Pauses during a Writing Session in two Typologically Different Languages

Ilmari Ivaska, Outi Toropainen & Sinikka Lahtinen

University of Turku, Turku | Finland

Abstract: This study investigates how pausing behaviour within a writing session is associated with the writer's language proficiency, focusing on Finnish and Swedish as both first language (L1) and learner language (L2). The data were collected through keyboard logging software and evaluated using CEFR-based assessments of the resulting texts. The relationship was analysed using ordinal mixed-effects logistic regression modelling, where proficiency is modelled as a function of various variables related to pausing behaviour. The results show that the L2 writing process reflects the writer's proficiency. However, there is a significant difference between L2 writers of Swedish and L2 writers of Finnish compared to L1 writers. The advanced L2 writers of Swedish behave similarly to the L1 Swedish writers. In contrast, even the most advanced L2 writers of Finnish have pause lengths and linguistic contexts that are more similar to the less advanced L2 writers than the L1 writers. In addition, the pauses between words do not indicate any clear proficiency-related patterning, leaving only within-word pauses as a robust indicator of proficiency, especially in Swedish. Unlike most writing process research, this study's parallel design allows for contrasting two typologically diverging languages while controlling for other contextual variables. Future studies could explore the grammatical nature of pause locations across the analysed languages.

Keywords: pauses, writing process, Finnish, Swedish, language proficiency

journal of WRITING RESEARCH

Ivaska, I., Toropainen, O., & Lahtinen, S. (2025). Pauses during a writing process in two typologically different languages. *Journal of Writing Research*, *16*(3), 407-433. https://doi.org/10.17239/jowr-2025.16.03.03

Contact: Outi Toropainen, University of Turku, Arcanuminkuja 1, 20500 Turku | Finland – outi.toropainen@utu.fi - https://orcid.org/0000-0002-6405-592X

1. Introduction

This article delves into pauses during the writing session in two typologically different languages, Finnish and Swedish, as well as the relationship between the pausing behaviour and the quality of the written product, i.e. an argumentative text. Writing in the first language (L1) and learner language (L2) are studied under cognitive, process-oriented models and theories. However, few studies have been conducted about the writing process in languages that differ typologically from English (Yiğitoğlu & Reichelt, 2019). Moreover, comparing L1 and L2 perspectives of the same language in a contrastive research design has so far remained scarce, and we are unaware of any previous studies where writing processes in two languages—let alone two typologically diverging languages—would be addressed in a uniform research design parallelly from both L1 and L2 perspectives. Furthermore, according to Barkaoui (2019, p. 550), few studies have been conducted on the relationship between pausing behaviour and text quality. Hence, this article provides new insights into the cross-linguistic generalizability of hypotheses and postulates regarding writing in L2 and the relationship between the writing process and the proficiency-related quality of the resulting text.

The shift from the written product to the writing process was actualised by Matsuhashi, who focuses on pauses during writing (1981, 1982, 1987). Since then, the interest has shifted to cognitive processes, leading researchers to present several process-oriented models of L1 writing and L2 (e.g. Bereiter & Scardamalia, 1987; Flower & Hayes, 1980; Kellogg, 1996). The methods used to investigate the writing process have evolved with technological development. Especially during the 2020s, writing process studies have benefitted from technological progress in online possibilities, including various keystroke logging programmes and eye-tracking software (see Wirtz, 2025, this issue). Such programmes, including GenoGraphiX-Log used in this project (abbreviated GGXLog; Leblay & Caporossi, 2015), generate quantifiable moment-by-moment data of the writing process by recording the activities that take place on the computer when creating a text.

The premise is that the writer's implicit cognitive sub-processes are made explicit in the form of pauses, whereby the writing is rendered as an interplay between pauses and physical acts of typing, i.e. bursts. According to Alamargot et al. (2007, p. 13), pauses might take 60–70% of the total writing time. To this end, a central question is to understand where and when pauses occur and for how long because pauses give indirect knowledge of the writer's cognitive processes and the pause duration gives a clue of the writer's cognitive load (Alamargot et al., 2007; Barkaoui, 2019).

2. Background

2.1 Writing process as a research subject

Writing is a complex cognitive activity that comprises various stages, regardless of the language produced or the language proficiency of the writer. Several models have been introduced to capture this activity, including that of Kellogg (1996; Kellogg et al., 2013), which has been adopted in many studies focusing on L2 writing (e.g. Michel et al., 2020; Mohsen, 2021; Révész et al., 2017; Révész et al., 2019). The model consists of three phases relevant to the working memory perspective.

- The first phase is formulation, which refers to planning and translating thought ideas into linguistic units. In other words, in this phase, the writer makes their implicit ideas explicit through written language.
- The second is execution, which concerns the physical act of writing using a pen or keyboard.
- The third, monitoring, involves the reviewing phase, in which the writer reads and edits the earlier written text.

These phases overlap in various ways during the text production, and several factors can influence a writer's pausing behaviour. Some pauses might be strategic choices (Alamargot et al., 2007) when the writer rereads an earlier produced text to revise it or repeat the content to be able to continue the writing (revising and planning). Moreover, physical and socio-psychological reasons also affect the patterning of pauses (e.g. Wengelin, 2006). Regarding L2 writing, language-related issues, such as word, orthographic, syntactic and morphological features, can cause pauses.

The writing process has often been approached from the point of view of pausing behaviour: the length and linguistic context of pauses during the process. Concerning pause length, an agreed threshold to separate cognitive and non-cognitive pauses does not exist (Hall et al., 2024). For instance, while Barkaoui (2019) uses the conventional 2000 milliseconds (ms) threshold of pauses to be included in the analysis, Révész et al. (2022) use a considerably lower threshold of 200 ms. Hence, direct comparability between the results remains somewhat limited. However, Mutta (2017, p. 514) suggests that 2000 ms is enough to tease cognitive pauses apart from the technical pauses primarily due to correcting typographical errors. Overall, however, earlier research can be summarised to suggest that the higher the language proficiency of the writer, the proportionally less there should be pauses within or between words.

The context, described often as pause location, indicates the textual boundary of the interruption and the potential reason for the pause, which can be associated with different cognitive sub-processes (Alamargot et al., 2007; Schilperoord, 1996). Pausing between clauses and sentences would correspond to a higher-order cognitive sub-process in writing (e.g. Guo et al., 2018). This contrasts with pauses between and within words

when the writer edits various lexical items and morphology, as they require less cognitive effort and are associated with the writer's lower-level cognitive processes. The pause length before various textual units is assumed to reflect the complexity and cognitive load of the forthcoming textual unit (e.g. Damian & Stadthagen-Gonzalez, 2009). Therefore, pause behaviour is not arbitrary. Through several investigations, L1 and L2 writers of different ages have reflected these assumptions (summarised in Barkaoui, 2019, p. 531).

2.2 Earlier research on the L2 writing process

In studies of the L2 writing process, the research has focused on a variety of issues, ranging from typological phenomena like gender (Zhang et al., 2019), text-level phenomena such as genre, revisions, complexity and argumentation (Barkaoui, 2016; Michel et al., 2020; Mohsen & Qassem, 2020; Révész et al., 2017; Tian et al., 2024), learners' overall proficiency (Gánem-Gutiérrez & Gilmore, 2018; Révész et al., 2022; Xu & Xia, 2019), as well as phenomena related to the medium of text production, such as keyboard typing (Barkaoui, 2016; Xu, 2018). Most commonly, the L2 writing process has been investigated in formal learning situations, mainly with English as L2 (Lindgren et. al, 2019; Yiğitoğlu & Reichelt, 2019).

