Fluency in writing : a multidimensional perspective on writing fluency applied to L1 and L2

Reference:
Van Waes Luuk, Leijten Mariëlle.- Fluency in w riting : a multidimensional perspective on w riting fluency applied to L1 and L2
Computers and composition - ISSN 8755-4615 - 38(2015), p. 79-95
Full text (Publisher's DOI): https://doi.org/10.1016/J.COMPCOM.2015.09.012
To cite this reference: https://hdl.handle.net/10067/1283110151162165141

Fluency in Writing:

A Multidimensional Perspective on Writing Fluency Applied to L1 and L2

Luuk Van Waes° & Mariëlle Leijten°*

° University of Antwerp, Belgium

* Research Foundation - Flanders, Belgium

Abstract

Some texts are easy to write, others are a real struggle. This article presents a brief review of how fluency has generally been measured in writing research. In addition to these 'traditional' measures, we define a wide range of complementary measures that might be diagnostic of fluency, by taking also more process-related characteristics into account.These complementary measures are derived from keystroke logging data, which were collected from an experiment among 68 students who wrote two descriptive texts, one in their mother tongue and the other in their second language. By using correlation and principal component analyses, we have reduced the set of variables and created a new multidimensional model to better address the complexity of fluency in writing. This model consists of four dimensions: (a) production, (b) process variation, (c) revision, and (d) pause behavior. These fourcomponents together create a multidimensional perspective on writing, which enables us todifferentiate between fluent and less fluent writers.

Keywords: writing fluency; keystroke logging; cognitive processes; multiple sources

Contact person

Luuk Van Waes
University of Antwerp
Department of Management
Prinsstraat 13
2000 Antwerpen
+32 3 265 51 59
luuk.vanwaes@uantwerpen.be

Second author

Mariëlle Leijten
marielle.leijten@uantwerpen.be

1 Introduction

Some texts hardly require any effort to write, whereas other texts—or even sentences—are the result of an intense struggle, with the blinking cursor and the backspace key as leading actors. Most of us also experience "writing fluency" in a foreign language as more problematic than writing in our first language. But, what exactly are the underlying concepts ofwriting fluency? Does it mean that some writers are able to finish a text of 500 words in a shorter period of time than others? That theyproduce more characters or words in a comparable amount of time? That theypause less during writing? Or that theydo not need to revise or edit theirfirst draft as intensively?

Fluency has been on the research agenda of language researchers for many years. Especially in speech and reading studies, the concept is quite clearly defined (e.g., Bosker, Pinget, Quené, Sanders, & de Jong, 2013; Segalowitz, 2010; Skehan, 2003, 2009). In writing, however, the definition of the concept is more vague. In contrast to spoken language, fluency in writing has no rhetoric function as such. If a person hesitates in phrasing a sentencein oral communication, or interrupts his or herdiscourse for a longer time, it will certainly be noticed by the listener and will probably influence the interaction. However, pausing during text production does not influence the reader-writer interaction, since a printed text normally does not (explicitly) reveal that the writer has paused at a certain instance in the text.

In writing research, fluency has been the topic of a myriadof studies, which focus for instance on developmental writing (Berninger, Cartwright, Yates, Swanson, & Abbott, 1994; McCutchen, Covill, Hoyne, & Mildes, 1994), writing modes, juxtaposing oral andwritten modes or handwriting andtyping (Olive, Favart, Beauvais, & Beauvais, 2009; Shanahan, 2006; Van Horenbeeck, Pauwaert, Van Waes, & Leijten, 2012), and especially L1 and L2 writing (Chenoweth & Hayes, 2001; Johnson, Mercado, & Acevedo, 2012; Kobayashi & Rinnert, 2013; Kormos, 2012; Latif, 2012; Lindgren, Sullivan, & Spelman Miller, 2008; Ong, 2014; Ong & Zhang, 2010; Snellings, Van Gelderen, & De Glopper, 2002; Tillema, 2012). In most of these studies, a distinctionis made between two or more groups of participants (e.g., 5th graders vs. 9th graders, L1 vs. L2), or in a within participants design thekind of tasks (e.g., narrative vs.argumentative tasks) or writing modes (e.g., handwriting vs.keyboarding) are compared. The relation between fluency and quality is also a recurring topic, although no clear conclusions could be drawn so far (see, for instance, Snellings et al., 2002; Yan et al., 2012).

In this article, we aim at describing an integrated approach to fluency, combining various perspectives. The approach that we present in this article is based on keystroke logging observations in a writing study that addressed the difference in fluency between L1 and L2 using two writing tasks.Starting from existing fluency measures,we will bring together traditionalprocess and product measures, both in isolation and in combination with each other. Moreover, we will introduce new perspectives at different levels, byfocusing on writer characteristics (individual measures), writing tasks (specific genres) and writing contexts (use of tools and sources). In a first stage, we will explore a wide range of potentialfluency indicators. This resultsin a list of about 200 variables. Subsequently, we will reduce the number of indicators using correlation and Principal Component Analyses (PCA) to finally reach a comprehensible set of ten fluency indicators, grouped in four underlying categories. Our mainaim is to construct a manageable set of components and variables that enable us to describe and measure writing fluency from a multidimensional perspective. In our point of view, separate measures and sub-dimensions of fluency are needed to painta more comprehensive and fine-grained picture of fluency performance in writing.

2 Fluency in speech and writing

Most language users experience and even get frustrated by the fact that they are less fluent in L2 than in L1, even if they master a foreign language at a high proficiency level. To get a better understanding of what this fluency gap underlies and how to overcome it, a large number of studies have been set up, both in oral and in written communication (see also, Segalowitz, 2010 for a comprehensive review).

In the development of measures to define fluency,writing researchershave largely built on insights from speech studies. Researchers in this research have long been concerned with identifying a large variety of critical features to adequately measure oral fluency (e.g., Goldman-Eisler, 1961; Lennon, 1990). Kormos (2006, p. 163), for instance, provides a summary of the most frequently used measures of fluency in oral studies and ends up with ten measures, which can be classified as follows: (a) Pauses (e.g., total pause time, silent and filled pauses per minute, length of pauses), (b) Disfluencies (i.e., breakdown of fluency, indicated by e.g., repetitions, repairs), (c) Rate (e.g., speech and articulation rate; length of runs), and (d) Pace and Stress (e.g., number and proportion of stress words) (see also Towell et al. (1996), Koponen & Riggenbach (2000) and Skehan (2003)). Iwashita et al. (2008) used a comparable set of variables as a starting point to better define the relationship between proficiency and fluency. Their analysis of more than 200 recordings of L2 speakers with different proficiency levels showed that total pause time, silent pause rate and speech rate correlate the highest with proficiency level.

The main findings on fluency in writing studies are very comparable. It is now commonly accepted that fluent writing processes are characterized by short pausing times, few revisions and a high production rate (MacArthur, Graham, & Fitzgerald, 2008).The use of the criterion last mentioned is particularlyfrequent in writing research. Kellogg's studies (1996, 2004) are typical examples of this approach to fluency. He demonstrated that initial planning leads to a decrease of cognitive effort in the transcription phase, positively influencing fluency. In his research,Kellogg refined the criterion of production rate by introducing two measures of production fluency: *Fluency I* refers to the mean number of words in the transcription phase (i.e., total time on task minus initial planning time); *Fluency II*is calculated on the basis of the gross time on task. This theme has inspired a lot of follow-up studies, and Kellogg'sfluency measures have beenwidely adopted (Graham & Perin, 2007; Johnson et al., 2012; Leijten, Van Waes, & Ransdell, 2010; Snellings et al., 2002; Van Waes, Leijten, & Quinlan, 2010). In general, the Fluency II-approach has resulted in more comprehensible and interpretable results, with 'number of word per minutes' being the most well-known measure of writing fluency nowadays(see also, Chenoweth & Hayes, 2001; Wolfe-Quintero, Inagaki, & Kim, 1998).

