Fluency across modes: An exploratory study of L1 and L2 spoken and written fluency

Maarit Mutta*, Pauliina Peltonen, Päivi Laine and Pekka Lintunen

***Corresponding author: Maarit Mutta**, School of Languages and Translation Studies, University of Turku, FI-20014 University of Turku, Finland, E-mail: maamut@utu.fi. https://orcid.org/0000-0003-1227-1484
Pauliina Peltonen, School of Languages and Translation Studies, University of Turku, FI-20014 University of Turku, Finland, E-mail: paupelt@utu.fi https://orcid.org/0000-0002-8743-0032
Päivi Laine, School of Languages and Translation Studies, University of Turku, FI-20014 University of Turku, Finland, E-mail: paipel@utu.fi https://orcid.org/0000-0001-7664-2091
Pekka Lintunen, School of Languages and Translation Studies, University of Turku, FI-20014 University of Turku, Finland, E-mail: peklin@utu.fi https://orcid.org/0000-0003-3448-9310

Fluency across modes: An exploratory study of L1 and L2 spoken and written fluency

Abstract: The article presents an exploratory cross-modal analysis of fluency profiles in spoken and written first (L1, Finnish) and second (L2, English) language production of the same language users. Our data come from two research projects, from which we identified 11 university students participating in both. The spoken tasks consisted of monologue picture description (analysed with Praat), and the written tasks were short argumentative essays (collected and analysed with keystroke logging software GGXLog). Based on commonly used measures to capture different aspects of spoken and written fluency, we used a set of 14 measures (seven for speech fluency, seven for writing fluency) to examine fluency across modes comprehensively. Four profiles were identified from the data: 1) fast and productive, 2) fast, 3) slow and productive, and 4) slow and reflective. Six speakers had the same profile in the L1 and L2, and seven writers had the same profile in the L1 and L2. Only one participant had the same profile in the L1 and L2 speaking and writing. The results suggest that the cross-modal differences are greater than the differences between languages. The modalities are inherently different, which is also reflected in individual variation between the modalities.

Keywords: fluency profiles, speaking, writing, language proficiency

1 Introduction

Fluency is a multifaceted concept used in cognitive sciences and second language (L2) teaching and learning research. It is also an important concept to all teachers and professionals working with language assessment, in addition to language users themselves. Fluency commonly refers to uninterrupted language production which progresses smoothly and with ease, but every researcher has to define the concept in their specific framework, which might present challenges when comparing the methods and results of studies conducted in various contexts. In general, L2 fluency studies related to language production focus either on spoken or written fluency, while cross-modal examinations are rare. In studies of speech fluency, the definitions can range from a broad understanding of fluency as language proficiency to a narrow understanding of fluency mainly as a temporal phenomenon measurable from speech samples (e.g., Lennon 1990; Tavakoli and Hunter 2018). Fluency in a narrow sense is generally used for research purposes, and recent studies largely

rely on Skehan's (2014) three-fold framework of utterance fluency, namely speed, breakdown, and repair fluency, where measures of fluency and disfluency are placed on a continuum. Recently, writing fluency has also received increasing attention (Lintunen et al. 2020b). It has been studied, for instance, by regarding how coherent, error-free, and well-structured a text is (Crossley et al. 2016). Furthermore, keystroke logging software allows the study of fluency during the writing process (e.g., Lindgren and Sullivan 2019) and, thus, enables combining measures for the writing process and the final product (Cislaru and Olive 2018; Mutta 2020). As with spoken fluency, fluency features during the writing process, including disfluency features, such as pauses and revisions (Baaijen and Galbraith 2018; Cislaru and Olive 2018; Ellis and Yuan 2004), have been investigated. In the present study, we focus on fluency-related features that can be identified and measured from both written and spoken first language (L1) and L2 productions.

Both written and spoken productions can be affected by individual variation: in speech, language users' speaking style varies according to individual characteristics, and during the writing process, there are corresponding differences and similarities related, for instance, to language proficiency (e.g., De Jong 2018; Olkkonen & Mutta 2020; Palviainen et al. 2012; Peltonen 2018). Language production in different modes might also be affected by other user-internal (e.g., personality and affective variables, such as emotions) and user-external (e.g., spontaneity of the production, power relations, test environment) factors (Lintunen et al. 2020b; Olkkonen et al. 2024). All these variables together, therefore, contribute to creating individual profiles for language users. Despite the importance of individual variation affecting L1 and L2 production across written and spoken modalities, previous cross-modal research on the topic is scarce. Thus, there is a gap in research regarding profiling language use from cross-modal perspectives to understand the complex picture of an individual's language use (Li et al. 2022). Furthermore, researchers have highlighted that there is a need for multiple-method studies (Révész et al. 2022: 2), especially focusing on language users' own linguistic profiles and skills in their L1 and L2 (e.g., Lintunen et al. 2020b). L2 studies have traditionally compared learners to native speakers, but comparing learner's production and fluency profile across different languages and modes in their own repertoire can reveal idiosyncratic patterns that are user-specific (Peltonen 2018).

To address these research gaps, we conducted a cross-modal examination of spoken and written productions from the same participants in their L1 Finnish and L2 English. Through our comparative analysis of fluency profiles across the two modalities, we provide a pioneering contribution to research on individual language users' spoken and written fluency profiles in their

L1 and L2. The analysis combined data from two large-scale research projects, one of which focused on spoken and the other on written fluency. From these data, 11 university students that participated in both projects were identified and their productions analysed for a range of spoken and written fluency measures.

2 Literature review

2.1 Spoken fluency: focus on utterance

L2 speech fluency has been of interest to researchers working in the domains of L2 learning, assessment, and teaching for several decades. Speech fluency is often approached as one dimension of broader L2 (oral) proficiency, such as in the complexity-accuracy-fluency (CAF) framework used for examining L2 proficiency and performance (e.g., Housen et al. 2012). Lennon (1990) referred to this sense of fluency as the "narrow" sense to distinguish it from a "broad" understanding of fluency as the level of general (oral) proficiency in a certain language. More recently, Tavakoli and Hunter (2018: 343) fine-tuned this division by distinguishing four approaches to fluency definitions, ranging from *very broad* (general proficiency) and *broad* (oral proficiency) to *narrow* (one dimension of oral proficiency associated with effortlessness and ease) and *very narrow* (the narrow sense specified to speed, breakdown, and repair dimensions, following e.g., Skehan 2014). Along with this distinction, Segalowitz' (2010) framework is widely cited in L2 speech fluency literature (see also Lintunen et al. 2020a). In his model (Segalowitz 2010), three dimensions of fluency can be distinguished: effectively functioning cognitive processes (*cognitive fluency*) enable the production of L2 speech, which can be analysed for various fluency-related features (e.g., based on the dimensions specified in the very narrow sense; *utterance fluency*). These features, in turn, create an impression of fluent speech for the listener (e.g., a rater in fluency assessment contexts; *perceived fluency*). In the present study, our focus for the spoken data is on utterance fluency and the very narrow sense, as we apply a set of temporal measures associated with the speed, breakdown, and repair dimensions to capture the participants' speech fluency in their L1 and L2 productions.

Various measures have been employed to examine L2 speech fluency. The speed dimension, referring to the rate of speech production, is most commonly measured with articulation rate (syllables per speaking time, excluding silent pauses) or its inverse measure, mean syllable duration

(e.g., De Jong 2018). Speech rate is another popular choice for capturing the speed of talk, but it is categorised as a composite measure in current literature (e.g., Tavakoli et al. 2020), because it includes information about pausing along with the pure speed of production (calculated as syllables per total time, including silent pauses). Due to the composite nature of this measure, (pruned) speech rate has been recommended as a good overall measure of fluency for studies where an overall estimate of fluency is sufficient (e.g., De Jong 2016). Along with speech rate, another composite measure, mean length of utterance (e.g., De Jong 2018) or mean length of run (e.g., Tavakoli et al. 2020), which combines information about pausing and the speed of talk, has been widely employed in fluency research, along with the present study. Regarding the breakdown fluency dimension, silent pauses (SPs) are usually examined for their frequency, duration, and/or location and distinguished from so-called filled pauses (FPs, e.g., *uh*, *um*). While especially more frequent mid-clause SPs have been associated with disfluency (e.g., Tavakoli et al. 2020), the link between FPs and fluency is less clear (e.g., Dumont 2018). In the present study, we only examined SPs, as they are more comparable with the cognitive pauses examined in the written production than FPs, which do not have an exact equivalent in written production. Finally, the repair dimension captures different types of corrections made to speech, such as false starts and reformulations. In the present study, we grouped false starts, repetitions, replacements, and reformulations (Foster and Skehan 1999) into one overall measure of repair fluency. This has been a relatively common practice in L2 speech fluency research due to, for instance, the low frequencies of repairs often found in speech samples and was adopted for the present study to maximise comparability with the writing fluency analyses.