In most of the earlier studies, the L2 language is English. The L1 is either another Indo-European language from the same genus, like Dutch or German, or a language using a non-alphabetic writing system, such as Chinese (e.g. Barkaoui, 2019; Leijten et al., 2019; Michel et al., 2020; Mohsen, 2021; Sasaki, 2002; Van Waes & Leijten, 2015; see also Vasylets & Marín, 2025, in this issue). An interesting exception is Chukharev-Hudilainen et al. (2019), where L2 English writers have as L1 Turkish, a synthetic language with rich suffixal morphology, much like Finnish, but with no genealogical relationship to Finnish. According to this study, the writers seemed to plan their text more in L1 Turkish than in L2 English because they paused longer at the beginning of subordinate clauses in L1 than in L2. Previous studies indicate that in morphologically rich languages, L2 comprehension occurs morpheme by morpheme (Durrant, 2013). This also resonates with our own experiences, whereby L2 writers of Finnish and L2 writers of Swedish seem to produce, in addition to bursts of multiword strings or multimorphemic units, separate morphemes, syllables and letters.

In L2, the complexity of the writing process increases as the writer's language proficiency plays a role in all phases of the writing process. For an L2 writer, lexical retrieval and morphosyntactic encoding take more time and effort than for an L1 writer—although cross-linguistic influences within the language repertoire of an individual could, in part, also ease lexical retrieval and morphosyntactic encoding for L2 writers. Because the L2 writer has to focus more on the word and sentence levels, the higher-level processes concerning, for example, the content of the text might be interrupted more frequently, compared to an L1 writer (e.g. Chenoweth & Hayes, 2001; Révész et al., 2022; Schoonen et al., 2009; Van Waes & Leijten, 2015).

Earlier research shows that L2 writers pause longer and make more revisions at low-level units than L1 writers or more proficient writers (e.g. Spelman Miller, 2000; Sasaki, 2004). Barkaoui (2019) also showed that L2 English writers at lower proficiency levels pause more frequently than higher-level writers. Moreover, Barkaoui (2019, p. 531) points out, by referring to several earlier studies, that both L1 and L2 writers tend to pause more frequently and for more extended periods before starting paragraphs and sentences, compared to within and between words and phrases (see also Michel et al., 2020). Moreover, proficient L2 writers spend more time planning, evaluating and revising the text. In contrast, L2 writers with lower language proficiency have more within-word pauses, or pauses at all lexical levels (i.e. within, before, after and between words) (Roca de Larios et al., 2008; Shen & Chen, 2021). This suggests various levels of processing at these text boundaries. In their study of L2 writers of English at proficiency levels B1–C1 in the Common European Framework of Reference for Languages (Council of Europe, 2020, abbreviated CEFR), Révész et al. (2022) found that writers at higher proficiency levels pause for shorter periods between words and more frequently between sentences. They interpreted the result to reflect that the better the language skills, the more automatised the linguistic and writing skills are, which is visible in shorter pauses.

2.3 Finnish and Swedish as objects of study

This article concerns writing in Finnish and Swedish, both as L1 and as L2. Finnish and Swedish are both official languages in the Republic of Finland, and the Language Act (6.6.2003/423) ascertains equal rights in both languages. Finnish is the language used by most of the 5.5 million inhabitants in the country, and a minority of approximately 6% uses Swedish. A parallel education system runs in both languages, from early childhood education to the tertiary level and all postgraduate degrees. The national core curriculum is followed by all schools in Finland, and it is independent from the language of instruction (Finnish National Agency for Education, 2014), and even the matriculation examination taken at the end of the upper secondary school can be taken in either language (with some limitations also in Sami, see the Act on the Matriculation Examination, 12.4.2019/502). All pupils receive instruction in both official languages from primary school onwards. All in all, in the context of this project, the data across the two languages are highly comparable in terms of the learner-related background variables.

Genealogically, the two languages stem from different language families and genera: Finnish belongs to the Finnic genus of the Uralic language family, while Swedish is part of the Germanic genus of the Indo-European language family. Typologically, Finnish is predominantly synthetic, employing extensive grammatically motivated agglutination, i.e. suffixal inflexion and derivation in verbs, nouns, adjectives, pronouns and numerals. Swedish, in turn, has agglutinative, inflected and analytic tendencies, for example regarding the marking of definiteness, grammatical gender and number of nouns, and some verb tenses. Swedish applies a strict, grammatically constrained constituent order

in main and subordinate clauses. In contrast, the Finnish clausal constituent order is relatively flexible, which is also motivated by the topical information of the clause. Both languages use left-to-right Latin script.

Regardless of the typology of a specific language, the CEFR makes it possible to compare different L2 texts under a uniform framework when trained and experienced evaluators interpret the descriptors of various criteria similarly. (e.g. Alanen, Huhta & Tarnanen, 2010.) In addition, the present study hails from Finland, where, as Ringbom (2007, p. 34) puts it, "[t]he two official languages [---] are linguistically different. Culturally and educationally, however, the two language groups are as close as can be found in any country anywhere in the world" (for an overview of the language scene in Finland, see Ringbom 2007, p. 34–39). Hence, all the participants stem from a culturally relatively homogenous context with highly comparable general backgrounds across the two languages. This allows for close comparability across languages and makes it possible to pinpoint potential linguistic differences rather than contextual differences.

2.4 Research objective and hypotheses

The objective/ aim of this article is to understand the relationship between language proficiency and the pauses that occur during the writing session – and the degree to which this relationship is language-independent. More particularly, this study answers the following research questions:

- In which ways is pausing behaviour associated with the overall language proficiency of the writer?
- To what extent are these possible associations language-dependent in nature?

Our general hypothesis is that the L2 writing session is conditioned by the language proficiency of the writer, and that this is visible in the pausing behaviour during the writing session. Moreover, we expect the sessions to be conditioned by the typological nature of written language. In the following section, we present three hypotheses and contextualize them in relation to earlier research.

Hypothesis 1: We hypothesize that less proficient writers have relatively longer pauses than the more proficient ones. This hypothesis is based on earlier research (e.g. Spelman Miller, 2000) regarding pause length in learner writing.

Hypothesis 2: We hypothesise that less proficient writers have relatively more pauses within and between words, whereas more proficient writers have relatively more pauses at the beginning and the end of sentences. Our hypothesis stems from earlier results suggesting that pause locations indicate the nature of the processing units and that those units are larger among more proficient users (e.g. Révész et al., 2022; Van Waes & Leijten, 2015).

Hypothesis 3: We hypothesise that more proficient L2 writers have relatively more pauses at the beginning and the end of the recorded writing session, while the pauses of less proficient L2 writers are distributed relatively more evenly across the recorded

session. Earlier research suggests that more proficient L2 writers spend more time planning before they start writing at the global text-level and revising their expressions more frequently (Roca de Larios et al., 2008; Sasaki, 2000, 2002, 2004; Sasaki et al., 2018).

Hypothesis 4: We hypothesise that due to the rich suffixal inflexion system in Finnish, the L2 writers of Finnish are relatively more likely to pause within words than the L2 writers of Swedish. Furthermore, we hypothesise that the effects motivated by typological differences may not result in a balanced outcome in the L2 writing session across the two languages. While the actual similarity between the two languages does not depend on the point of view, the perceived similarity may well be imbalanced, whereby "[s]peakers of language X may find it easier to understand language Y than speakers of language Y to understand language X" (Ringbom, 2007, p. 7). To this end, we do not have specific hypotheses regarding the nature of such possible differences. On the one hand, L1 users of Swedish in Finland might be more exposed to Finnish than L1 users of Finnish are to Swedish, suggesting that it could provide ease to L1 writers of Swedish when writing in L2 Finnish. On the other hand, virtually all the younger generations in Finland have studied English as a foreign language from early on, and together with the constant overall exposure to English in their present-day surroundings, this could ease the use of L2 Swedish by L1 writers of Finnish.

3. Data and Methods

3.1 Participants and Data Collection

The 26 participants in this study is a sample of a larger dataset that comprises approximately 130 participants collected in the research project KISUVI (Multilingual writers' writing processes: graph-theory based visualisation of formulaic sequences and fluency patterns) at the University of Turku (202*2*–202*6*), funded by the Kone Foundation. We used the following sampling criteria:

1. Both the recorded writing session and the resulting text product should be available in L1 for all the participants, either in Finnish or Swedish.
2. The participants should have a recorded writing session and the resulting text product in L2 Finnish or Swedish.