Recent introductions of more advanced observation and analysis tools have had an impact on the kind of fluency measures that are being used, primarily because they make it easier to calculate more complex indicators. In speech studies, software programs such as *Praat*(De Jong & Wempe, 2009) allow researchers to automatically obtain complex fluency measures from larger corpora. In writing studies,improved keystroke logging tools —and particularlytheir analyzing tools—make it possible to focus on more fine-grained writing-process data (Leijten & Van Waes, 2013; Sullivan & Lindgren, 2006;

Van Waes, Leijten, Wengelin, & Lindgren, 2012). Keystroke logging studies have shown, for instance, that not only the number of words per minute in the final text is an interesting measure of fluency, but also the number of words *produced* per minute (including revised words and characters). The former is a *product-based* measure, whereas the latter takes the *process* of writing into account, including revisions. Taking on a process perspective, fluencycan also be interpreted as the average speed between two characters of a word, i.e.,the interkey-transition times or pauses, which is a measure that can also be easily derived from keystroke logging(Johansson, Wengelin, Johansson, & Holmqvist, 2010; Lindgren, Sullivan, & Miller, 2012; Wengelin, 2006).

Another concept of fluency, viz. 'length of run' (Kormos, 2006; Skehan, 2003), is also transferred from speech to (hand)writing research. In writing research production bursts are measured by number of bursts and average burst length(Alves, Castro, & Olive, 2008; Chenoweth & Hayes, 2001; Olive, Alves, & Castro, 2009). Writing processes that are more often interrupted by pauses and/or revisions (respectively P-bursts and R-bursts) are considered to be less fluent than processes that are less frequently interrupted. Chenoweth and Hayes (2001, 2003; 2006) introduced bursts in a series of studies in which they compared L1 and L2 writing processes. They showed that burst length "is fundamentally related to fluency", since it reflects "the capacity of the translator to handle complex language structures" (p. 94). The length of pause—and revision—bursts drops significantly when(young) writers compose text in a second language (L2), as opposed to text production in their mother tongue.

In thewriting research mentioned above, a pause threshold of 2 secondsis most frequently used to define pause bursts. However, a discussion has started recently about the appropriateness of this threshold(Sullivan & Lindgren, 2006; Wengelin, 2006). Researchers point to the importance of inter and intrapersonal variation and the need to define a more robust fluency measure in relation to specific research goals(Leijten, De Maeyer, & Van Waes, 2011; Van Waes & Leijten, 2011; Van Waes et al., 2012; Wengelin, 2006).This discussion in writing research is quite comparable to the situation in speech studies,where, for instance, silent pauses have been defined ranging from 200 to 1000 milliseconds (ms). This makes it rather complex to compare fluency results across different studies (Segalowitz, 2010, p. 39). We will therefore explore in the current study other, and smaller pause thresholds in combination with writers' and task characteristics. As studies that focus on this kind of individual differences in writing are still scarce (Grabowski, 2008; Kormos, 2012), we will try to propose a more balanced approach to define writing disfluencies by explicitly addressing the role of individual differences and the task context.

In sum, the current study aims at approaching fluency from different perspectives using keystroke logging data. We contend that it is impossible to characterize a person's fluency in a unidimensional fashion. An integrated, multidimensional perspective is needed. Therefore, we start by summing up and describing the measures commonly used and then gradually add more process-related and writer/task-related characteristics. How stable is a writer in his or her writing task? What is his or her 'maximal writing speed'? How fluent is a particular writing process in relation to this maximum? Finally, we will propose an integrated view of fluency, based on a limited yet varied set of indicators, using statistical techniques like Correlation and Principal Component Analysis (PCA).

3 Fluency revisited

In order to define a wide range of fluency indicators, a quasi-experimental study was set up in which the participants produced two expository texts: one in L1 and the other in L2. The collected data were analyzed both from a product and a process perspective.

3.1 Participants

The study was conducted at the University of Antwerp among 84 students of the Master of Multilingual Professional Communication.The master students took an advanced course in one or more foreign languages (English, French, Spanish and/or German),in additionto courses in their mother tongue (Dutch). The students wereallowed to enroll in a language course if their level was B2 or higheraccording to the Common European Framework of Reference for Languages (CEFR). In other words these students are considered to be independent, upperintermediate L2 language users who "can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options" (Council of Europe, 2011). The group consisted of 67 female and 17 male students. The mean age of the participants was 22.55 (*SD* = 1.52).

3.2 Writing tasks

The participants each wrote two texts, one in L1 and the otherin L2. The writing task consisted of a simple expository task (knowledge telling) to create a situation in which the participants could write as fluently as possible, with as little cognitive loadas possible(Bereiter & Scardamalia, 1987).The participants were invited to write a short text about a recent experience, viz.,'describe your last holiday' and 'describe your last weekend'. For each story a few clues were given: 'where did you go to?';'with whom?';'what did you do?'.Given the limited cognitive complexity of these tasks, it is easier to attribute differences in text production to the language factor (Ong & Zhang, 2010).

3.3 Design and procedure

The data were collected in a computer classroom during three separate sessions, each with approximately30 students participating. The experiment startedwith anoral instruction about the organization of the experiment, after which a brief written instruction was distributed about the two writing tasks. A within-subject design was used, in which each student wrote both a text in L1 and in L2. The task-order and the language-order were both counter-balanced in a Latin square design.

The participants first read the assignment and then had two minutes of preparation time to mentally plan what they wanted to write. The writing task itself tookeight minutesmaximum. Two minutes prior tothe end,the participants received an auditory signal. The post-hoc analysis showed no difference between time-on-task,neither with respect to the task (T1: 7m43s (32*s*) vs. T2: 7m53s (*24s*); $F(1,134)=0.20$, $p=.889$)or the language (L1: 7m46s (*29s*) vs. L2: 7m50s (*28s*); $F(1,67)=0.91$, $p=.343$).

Between the two writing tasks, we included a distraction task, viz., a visual estimation task to test numerical intuition (see Halberda, Mazzocco, & Feigenson, 2008). The session ended with a questionnaire about the students' 'language context', typing skills and some questions about the way they experienced the two tasks. This information was partly used to exclude some participants from the final data set: e.g., by taking into account reported language or learning disabilities, familiarity with the Belgian azerty-keyboard, and not havingDutch as L1.

3.4 Data coding and analysis

The process data in the current studywerecollected with Inputlog 5 (www.inputlog.net - Leijten & Van Waes, 2013). Inputlog is a keystroke-logging program that has been developed to collect all keyboard and mouse activities during text production. Since this program also addsa time stamp (ms) to each logging event, the resulting log files can be analyzed to explore the dynamics of the writing process from different perspectives. Inputlog offers by default some predefined analyses. In the context of this study, mainly the so-called general, summary and pause analyses were used(Leijten & Van Waes, 2014), which were generated using variouspause thresholds and interval lengths (see infra).