The different dimensions of fluency, including the various fluency measures discussed above, have mostly been examined based on monologic L2 speech. Traditionally, L2 speech has been contrasted with native speakers' fluency, either implicitly or by measuring fluency-related features from L1 speech samples collected from a native speaker control group (e.g., Peltonen and Lintunen 2016). While this approach can provide important information about the learners' fluency level in relation to native speakers' fluency, recently, researchers have begun to draw attention to speakers' individual styles affecting L2 speech, resulting in research designs including L1 and L2 (even L3, see Peltonen and Lintunen 2022) from the same participants to enable within-participant examinations of fluency across two or more languages in a speaker's repertoire. In the past 10–20 years, the field has witnessed a rapid growth in research exploring these L1 personal speaking style effects on L2 speech fluency (for a recent meta-analysis, see Gao and Sun 2024), demonstrating heightened interest in individual variation regarding fluency-related phenomena. However, to our

knowledge, these examinations have been limited to either spoken or written fluency (see Section 2.2), while within-participant studies on cross-modal fluency have been rare (but see Lintunen and Mäkilä 2014 on syntactic complexity among L2 written and spoken production).

2.2 Written fluency: focus on process

Research on writing as a final product has a long history, while studies focusing on writing processes have become more popular in the last 20 years following the increased use of technological tools, such as keystroke logging (Gunnarsson-Largy 2012; Lindgren and Sullivan 2019; Mohsen 2021; Spelman Miller et al. 2008; Van Waes and Leijten 2015). Furthermore, writing processes have been more widely examined in L1 than L2. Therefore, knowledge on L1 writing processes and process-oriented models have influenced studies on L2 writing processes (Hayes 1996; Hayes and Flower 1980) to describe L2 writers' mental processes (e.g., Roca de Larios et al. 2016). Studying L2 writing processes can provide explicit information about language users' choices and strategies, and this information can be further used for pedagogical purposes, for instance, by examining similarities and differences between proficient and less proficient writers or between L1 and L2 writing habits or styles (Davoodifard 2022).

L2 writing fluency has also been the focus of more recent studies (Lintunen et al. 2020b). In writing fluency research, there are different measures for the writing process and the final product (Cislaru and Olive 2018; Kowal 2014; Mohsen 2021; Mutta 2020; Spelman Miller et al. 2008; Van Waes and Leijten 2015): in product-oriented studies, fluency illustrates the reader experience, that is, how the text flows from the reader's point of view (reader perception; comparable to perceived fluency, see Section 2.1) in terms of cohesion and linguistic correctness (Crossley et al. 2016). On the other hand, fluency in the process-oriented studies is related to the main cognitive processes (planning, formulation, and revision), which are dynamic and recursive in nature (Ellis and Yan 2004; Roca de Larios et al. 2016) and focus on the interplay between automatically processed and more controlled cognitive processes, which are reflected in writing fluency measures (Cislaru and Olive 2018; Ellis and Yuan 2004; Olkkonen and Mutta 2020; cf. utterance and cognitive fluency in Section 2.1). Automatic processes (i.e., lower-level processes, such as orthographic processing) allow more working memory space for more controlled processes (i.e., higher-level processes, such as strategic planning and revision) (Baaijen and Galbraith 2018; Cislaru and Olive 2018; Ellis and Yan 2004; Olkkonen and Mutta 2020).

The automaticity of cognitive processes can be manifested through bursts, which are uninterrupted flows of producing text (Chenoweth and Hayes 2001; Cislaru and Olive 2018; Kowal 2014). Bursts represent one measure of fluent writing, defined as linguistic material produced between cognitive pauses (≥2 secs), revisions, and other tool-related actions (Baaijen and Galbraith 2018; Chenoweth and Hayes 2001; Conijn et al. 2024). Bursts can include linguistic elements from single letters to several sentences. The longer and more frequent the bursts, the more fluent the writing process is assumed to be (Chenoweth and Hayes 2001). The fluent writing process can be interrupted by pauses and revisions, and when writing on a computer, by event transitions, including scrolling, mouse movement, and other operations that occur before actual text production (Baaijen and Galbraith 2018). Pauses are markers of spontaneous segmentation of the production flow and basic units to study cognitive activity: short pauses reflect automatised lower-level processes, whereas non-automatised lower-level and higher-level processes induce longer pauses (Cislaru and Olive 2018; Olkkonen and Mutta 2020). Pauses, revisions, and scrolling are usually considered disfluency features, but they do not always affect writing fluency similarly, as they might occur due to various reasons; for instance, pause length can be related to orthographic (e.g., use of shorter pauses) or content revisions and mental planning (e.g., longer pauses) (Mutta 2020; Xu 2018). Bursts, pauses, revisions, and scrollings are situated on a fluency–disfluency continuum and their exact position depends on the context.

In addition to bursts and pauses, various other measures can also be used when examining writing fluency. Compared to spoken language, where syllables per speaking or total time are often used as a fluency measure, the most general measure is to calculate the words produced per unit of composition time to give an average measure of overall fluency (Chenoweth and Hayes 2001; Ellis and Yuan 2004). To yield a more comprehensive understanding of writing processes, fluency has also been studied with more detailed measures and methods. Ellis and Yuan (2004), Gunnarsson-Largy (2012) and Spelman Miller et al. (2008) studied L2 productions by combining product and process-oriented approaches in the CAF framework. Interestingly, their measures reflect the same measures as in spoken studies. Ellis and Yuan (2004) studied the effects of planning on fluency (e.g., syllables per minute and number of disfluencies), complexity (e.g., the ratio of clauses to T-units, syntactic variety, and mean segmental type-token ratio), and accuracy (e.g., error-free clauses and correct verb forms) in a paper-and-pencil task in L2 English among 42 Chinese learners' written narratives. The results showed that pre-task planning affected fluency and syntactic variety, whereas accuracy was greater when there was an opportunity for unpressed on-line planning. By contrast, Gunnarsson-Largy (2012) analysed L2 French competence by means of a keystroke

logging program (*ScriptLog*) and video-filmed thinking-aloud protocols. For fluency, the measures were similar to Ellis and Yuan (2004). In her case study, fluency and syntactic complexity seemed not to be related when using product- and process-oriented measures (see also in Section 2.3). Similarly, Spelman Miller et al. (2008) studied the L2 English production of 17 L1 Swedish high school students over a 3-year period by using a computer program JEdit. They compared text quality variables (CAF) to fluency measures such as bursts, number of pauses, and revisions and found that all writers increased their fluency and made longer bursts, but they did not increase their text quality over time. There was a lot of individual variation, and the researchers suggested that L1 writing processes could explain some patterns.

For his part, Mohsen (2021) studied 30 writers' argumentative text in L1 Arabic and L2 English with a keystroke logging program (Inputlog) and compared writers' writing behaviour (fluency, pauses, and revision) with text quality (six-point scale based on, for instance, organisation and development, coherence, and language use). He combined log files, screened video recordings, and stimulated recall interviews in a mixed-method analysis and found that writers were more fluent in their L1 than their L2 (i.e., more characters, fewer pauses at word boundaries; Van Waes & Leijten 2015) and needed time in the initial phase of writing to generate ideas. Interestingly there was no significant difference in time spent on revisions in L1 and L2. His results also showed that the fluency indicators could predict a good quality of L1 writing.

Examining a broader range of L2s, Van Waes and Leijten (2015) created a model to describe the complexity of fluency based on process-oriented characteristics. They collected data through keystroke logging (Inputlog) among 68 students writing in L1 Dutch and in L2 (English, French, Spanish, or German at level B2 or above according to CEFR scale (Council of Europe 2001)) to make a principal component analysis and identify the main variables of fluency. They identified four main components and corresponding fluency variables as follows: production (mean number of characters), process variance (standard deviation of characters), revision (process vs. product ratio and length of R-bursts), and pause behaviour (mean pause length and proportion of pause time). They concluded that there is a need for a standardised analysis procedure and for a multidimensional perspective of fluency to understand the complexity of fluency.

More recently, Révész et al. (2022) used a mixed-methods approach to study the relationship between proficiency and speed fluency, pausing, and eye-gaze behaviours in L2 writing in English by 60 Chinese participants. They measured speed fluency with two measures: the number of

characters typed between pauses (≥2 secs) and mean duration of character production (i.e., total time spent on writing excluding pauses divided by the number of characters), whereas pausing behaviour was examined in terms of pause frequency and length. For eye-gaze behaviours, they relied on, for instance, total fixation count and mean fixation length (Révész et al. 2022: 6). Their results showed that proficiency was the strongest predictor of speed fluency. For our analysis using a cross-modal approach, we have chosen measures from product- and process-oriented approaches to be able to compare the measures with spoken fluency variables (for details, see Section 3.2).