Using these criteria, we had 26 writers and 52 recorded writing sessions. Half of the texts were written in L1 (n = 26) and half in L2 (n = 26). Each participant wrote two argumentative texts—one in their L1 and another in their L2—consistent with the project design. Eight of the 26 writers had Swedish as their L1 and Finnish as their L2. The remaining writers were L1 Finnish, with Swedish as their L2 (Table 1). In other words, L2 writers of Swedish also wrote in L1 Finnish, and L2 writers of Finnish also wrote in L1 Swedish.

All 26 writers are Finnish university students majoring in various languages at universities in Southern Finland, with Finnish or Swedish as their L1 and the other language as their L2. Participating students had 30 minutes to write each argumentative text, one in their L1 and one in their L2, as part of a course they participated in at their home university. The writing prompts were the same for L1 and L2 writers but varied depending on the used language. The prompts were as follows:

- L1/L2 Finnish: *Mitä mieltä olet töiden tekemisestä opiskelun ohessa? Mitä hyviä ja huonoja puolia siinä on?* [What do you think about working while studying? What are the pros and cons?]
- L1/L2 Swedish: *Tycker du att det är viktigt att ha körkort? Vilka för- och nackdelar finns det med att ha bil?* [Do you think it is important to have a driver's licence? What are the advantages and disadvantages of having a car?]

The writing order varied in the analysed sample, but the order distributed relatively similarly in both languages and among L1 and L2 writers: 11 out of 26 writers wrote their L1 as the first text (3/8 of the L1 writers of Swedish and 8/18 of the L1 writers of Finnish). The writing sessions of these 52 texts were recorded using the keyboard recording software GenoGraphiX-Log (abbreviated GGXLog), which allows for the analysis and visualisation of the writing process (Leblay & Caporossi, 2015).

All resulting L2 text products (n = 26) were separately evaluated by two trained and experienced evaluators according to the descriptors of the six proficiency levels in CEFR. To ensure reliability, both evaluators had to agree on the CEFR proficiency level, and a third evaluator was engaged if the two evaluators disagreed. If all three evaluators had disagreed, the text would not have been included for further analysis. These proficiency levels give an overall language proficiency, allowing for comparisons between the data from the recorded sessions and the products. The L1 text products (n = 26) were not evaluated.

However, this study's primary data originate from the 52 recorded writing sessions and is the raw output of keystroke logs, including numeral information of the empty time (pauses, inactive time) and filled time (active time, using keyboard or mouse). In this study, we applied a ≥ 2000 ms limit to pauses. The secondary data consist of GGXLogs categorising pause locations (e.g. between/within a word, sentence beginning/end), which took place during the writing session and are manually checked. In the early stage of the study, we recognised that the programme could not do reliable categorisation due to the morphosyntactic features of Finnish and Swedish. Therefore, the data were moved to an Excel spreadsheet in the early stages, allowing us to fine-tune the data manually.

However, as there were no data for Finnish as L2 in the A1 level or for Swedish as L2 in the C2 level, we decided to merge the data into four categories. Therefore, the response variable is an ordered categorical variable with four possible values (A < B < C < L1). Table 1 summarises the number of writing sessions across the proficiency levels, the number of L1 writers, and the time used in the sessions in both languages.

The data collection in this study was conducted according to the European data protection rules (The General Data Protection Regulation, GDPR), which guarantees the anonymity of the participants. All students participated voluntarily in the data collection. The participants were anonymised by individual codes. The data are stored on a university server and are only accessible to research team members.

Table 1. The used data, number of writers in both L1/L2 languages and sum-minutes of the writing session

Proficiency	Swedish		Finnish		Total	
	sessions	minutes	sessions	minutes	sessions	minutes
A	4	95	1	28	5	124
B	8	185	3	81	11	265
C	6	153	4	89	10	242
L1	8	185	18	466	26	651
Total	26	618	26	664	52	1282

3.2 Data processing and variable operationalisation

As described above, the data stem from 52 recorded writing sessions. However, to adequately address this study's objective, in our analysis, each occurring pause—rather than each recorded writing session—constitutes a unit of observation. In other words, we analyse each pause in relation to the background variables that describe that particular pause. This solution makes it possible to include variables pertaining to the writer, the writing session, the resulting text and pause context in one multivariate model. Moreover, the solution makes it possible to operationalise and describe the units of observation (i.e. pauses ≥ 2000 ms), in a more language-agnostic fashion, allowing for a more direct cross-linguistic comparison between Swedish and Finnish.

This solution makes it possible to control for the relative effects of different variables and still account for the structured nature of the data at hand. Next, we describe in detail all the included variables and how they were operationalised for this study. We describe the details of the statistical modelling in Section 3.3, also addressing how the model captures dependencies across observations. The basic logic, however, can be thought of as follows: for each pause in our data, the goal is to predict the evaluated proficiency of the text product from which each pause stems by looking at a range of variables associated with that relationship in earlier research.

Unit of observation: pauses over 1999 ms. In the analysis, each pause of 2000 ms or longer constitutes a unit of observation (e.g. Chenu et al., 2014). The two-second threshold is frequently used in writing process studies conducted by keyboard logging, where pauses ≥ 2000 ms are defined as cognitive pauses (Chenoweth & Hayes, 2001; Cislaru & Olive, 2018, p. 43; Kowal, 2014; Wengelin, 2006), to tease them apart from the technical pauses that occur primarily due to correcting typographical errors (Mutta, 2017). With these criteria, our final data comprised 2,749 observations. Consequently, the variables described below have been used to describe each observation.

Response variable: evaluated proficiency of the text product. We used CEFR-based proficiency evaluations of the text products as the response variable (see Table 2 for the distribution). Consequently, the response variable is an ordered categorical variable with four possible values (A < B < C < L1). Each observation from the same text is assigned the same proficiency level[1].

Table 2. Distribution of observations across proficiency levels and languages

Proficiency	Swedish *(N of obs)*	Finnish *(N of obs)*	Total
A	209	59	268
B	446	161	607
C	392	155	547
L1	384	943	1327
Total	1431	1318	2749

Predicting variable 1: pause duration. The length of the pause is a numeric predicting variable. Because we limited the analysis to pauses of ≥ 2000 ms, the pauses varied from 2–195 seconds (mean: 7.5 s; median: 4.1 s). We log-transformed the values for statistical modelling. Table 3 summarises the distribution across the two languages.

Table 3. Means and standard deviations for pause lengths in milliseconds (ms and log-transformed) across languages

	Swedish		Finnish	
	ms	logged	ms	logged
Mean	7547	8.53	7419	8.49
SD	11246	0.77	12248	0.77

Predicting variable 2: linguistic context of the pause. Each pause is assigned a contextual description of the pause location, constituting a categorical predicting variable. The software used in the data recording assigned a tentative value, which we manually checked and corrected where needed. We limit the analysis to five contexts of occurrence: between sentences (example 1), beginning-of-sentence (example 2), end-of-

sentence (example 3), between words (example 4), and within words (example 5)[2]. The numbers in the examples show the pause length in milliseconds. Table 4 shows the distribution across the data.

(1) [---] *det.<3857> Även bilarna* [---]
[---] it.<3857> Even cars [---]
(Swedish, ID 3014, L1-swe, between sentences)
(2) [---] <37341>*Minun mielestä* [---]
[---] <37341>In my opinion [---]
(Finnish, ID 3007, L1-swe, beginning-of-sentence)
(3) [---] *för en längre tid*<10684>.
[---] for a longer time<10684>.
(Swedish, ID 2076, L1-swe, end-of-sentence)
(4) [---] *että* <5019>*opiskelet* [---]
[---] that <5019>you study [---]
(Finnish, ID 2076, L1-swe, between words)
(5) [---] *Saa*<2082>*t enemmän työkokemusta* [---]
[---] get<2082>SG2 more work experience [---]
(Finnish, ID3007, L1-swe, within word)

The difference in pauses between sentences on the one hand and pauses at the beginning or the end of sentences on the other is defined as follows: pauses that are coded to occur between sentences take place directly after the punctuation that closes the previous sentence; thus, the pause is followed by a space stroke. In contrast, pauses coded to occur at the beginning of the sentence are preceded by a space stroke, followed directly by the first letters of the new sentence. Finally, pauses coded to occur at the end of a sentence take place after the last letter of the sentence but before punctuation.