Asa first step in the data-analysis, we checked the technical quality of the logged data (in combination with the selection criteria derived from the questionnaire, cf. supra). Takingcompleteness and an error-free log diagnosis as selection criteria, the total dataset was subsequently reduced to 136 log files (68 for each task). These log files were filtered individually (usingthe *pre-processing event filter* of Inputlog) to remove noise events at the beginning and the end of each process, delimiting the analyses to actual text production, i.e., from the first to the last character (or revision). The resulting data set was used as input for the following analyses:

- *General analysis*: Apart from some metadata, this analysis represents the logging process as a fine-grained dataset containing all keystrokes, mouse movements and their respectivetime stamps in linear order. See Table 1 for an example: here, the person typed "[On] Saturday ~~we~~ night I (...)". Each line in Table 1 represents a keystroke (or mouse action), which is documented with position and time-based information.
- *Summary analysis*: This analysis provides a statistical summary ofaggregated data. For example: total process time, total pausing time, number of and mean pausetime, number and length of P-bursts. In order to compare the inter and intrapersonal variation, we have conducted the summary analysis usingfivedifferent pause thresholds: 0, 500, 1000, 2000 and 5000 ms. This implies that only those pauses that are longer than the defined threshold are taken into account in the analysis run.

- *Pause analysis*: This is a summarizing representation of the pausing behavior in relation tothe number of pauses, mean pause length, the location of pausesin the writing task (e.g., within words, between words) and the phase in the writing task (e.g.,intervals, which means that each process has been divided into ten comparable intervals).

Finally, the analyses files were merged in the *post-processing module*of Inputlog. Additional variables (e.g., related to revision) were calculated semi-automatically to complement the spectrum of product and process measures.

Table 1
Example of output as generated by the General analysis module in Inputlog

#Id	Event Type	Output	Position	Character Production	StartTime	StartClock	EndTime	EndClock	ActionTime	PauseTime	PauseLocation
27	keyboard	S	3	4	23697	0:00:23	23743	0:00:23	46	203	WITHIN WORDS
28	keyboard	a	4	4	24024	0:00:24	24102	0:00:24	78	327	WITHIN WORDS
29	keyboard	t	5	5	24087	0:00:24	24149	0:00:24	62	63	WITHIN WORDS
30	keyboard	u	6	6	24664	0:00:24	24711	0:00:24	47	577	WITHIN WORDS
31	keyboard	r	7	7	24820	0:00:24	24898	0:00:24	78	156	WITHIN WORDS
32	keyboard	d	8	8	25038	0:00:25	25101	0:00:25	63	218	WITHIN WORDS
33	keyboard	a	9	9	25288	0:00:25	25350	0:00:25	62	250	WITHIN WORDS
34	keyboard	y	10	10	25803	0:00:25	25865	0:00:25	62	515	WITHIN WORDS
35	keyboard	SPACE	11	11	26224	0:00:26	26302	0:00:26	78	421	AFTER WORDS
36	keyboard	w	12	12	26723	0:00:26	26801	0:00:26	78	499	BEFORE WORDS
37	keyboard	e	13	13	26988	0:00:26	27051	0:00:27	63	265	WITHIN WORDS
38	keyboard	BACK	14	14	27394	0:00:27	27441	0:00:27	47	406	WITHIN WORDS
39	keyboard	BACK	13	15	27534	0:00:27	27597	0:00:27	63	140	REVISION
40	keyboard	n	12	15	29063	0:00:29	29110	0:00:29	47	1529	WITHIN WORDS
41	keyboard	i	13	15	29453	0:00:29	29500	0:00:29	47	390	WITHIN WORDS
42	keyboard	g	14	16	29827	0:00:29	29890	0:00:29	63	374	WITHIN WORDS
43	keyboard	h	15	17	29999	0:00:29	30061	0:00:30	62	172	WITHIN WORDS
44	keyboard	t	16	18	30311	0:00:30	30373	0:00:30	62	312	WITHIN WORDS
45	keyboard	SPACE	17	19	30685	0:00:30	30748	0:00:30	63	374	AFTER WORDS
46	keyboard	LSHIFT	18	20	31029	0:00:31	31294	0:00:31	265	344	BEFORE WORDS
47	keyboard	I	18	21	31200	0:00:31	31263	0:00:31	63	171	WITHIN WORDS
48	keyboard	SPACE	19	21	31497	0:00:31	31559	0:00:31	62	297	AFTER WORDS

Finally, all the data were imported in SPSS to statistically analyze them. Differences between L1 and L2 were calculated with General Linear Models (within-subjects repeated measures). Complementary analyses involved correlation analysis (Pearson) and Principal Component Analysis (PCA). SincePCA is central in this study, we briefly elaborate on the characteristics and the objectives of this analysis technique.

3.5 Principal Component Analysis (PCA)

The main aim of the data analysis was to identify and compute the underlying structure in the whole set of selected variables that were used as indicators of writing fluency. For this purpose, a Principal Component Analysis (PCA) was conducted (Field, 2009). We systematically evaluated the factorability of the variables by applying the analysis stepwise and iteratively, drawing on the general criteria of PCA(Field, 2009). Prior to the analysis, all variables were converted to log-values.

In the first stage of this process, it was important to reduce the large collection of variables (more than 200) into a manageable and practicable set. This was important for two reasons. First, we want to propose a fluency approach that is optimally comprehensive on the one hand, but at the same time easy to apply by researchers and practitioners alike. Second, we had to take into account a number of technical limitations with respect to the PCA technique used. One of the commonly quoted rules of thumb in PCA analysis isthat the proportion of participants versus items (variables) preferably does not exceed the ratio 10:1. A ratio of 5:1 is considered to be the absolute minimum (Hatcher, 1994; Nunnally, 1978). Therefore, it was important to reduce the total amount of possible fluency indicators to a limited set of variables that could be used as input for the PCA.

A first and logical step in reducing the total number of variables was to limit the inclusion of variables to those that are generalizable over a wide range of writing processes. To illustrate this line of thought: a variable that refers to the 'absolute length of the writing process' was excluded, because this type of variable does not allow us to compare writing processes of different lengths. Consequently, only variables that are based on proportions, ratios or means were retained. Using this criteria, we were able to retain approximately fifty variables to include in the further procedure. In a second step, we evaluated the degree of correlation between these variables. We divided them in eight sub-groups (see Table 4) based on product measures that were derived from product-related fluency research on the one hand and process research on the other hand (Chenoweth & Hayes, 2001). For those variables that correlated .9 or higher,each time a representative variable was selected. Representativity was defined on the basis of the following criteria: most marked correlation with other variables in the sub-group;correspondence with other fluency studies; and the ease to derive or calculate the variable through data analysis. This resulted in a basic set of 13 variables that were used as input for the PCA analysis. Finally, from that set those variables were removed that (a) had a loading of less than |.5|; (b) did not contribute to a simple factor structure; and (c) failed to meet a minimum criteria of having a primary factor loading of |.6| or above, and no cross-loading of |.4| or above on other factors (Tabachnick & Fidell, 2013).[1]

Finally, the reliability of the scores for the established scales were checked by means of their Cronbach's Alpha values.

4 Results

In this section, we first present some product and process-based results from a L1-L2 contrastive perspective. As stated before, the results of this study isnot themain focus, but the study is a means to describe *fluency* as a multimodal construct. By exploring the contrasts in a systematic way, we want to build a set of variables that canbe used as potential fluency indicators. We start from a more traditional approach on fluency, and then extend our scope by taking into account the dynamics of the writing process more explicitly (time and personrelated). In the final part, we describe how the use of PCA analysis leads us to a multidimensional characterization of writing fluency.

4.1 Traditional Measures and Process measures

As mentioned in Section 2, several studies have clearly demonstrated that writers produce L1 texts more fluently than L2 texts (Lindgren et al., 2008; Ong & Zhang, 2010; Segalowitz, 2010).This is perhaps not an astonishing finding in itself, but it begs the interesting question which variables really enable us to better

[1] One variable, viz., 'proportion of pause time', did not fully meet all these selection requirements and had a cross-loading on a secondary factor of .45 (>|.4|; "|" indicates 'absolute value'). Nevertheless, we decided to retain this variable in the final model, because it had an acceptable primary loading and complemented the other pause factor variable in the same factor. Moreover, from an exploratory point of view, we know from previous studies that this variable showed to be important to study fluency from a developmental perspective.

describe and understand these differences. To our knowledge, consensus amongresearchers on this issue is scarce.