2.3 Fluency profiles in L2 production

In previous L2 fluency studies, identifying profiles based on individual differences is rarer than in psychological and educational research (see Li et al. 2022; Sparks et al. 2012). However, those spoken fluency studies that have examined learner profiles have often related individual differences underlying fluency profiles to speech styles and the efficiency of cognitive processing (e.g., De Jong 2018; Olkkonen and Mutta 2020; Olkkonen et al. 2024; Peltonen 2018). In written language studies, there is sometimes a distinction between writing profiles describing structural variability and the organisation of processes, and writer profiles that are specific to individual writers, but on several occasions, these profiles are intertwined; writer profiles have often been studied based on writing processes and on certain product features (Van Waes 1992; see also Mutta 2020; Van Waes and Schellens 2003).

L2 speech fluency is affected by user-internal (e.g., affective variables like linguistic anxiety) and user-external variables (e.g., formal vs. informal context) (Lintunen et al. 2020b; Szyszka and Lintunen 2023). To recognise L2 fluency profiles, objective measures need to be defined (for measures, see Sections 2.1 and 2.2). In a recent study, Olkkonen et al. (2024) examined fluency profiles in relation to the cognitive fluency dimension. The participants were native Finnish-speaking advanced university level English users (n = 64). Their tasks and measures were (profile dimensions in parentheses): LexTale in L2: lexical decision (language proficiency); monologue in L1 and L2: speech rate, syllables / minute (speech fluency), false starts, reformulations, and repetitions / minute (disfluency); rapid word reading L1 and L2: percentage of correct answers (automaticity); Stroop task in L1 and L2: amount of interference (additional reaction time), number of repairs (control of attention). To identify the profiles, they conducted a cluster analysis. As a result, four fluency profile groups were identified: 1) Fast and fluent, 2) Accuracy-oriented, 3)

Proficient but disfluent, and 4) Less proficient but fast. Olkkonen et al. (2024) argue that identifying cognitive fluency profiles helps to understand mechanisms behind speech production.

Dumont (2018) also relied on a cluster analysis to examine utterance fluency to create non-native (French speakers of L2 English) and native (British English) speaker (dis-)fluency profiles in her thesis. She used a versatile set of measures to capture different aspects of utterance fluency, including breakdown measures (unfilled and filled pauses), composite measures capturing speed and breakdown (e.g., speech rate, mean length of run), repair measures (e.g., false starts and restarts), and other fluency-related phenomena, such as lengthenings and discourse markers. Dumont (2018) did not name the profiles, but the results show that temporal (dis)fluency measures are crucial for identifying each profile. For instance, the most fluent native and non-native speaker profiles shared very high temporal fluency (e.g., few unfilled pauses, high phonation-time ratio, and speech rate), but there were differences in the use of discourse markers and lengthening (Dumont 2018: 418–419). She also found that there was no relationship between the learners' fluency profiles and their assessed CEFR levels.

When examining fluency profiles in written production, studies often combine product-oriented (e.g., productivity in number of words and paragraphs, error-free sentences) and process-oriented measures (e.g., writing time, bursts, pausing behaviour, revisions, and corrections). For instance, Van Waes (1992) and Van Waes and Schellens (2003: 836–837) examined three aspects of writing, namely writing time and final product (e.g., total duration of the writing process, duration of each phase, number of words), pausing behaviour (number, duration, types of pauses) and revision behaviour (number, level, purpose, location, remoteness, and temporal location of revision). They identified five writer profiles: initial planners, fragmentary first-phase writers, second-phase writers, non-stop writers, and average writers (related to examined variables of the total group). Based on these studies, Mutta (2020) used 10 process-oriented and 6 product-oriented measures including evaluation of the final product by native speakers. The results showed that product-oriented fluency was related to values such as number of words, clauses, and sentences, the average length of sentences, the number of words in relation to the total writing time, type/token ratio (TTR, indicating lexical variation), and evaluation of the product by native teachers. However, there was no significant correlation between measures of the process and the end product and, therefore, no direct link between the writer profiles and the evaluation (evaluated proficiency level).

In the present study, the aim is to fill a gap regarding cross-modal fluency research by examining the fluency profiles of university students that produced spoken and written samples in their L1 (Finnish) and L2 (English). We compare the participants' fluency profiles in their L1 and L2 writing and speech to reveal idiosyncratic patterns that are user-specific (Peltonen 2018).

The following research questions guide the present study:

(1) What kinds of fluency profiles can be identified from spoken and written productions?
(2) What kinds of differences and similarities exist between L1 and L2 profiles?

3 Methodology

To study individual language users' way of speaking and writing both in their L1 (Finnish) and L2 (English) and to recognise fluency profiles across spoken and written production by the same individuals, we combined data from two larger research projects on spoken (Fluency and Disfluency Features in L2 Speech, FDF2) and written fluency (KISUVI).[1] We identified 11 university students who participated in both projects, and the following analysis is based on their productions. Both production tasks were performed in test environments; the spoken task contained a freely produced short monologue, and the written task was a short spontaneously written argumentative text with no extra tools (e.g., dictionaries). In both cases, the participants could only rely on their mental lexicon and acquired vocabulary and structures in the L1 and L2.

3.1 Data collection

The spoken and written data were collected at a university in Southern Finland in 2021 and 2022 following the EU's General Data Protection Regulation (GDPR) and with consideration for the anonymity of the students. The participants gave their informed consent to participate in the respective studies. Background information used in this report is based on the 2022 data collection. Figure 1 illustrates the study procedure combining the two projects in a flowchart.

[1] Fluency and Disfluency Features in L2 Speech (FDF2) is funded by the Research Council of Finland (2020–2024), The project is led by the Research Group FlowLang (https://sites.utu.fi/flowlang/) and focuses on fluency and disfluency in second language speech. The KISUVI project is funded by Kone Foundation (2022–2025) and focuses on multilingual writers' writing processes and especially on formulaic sequences and fluency patterns in writing (https://sites.utu.fi/kisuvi/en/).

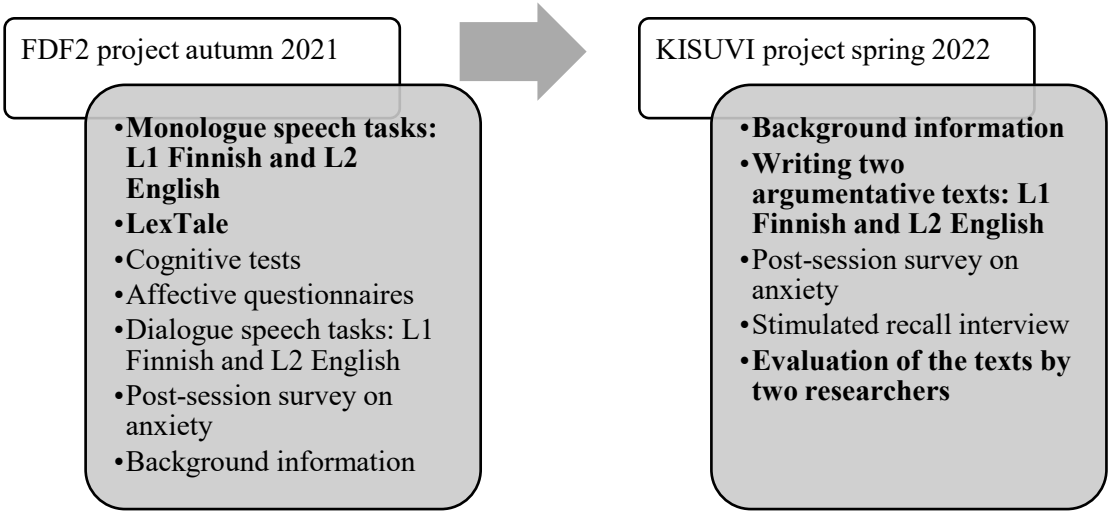


Figure 1: The study procedure combining the two projects (Data studied in the present article are in bold).

The 11 participants were native Finnish-speaking university students studying language subjects (median age 23; 5 females, 5 males and 1 did not want to say). All participants were advanced English users, either having spent a relatively long time abroad (n = 5) and/or studied English at university level (n = 9). On average, the participants had studied English at school for 10 years (SD = 1.57). Five participants self-evaluated their written proficiency as good (B1/B2 in the CEFR) and six as excellent (C1/C2 in the CEFR). In Finland, English is studied by over 90% of learners as their first L2. While we focus on L1 Finnish and L2 English in the present study, overall, the participants can be considered multilingual, as their language repertoires included one to three L2s in addition to their L1. The participants' data were anonymised and coded (participants IDs 2007–2017).

The spoken data collection for the FDF2 project was executed in the university's language and computer laboratories, and the participants performed the tasks in both their L1 and L2 English. Freely produced monologue speech samples were based on two similar comic strip prompts with six pictures. There was no time limit, and the participants had two minutes of planning time before the task. The task was to tell the story in their own words. The cartoon strips and the order of languages (L1 and L2) were counterbalanced (for more on the procedure, see Peltonen et al. 2024). Furthermore, the receptive vocabulary test LexTale (see Lemhöfer & Broersma 2012 for the test and its validation) was used to assess the participants' overall L2 proficiency level, indicating that nine participants (89.8%) represented the C1/C2 level and two the B2 level.