Table 4. Distribution of observations (obs) across pause locations and languages

Pause location	Swedish (N of obs)	Finnish (N of obs)	Total
between sentences	37	33	70
beginning-of-sentence	304	290	594
end-of-sentence	111	86	197
between words	903	794	1697
within word	76	115	191
Total	1431	1318	2749

Predicting variable 3: pause location within the writing session. The location of each pause, measured as milliseconds from the beginning of the writing session, is a numeric

predicting variable. We standardised the values to reflect the relative positioning of the pause within the recorded session. Therefore, the values ranged between 0 (the beginning of the session) and 1 (the end of the session) (see Table 5).

Table 5. Standardised means and standard deviations for pause location within the writing session across languages

	Swedish	Finnish
Mean (standardised)	0.485	0.478
SD (standardised)	0.290	0.292

Predicting variable 4: language. We also included a categorical variable on the language of the text. With two possible values (Swedish or Finnish), the variable serves two purposes. We have included an interaction term between the language and each of the three process-related variables described above to capture possible effects of pause duration, the linguistic context of the pause, or pause location that lend themselves to differences across the two languages studied. Aside from the theoretical motivation for including the language produced as a variable in the model, it is also important as a control variable, as the distribution across different proficiency levels differs between the two languages. In other words, it must be included as a separate variable because the uneven distribution of data could otherwise confound the effect of other included variables.

Predicting variable 5: writer. Finally, we have included random intercepts for each writer. While the effects of individual writers are not a primary focus of this study, this random variable makes it possible to control for dependencies across observations (for a linguistically oriented take on the importance of controlling for the dependencies across observations, see Winter and Grice, 2021), including but not limited to individual typing skills or idiosyncratic preferences related to text production.

3.3 Statistical modelling

For the statistical modelling, we used an ordinal logistic regression mixed-effects model, with each pause of ≥ 2000 ms constituting an observation. Ordinal logistic regression is suitable for variables like proficiency, where the nature of the variable is qualitative rather than quantifying in nature, but where the values of the variable have an ordered structure. As discussed above, the response variable was the proficiency level of the text product associated with the writing session, where the respective pause occurred. In the response variable of the present study, there are three levels for L2 writers and one level for L1 writers (A < B < C < L1). While the computational details of ordinal logistic regression fall beyond the scope of the present study (for a linguistically oriented example, see Gries 2021, p. 353–361), the basic underlying logic is similar to binary logistic regression, with a tweak that the different category splits are compared sequentially (A vs. B-or-above, B-or-lower vs. C-or-above, etc.) to form a uniform model. Including both L2 and L1 texts in

the same model with ordinal regression also inherently captures the general differences between L2 and L1 writing (such differences would be reflected in the results in contrasts between the C-or-lower vs. L1). We modelled the level of proficiency as a function of pause duration, the linguistic context of the pause, the location of the pause within the writing session, and the language of the text. To capture the possible differences between Swedish and Finnish, we also included an interaction term between the language of the text and each of the three process-related variables. In other words, the interaction term controls for the nested structure of the data pertaining to the language written.

The writer was included as a random variable to capture the nested structure of the data regarding individual writers and to control for any idiosyncratic behaviour. All statistical analyses were conducted in an R programming environment (R Core Team, 2022). We fitted the model using the `ordinal` package (Christensen, 2022) and used the functions from the `performance` package (Lüdecke et al., 2021) to calculate the conditional and marginal pseudo-R^2 values to evaluate the model fit.

The research design was confirmatory, and all the included variables were motivated based on earlier research. Hence, to avoid issues related to overfitting and reporting overly optimistic results, we did not conduct model optimisation but presented the results of the full model (for rationale, see, e.g. Barr et al., 2013; Whittingham et al., 2006). We evaluated the statistical significance of each predicting variable or interaction using a likelihood ratio test (LRT), comparing the entire model to a model in which that variable was left out. When interpreting the effects[3] of the predicting variables, we conceptually followed the procedure described in Gries (2021, p. 353–358) and used the predicted probabilities of different variable levels instead of the variable coefficients.[4]

4. Analysis and results

4.1 Model Summary and Diagnostics

The language proficiency level of the writers could be modelled relatively well ($R^2_{marginal}$ = 0.389, $R^2_{conditional}$ = 0.493), and the variance inflation factors of the individual predicting variables are all well below 5, suggesting that multicollinearity between variables does not question the reliability of the results obtained using the model. Hence, we felt confident in proceeding with the analysis. Table 6 reports the results of the LRT. Pause duration and the linguistic context in which the pauses occur had a statistically significant effect when modelling the writer's proficiency level. As for the location of the pause within the writing session, it did not contribute to the model in a statistically significant fashion. Interestingly, the nature of the effect of all three variables differs between the two studied languages, as indicated by the statistically significant effect of the interaction terms.

Tentatively, these observations corroborate our hypotheses regarding pause duration and the linguistic context of pauses. As the effect of the pause location differs between the studied languages but does not reach statistical significance, the results do not fully

support hypothesis 3 that more proficient users would compose their texts differently regarding text-level pausing behaviour. However, this supports hypothesis 4 of proficiency-related differences between writing sessions in the two languages. Next, we will analyse the three variables in greater detail.

Table 6. Scores of the likelihood ratio test. P-values below 0.05 are considered statistically significant and are marked with*

Variable	df	LR statistic	p-value
Duration: Language	1	26.3	< 0.0001*
Linguistic Context: Language	4	107.3	< 0.0001*
Location: Language	1	2020.1	< 0.0001*
Language	1	2075.1	< 0.0001*
Duration	1 (numeric)	39.1	< 0.0001*
Linguistic Context	4	211.6	< 0.0001*
Location	1 (numeric)	2.6	0.1094
ID (random)	26 individual values, sd = 0.7526		< 0.0001*

4.2 Pause duration

As indicated in Table 6, pause duration has a statistically significant effect on the writer's proficiency level. Figure 1 visualises the predicted probabilities of different proficiency levels as a function of pause duration in interaction with the language. Note that as the distribution of pauses across the different proficiency levels is uneven, we have standardised the predicted probabilities for these visualisations using the so-called Z-standardisation. This means that values over zero indicate probabilities above chance, while values below zero indicate probabilities below chance, given the data distribution. For example, in the case of Finnish, the longer the pause, the more probable it is that it stems from a text that has been evaluated as representing a lower level of proficiency—and the less probable it is that it stems from a text that has been written by an L1 writer of Finnish. We used the same standardisation procedure in all the visualisations.

A closer examination of these results confirms hypothesis 1: The more proficient the writer, the shorter the pauses when writing. However, there is a drastic (and statistically significant) difference between the two languages: in Swedish, the tendency is similar among L1 writers and more proficient L2 writers insofar as their pause lengths are concerned – effectively distinguishing them from both A- and B-level L2 writers who are relatively more likely to have longer pauses. This is visible in the left panel of Figure 1, where the lines indicating the relative probabilities of a pause stemming from a text

evaluated to represent a C-level L2 writer or an L1 writer decrease (y-axis) when the pauses are longer (x-axis).