This paragraph briefly presents the first study's results of the most commonly used fluency measures (see Table 2): e.g., words or characters per minute in the final text; words or characters per minute during the writing process.The within-subject variance analysis shows that in general L1 and L2 writing differ from a product as well asa process perspective. This is true in isolation, but also combined as calculated in the ratio factors: e.g., the ratio between the length of the final textand the characters and words produced during writing. The effect size is larger when measuring product or process characteristics at the character level (excl. spaces). At the character level, writers seem to be 20-25% more productive in L1 than in L2; at the word level, we found a difference of approximately14%. A comparison of the ratios also show that the participants revised relatively more in their L2 than in their L1 texts, as this latter percentage is closer to 100%[2].

Table 2
Comparison of L1 and L2 production from a product and process perspective

	L1		**L2**				
	Mean	SD	Mean	SD	*F*	*p*	η_p^2
Product							
words per minute	31.5	10.1	27.5	9.6	22.90	**.000**	.255
characters per minute (excl. spaces)	148.6	44.8	118.2	39.6	73.94	**.000**	.525
characters per minute (incl. spaces)	180.0	54.7	145.7	49.0	63.03	**.000**	.485
Process							
words per minute	34.5	10.0	30.2	9.7	27.14	**.000**	.288
characters per minute (excl. spaces)	190.4	49.4	156.7	44.4	68.25	**.000**	.505
characters per minute (incl. spaces)	225.5	59.1	187.7	53.6	61.16	**.000**	.477
Ratio product/process							
words (%)	79.56	9.80	77.39	10.95	5.09	**.027**	.071
characters (excl. spaces; %)	77.88	10.31	75.32	11.55	6.51	**.013**	.089
characters (incl. spaces; %)	79.56	9.79	77.38	10.95	5.09	**.027**	.071

The pause analysis allows us to explore a complementary approach ofcharacterizing the writing process. The pauses were defined at several levels: number, mean length, pause location, pause threshold (between 200 ms and 5000 ms), and P-bursts (text production between two significant pauses).

[2] A total of 100% implies that the produced text and the final text are completely similar and that no revisions occurred.

Table 3
Comparison of pausing behavior between L1 and L2

	L1		L2				
	Mean	SD	Mean	SD	*F*	*p*	η_p^2
Number of pauses per minute							
> 200 ms	60.4	15.6	68.1	14.8	31.06	**.000**	0.327
> 500 ms	17.9	3.9	20.8	4.5	58.43	**.000**	0.477
> 1000 ms	6.8	1.8	7.9	2.1	24.29	**.000**	0.275
> 2000 ms	2.8	1.0	3.3	1.3	18.62	**.000**	0.225
Mean pause length within words (s)							
> 200 ms	0.375	0.060	0.396	0.068	12.79	**.001**	0.167
> 500 ms	0.894	0.188	0.926	0.172	2.33	.132	0.035
> 1000 ms	1.709	0.708	1.752	0.545	0.42	.518	0.007
> 2000 ms	2.079	1.830	2.428	1.610	2.52	.117	0.038
Mean pause length between words (s)							
> 200 ms	0.643	0.170	0.681	0.198	4.53	**.041**	0.064
> 500 ms	1.299	0.311	1.372	0.357	3.06	.085	0.046
> 1000 ms	2.105	0.490	2.316	0.636	6.25	**.015**	0.089
> 2000 ms	3.320	0.755	3.595	1.162	3.40	.070	0.050
Mean pause length between sentences (s)							
> 200 ms	1.292	0.801	1.421	1.046	1.01	.318	0.016
> 500 ms	2.284	1.459	2.489	1.716	0.91	.343	0.014
> 1000 ms	3.112	2.429	3.446	2.281	0.81	.372	0.012
> 2000 ms	4.223	4.294	4.342	3.084	0.05	.828	0.001
Mean length of P-Bursts (in characters)							
> 2000 ms	80.0	60.8	55.1	39.1	24.74	**.000**	0.270
> 5000 ms	284.0	236.1	231.9	220.7	4.54	**.037**	0.063

Table 3 shows that writers pause significantly more in L2 thanL1. There is a quite consistent increaseof approximately10-15%, measured at the different pause thresholds. Especially the number of short pauses seems to increase when writing a text in L2. The analyses of the mean pause length at the different pause locations (i.e., within words, between words and sentences) show a more erratic pattern. In particular, pauses within and between words differsignificantly between L1 and L2 when a low threshold (>200 ms) is used. At the sentence level, no differences are observed in this study. This indicates that when writing in L2, relatively more attention is devoted to lower-level cognitive processes, such as those relatedto lexical issuesand spelling. A less pronounced difference is found for longer pausing time and for pauses situated above the word level. These pauses aremainlyconnected to more complex cognitive activities, for instance oriented at planning and (structural) revision.

A similar finding can be deducedfrom the P-Burst analysis. When applying lower thresholds in this analysis, the amount of text between pauses above the threshold decreases more than 40% when we compare L1 and L2 text production: L1: M= 80.0 (*60.8)* vs. L2: M=55.1 (*39.1)*; $F(1,67)=24.74$; $p < .01$). This illustrates again that text production in L2 is more fragmented and evolves in smaller, fluently produced text units.

Finally,Figure 1shows the proportion of the total pausing time (as opposed to the total active writing time). At each pause threshold level, the variance analysis indicates that L1 text production is characterized by significantly less pausing time than L2 text production. As expected, when increasing the pause threshold, the proportion of total pausing timedecreases systematically.

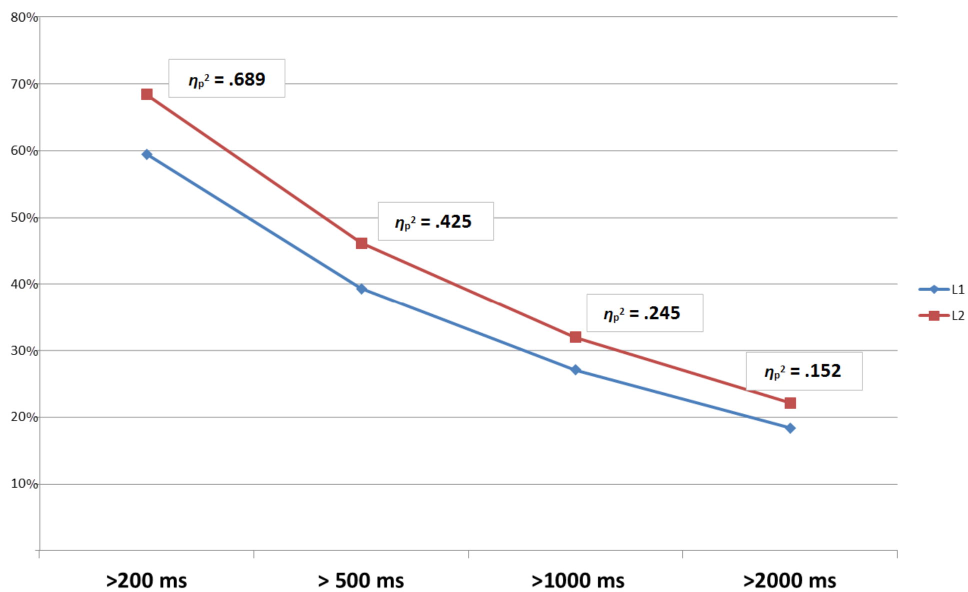

Figure 1. Proportion of total pausing time vs. total active writing time calculated at different pause thresholds.

