seem to be important factors.

**Keywords:** fluency, keystroke logging, retrospective interview, visualisation, writing process