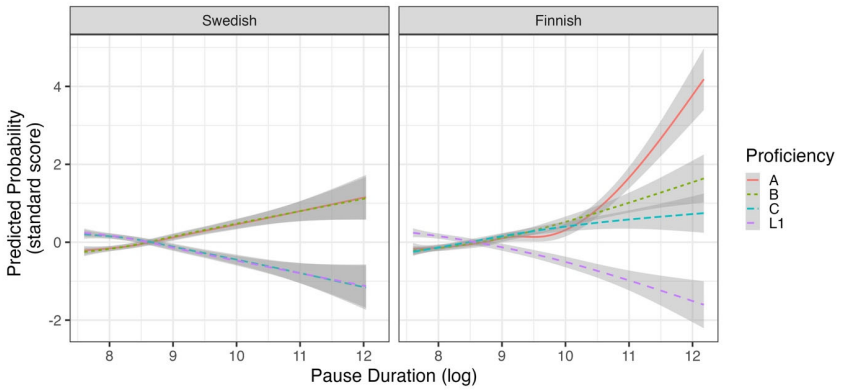

Figure 1. Predicted probability of proficiency level as a function of pause duration and language.

Regarding Finnish, there is a clear difference between L1 writers and all L2 writers, regardless of their proficiency level: longer pauses are more likely in texts written by L2 writers than in texts written by L1 writers. This tendency is the clearest among A-level L2 writers. However, the probability of the writer being of L2 status increases across all L2 proficiency levels; the probability of the writer being of L1 status decreases when the pauses are longer.

We interpret this result to mean that, in Swedish, the pause duration in the writing session correlates with the overall language proficiency evaluated based on the product: the gradual change in the pause duration is conversely correlated with the predicted probabilities of writers' evaluated proficiency. In Finnish, in contrast, the pauses do not seem to correspond to product-based proficiency in the same way. Instead, the clear difference between L1 writers and even the most proficient L2 writers suggests that the writing session in Finnish is associated more with the L1 status than with the language proficiency evaluated based on the product.

4.3 Linguistic context of the pause

The writer's proficiency level was also conditioned by the linguistic context of the occurrence of pauses in a statistically significant fashion, both in and of itself and in interaction with the language observed (Table 6). The predicted probabilities of different contexts across proficiency levels in the two studied languages are shown in Figure 2. Again, the values are standardised; they reflect the probability of the given linguistic context relative to other contexts among the pauses from the texts written in the same language and are evaluated to reflect the same proficiency level. For instance, looking at A and B levels in the Swedish data, the pauses that occur within a word are a relatively

more typical context for a pause than any other context, as indicated by the higher values of that context than by the other contexts within the A and B levels of Swedish data.

A closer look at the Swedish part of the data in Figure 2 reveals that pauses within words are primarily associated with lower proficiency levels. This relative preference then fades away, and at the C and L1 levels, pauses are, in fact, least likely to occur within words. As for the higher proficiency levels, pauses at the beginning and end of sentences are relatively more typical than C-level L2 writers and L1 writers in other contexts. Pauses between words and between sentences, then, do not indicate any notable difference across proficiency levels.

The Finnish part of the data reveals a partially different trend: most notably, the pauses within words remain the relatively most likely context for pauses among all L2 writers across all proficiency levels. However, the difference among pauses stemming from C-level texts was smaller than those from other levels (Figure 2). This is in stark contrast to L1 writers, for whom pauses within words are drastically less likely than in any other context of occurrence. The pauses at the beginning of a sentence are, then, the most likely context for pauses among L1 writers of Finnish, and the least likely context across all proficiency levels of L2 writers.

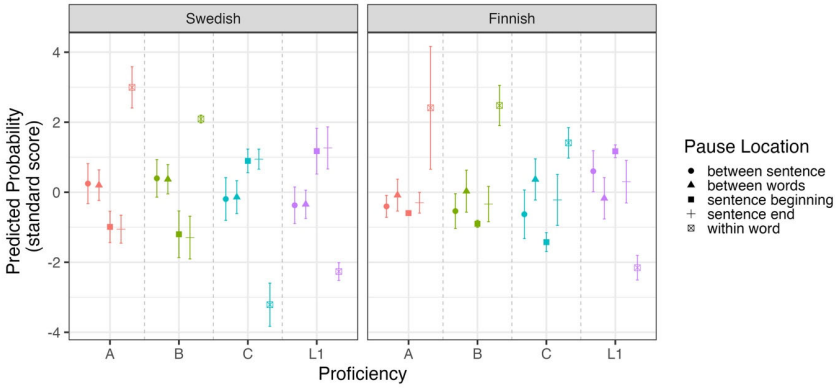

Figure 2: Predicted probability of proficiency level as a function of the linguistic context of pause and language

The linguistic context of pauses has earlier been linked to overall proficiency, whereby more proficient writers process and produce larger units and, consequently, pause between such larger units relatively more often than within such units (e.g. Roca de Larios et al., 2008; Sasaki et al., 2018). In the present study, sentence boundaries served as an operationalisation of such a larger unit, whereas pauses between words suggested smaller processing units and pauses within words, even smaller units. In the case of Swedish, the pauses seem to pattern precisely according to hypothesis 2: pauses within words are the relatively most probable context until proficiency level B, whereas from the C-level

upwards, the relative probability of pauses within words decreases. This tendency is mirrored by the relative probability of pauses at the beginning and end of sentences, further corroborating this interpretation. While it is also interesting to note that pauses between words and pauses between sentences pattern together, one should bear in mind that there were all in all only few pauses between sentences (n = 37 in Swedish data), and that the difference between pauses between sentences and those located in the beginning or at the end of a sentence may be rather due to preferences related to individual writers, and not to linguistic structure.

All in all, contrary to our hypotheses, pauses between words seem to reflect a different kind of relationship with language proficiency and, possibly, a different processing level than pauses within words.

Regarding the Finnish data, the divergence is primarily between L1 and L2 writers and only secondarily between different proficiency levels among L2 writers. We interpret this to stem from the profound typological difference between Swedish and Finnish: a structural mechanism that differs from that of the L1 affects writing even when the product-based language proficiency is extremely high. Furthermore, contrary to Swedish, in the Finnish data, pauses between words tend to pattern together with pauses within words. However, the difference from other linguistic contexts for pauses is less drastic than for pauses within words. In this regard, the results for the Finnish data confirm hypothesis 4.

Tentatively, we interpret the overall imbalance between Swedish and Finnish as L2s to indicate that the two languages differ in their perceived similarity (for the concept, see, e.g. Ringbom 2007, p. 7–8) in that L2 Finnish is perceived structurally to differ more from L1 Swedish than L2 Swedish from L1 Finnish. This may be related to the linguistic circumstances in Finland, for instance, so that the overall language repertoire of the L1 writers of Finnish—most notably English—provides more support for their L2 Swedish writing than the other way around. In addition, the difference may be related to structural differences across languages and the ways in which they interact with phenomena like salience, although this still warrants further investigation. Earlier research has shown that L2 users first tend to focus on open-class words at the expense of grammatical morphemes, so that grammatical morphemes tend to receive more attention later in the learning process (Dietrich et al., 1995, p. 261–280; for an overview on salience in L2, see also Cintrón-Valentin & N. Ellis, 2016).

Hence, typological differences between Swedish and Finnish would suggest that the multitude of bound grammatical morphemes in Finnish might distinguish the L2 Finnish writing process from the L1 writing process, even when the product-based evaluations of proficiency indicate an already exceedingly high level of proficiency. It is also important to note that proficiency-related interpretation and that related to L1 status are not mutually exclusive. While the pauses within words clearly distinguish L1 writers of Finnish from all the L2 writers, the relative differences between contexts for pauses are already much smaller at the C-level than at the lower proficiency levels.

4.4 Pause Location within the Writing Process

There was no statistically significant relationship between proficiency level and pause location within the writing session. However, the two languages observed diverged from one another in a statistically significant fashion (Table 6). The predicted probabilities of the pause location across proficiency levels are shown in Figure 3. The results depict an interesting contrast: In the texts written in Finnish, pauses that occur early in the writing process are relatively more likely to stem from texts by L1 writers. In contrast, the relative probability of the pauses stemming from texts written by L2 writers increases as the writing session progresses; the tendency is similar across the proficiency levels. In contrast, in the texts written in Swedish, there was no difference between different proficiency levels.