4.2 Variability of the Writing Process

To date, fluency has been mainly described as a general measure referring to the writing product and process as a whole. However, in our opinion it is also important to take into account the dynamics of the writing process in relation to time, e.g., byaddressing the dominance of certain subprocesses at different stages in the writing process(Leijten, Van Waes, & Janssen, 2010; Van den Bergh & Rijlaarsdam, 1996; Van

Weijen, van den Bergh, Rijlaarsdam, & Sanders, 2009). We want to illustrate in this study that it might be important to explore the variability of the writing process as an indicator for writing fluency, because fluency is not a stable factor throughout the process and might differ at different stages. Therefore, we opted to calculate a process fluency measure based on ten intervals. We divided the writing process into ten equal intervals and calculated the number of characters (and the standard deviation) per interval. In addition to taking into account the process variance, we deem it important to consider text production in light of the personal typing skills of a writer.

The first approach we would like to introduce to complement the static fluency indicators is represented inFigure 2 (see also Formula 1). In this figure, the writing process for each writing task is divided in ten equal process intervals (*int*), and per interval the number of characters produced (including spaces-$char_{int}$)is calculated (normalized to the average number of characters per minute). Moreover, the proportion of this amount of characters is calculated in relation to a so-called *theoretical maximum* ($char_{opt}$), which in the current study was defined at 400 characters per minute.[3]We hypothesize that writers will not surpass this maximum during 'normal' text production, e.g., when writing an expository text. Explained in more technical terms, the proportion represented inFigure 2, is calculated on the basis of the following formula:

$$\sum_{i=1}^{\#_{int}} \left(\frac{\mathrm{char}_{int}}{\mathrm{char}_{opt}}\right) \cdot \frac{1\ min}{t_{proc\ .\ \#int}}$$

[Formula 1]

This approach enables us to represent the dynamics of the writing process in function of time (interval) and in relation to a fixed theoretical maximum. The constructed scale shows the variability of the writing process in such a way that it is comparable across different writers and writing tasks.

[3] This theoretical maximum is based on the reference speed for an experienced typist during a copy task (http://en.wikipedia.org/wiki/Typing). In future studies, we aim at defining this maximum on a more solid basis. Therefore, we have collected data of typing experts at the international InterSteno conference, during which a world competition is organized to select the fastest typists. We would like Danny Devriendt for facilitating these observations.

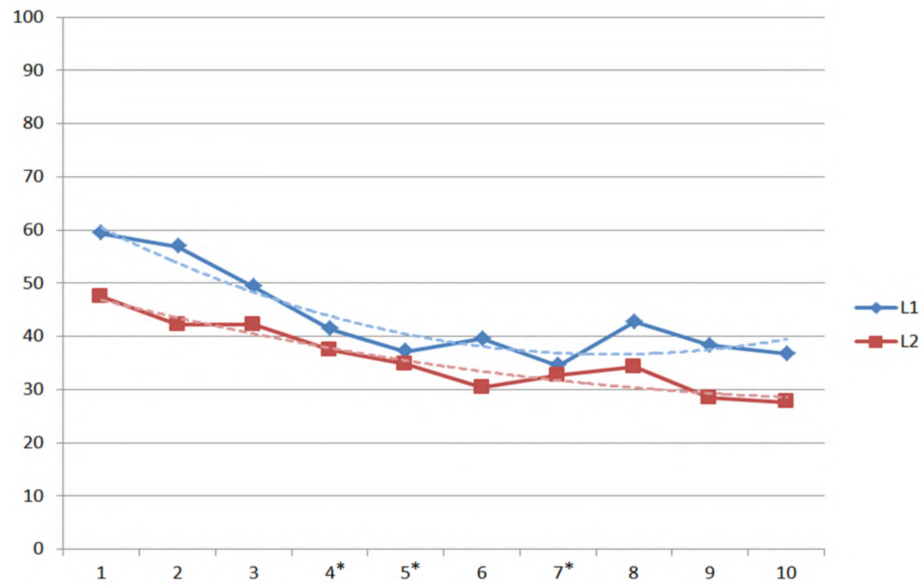

Figure 2. Proportional distribution per interval of the total number of characters in function of the theoretical maximum in L1 and L2. (Remark: The dotted line represents the polynomial trend line for both L1 and L2.)

Figure 2 shows that on average the proportional number of characters per interval is higher in L1 than in L2. There is a main effect of language (L1: 40.4 *(1.2)* vs. L2: 38.8 (*1.2*); $F(1,67)= 7.54$; $p<.01$; $\eta_p^2 =.101$). Additionally, there is a main effect of time (interval) ($F(9,59)= 47.86$; $p<.01$; $\eta_p^2 =.880$). However, when testing the differences between separated intervals (Bonferroni post-hoc), we could observethat in the middle part of the writing process (interval 4, 5 and 7) no significant differences in fluency between L1 and L2 could be found (p = .086, .319, and .431resp.). During these mid-intervals of the writing process, text production in both languages is comparable and fluency does not differ.

In this approach to fluency, in which we relate fluency to the dynamics of the process and a theoretical maximum, an important factor is still missing. Personal writing characteristics that typifytyping skills for instance, are not taken into account as such. However, recent research has shown that motoric pause thresholds differ considerably at the individual level(Leijten et al., 2011; Sullivan & Lindgren, 2006; Van Weijen et al., 2009; Wengelin, 2006). To partly meet these objections, we would like to propose a second approach to describe intra and interpersonal variability in the dynamics of text production. Based on a specific writing task (viz., writing a Dutch expository text), we calculated a so-called *taskmaximum* to express an individual text-production measure(Leijten & Van Waes, 2014)

To calculate the participants' highest writing speed in the current writing tasks, we divided the writing process in ten equal intervals once again. For each interval, we determined the number of characters produced during that particular time interval (incl. spaces) and converted it to the number of charactersproduced per minute. In contrast with the previous approach, we do not juxtapose this process dynamicsto a *theoretical*maximum, but to a *task*maximum. This maximum is calculated for each participant individually (L1), on the basis of a triple moving average applied to consecutive 10 seconds text-production intervals in the targeted writing task. By using this method, we determined the maximum text-production rate for each individual during a time slot of half a minute. Finally, the result of this calculation, converted per minute, is contrasted with the text production per interval. The resulting score represents the proportion of the actual text production per interval compared to the maximum text production in the reference writing task. Figure 3shows the text production per interval in relation to the taskmaximum.[4]

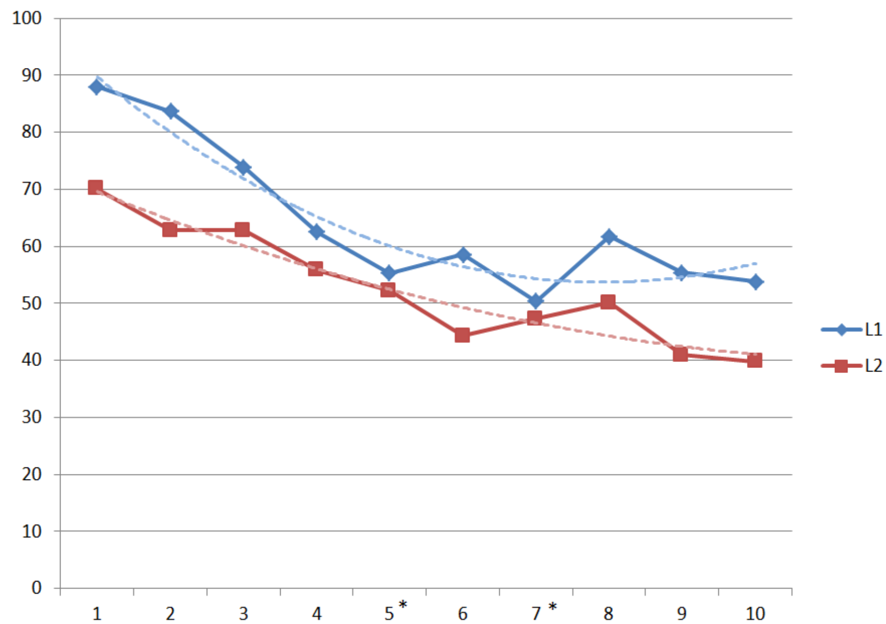

Figure 3. Proportional distribution of the amount of characters per interval in function of the task maximum in L1 and L2.