The speech samples were transcribed and double-checked by research assistants. The speech analysis software Praat (Boersma and Weenink 2018) was used for fluency analyses: SPs were identified with a script (De Jong and Wempe 2009) and checked by research assistants. The research assistants also annotated the repair phenomena, followed by checks by the FDF2 project researchers. Relevant information was extracted with a script (Lennes 2002) from the Praat annotations and used as the basis for fluency measurement calculations. Frequency-based fluency measures were standardised per minute of speaking time (excluding SP time), following a common practice in the field (De Jong 2016).

The written data were collected and analysed with the keystroke logging software GenoGraphiX-Log (GGXLog; Caporossi and Leblay 2015).[2] The data are a part of the KISUVI project, a larger body of research on writing fluency and the use of multiword sequences. The participants wrote an argumentative text both in L1 and L2; the topics, "Working beside studying" (L1) and "Opinion on student exchange" (L2), were chosen for their familiarity to the participants so that they could write a short argumentative text without any extra means such as dictionaries. The task was conducted in the university language and computer laboratories for 30 minutes each. The order of the languages was counterbalanced in the test (for more on the procedure, see Mutta et al. accepted; see also Figures A.1 in the Appendix for examples of the software outputs). Furthermore, the written texts were evaluated by two project researchers, who assessed the texts separately according to the CEFR scale (Council of Europe 2001). A third evaluator was available in case of disagreement. All the texts were evaluated on levels B1–C2 with the result: B1 (1), B2 (5), C1 (4) and C2 (1).

3.2 Fluency and productivity measures and data analysis

To study profiles, we need to define which objective measures from the fluency-disfluency continuum are comparable for cross-modal examination. There is no ready-made list of features to be studied, as to our knowledge, there are no previous studies where these two modes have been combined in analysis.

To identify and compare fluency profiles, we defined multiple measures in spoken and written production that could be considered comparable with respect to conventional aspects in fluency

[2] The GenoGraphiX-Log 2.0 (GGXLog) is a keystroke logging software which is a collaborative development between HEC Montréal, University of Turku, ITEM and GERAD. This project is mainly focused on combining text genetics and graph-theory with keystroke logging in writing studies (https://www.ggxlog.net).

studies. We did not use more detailed revisions measures, which have been used, for instance, to study differences between immediate, distant, and end revisions (Xu 2018: 109–110), as in spoken language the revision is always immediate. Instead, we relied on values, such as percentage of deleted characters and product to process ratio, indicating how much deletion has occurred during the writing process. We chose 14 comparable spoken and written fluency measures:

Spoken fluency

- Duration of production (total time)
- Speaking time (ST; total time excluding SPs)
- Syllable count
- Syllables / minute of total time (speech rate)
- Mean length of utterance (MLU; syllables / number of utterances)
- Silent pauses (SP; min. 0.25 sec.) / minute of ST
- Repair / minute of ST

Written fluency

- Duration of production (total writing time)
- Word count
- Words per minute
- Mean length of bursts (overall)
- Pause time (% of cognitive pause ≥ 2 seconds)
- Deletions (percentage of deleted characters)
- Product to process ratio

In spoken language, we examined both the overall productivity and speech fluency measures that have commonly been applied in research literature. In addition to the so-called "pure" fluency measures, we included three overall production measures in our analysis – duration of production, syllable count, and speaking time – to contextualise the fluency measurements and to give an overall estimate of the productivity of the speakers (cf. writing fluency measurements below). The two fluency measures, speech rate and mean length of utterance (MLU, where an utterance was defined as a speech segment between silent pauses of a minimum of 0.25 seconds in duration), represent composite measures combining the speed and breakdown fluency dimensions. The two disfluency measures, silent pauses (minimum duration 0.25 seconds, no maximum duration) and repair per minute of speaking time, then again, represent the breakdown and repair dimensions,

respectively. When analysing the fluency measures in spoken production, the values from each participant were considered in relation to the other participants' values: fluency/productivity is considered higher when the values are above the group average for duration of production, syllable count, speaking time, speech rate, and mean length of utterance, or below the group average for silent pauses and repairs. However, it should be noted that repair phenomena are less straightforwardly linked with fluency/disfluency than the other measures (see, e.g., Peltonen et al. 2024).

In written data, comparable measures were selected (the corresponding spoken fluency measure in parentheses): time of writing process (duration of production and speaking time), word count (syllable count[3]), words per minute (speech rate) and mean length of bursts (mean length of utterance); for pause time, the percentage of cognitive pauses ≥ 2 seconds (silent pauses); for repair (i.e. revision), percentage of deleted characters and product to process ratio (repair per minute of speaking time). Similarly to the spoken fluency analysis, the written fluency values were considered in relation to the other participants' values: fluency is considered higher when the values are above the group average for number of words, words per minute, and long burst length, or below the group average for time of production, percentage of cognitive pauses, and product to process ratio. The identified profiles (see Section 4.2) were also compared to participants' proficiency levels, that is, to the LexTale and CEFR values, to see if there was a relationship between these phenomena. The analysis was mainly qualitative, with quantification based on descriptive statistics.

4 Results

To address the two research questions examining the nature of fluency profiles in spoken and written productions and differences and similarities between the L1 and L2 profiles of the participants, we first provide descriptive values of the two data sets. Then we present the speaking and writing fluency profiles identified on the basis of the measures presented above (see Section 3.2).

4.1 Descriptive data based on spoken and written samples

[3] While word count in speech could be viewed as a more straightforward point of comparison for word count in writing, due to the typological differences between L1 Finnish and L2 English, syllables were preferred in the analysis of spoken data (for further discussion, see, e.g., Peltonen 2018).

Based on the above-mentioned productivity and fluency measures, we calculated the mean values in the spoken and written data in the L1 and L2. Table 1 demonstrates an overview of the spoken data.

Table 1: Mean values (range in parentheses) of the spoken data in the L1 and L2.

Lang.	Length of production (total time)	Speaking time (total time excluding SPs)	Syllable count	Speech rate (Syllables / minute of total time)	Mean length of utterance (syllables / number of utterances)	Silent pauses / minute of speaking time	Repair / minute of speaking time
L1 (Fi)	67.68 (22.05-124.62)	43.46 (19.36-73.27)	241 (89-445)	215.70 (123.78-295.72)	9.95 (4.56-22.25)	36.87 (9.30-54.12)	3.76 (0-8.25)
L2 (En)	75.20 (21.90-131.25)	48.28 (18.32-73.11)	185 (56-322)	145.95 (108.04-208.93)	6.07 (3.50-10.40)	37.07 (22.39-56.48)	4.18 (0-17.58)

Note: Length of production is indicated in seconds. Range in parentheses shows the minimum and maximum values in the data. SP = silent pause.

The overall production measurements for the L1 and L2 speech samples demonstrate that, on average, the participants produced somewhat longer samples (duration of production) and spent somewhat more time speaking in their L2 English compared to L1 Finnish. In contrast, more syllables were produced in the L1 samples. Of the fluency measures, speech rate was higher in the L1 Finnish samples and, similarly, the MLU, on average, was higher in the L1 samples, indicating longer utterances in the L1. This is potentially related to the morphological structure/linguistic typology of the Finnish language (agglutinative language), where words are longer. The standardised frequencies of silent pauses showed little difference, on average, across L1 and L2 speech. Finally, repair phenomena were used somewhat more frequently in L2 English, but the difference in means across the two languages was relatively minor.

Table 2 illustrates the individual values of the L1 Finnish speech samples.

Table 2: Individual values of spoken data in L1.

ID	Length of production	Speaking time	Syllables	Speech rate	MLU	SPs	Repair
2007	93.81	73.27	445	284.62	13.48	26.20	0.82
2008	42.88	30.78	143	200.09	8.41	31.19	3.90
2009	57.01	37.53	281	295.72	13.38	31.98	4.80
2010	80.98	57.85	356	263.78	12.28	30.08	4.15
2011	73.36	47.31	243	198.75	6.39	48.19	2.54

2012	53.09	29.08	138	155.97	6.00	45.39	8.25
2013	41.74	28.34	163	234.31	9.06	36.00	6.35
2014	79.81	45.52	249	187.20	6.38	50.09	3.95
2015	124.62	72.45	387	186.33	7.30	43.06	6.63
2016	75.14	36.58	155	123.78	4.56	54.12	0.00
2017	22.05	19.36	89	242.20	22.25	9.30	0.00

Note: Length of production is indicated in seconds. MLU = mean length of utterance, SP = silent pause.