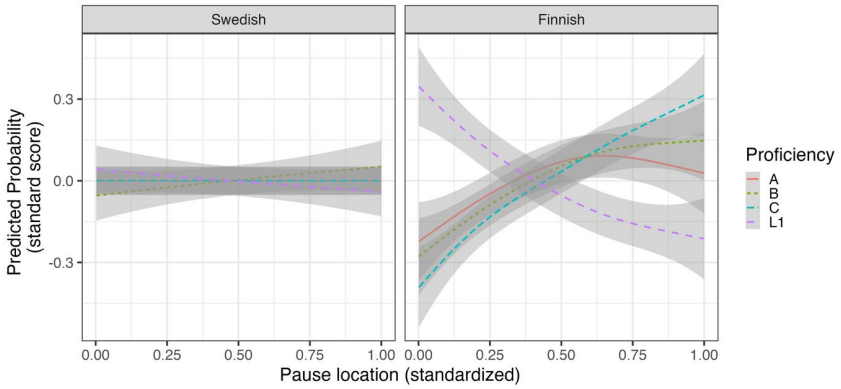

Figure 3. Predicted probability of proficiency level as a function of pause location within the writing session and language

Regarding pauses from texts written in Finnish, the results relate to hypothesis 3. However, they do not fully corroborate it. We hypothesised that the more proficient writers would have relatively more pauses at the beginning and end of the writing session, indicating planning and text revision at a global level. This corresponds to a U-shaped probability curve for pauses stemming from texts evaluated to reflect higher proficiency, while the lower proficiency levels should mirror that trend. However, as shown in Figure 3, the more proficient writers pause more at the beginning of the writing session, but there is no similar effect at the end.

When reviewing pauses in Swedish texts, the results do not support the hypothesis. There is no difference between the probability of pauses across proficiency levels at different points in the writing process.

5. Discussion and Conclusions

This study aimed to increase the understanding of the relationship between pauses during the writing session and the proficiency level of the written product, an argumentative text. We were interested in seeing to what extent these possible associations were language-dependent – in other words, whether L1 and L2 users of Finnish and Swedish as typologically diverging languages diverged from one another.

5.1 Summary of findings

Based on the obtained results, we answer our research questions as follows: The text quality, i.e., the written products' evaluated proficiency, could be modelled relatively well as a function of process-related variables that captured various aspects of pauses during the writing session. Regarding the language proficiency of the writers and their pause behaviour, the results showed that more proficient writers generally had shorter pauses, as suggested by hypothesis 1. Moreover, in line with hypothesis 2, less proficient writers pause relatively more within words, reflecting smaller processing units than pauses corresponding to larger processing units, such as sentence beginnings and sentence ends. These results thus confirm our general assumption that the writer's proficiency conditions L2 writing, which is visible in their pausing behaviour during the writing session.

However, our results only partially supported our hypothesis 3 that product-based proficiency would correlate with pausing behaviour regarding pause location within the writing session. Moreover, in line with hypothesis 4, in many ways our results partially question the cross-linguistic generalizability of these superficially intuitive results. There was a clear difference between L2 writers of Swedish and L2 writers of Finnish regarding their positioning in relation to L1 writers of the respective languages. That is, while the advanced L2 writers of Swedish behaved similarly to the L1 writers of the same language, in Finnish, the behaviour of even the most advanced L2 writers was closer to the less advanced L2 writers than it was to the L1 writers. While we could identify a gradual secondary effect related to proficiency rather than L1 status, it remained considerably weaker, combining even the C-level L2 writers of Finnish with the other L2 writers rather than the L1 writers.

5.2 Implications

We interpret our results to suggest that the similarity—and the difference—across languages may affect L2 writing even more than the proficiency in the L2. To this end, it is crucial to distinguish between actual and perceived similarity, as the former refers to the typological distance between languages and is symmetric in our study design, where the same two languages serve as both L1 and L2. Because neither typological distance nor product-based proficiency explains the difference, we believe the explanation might be partly due to differences in the perceived similarity across the languages.

All L2 writers of Swedish have studied and been otherwise exposed to English, which may facilitate the writing of Swedish as a typologically similar language. In contrast, L2 writers of Finnish did not have such support in their language repertoires. Moreover, the rich suffixal morphology of Finnish might be less salient than the more analytical Swedish, affecting writing even among very proficient L2 writers of Finnish. While the nature of these underlying mechanisms remains open, the results could be interpreted to suggest that L2 writers of Finnish with Swedish as L1 might experience Finnish as being more different from their L1 than L2 writers of Swedish with Finnish as L1 do with Swedish. Notably, such potential asymmetries should somehow be accounted for in any model of L2 writing for such a model to be cross-linguistically generalisable.

Turning to other unexpected results, we find it interesting that while the pauses located within words were robust indicators of language proficiency in Swedish, as suggested in earlier research, pauses between words did not attest to the same effect. The two types of pauses have often been treated as indicators of smaller processing units. In these data, this was the case only with Finnish data, whereas among writers of Swedish, the pauses between words instead grouped with pauses between sentences. We believe this might be partly due to the conservative threshold of pause length of 2000 ms used in the present study, but the observation warrants further investigation. Similarly, we were surprised by the observation that in Swedish, the pauses between sentences do indeed pattern differently from pauses at the beginning or the end of a sentence. Because this difference is clear and at least superficially challenges hypothesis 2 related to pause locations defining processing units, we raise it for any future enquiries to consider.

The effect of pause length threshold requires further investigation in terms of the exact nature of this effect. Many studies, including this one, have operationalised them by using the threshold level of 2000 ms to distinguish between cognitive pauses—typically the primary research interest—and technical pauses related to typing as a production medium. Others have employed a considerably lower threshold (e.g. in Révész et al., 2022, the threshold is 200 ms), while yet others (e.g. Van Waes & Leijten, 2015) have assessed the effect of various thresholds as a part of more extensive profiling of different facets of writing fluency.

A core argument in favour of lower thresholds is that it makes it possible to capture lower-level writing units that would otherwise remain unnoticed. As we see it, this is closely related to the overall aim of the study. In this study, the difference in the proportions of pauses occurring within a word—the prime example of an indicator of lower-level writing units—was highly effective in distinguishing between the different proficiency levels, suggesting that a lower threshold would not be needed for this purpose.

5.3 Limitations

The most apparent limitation of this study is that we deliberately avoided digging deeper into the linguistic nature of the pause locations. It is doubtful that pauses within words or

between words are alike. We expect the differences to be related to the grammatical nature of the pause locations in the analysed languages. Pauses can, for example, occur within a noun or adpositional phrase, between phrases or between clauses of distinct types. This is also one of the contexts where the typological differences between the two languages can be observed: Swedish has both indefinite and definite articles and adverbial functions are typically expressed by means of prepositions (e.g. ***vid den grön-a älv-en*** 'by the green-DEF river-DEF'). Finnish, in turn, has no articles and adverbial functions can be expressed either by means of case marking (***vihreä-llä joe-lla***, 'green-ADE river-ADE' for 'by the green river') or as adpositional phrases (***vihreä-n joe-n ääre-llä*** 'green-GEN river-GEN by-ADE' for 'by the green river'). Because of these differences, there are proportionally less words – and hence, proportionally less possible between-word contexts – in nominal phrases in Finnish than in Swedish. Furthermore, pause locations within words can, for instance, occur on morpheme boundaries or within morphemes – and also sheer amount of such contexts differs across Finnish and Swedish (as can be seen in the above examples). The question is, how does this all correspond to the proficiency of the writers. Either of these perspectives is worthy of an independent piece of research and cannot be included in one study.