The personal distribution of the amount of characters in relation to the taskmaximum shows a higher percentage than it showed in relation to the absolute maximum. Thus, the participants hada much lower

[4] The fluency approach based on the theoretical (or absolute), the personal and task maximum are integrated in the new Fluency analysis of Inputlog 6 (Generate > Analysis > Fluency; www.inputlog.net) in order to allow researchers to automatically generate this fluency indicator from keystroke-logging data.

typing speed than the predefined absolute maximum. From Figure 3, it can be observedthat the beginning of the writing process is characterized by the highest text production: L1 writers reacha text-production rate of roughly80 to 90% of their individual taskmaximum.[5] In L2,the average text production reaches only 60 to 70% in the first intervals. The text production gradually decreases from the start of the writing process to the middle part. A main statistic effect was found for both language use (L1: 64.3 *(1.3)* vs. L2: 52.7 (*1.2*); $F(1,67)= 69.7$; $p<.001$; $\eta_p^2 =.510$) and time interval ($F(9,59)= 26.8$; $p<.001$; $\eta_p^2 =.804$). Similar to the results of the proportional distribution in function of the theoretical maximum (Figure 2), the post-hoc analysis revealed also in the case of this approach no significant difference for intervals 5 and 7 in the middle phase of the writing process. In interval 5 and 7,the fluency in L1 and L2 is similar(resp. $p = .319$ and $p = .431$).

To complement this representation of the fluency measures described above, we also calculated the standard deviation per person between intervals. This measure could be used to define the fluency variance (over time)more explicitly. Figure 4 visualizes this fluency variance in L1 and L2 by plotting the proportional mean taskmaximum against the standard deviation (between intervals).

[5] This proportion is relatively high, because the task at hand was relatively easy. In the discussion section, we will suggest alternative ways to calculate the task maximum.

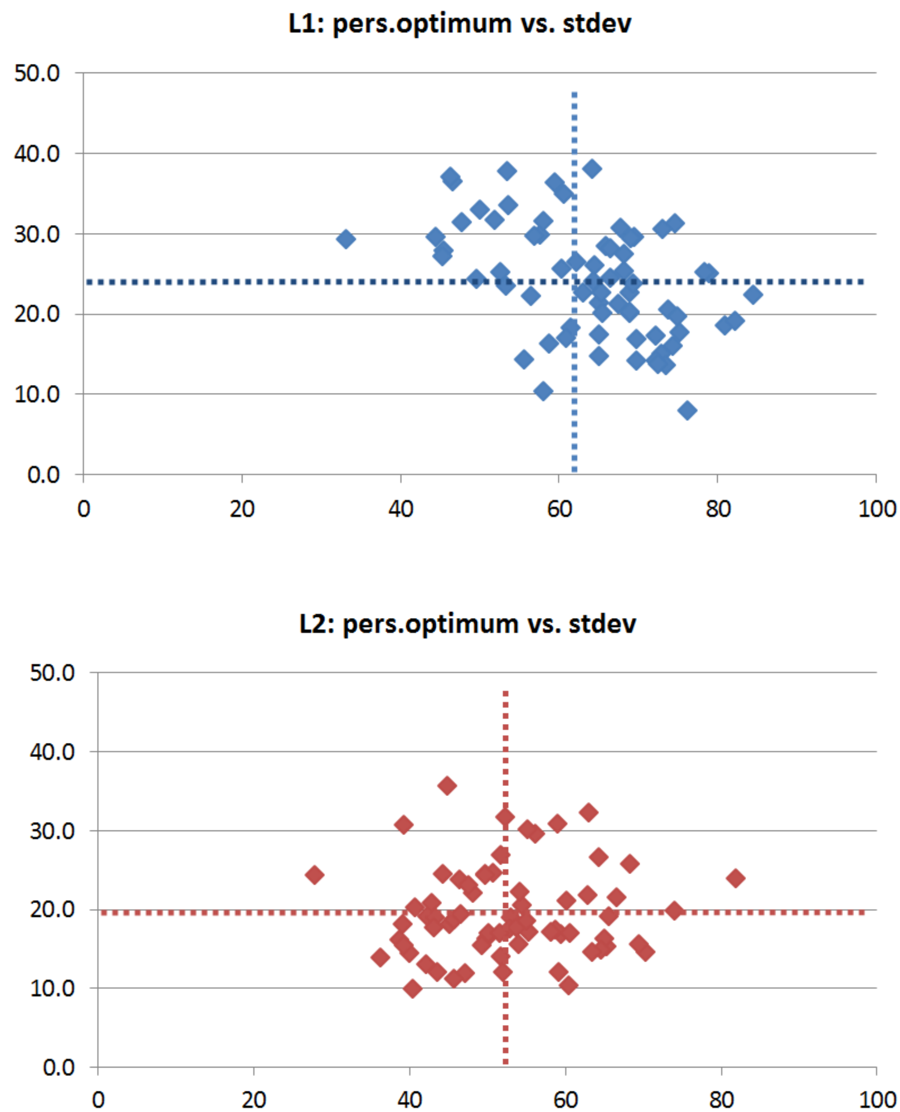

Figure 4. Comparison of the mean proportional task maximum (x-axis) and standard deviation (y-axis) in L1 and L2.

Figure 4 shows a clear trend.The average text production is higher in L1 than in L2: 62% of the taskmaximum in L1 to 54% in L2. Also the standard deviation is higher: 23.9 in L1 and 19.6in L2. In other words, the dispersion is higher in L1, pointing at the fact that the mother tongue writing process is characterized by a higher variance.

4.3 Constructing Fluency Measures

The exploration described in Section 4.1 and 4.2allowedus to describe writing fluency from different perspectives, combining simultaneously indicators at different levels. Table 4 shows an overview of the selected variables that could also be used as indicators of writing fluency. These variables were used as a basis for the Principal Component Analysis.

Table 4
Perspectives to fluency

Perspective	Explanation
1. Product	Length of the final text in words and characters (incl. and excl. spaces)
2. Process	Length of the process in msand number of words and characters (incl. and excl. spaces) produced (incl. deleted words and characters).
3. Ratio product/process	Proportion between product and process measures.
4. Pause length	Mean pause length (i.e., latency between the previous and the current action).
5. Number of pauses	Number of pauses (in total and per minute).
P-bursts	A P-burst is defined as the string of actions delimited by an initial and end pause exceeding the defined pause threshold. The length of a P-burst could be defined in ms or in characters.
Pause location	Location of the pause: within words, between words, and between sentences.
Pause threshold	Minimum pause threshold that has been used to define a pause: 200, 500, 1000, 2000 and 5000 ms. Small pause thresholds are usually applied when low cognitive processes are important to take into account; larger thresholds are more oriented at higher cognitive processes. Pauses smaller than 200 ms mostly refer to motoric interkey-transitions and are not included in the current study. Additionally, we also defined two pause intervals: (1) between 200 and 500 ms, and (2) between 200 and 1000 ms.
6. Revisions	Number of insertions and deletions in the writing process (per minute).
R-bursts	An R-burst is defined as the string of actions between two revisions. Both the number of R-burst per minute and the duration of an R-burst are computed.
7. Interval variance	Each writing process is divided into 10 equal time parts, called intervals. This enables us to proportionally compare writing processes that differ in duration. The dynamics and the variance of the process is the main focus. For instance, the lower the standard deviation, the more constant the process.
8. Theoretical maximum	Number of characters produced in the writing process divided per interval in relation to a theoretical maximum, currently defined at 400 characters per minute (cpm) (see Section4.2).
Taskmaximum	Number of characters produced in the writing process divided per interval in function of a taskmaximum, currently defined on the basis of the L1 writing task (see Section4.2).