As illustrated in Table 2, the most productive and fluent participant based on L1 speech was participant 2007: the three productivity measures were above the group average, speech rate was faster and MLU longer than average, and the participant produced fewer SPs and repairs than the participants on average. Participant 2009 can also be considered among the most fluent speakers in their L1, as their speech rate and MLU were clearly above the group average and SPs below average (repairs slightly above average). Overall, it should be noted that several of the measures demonstrate substantial variation across individuals. Table 3 shows the individual values of spoken data in L2.

Table 3: Individual values of spoken data in L2.

ID	Length of production	Speaking time	Syllables	Speech rate	MLU	SPs	Repair
2007	89.60	73.11	312	208.93	10.40	23.80	1.64
2008	21.90	18.32	56	153.41	7.00	22.93	0.00
2009	131.25	68.41	322	147.20	6.19	44.73	3.51
2010	79.74	54.48	225	169.29	7.03	34.14	4.41
2011	55.84	40.83	147	157.95	7.00	29.39	1.47
2012	103.98	63.07	246	141.94	5.35	42.81	0.00
2013	42.03	27.31	90	128.48	4.29	43.95	17.58
2014	73.06	47.65	188	154.39	5.37	42.81	8.81
2015	113.73	65.41	242	127.67	5.15	42.20	0.00
2016	69.97	38.24	126	108.04	3.50	56.48	1.57
2017	46.03	34.31	83	108.19	5.53	24.48	6.99

Note: Length of production is indicated in seconds. MLU = mean length of utterance, SP = silent pause.

As can be seen from Table 3, the individual values reveal, again, that participant 2007 was also among the most productive and fluent participants in their L2 production: the production duration and speaking times were above the group averages, and they produced more syllables. Speech rate was also higher and MLU longer, coupled with fewer SPs and repairs than the group average. In contrast to the L1 production, based on the L2 productions, participant 2010 can be included among the most fluent participants: they produced a sample longer than the group average, with longer speaking time and more syllables. Furthermore, their speech rate was higher and MLU longer than

the group average. They also produced fewer SPs, and repairs were close to the group average. Notably, one of the fastest speakers in L1, participant 2009, was close to average fluency according to their L2 production and, in particular, the measures of speech rate and MLU. However, it should also be noted from the L2 productions that there was substantial individual variation associated with certain fluency measures, especially the repairs, similarly to the L1 productions.

For comparison of the participants' language production in different modes, Table 4 illustrates the mean values of the written data in L1 and L2.

Table 4: Mean values of the written data (range in parentheses) in L1 and L2.

Lang.	Writing time	Word count	Word count / minute	Mean burst length (overall)	Percentage of cognitive pauses (vs. session duration)	Deletions (percentage of deleted characters)	Product to process ratio
L1 (Fi)	24:22 (18:09-29:42)	315.55 (244-476)	13.34 (9.22-25.16)	21.15 (15.85-30.23)	41.50 (20.08-59.71)	17.46 (8.29-25.02)	0.83 (0.75-0.92)
L2 (En)	26:17 (20:31-29:59)	428.55 (310-653)	16.70 (11.04-27.84)	17.24 (11.23-23.02)	42.03 (21.96-56.12)	17.46 (8.34-28.72)	0.82 (0.71-0.96)

Note: Duration of production is indicated in minutes and seconds. Range in parentheses shows the minimum and maximum values in the data.

In the written data, there is not much difference between L1 and L2 writing, except that the participants wrote more words in L2 English and the mean burst length was longer in L1 Finnish than in L2 English. This is partly due to linguistic differences (i.e., typologically different languages) as words are longer in Finnish, but also the length of bursts is generally longer in L1 than in L2 (Spelman Miller et al. 2008). The measures are, therefore, surprisingly similar in both languages, which might be related to the participants' English language proficiency and the presence of English in their everyday life. Table 5 illustrates the individual values in L1 written production.

Table 5: Individual values of written data in L1.

ID	Writing time	Word count	Word count / minute	Mean burst length (overall)	Percentage of cognitive pauses (vs. session duration)	Deletions (percentage of characters)	Product to process ratio

2007	25:58	251	9.66	19.40	58.11	17.30	0.83
2008	26:10	307	11.73	17.47	57.79	22.37	0.78
2009	18:09	244	13.43	22.74	27.70	19.69	0.80
2010	27:02	346	12.80	23.52	32.39	25.02	0.75
2011	20:28	311	15.19	25.14	37.89	7.76	0.92
2012	23:49	314	13.18	21.48	46.39	13.50	0.86
2013	28:03	361	12.87	19.24	28.43	18.47	0.82
2014	22:48	257	11.26	16.39	47.16	22.05	0.78
2015	26:57	330	12.24	-	40.85	22.56	0.79
2016	29:42	274	9.22	15.85	59.71	15.05	0.85
2017	18:55	476	25.16	30.23	20.08	8.29	0.92

Note: Duration of production is indicated in minutes and seconds. The mean burst length value is missing for participant 2015 due to technical failure in the log file.

As seen from Table 5, the results show that the two most fluent writers in L1 were participants 2011 and 2017. They wrote a great number of words in relation to time spent on writing, had longer bursts, and deleted fewer words than the others. It should be noted that there is no mean burst length available for participant 2015 due to a technical failure in the log file. To compare these results with the participants' production in L2, Table 6 presents these same values with the evaluation according to the CEFR scale.

Table 6: Individual values of written data in L2.

ID	Writing time	Word count	Word count/ minute	Mean burst length (overall)	Percentage of cognitive pauses (vs. session duration)	Deletions (percentage of characters)	Product to process ratio	CEFR
2007	25:36	367	14.33	16.48	49.41	24.27	0.76	C1
2008	27:31	339	12.31	-	55.57	19.61	0.80	B2
2009	20:31	467	22.75	21.98	37.21	10.68	0.89	C2
2010	28:50	452	15.67	19.41	36.09	28.72	0.71	C1
2011	20:58	415	19.79	-	36.26	8.64	0.96	B2
2012	25:21	389	15.34	17.95	44.85	12.43	0.88	B2
2013	29:48	455	15.26	14.29	37.03	15.21	0.85	B2
2014	28:04	310	11.04	11.23	56.12	25.25	0.73	B2
2015	29:05	452	15.54	16.30	32.70	21.98	0.77	C1
2016	29:59	415	13.84	14.49	55.09	16.96	0.83	C1
2017	23:27	653	27.84	23.02	21.96	8.34	0.89	B1

Note: Duration of production is indicated in minutes and seconds. The mean burst length values are missing for participants 2008 and 2011 due to technical failure in the log file. CEFR = CEFR scale (Council of Europe 2001).

Table 6 shows that, as in spoken data, there was a lot of individual variation, even if the participants seemed quite familiar with English language use. According to the values, the most fluent and productive writer was participant 2017. However, if we study the evaluation of the end-production (B1 level), we can see that the content of this participant's writing does not reach the expectations of a higher proficiency level. On the other hand, participant 2009 was fluent and productive with an excellent CEFR level (C2).

In Section 4.2, we present the spoken and written fluency profiles identified on the basis of the measures and compare each participant's cross-modal profiles.

4.2 Spoken and written fluency profiles

Based on the chosen measures, we identified four profiles in spoken production: fast and productive speaker, fast speaker, slow and productive speaker, and slow and reflective speaker. We also observed that repair profiles varied substantially within each fluency profile, and thus emphasised the importance of the other fluency measures over repair measures when creating these profiles. To give an example of categorisation of the spoken productions, participant 2011 had the following values in L1 and L2 speech, respectively (see also Tables 2 and 3):

- Duration of production (total time in seconds): 73.36 (L1) and 55.84 (L2)
- Speaking time (ST; total time excluding SPs in seconds): 47.31 and 40.83
- Syllable count: 243 and 147
- Syllables / minute of total time (speech rate): 198.75 and 157.95
- Mean length of utterance (MLU; syllables / number of utterances): 6.39 and 7.00
- Silent pauses (SP; min. 0.25 sec.) / minute of ST: 48.19 and 29.39
- Repair / minute of ST: 2.54 and 1.47

This participant was categorised, based on group averages in L1 and L2 spoken productions, as slow and productive in L1 but a fast speaker in L2.

Table 7 illustrates the spoken fluency profiles, their descriptions, number of speakers in each category in L1 and L2, and the proficiency level estimate based on the LexTale scores.

Table 7: Spoken fluency profiles.

Profiles	Description	L1	L2	LexTale

		N=11	N=11	L2
1. Fast and productive speaker	Produces a relatively long sample with lengthy speaking time and many syllables (high productivity). Speaks fast (high SR) with average to long MLUs and relatively few pauses. Repair profile can vary.	2	3	C1/C2
2. Fast speaker	Produces a relatively brief sample with short speaking time and few syllables (low productivity). Speaks fast (high SR) with average to long MLUs and relatively few pauses. Repair profile can vary.	3	2	1 x B2 1 x C1/C2
3. Slow and productive speaker	Produces a relatively long sample with lengthy speaking time and quite many syllables (average to high productivity). Speaks slowly (slow SR) with short MLUs and relatively many pauses. Repair profile can vary.	3	2	C1/C2
4. Slow and reflective speaker	Produces a relatively brief sample with short speaking time and few syllables (low productivity). Speaks slowly (slow SR) with short MLUs and relatively many pauses. Repair profile can vary.	3	4	1 x B2 3 x C1/C2

Note: SR = speech rate, MLU = mean length of utterance.