Another limitation is that we did not explore the potentially and probably structured relationship between pause lengths and the linguistic contexts of pauses. Such relationships have been reported in earlier studies (e.g. Révész et al., 2022; Van Waes & Leijten, 2015;), and we have no reason to believe that would not be the case in our data. However, as our study design was already arguably very complex—with two languages studied from both L1 and L2 perspectives and with a focus on interactions between the languages and included variables—we decided to refrain from an even more complex design. Keeping in mind that L2 English is the most researched language, even in studies concerning the writing process, this study offers a broader perspective on writing process research by analysing pause behaviour in two typologically different languages, Finnish and Swedish.

5.4 Methodological Contribution

Finally, we want to emphasise that our research design comprises two languages. Consequently, the results between them are at least internally fully comparable, and similar tendencies found in both language pairs provide further evidence in favour of the chosen threshold's suitability for the aims of this study. However, it is impossible to say how the study's overall results would have changed if we included shorter pauses. This could have affected the proportions of different contexts of pauses, and it would undoubtedly have affected the distribution of average pause lengths. But how exactly? Instead of speculating on these possibilities, we invite further researchers—ourselves included—to focus specifically on this effect by running fully parallel experiments with only changing the threshold level of the pause.

Regarding this study's methodological contribution, we offer two points: cross-linguistic generalisability and the ecological validity of the research design. First, the observed differences between two typologically diverging languages when all learner-related and contextual background variables were constant suggest that a broader range of different languages should be considered—both as L1 and as L2. In our opinion, language-agnostic tendencies of L2 writing can only be revealed when results studying different language combinations in comparable research designs corroborate comparable results.

Second, we drew attention to the type of statistical model applied: Language proficiency is by nature an incremental yet holistic construct. Hence, we approached it using a mixed-effects ordinal logistic regression, where each pause above the threshold level constituted a unit of observation. Moreover, the proficiency evaluation of the text where the pause served as the response variable, and various measures related to the pause, the writing session, and the writer served as the predicting variables. Such a design arguably tightens the connection between overall proficiency and process-related observations while controlling for dependencies across observations from the same participants.

Notes

1. Even though the evaluations were conducted at the text level, in the interest of readability, we often refer to proficiency of the writers.
2. We have excluded process-initial pauses and those associated to either deleting or navigating within the produced text. Furthermore, we have treated paragraph boundaries as sentence boundaries.
3. Here, we follow the established convention in multivariate modelling and use the term ***effect*** to refer to the relationships between the response variable and the predicting variables. The term does not imply any causal relationship between the variables.
4. All the scripts used in the statistical analyses, as well as the anonymized data, are available here: https://osf.io/n2kux/

References

Act on the Matriculation Examination (12.4.2019/502). Available at: https://www.finlex.fi/en/laki/kaannokset/2019/20190502 (Consulted October 29, 2024).

Alamargot, D., Dansac, C., Chesnet, D., & Fayol, M. (2007). Parallel processing before and after pauses: A combined analysis of graphomotor and eye movements during procedural text production. In M. Torrance, L. Van Waes, & D. Galbraith (Eds.), ***Writing and cognition: Research and applications*** (pp. 13–29). Amsterdam: Elsevier. https://doi.org/10.1163/9781849508223_003

Alanen, R., Huhta, A., & Tarnanen, M. (2010). Designing and assessing L2 writing tasks across CEFR proficiency levels. In I. Barting, M. Martin & I. Vedder (Eds.), ***Communicative proficiency and***

linguistic development: intersections between SLA and language testing research. EUROSLA MONOGRAPHS SERIES. (pp. 21–56). European Second Language Association.

Barkaoui, K. (2016). What and When Second-Language Learners Revise When Responding to Timed Writing Tasks on the Computer: The Roles of Task Type, Second Language Proficiency, and Keyboarding Skills. *The Modern Language Journal*, *100*(1), 229–400. https://doi.org/10.1111/modl.12316

Barkaoui, K. (2019). What can L2 writers' pausing behavior tell us about their L2 writing processes? *Studies in Second Language Acquisition*, *41*, 529–554. https://doi.org/10.1017/S027226311900010X

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Lawrence Erlbaum.

Chenoweth, N., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, *18*, 80–98. https://doi.org/10.1177/0741088301018001004

Chenu, F., Pellegrino, F., Jisa, H., & Fayol, M. (2014). Interword and intraword pause threshold in the writing of texts by children and adolescents: A methodological approach. *Frontiers in Psychology*, *5*(182), 1–7. https://doi.org/10.3389/fpsyg.2014.00182

Christensen, R. H. B. (2022). ordinal—Regression Models for Ordinal Data.

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H-H. (2019). Combined deployable keystroke logging and eye-tracking for investigating L2 writing fluency. *Studies in Second Language Acquisition*, *41*(3), 583–604. https://doi.org/10.1017/S027226311900007X

Cintrón-Valentín, M. C., & Ellis, N. C. (2016). Salience in Second Language Acquisition: Physical Form, Learner Attention, and Instructional Focus. *Frontiers in Psychology 7*. 1284. https://doi.org/10.3389/fpsyg.2016.01284

Cislaru, G., & Olive, T. (2018). *Le processus de textualisation : analyse des unités linguistiques de performance écrite [The textualization process : Analysis of written performance units].* De Boeck. https://doi.org/10.3917/dbu.cisla.2018.01

Council of Europe (Ed.), (2020). *Common European framework of reference for languages: learning, teaching, assessment – companion volume*. Strasbourg: Council of Europe Publishing. http://www.coe.int/lang-cefr.

Damian, M. F., & Stadthagen-Gonzalez, H. (2009). Advance planning of form properties in the written production of single and multiple words. *Language and Cognitive Processes*, *24*(4), 555–579. https://doi.org/10.1080/01690960802346500

Dietrich, R., Klein, W., & Noyau, C. (1995). *The Acquisition of Temporality in a Second Language.* John Benjamins. https://doi.org/10.1075/sibil.7

Durrant, P. (2013). Formulaicity in an agglutinating language: the case of Turkish. – *Corpus Linguistics and Linguistic Theory*, *9*(1), 1–38. https://doi.org/10.1515/cllt-2013-0009

Finnish National Agency for Education (2014). *Perusopetuksen opetussuunnitelman perusteet* 2014 [The National Core Curriculum for Basic Education 2014]. Finnish National Agency for Education.

Flower, L., & Hayes, J. R. (1980). A Cognitive Process Theory of Writing. *College Composition and Communication*, *32*(4), 365–387. https://doi.org/10.2307/356600

Gánem-Gutiérrez, A., & Gilmore, A. (2018). Tracking the real-time evolution of a writing event: Second language writers at different proficiency levels. *Language Learning*, *68*(2), 469–506. https://doi.org/10.1111/lang.12280

Gries, S. Th. (2021). *Statistics for Linguistics with R.* 3rd edition. De Gruyter Mouton. https://doi.org/10.1515/9783110718256

Guo, H., Deane, P. D., van Rijn, P. W., Zhang, M., & Bennett, R. E. (2018). Modelling Basic Writing Processes from Keystroke Logs. *Journal of Educational Measurement*, *55*(2), 194–216. https://doi.org/10.1111/jedm.12172

Hall, S., Baaijen, V.M., & Galbraith, D. (2024). Constructing theoretically informed measures of pause duration in experimentally manipulated writing. *Reading and Writing*, *37*, 329–357. https://doi.org/10.1007/s11145-022-10284-4

Kellogg, R.T. (1996). The model of Working Memory in Writing. In C. M. Levy & S. Ransdell (Eds.) *The Science of Writing: Theories, Methods, Individual Differences and Applications* (pp. 57–72). Routledge.

Kellogg, R.T., Whiteford, A.P., Turner, C.E., Cahill, M., & Mertens, A. (2013). Working Memory in Written Composition: An Evaluation of the 1996 Model. *Journal of Writing Research*, *5*(2), 159–190. http://dx.doi.org/10.17239/jowr-2013.05.02.1

KISUVI (2022–2025). Multilingual writers' writing processes: graph-theory based visualisation of formulaic sequences and fluency patterns. Available at: https://sites.utu.fi/kisuvi/en/ (Consulted October 21, 2024).