4.4 Principal Component Analysis (PCA)

The PCA aimed at identifying a limited and classified set of indicators to describe fluency from a multidimensional perspective. In order to reduce and further classify the variables described above, we followed the data analysis procedure described in Section 3. The analysis is primarily based on the L1 writing task and then elaborated to the L2 writing task. The stepwise and iterative procedure resulted in a basic set of ten variables that was used as input for the Principal Component Analysis (PCA). These ten variables (selected from a set larger than 200) all met the factorability criteria. The Kaiser-Meyer-Oklin (KMO) measure of sampling adequacy was .627, exceeding the recommended value of .600. Bartlett's Test of sphericity reached statistical significance ($X^2(45)= 709.39$, $p<.001$), indicating that there is sufficient commonality in the selected variables to perform the PCA.The PCA elicits the presence of four components, whichtaken together explain 92.1% of the total variance (before rotation) in this study. The visual inspection of the screeplot confirms the presence of these four components and shows a clear break after the fourth component. On the basis of these observations, we decided on a solution with four components, explaining 45.3%, 20.1%, 16.2%, and 10.5% of the total cumulative variancerespectively (Table 6).

To better interpret these four components, we applied an oblique rotation (Oblimin with Kaiser Nominalization) for the final solution, assuming that the factors could be correlated to a certain extent. Table 5shows the final pattern matrix with the factor loadings after rotation (7 iterations). InTable 5the correlations between the components are presented. Both tables show that the four components clearly represent four different perspectives on fluency.

Internal consistency for each of the four established factors were examined using Cronbach's Alpha. An alpha value of .7 is considered reliable (Field, 2009, p. 675). All factors had good to excellent reliabilities (Table 6).

Table 5
Pattern matrix of the factor loadings based on a principle component analysis with Oblimin rotation for ten selected fluency variables (N=68); combined with a correlation matrix of the four PCA-components

	Factor loadings				**Correlations**		
	1	**2**	**3**	**4**	**2**	**3**	**4**
1. Production: mean number of characters (incl. spaces)					.191	-.138	-.329
during the process (per minute)	**1.02***	-.01	-.15	.04			
in the final product (per minute)	**.94**	-.04	.12	.05			
per .10 interval (corrected for abs. opt. of 400 cpm)	**.86**	-.04	.29	-.06			
per P-burst (per minute; threshold 2000 ms)	**.72**	.07	-.07	**-.48**			
2. Process variance: st. dev. of characters (incl. spaces)						-.096	.046
per .10 interval (per minute)	.19	**.99**	.07	.03			
per .10 interval (corrected for taskmaximum)	-.28	**.88**	-.09	-.07			
3. Revision: mean number of characters (incl. spaces)							.014
product vs. process ratio	-.05	-.08	**.94**	-.22			
length of R-burst	.10	.07	**.89**	.21			
4. Pausing behavior							
mean pause time length between words (th > 200 ms)	.09	-.03	-.07	**.96**			
proportion of total pause time (th > 2000 ms)	-.41	.06	.13	**.67**			

**Note: Factor loadings > |.48| are formatted bold (Rietveld & Van Hout, 1993).*

Table 6
Reliability scores for PCA components

Component	**Cronbach's Alpha**	**Explained variance**
Production	.95	45.3%
Process Variance	.85	20.1%
Revision	.83	16.2%
Pausing behavior	.70	10.5%

Figure 5schematically visualizesthe four components that - together with the underlying variables - form the basis of the multi-dimensional fluency model.

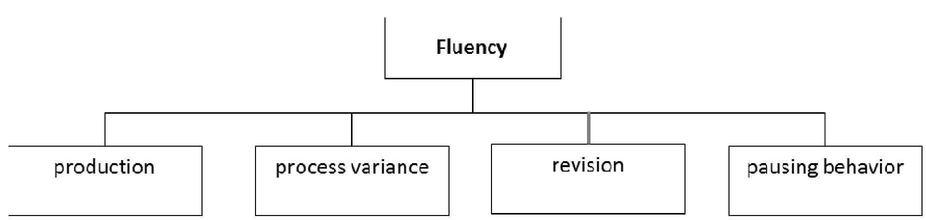

Figure 5. Four factor model describing writing fluency.

4.5 Fluency: L1 versus L2

As an example of how this framework can be instrumentalized in writing process research we will illustrate this via the data of the study at hand. Therefore, we conducted a General Linear Modelling (GLM) repeated-measure test to compare the fluency between theprocesses of writing expository texts in L1 and L2. This analysis is based on the componentsand variables described above.Table 7shows that although the participants were quite proficient L2 users, the production, process variance and revision parameters at the component level differ significantly between L1 and L2; the pause behavior component does not seem to yield a significant differencein this study. At the lower level,all underlying variables indicate a significant difference between L1 and L2 writing processes[6]. As expected, writers producing texts in L1 are more productive (process and product); their process is characterized by more variance over time; they revise more extensively; P-burst are longer;they need less pausing time than when they produce L2 texts, and the pausing time between words is shorter in L1 text production.

Table 7
Comparison of underlying factor variables for L1 and L2 (GLM repeated measures)

	L1		**L2**				
	Mean	SD	Mean	SD	*F*	*p*	η^2
Production: mean number of characters (incl. spaces)*	**58.7**	***25.3***	**39.9**	***27.3***	**40.9**	**.001**	**.379**
during the process (per minute)	225.2	*59.1*	186.9	*53.7*	57.6	.001	.474
in the final product (per minute)	178.6	*54.1*	144.3	*48.3*	57.9	.001	.475
per .10 interval (corrected for absolute maximum of 400 cpm)	43.5	*11.3*	35.6	*10.4*	59.3	.001	.481
per P-burst (per minute; threshold 2000 ms)	80.4	*62.0*	53.8	*38.6*	27.3	.001	.299
Process variance: standard deviation of characters (incl. spaces)	**60.5**	***28.1***	**41.1**	***272***	**31.4**	**.001**	**.319**
per .10 interval (per minute)	50.3	*15.7*	41.1	*12.2*	27.4	.001	.300
per .10 interval (corrected for taskmaximum)	24.2	*7.2*	19.7	*5.8*	28.5	.001	.308
Revision: mean number of characters (incl. spaces)	**54.8**	***26.8***	**46.0**	***30.3***	**13.8**	**.001**	**.171**
product vs. process ratio	0.79	*0.10*	0.77	*0.11*	4.2	.045	.061
length of R-burst	30.4	*12.8*	24.5	*11.3*	37.6	.001	.370
Pause behavior	**49.7**	***29.4***	**52.9**	***29.8***	**1.04**	**.313**	**.016**
mean pause time length between words (threshold > 200 ms)	0.643	*0.170*	0.688	*0.198*	4.4	.041	.064
proportion of pause time (threshold > 2000 ms)	18.5	*9.6*	22.4	*11.4*	12.1	.001	.159

*The values for the four factors represent calculated percentages based on a standard normal distribution (normsdist) conversion of the component scores for L1 and L2 (combined).

[6] A multivariate analysis (Manova) showed no main effect for the selected fluency parameters between the different foreign languages that were used by the participants (mainly L2-English: n=54 and L2-French: n=11): *F*(9,55)=4.655; *p*=.224).