Table 7 shows that the number of speakers in each profile varies. Our analysis revealed that six speakers had the same profile in L1 and L2: participants 2007 and 2010 were fast and productive speakers, 2014 and 2015 were slow and productive speakers, and 2012 and 2016 were slow and reflective speakers in both languages. The profiles were not directly connected with the proficiency level estimates derived from the LexTale scores: the three fast and productive speakers were all at the C1/C2 level, but so were the two slow and productive speakers.

In writing, we identified four similar profiles: fast and productive writer, fast writer, slow and productive writer, and slow and reflective writer. To give an example of categorisation of the written productions, participant 2007 (C1 level) had the following values in L1 and L2 writing, respectively (see also Tables 5 and 6):

- Writing time (minutes:seconds): 25:58 (L1) and 25:36 (L2)
- Word count: 251 and 367
- Word count / minute: 9.66 and 14.33
- Mean burst length (overall): 19.40 and 16.48
- Percentage of cognitive pauses: 58.11 and 49.41
- Deletions (percentage of deleted characters): 17.30 and 24.27

- Product to process ratio: 0.83 and 0.76

This participant was categorised, based on group averages in L1 and L2 written productions, as slow and reflective in L1 but a fast writer in L2. Table 8 demonstrates the profiles, their descriptions, the number of writers in each category in L1 and L2, and finally the writers' CEFR proficiency levels.

Table 8: Written fluency profiles.

Profiles	Description	L1 N=11	L2 N=11	CEFR L2
1. Fast and productive writer	Productive in a fairly short time. Writes quickly, produces a lot of text and has long bursts. There are few corrections and pauses.	2	3	B1 B2 C2
2. Fast writer	Writes quickly but is not very productive. Bursts are of average length and the number of pauses is higher than the average. The number of revisions varies.	3	2	B2 C1
3. Slow and productive writer	Writes slowly but is productive and has long bursts. There are few pauses and lengthy revisions.	2	3	1 x B2 2 x C1
4. Slow and reflective writer	There are a considerable number of pauses and lengthy revisions, and therefore the output is scarce, even though a lot of time is spent.	4	3	2 x B2 1 x C1

Note: CEFR = CEFR scale (Council of Europe 2001).

Table 8 reveals that the number of writers in each profile varies. Our analysis revealed that seven writers had the same profile in L1 and L2: participants 2011 and 2017 were fast and productive writers, participant 2012 was a fast writer, participants 2010 and 2013 were slow and productive writers, and participants 2008 and 2016 were slow and reflective writers in their L1 and L2 productions. The other writers were assigned to different categories: 2009 was a fast writer in L1 but fast and productive in L2, 2007 was slow and reflective in L1 but a fast writer in L2, and 2015 was slow and reflective in L1 and slow and productive in L2, whereas 2014 was a fast writer in L1 but slow and reflective in L2. These results show that the writers were often faster and more productive in L2 than in L1. When comparing these profiles to the proficiency levels, it is noteworthy that all fluency profiles include participants from all proficiency levels, and that participant 2017 is a fast and productive writer with the lowest proficiency level (B1) in the group. Some writers seem to have a more fluent writing profile in L2 than L1 (e.g., 2007 and 2009).

When we compare the speaking and writing fluency profiles, the same participants are not always fast and productive, which might be due to the mode-specific production characteristics. However, we found three cross-modal profiles in L1: participants 2008 and 2016 were slow and reflective and 2015 was slow and productive in both speaking and writing. We also identified three cross-modal profiles in L2: 2009 was fast and productive, 2015 was slow and productive, and 2016 was slow and reflective in both modes. Participants 2015 and 2016 shared a similar speaking profile across their L1 and L2, but only 2016 also had the same writing profile in L1 and L2. This means that only one participant, participant 2016, had the same cross-modal profile in L1 and L2: slow and reflective speaker and writer. Participant 2015, on the other hand, had the same profile in L1 and L2 speaking and L2 writing (slow and productive), but in L1 they were a more reflective writer. These participants had a high proficiency level in speaking (LexTale C1/C2) and writing (CEFR level C1). They seem to have clear personal styles of speaking and writing, which might be related to high language proficiency or relative ease in language use. However, based on our exploratory analysis, it seems that, overall, cross-modal similarities in fluency profiles were relatively rare.

5 Discussion and conclusion

In this exploratory study, we analysed cross-modal fluency profiles in spoken and written L1 (Finnish) and L2 (English) language production of the same 11 university students. In general, fluency studies concentrate only on spoken (Lennon 1990; Lintunen et al. 2020a; Segalowitz 2010; Skehan 2014; Tavakoli and Hunter 2018) or written fluency (Cislaru and Olive 2018; Kowal 2014; Mutta 2020). The present study aimed to fill the existing gap in research regarding language users' cross-modal fluency to understand the comprehensive picture of individual language production in L1 and L2 (Li et al. 2022; Peltonen and Lintunen 2016). We combined aspects of spoken and written fluency studies and chose comparative measures to identify fluency profiles.

The findings from the present study resonate with other studies on spoken and written fluency. Based on previous studies (on spoken fluency, see, e.g., De Jong 2016; Foster and Skehan 1999; Tavakoli et al. 2020; on written fluency, see, e.g., Baaijen and Galbraith 2018; Cislaru and Olive 2018; Ellis and Yuan 2004; Gunnarsson-Largy 2012; Mutta 2020; Roca de Larios et al. 2016), we chose 14 comparable spoken and written fluency measures, situated on a continuum from fluency to disfluency, to examine fluency comprehensively across the two modes. The measures were based

on process- and product-oriented approaches, which are intertwined in spoken language due to the characteristics of the mode. The measures have commonly been used to study overall productivity and time-related fluency in the respective fields. Despite our aim to select as comparable measures as possible, it should be noted that due to the inherent differences between the two modes, some spoken and written fluency measures are more closely linked than others. For instance, as the treatment of (silent) pauses followed the respective traditions of speaking and writing fluency research, this resulted in somewhat different operationalizations of pause measurements in the present study. We hope that our exploratory examination will spark further research in the field, including methodological discussion on the comparability of measurements used in spoken and written fluency research.

In creating spoken profiles, we observed great variation in relation to repair profiles (as reported also in Peltonen et al. 2024 examining a larger sample of the FDF2 project data), and therefore emphasised other measures when identifying the profiles. Based on the analysis, we identified four spoken and written fluency profiles: fast and productive, fast, slow and productive, and slow and reflective. Six speakers had the same profile in L1 and L2, and seven writers had the same profile in L1 and L2. Our findings revealed that there were no distinctive cross-modal patterns, except for one participant who had the same profile across modes and L1 and L2 productions. The results, therefore, suggest that the difference across modes is greater than between languages. The inherent difference between the modalities also affects the individual variation across the participants.

The participants' proficiency level was quite high according to the LexTale overall measure - nine participants on the C1/C2 level and two on the B2 level – whereas the writing proficiency level, based on CEFR assessments of the texts, was more varied (between B1–C2). The analysis showed that the fluency profiles in production were not linearly linked to the proficiency levels. For instance, participant 2017 was a fast and productive writer in L1 and L2 and a fast speaker in L1, but a slow and reflective speaker in L2, despite representing a lower proficiency level compared to many other participants (LexTale B2 and CEFR level B1). This result is in line with findings of earlier studies (e.g., Dumont 2018; Mutta 2020). Based on our analysis, fluency profiles seem to be subject to influence from both user-internal and user-external variables (Lintunen et al. 2020b; Olkkonen et al. 2024), which also have an impact on affecting cross-modal fluency profiles. While we did not relate the profiles systematically to all background information or the anxiety questionnaire data (see Szyszka and Lintunen 2023; Szyszka et al. 2024) collected in the two

projects, further studies could examine the effects of various user-internal and user-external variables on fluency profiles in the future.