Kowal, I. (2014). Fluency in Second Language Writing: A Developmental Perspective. *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, *131*, 229–246. https://doi.org/10.4467/20834624SL.14.013.2321

Language Act (6.6.2003/423). Available at: https://www.finlex.fi/en/laki/kaannokset/2003/en20030423 (Consulted October 29, 2024).

Leblay, C., & Caporossi, G. (2015). A graph theory approach to online writing data visualization. In G. Cislaru (Ed.), *Writing(s) at the Crossroads: The Process-Product Interface* (pp. 171–181). John Benjamins. https://doi.org/10.1075/z.194.09leb

Leijten, M., Van Waes, L., Schrijver, I., Bernolet, S., & Vangehuchten, L. (2019). Mapping master´s students´ use of external sources in source-based writing in L1 and L2. *Studies in Second Language Acquisition*, *41*(3), 555–582. https://doi.org/10.1017/S0272263119000251

Lindgren, E., Knospe, Y., & Sullivan, K. P. H. (2019). Researching Writing with Observational Logging Tools from 2006 to the Present. In E. Lindgren & K. P. H. Sullivan (Eds.), *Researching Writing with Observational Logging Tools from 2006 to the Present* (pp. 1–29). Leiden: Brill. https://doi.org/10.1163/9789004392526_002

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Matsuhashi, A. (1981). Pausing and planning: the tempo of written discourse production. *Research in the Teaching of English*, *15*(2), 113–134. https://doi.org/10.58680/rte198115773

Matsuhashi, A. (1982). Explorations in the real-time production of written discourse. In M. Nystrand (Ed.), *What writers know: the language, process, and structure of written discourse.* (pp. 269–290). Academic Press. https://doi.org/10.1163/9789004454118_015

Matsuhashi, A. (1987). *Writing in real time: modelling production processes.* Ablex.

Michel, M., Révész, A., Lu, X., Kourtali, N-E., Lee, M., & Borges, L. (2020). Investigating L2 writing processes across independent and integrated tasks: A mixed-methods study. *Second Language Research*, *36*(3), 307–334. https://doi.org/10.1177/0267658320915501

Mohsen, M. A. (2021). L1 versus L2 writing processes: What insight can we obtain from a keystroke logging program? *Language Teaching Research*, 1–25. https://doi.org/10.1177/13621688211041292

Mohsen, M.A., & Qassem, M. (2020). Analyses of L2 learners' text writing strategy: Process-oriented perspective. *Journal of Psycholinguistic Research*, *49*, 435–451. https://doi.org/10.1007/s10936-020-09693-9

Mutta, M. (2017). Pausal Behavior in the Writing Processes of Foreign and Native Language Writers: The Importance of Defining the Individual Pause Length. In S. Plane, C. Bazerman, F. Rondelli, C. Donahue, A. N. Applebee, C. Boré, P. Carlino, M. M. Larruy, P. Rogers & D. R. Russell (Eds.),

Research on Writing: Multiple Perspectives. (pp. 511– 529). The WAC Clearinghouse. https://doi.org/10.37514/INT-B.2017.0919.2.28

R Core Team. 2022. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Révész, A., Kourtali, N., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, *67*(1), 208–241. https://doi.org/10.1111/lang.12205

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. https://doi.org/10.1017/S027226311900024X

Révész, A., Michel, M., Lu, X., Kourtali, N., Lee, M., & Borges, L. (2022). The relationship of proficiency to speed fluency, pausing, and eye-gaze behaviours in L2 writing. *Journal of Second Language Writing*, *58*. https://doi.org/10.1016/j.jslw.2022.100927

Ringbom, H. 2007. *Cross-linguistic Similarity in Foreign Language Learning.* Multilingual Matters. https://doi.org/10.21832/9781853599361

Roca de Larios, J., Manchyn, R., Murphy, L., & Marнn, J. (2008). The foreign language writer's strategic behaviour in the allocation of time to writing processes. *Journal of Second Language Writing*, *17*(1), 30–47. https://doi.org/10.1016/j.jslw.2007.08.005

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, *9*(3), 259–291. https://doi.org/10.1016/S1060-3743(00)00028-X

Sasaki, M. (2002). Building an empirically-based model of EFL learners' writing processes. In G. Rijlaarsdam, S. Ransdell, & M. Barbier (Eds.), *Studies in writing, volume 11: New directions for research in L2 writing* (pp. 49–80). Kluwer. https://doi.org/10.1007/978-94-010-0363-6_3

Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning*, *54*(3), 525–582. https://doi.org/10.1111/j.0023-8333.2004.00264.x

Sasaki, M., Mizumoto, A., & Murakami, A. (2018). Developmental trajectories in L2 writing strategy use: A self-regulation perspective. *The Modern Language Journal*, *102*(2), 292–330. https://doi.org/10.1111/modl.12469

Schilperoord, J. (1996) The distribution of pause time in written text production. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Current research in writing: theories, models and methodology.* (pp. 21–35). Amsterdam University Press. https://doi.org/10.1163/9789004458598_004

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. In R. M. Manchyn (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 77–101). Multilingual Matters. https://doi.org/10.21832/9781847691859-007

Shen, C., & Chen, N. (2021). Profiling the pausing behaviour of EFL learners in real-time computer-aided writing: a multi-method case study. *Asian-Pacific Journal of Second and Foreign Language Education*, *6*(15), 26 pages. https://doi.org/10.1186/s40862-021-00118-1

Spelman Miller, K. (2000). Academic writers on-line: Investigating pausing in the production of text. *Language Teaching Research*, *4*(2), 123–148. https://doi.org/10.1177/136216880000400203

Tian, y., Kim, M., & Crossley, S. (2024). Making sense of L2 written argumentation with keystroke logging. *Journal of Writing Research*, *15*(3), 435-461. https://doi.org/10.17239/jowr-2024.15.03.01

Van Waes, L., & Leijten, M. (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, 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

Vasylets, O., & Marin, J. (2025) Linguistic and behavioral alignment in writing: A scoping review. *Journal of Writing Research*, *16*(3), 377-406. https://doi.org/10.17239/jowr-2025.16.03.02

Wirtz, J. (2025). Written performance indicators from experienced creative writers. *Journal of Writing Research*, *16*(3), 499-520. https://doi.org/10.17239/jowr-2025.16.03.06

Wengelin, E. (2006). Examining pauses in writing: Theory, methods and empirical data. *Computer Key-Stroke Logging and Writing: Methods and Applications* 18, 107–130. https://doi.org/10.1163/9780080460932_008

Whittingham, M. J., Stephens, P. A., Bradbury, R. B., & Freckleton, R. P. (2006). Why Do We Still Use Stepwise Modelling in Ecology and Behaviour? *The Journal of animal ecology*, *75*(5). 1182–1189. https://doi.org/10.1111/j.1365-2656.2006.01141.x

Winter, B., & Grice, M. (2021). Independence and generalizability in linguistics. *Linguistics*, *59*(5). 1251–1277. https://doi.org/10.1515/ling-2019-0049

Xu, C. (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

Xu, C., & Xia, J. (2019). Scaffolding process knowledge in L2 writing development: Insights from computer keystroke log and process graph. *Computer Assisted Language Learning*, *34*(4), 583–608. https://doi.org/10.1080/09588221.2019.1632901

Yiğitoğlu, N., & Reichelt, M. (2019). Introduction: L2 Writing in Non-English Languages: Toward a Fuller Understanding of L2 Writing. In N. Yiğitoğlu & M. Reichelt (Eds.), *L2 Writing Beyond English* (pp. xiii–xix). Multilingual Matters. https://doi.org/10.21832/9781788923132-003

Zhang, M., Bennett, E.R., Deane, P., & van Rijn, P. (2019). Are there gender differences in how students write their essays? An analysis of writing processes. *Educational Measurement: Issues and Practice*, *38*(2), 14–26. https://doi.org/10.1111/emip.12249