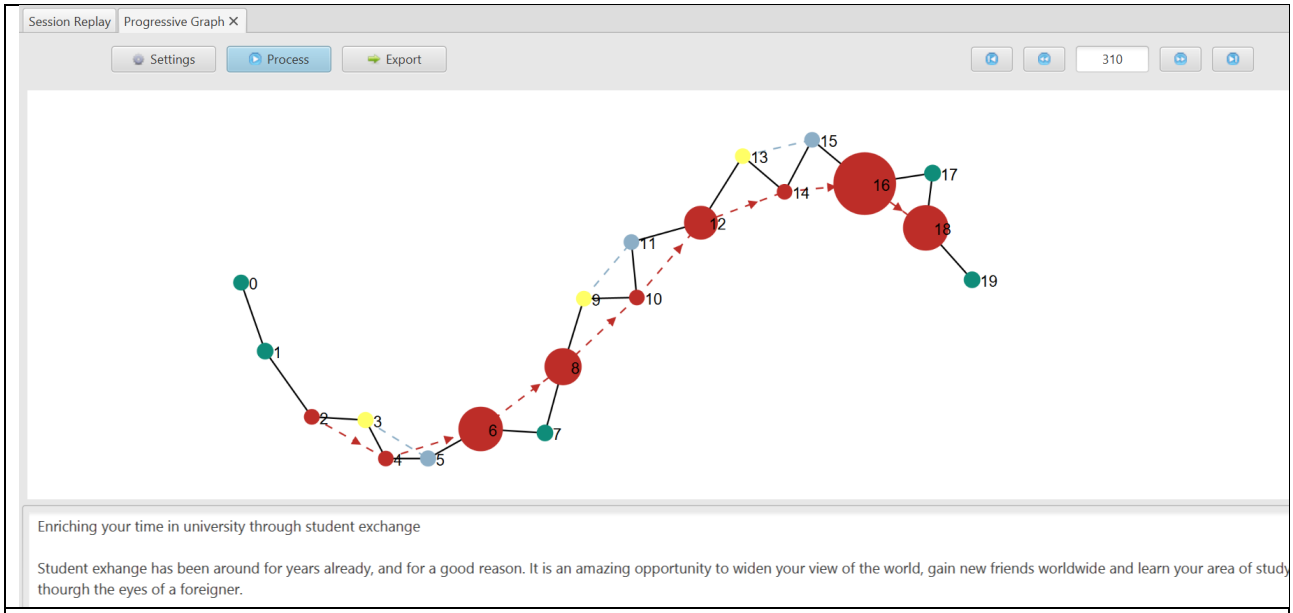
While our study can be considered pioneering in having combined L1 and L2 data from the same participants in two modes, speaking and writing, some limitations of present study should be acknowledged. First, the number of participants that participated in both projects and, thus, could be selected for the present study is relatively small, allowing only for a qualitative examination focusing on individual differences. In the future, the tendencies discovered here regarding the relatively rare combinations of cross-modal and cross-linguistic profiles should, ideally, be confirmed with larger samples and other L1–L2 combinations to enhance the reliability of the fluency profile categorisation. In addition, other speaking and writing tasks could be used to explore the extent to which the findings obtained with our two tasks can be extended to other task types. Furthermore, while we used a relatively comprehensive and versatile set of measurements, informed by previous research on written and spoken L2 fluency, the chosen measurements were limited to seven measures in each modality. Our set of measurements should, ideally, be validated in future research focusing on other learner populations and contexts, and the set of measurements could even be complemented with other measures in future studies, for example, by including detailed pause location analysis. Furthermore, alternative analytical approaches, such as a cluster analysis, could be considered to complement our qualitative approach to potentially provide more defined categories of the fluency profiles (see, e.g., Dumont 2018; Olkkonen et al. 2024).

To conclude, our findings add a more comprehensive approach to fluency studies by examining fluency profiles across spoken and written production by the same users of L2 English at the university level. The results can enhance an understanding of individual differences and similarities between these users in a Finnish context but can also generate pedagogical implications for various educational environments. While especially in quantitative fluency research learners tend to be divided into “fluent” or “disfluent” speakers or writers, our study shows that the same participants can exhibit quite varying profiles depending on the language or mode they are using. In pedagogical contexts, it is, therefore, important to acknowledge learners’ different profiles when developing writing and speaking skills, and embrace individual variation as a strength, rather than as a weakness.

Appendix

The following screenshot from the software (GGXLog) illustrates the outputs of the writing process. To learn more about the analyses, see for instance Mutta et al. (sub.).

The screenshot in Figure A.1 illustrates the GGXLog editor window with a progressive graph at the beginning of the text, the written text so far (produced text) and its linear version. This participant makes only a few corrections at this point. In the linear text, the numbers in angle brackets indicate the length of pauses in milliseconds, and in the visualisation, green nodes represent pauses, red nodes insertion of text, and yellow and blue nodes represent removed text and deletion, respectively.



Linear text:
<MOUSE EVENT><6721>Enruching <BACKSPACE7>iching your time in university through student exchange<RETURN><RETURN><9248>Student exhange has been around for years already. <BACKSPACE2>, and for a good reason. It is an amazing oppp<BACKSPACE>ortunity to widen your view of the world, gain new friends worldwide and <19393>learn your area of study thourgh the eyes of a foreigner.

Produced text:
Enriching your time in university through student exchange

Student exhange has been around for years already, and for a good reason. It is an amazing opportunity to widen your view of the world, gain new friends worldwide and learn your area of study thourgh the eyes of a foreigner.

Figure A.1: GGXLog editor window with the progressive graph process, the linear text, and the produced text in the beginning of the writing (ID 2010, C1).

References

Boersma, Paul & David Weenink. 2018. Praat: doing phonetics by computer [Computer program].

http://www.praat.org/. (accessed 17 February 2021).
Caporossi, Gilles & Christophe Leblay. 2015. A graph theory approach to online writing data visualization. In Georgeta Cislaru (ed.), *Writing(s) at the crossroads: The process–product interface*, 171–181. Amsterdam: John Benjamin.
Chenoweth, N. Ann & John R. Hayes. 2001. Fluency in writing: Generating text in L1 and L2. *Written Communication* 18(1). 80–98. https://doi.org/10.1177/0741088301018001004
Cislaru, Georgeta & Thierry Olive. 2018. *Le processus de textualisation. Analyse des unités linguistiques de performance écrite*. Louvain-la-Neuve: De Boeck Supérieur.
Conijn, Rianne, Emily Dux Speltz & Evgeny Chukharev-Hudilainen. 2024. Automated extraction of revision events from keystroke data. *Reading and Writing* 37. 483–508. https://doi.org/10.1007/s11145-021-10222-w
Council of Europe (CEFR). 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Strasbourg: Council of Europe. https://rm.coe.int/1680459f97. (accessed June 2023).
Crossley, Scott A., Kristopher Kyle & Danielle S. McNamara. 2016. The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing* 32. 1–16. https://doi.org/10.1016/j.jslw.2016.01.003
Davoodifard, Mahshad. 2022. An overview of writing process research: Towards a better understanding of L2 writing process. *Studies in Applied Linguistics & TESOL at Teachers College, Columbia University* 21(2). 1–20. https://eric.ed.gov/?id=EJ1340044 (accessed November 2023).
De Jong, Nivja H. 2016. Fluency in second language assessment. In Dina Tsagari & Jayanti Banerjee (eds.), *Handbook of second language assessment*, 203–218. Berlin: Mouton de Gruyter.
De Jong, Nivja. H. 2018. Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly* 15(3). 237–254. https://doi.org/10.1080/15434303.2018.1477780
De Jong, Nivja H. & Ton Wempe. 2009. Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods* 41(2). 385–390. https://doi.org/10.3758/BRM.41.2.385
Dumont, Amandine. 2018. *Fluency and disfluency. A corpus study of non-native and native speaker (dis)fluency profile*s. Louvain: Université Catholique de Louvain dissertation. http://hdl.handle.net/2078.1/198393
Ellis, Rod & Fangyuan Yuan. 2004. The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition* 26(1). 59–84. https://doi.org/10.1017/S0272263104026130

Foster, Pauline & Peter Skehan. 1999. The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research* 3(3). 215–247. https://doi.org/10.1177/136216889900300303

Galbraith, David & Veerle M. Baaijen. 2019. Aligning keystrokes with cognitive processes in writing. In Eva Lindgren & Kirk P. H. Sullivan (eds.), *Observing writing: Insights from keystroke logging and handwriting* (Studies in Writing 38), 306–325. Leiden & Boston: Brill. https://doi.org/10.1163/9789004392526_015

Gao, Jianmin & Peijian Paul Sun 2024. How does learners' L2 utterance fluency relate to their L1? A meta-analysis. *International Journal of Applied Linguistics* 34(1). 276–291. https://doi.org/10.1111/ijal.12493

Gunnarsson-Largy, Cecilia. 2012. The development of complexity, accuracy and fluency in the written production of L2 French. In Alex Housen, Folkert Kuiken & Ineke Vedder (eds.), *Dimensions of L2 performance and proficiency. Complexity, Accuracy and Fluency in SLA*, 247–276. Amsterdam: Benjamins Publishing Company. https://doi.org/10.1075/lllt.32.11gun

Hayes, John R. 1996. A new framework for understanding cognition and affect in writing. In C. Michael Levy & Sarah Ransdell (eds.), *The science of writing: Theories, methods, individual differences, and applications*, 6–44. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hayes, John R. & Linda S. Flower. 1980. Identifying the organization of writing processes. In Lee W. Gregg & Erwin R. Steinberg (eds.), *Cognitive processes in writing*, 3–30. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Housen, Alex, Folkert Kuiken & Ineke Vedder (eds.). 2012. *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. Amsterdam: John Benjamins.

Kowal, Iwona. 2014. Fluency in second language writing: A developmental perspective. *Studia Linguistica Universitatis Iagellonicae Cracoviensis* 131. 229–246.

Lemhöfer, Kristin & Mirjam Broersma. 2012. Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods* 44. 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lennes, Mietta 2002. Total duration of labeled segments. Praat script. https://github.com/lennes/spect/blob/master/scripts/total_duration_of_labeled_segments.praat (accessed 12 June 2024).

Lennon, Paul. 1990. Investigating fluency in EFL: A quantitative approach. *Language Learning* 40(3). 387–417.

Li, Shaofeng, Phil Hiver & Mostafa Papi (eds.). 2022. *The Routledge handbook of second language acquisition and individual differences*. New York & London: Routledge.
Lindgren, Eva & Kirk Sullivan (eds.). 2019. *Observing writing: Insights from keystroke logging and handwriting* (Studies in Writing 38). Leiden & Boston: Brill.
https://doi.org/10.1163/9789004392526
Lintunen, Pekka & Mari Mäkilä. 2014. Measuring syntactic complexity in spoken and written learner language: Comparing the incomparable? *Research in Language* 12(4). 377–399.
https://doi.org/10.1515/rela-2015-0005
Lintunen, Pekka, Maarit Mutta & Pauliina Peltonen. 2020a. Defining fluency in L2 learning and use. In Pekka Lintunen, Maarit Mutta & Pauliina Peltonen (eds.), *Fluency in L2 learning and use*, 1–15. Bristol: Multilingual Matters.
Lintunen, Pekka, Maarit Mutta & Pauliina Peltonen. 2020b. Synthesising approaches to second language fluency: Implications and future directions. In Pekka Lintunen, Maarit Mutta & Pauliina Peltonen (eds.), *Fluency in L2 learning and use*, 186–201. Bristol: Multilingual Matters.
Mohsen, Mohammed Ali. 2021. L1 versus L2 writing processes: What insight can we obtain from a keystroke logging program? *Language Teaching Research* 0(0).
https://doi.org/10.1177/13621688211041292
Mutta, Maarit. 2020. L2 fluency and writer profiles. In Pekka Lintunen, Maarit Mutta & Pauliina Peltonen (eds.), *Fluency in L2 learning and use*, 63–80. Bristol: Multilingual Matters.
Mutta, Maarit, Päivi Laine, Sinikka Lahtinen, Mari Mäkilä, Anne-Maj Åberg & Pekka Lintunen (accepted). Writers' L1 and L2 fluency: The use of multiword sequences in the initial writing phase and the impact of proficiency level. In Gustaf B. Skar, Paul Rogers & Jonathan Marine (eds.), *Writing research across borders: From early literacy learning to writing in professional life*. Fort Collins, Colorado: The WAC Clearinghouse.
Mutta, Maarit, Päivi Laine, Anne-Maj Åberg & Mari Mäkilä (sub.). Visualising multilingual writers' bursts and profiles in the initial writing phase. In Christophe Leblay, Gilles Caporossi & Hakim Usoof (eds.), *An introduction to data visualisation of the writing process* (Studies in writing series). Leiden & Boston: Brill.
Olkkonen, Sanna & Maarit Mutta. 2020. Cognitive fluency in L2: What inaccuracies can reveal about processing and proficiency. In Pekka Lintunen, Maarit Mutta & Pauliina Peltonen (eds.), *Fluency in L2 learning and use*, 34–48. Bristol: Multilingual Matters.
Olkkonen, Sanna, Maarit Mutta & Pekka Lintunen. 2024. Fluency profiles in L2 monologue production: The role of underlying cognitive factors. In Minnaleena Toivola, Pekka Lintunen & Leena Maria Heikkola (eds.), *New directions in speech research – Freely produced speech as data*

(AFinLA-teema n:o 17), 167–190. Jyväskylä: The Finnish Association of Applied Linguistics. https://doi.org/10.30660/afinla.136716
Palviainen, Åsa, Paula Kalaja & Katja Mäntylä. 2012. Development of L2 writing fluency and proficiency. In Lea Meriläinen, Leena Kolehmainen & Tommi Nieminen (eds.), *AFinLA-teema* (Soveltavan Kielitieteen Tutkimuksia 4), 47–59. https://journal.fi/afinla/article/view/7037
Peltonen, Pauliina. 2018. Exploring connections between first and second language fluency: A mixed methods approach. *The Modern Language Journal* 102(4). 676–692. https://doi.org/10.1111/modl.12516
Peltonen, Pauliina & Pekka Lintunen. 2016. Integrating quantitative and qualitative approaches in L2 fluency analysis: A study of Finnish-speaking and Swedish-speaking learners of English at two school levels. *European Journal of Applied Linguistics* 4(2). 209–238.
Peltonen, Pauliina & Pekka Lintunen. 2022. Multilingual speakers' L1, L2, and L3 fluency across languages: A study of Finnish, Swedish, and English. *Nordand* 17(1). 48–63. https://doi.org/10.18261/nordand.17.1.4
Peltonen, Pauliina, Sanna Olkkonen, Magdalena Szyszka & Pekka Lintunen. 2024. L2 repair fluency through the lenses of L1 repair fluency, cognitive fluency, and language anxiety. *Applied Linguistics Review* (ahead of print). https://doi.org/10.1515/applirev-2023-0011
Révész, Andrea, Marije Michel, Xiaojun Lu, Nektaria Kourtali, Minjin Lee & Laís Borges. 2022. The relationship of proficiency to speed fluency, pausing, and eye-gaze behaviours in L2 writing. *Journal of Second Language Writing* 58(December 2022) 100927. https://doi.org/10.1016/j.jslw.2022.100927
Roca de Larios, Julio, Florentina Nicolás-Conesa & Yvette Coyle. 2016. Focus on writers: Processes and strategies. In Rosa M. Manchón & Paul Kei Matsuda (eds.), *Handbook of second and foreign language writing*, 267–286. Berlin & Boston: De Gruyter Mouton. https://doi.org/10.1515/9781614511335-015
Segalowitz, Norman. 2010. *The cognitive bases of second language fluency.* New York: Routledge.
Skehan, Peter. 2014. The context for researching a processing perspective on task performance. In Peter Skehan (ed.), *Processing perspectives on task performance*, 1–26. Amsterdam: John Benjamins.
Sparks, Richard L., Jon Patton & Leonore Ganschow. 2012. Profiles of more and less successful L2 learners: A cluster analysis study. *Learning and Individual Differences* 22(4). 463–472. https://doi.org/10.1016/j.lindif.2012.03.009
Spelman Miller, Kristyan, Eva Lindgren & Kirk Sullivan. 2008. The psycholinguistic dimension in second language writing: Opportunities for research and pedagogy using computer keystroke

logging. *TESOL Quarterly* 42(3). 433–454. https://www.jstor.org/stable/40264477. (accessed 23 November 2023).

Szyszka, Magdalena & Pekka Lintunen. 2023. Zooming into the L2 speech fluency markers of anxious and non-anxious advanced L2 learners – an extreme case sampling report. *Research in Language* 21 (4). 357–376. https://doi.org/10.18778/1731-7533.21.4.02

Szyszka, Magdalena, Pauliina Peltonen & Pekka Lintunen. 2024. Unravelling the relationship between language anxiety and foreign language speech fluency in a monologue production. *Journal of Multilingual and Multicultural Development*. 1–15. https://doi.org/10.1080/01434632.2024.2387149

Tavakoli, Parvaneh & Ann-Marie Hunter. 2018. Is fluency being 'neglected' in the classroom? Teacher understanding of fluency and related classroom practices. *Language Teaching Research* 22(3). 330–349. https://doi.org/10.1177/136216881770846

Tavakoli, Parvaneh, Fumiyo Nakatsuhara & Ann-Marie Hunter. 2020. Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal* 104(1). 169–191. https://doi.org/10.1111/modl.12620

Van Waes, Luuk. 1992. The influence of the computer on writer profiles. In Henk Pander Maat and Michaël Franciscus Steehouder (eds.), *Studies of functional text quality*, 173–186. Amsterdam & Atlanta, GA: Editions Rodopi B.V.

Van Waes, Luuk & Mariëlle Leijten. 2015. Fluency in writing: A multidimensional perspective on writing fluency applied to L1 and L2. *Computers and Composition* 38(Part A). 79–95. https://doi.org/10.1016/j.compcom.2015.09.012

Van Waes, Luuk & Peter Jan Schellens. 2003. Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics* 35. 829–853.

Xu, Cuiquin. 2018. Understanding online revisions in L2 writing: A computer keystroke-log perspective. *System* 78. 104–114. https://doi.org/10.1016/j.system.2018.08.007