5 Conclusion: Towards an Integrated Multidimensional Approach to Writing Fluency

This article aims at taking ona new approachtowriting fluency. A multidimensionalperspectiveof fluency should allow researchers to measure this concept by making the necessary differentiations and respecting the inherent stratification. We started from the traditionally used fluency measures and complemented them with variables derived from more recent writing-process research. Process characteristics (pauses and revisions) and aspects related towriting dynamics in relation to time were the primary measures added. Since we opted for a data-driven exploration, we could use PCA-analysis to systematically uncover and describe the underlying concepts of a set of variables that were selected as suitable indicators to describe the writing process from different perspectives.

The analysis presented in this study creates a comprehensive framework for a multidimensional model, in which four main components can be distinguished: (1) production, (2) process variance, (3) revision, and (4) pausing behavior. The combination of and relation between these components should allow a more nuanced description of writing fluency in all its aspects. In this case, the proportional measures shed a new light on absolute differences between L1-L2 writing. Previous research showed, for instance, that writers are more fluent in L1 than in L2 (e.g., Angelone, 2013), but a multidimensional fluency approach allows us to further refine and define this difference. Our study, for instance, points out that in the current L1-L2 expository writing tasks L2-writers are only less fluent in the beginning and the end of the writing process, and that their fluency hardly differs in the middle of the process.

This kind of observations indicate that this approach to fluency could serve as a useful instrument in further research that aims at describing differences in writing fluency on a personal or group level. Possible areas of research are: developmental studies in educational settings; intervention studies, for instance, on the effect of initial planning strategies or instructions on writing dynamics; professional writing studies on digital communication and multiple sources; translation studies; correlation studies on writing fluency and text quality(Hurlebusch, 2000).

However, at this stage of the research, it is warrantedto point out that the explorative model is based on two short, expository texts, written by advanced language learners (master students in professional communication).Note also that, due to the limited size of the sample (68 participants), the results of the PCA should always be treated with caution. In writing research, relatively small sample sizes are common, butto optimally interpret a PCA, sample sizes between 200 and 300 are often recommended (Hatcher, 1994). Therefore, we have planned to set up further research and adapt this method to obtain data on a larger variety of text genres and of multiple texts per genre per participant(Van Steendam, Tillema, Rijlaarsdam, & Van den Bergh, 2012). Moreover, we would like to differ in writer populations(see also discussion on sample size in PCA in Hatcher, 1994) and investigate whether the described model is applicable to other genres and writing contexts.

Since there is a clear need to further standardize analysis procedures, follow-up studies are important both from a methodological perspective (e.g., for selection and exchangeability of indicators) and with respect to content. Therefore, we strive to integrate an automated fluency analysis in Inputlog to facilitate these lines of research and provide peer researchers with a set of instruments to easily and consistently calculate the proposed fluency indicators. In the latest version of this keystroke-logging program (Inputlog 6.0 - www.inputlog.net), a basic fluency analysis module is already available. However, the different fluency indicators have not yet been integrated in a single, combined analysis. Therefore, further steps are undertaken to facilitate the calculationsunderlying the multidimensional approach to fluency and to make it available on a larger scale. With respect to the content of the research subject, we aim at comparing within-subject fluency in other composing tasks. The fluency analysis described above is based on a so-called isolated writing task. However, contemporary writing is almost always characterized by an interaction with multiple (digital) sources during text production. Writers often start their texts from existing documents, they integrate and adapt copied text from mail conversations or blogs, look for extra information on the internet, integrate illustrations they design in Photoshop, relate their text to a project planner etc. (Leijten et al., 2014). Moreover, their writing process is often interrupted by incoming emails, tweets or Facebook alerts. These kinds of interruptions have become a natural part of most digital writing processes, and—in our view—also influence writing fluency when interpreted from a broader perspective.Therefore, and to complement the approach described above, in a follow-up study we would like to focus on the interaction with digital sources while writing. This should enable us to also take into account the interaction with sources in our multimodal perspective to fluency. We are convinced that this interaction and the related fragmentation of the process have an important impact on the writing dynamics (Leijten & Van Waes, 2012; Leijten, Van Waes, Schriver, & Hayes, 2014; O'Hara, Taylor, Newman, & Sellen, 2002).

On top of that, it is also important to compare text genres like argumentative, instructive and descriptive texts (L1; different L2/L3 languages) and short versus longer texts. Moreover,the component structure should be further validated in the light of writing development (children, young adults, adults, healthy elderly people, aphasic or dementia patients), and regarding text production forms such as translation and journalistic processes etc.

By following this line of research, as well as by aggregating the writing data for confirmatory factor analysis, we want to further validate the proposed multidimensional model. Follow-up studies on the relation between writing fluency, text quality and perceived fluency (Bosker et al., 2013; Segalowitz, 2010) could also complement the research agenda. Furthermore, more specific research on procedures to define a personal base line for (motoric) typing skills in different languages by means ofa validated typing task would certainly create an added value(Wallot & Grabowski, 2013). In this study, we calculated the taskmaximum on the basis of the writing task itself. A specificallydesigned copy task with extreme low cognitive demands might be a better and more stable instrument to measure motoric skills in a controlled and task-independent manner, indicating a so-called *personal maximum*. In the long term, it would be fruitful to conduct a meta-analysis of similar research studies conducted by various researchers on the research topics described above: writing tasks, target audiences and multiple languages.

References

Alves, Rui Alexandre, Castro, São Luís, & Olive, Thierry (2008). Execution and pauses in writing narratives: Processing time, cognitive effort and typing skill. *International Journal of Psychology, 43*(6), 969-979.

Angelone, Erik (2013). Watching and Learning from'Virtual Professionals': Utilising Screen Recording in Process-oriented. In D. Kiraly, S. Hansen Shirra & S. M. Karin (Eds.), *New Prospects and Perspectives for Educating Language Mediators* (pp. 121-139). Tübingen, Germany: Gunter-Narr Verlag.

Bereiter, Carl, & Scardamalia, Marlene (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Berninger, Virginia W., Cartwright, Ana C., Yates, Cheryl M., Swanson, H. Lee, & Abbott, Robert D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades. *Reading and Writing: An Interdisciplinary Journal, 6,* 161-196.

Bosker, Hans Rutger, Pinget, Anne-France, Quené, Hugo, Sanders, Ted, & de Jong, Nivja H (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing, 30*(2), 159-175. doi: 10.1177/0265532212455394

Chenoweth, N. Ann, & Hayes, John R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication, 18*(1), 80-98. doi: 10.1177/0741088301018001004

Chenoweth, N. Ann, & Hayes, John R. (2003). The inner voice in writing. *Written communication, 20*(1), 99-118.

Council of Europe (2011). *Common European Framework of Reference for Learning, Teaching, Assessment*. Retrieved from http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp.

De Jong, Nivja H., & Wempe, Ton (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior research methods, 41*(2), 385-390.

Field, Andie (2009). *Discovering statistics using SPSS* (3 ed.). London/Thousand Oaks, CA: SAGE Publications Ltd.

Goldman-Eisler, Frieda (1961). The distribution of pause durations in speech. *Language and Speech, 4*, 232-237.

Grabowski, Joachim (2008). The internal structure of university students' keyboard skills. *Journal of Writing Research, 1*(1), 27-52.

Graham, Steve, & Perin, Dolores (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools.* New York, NY: Carnegie Corp.

Halberda, Justin , Mazzocco, Michèle M. M., & Feigenson, Lisa (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665-668. doi: 10.1038/nature07246

Hatcher, Larry (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*: Sas Institute.

Hayes, John R., & Chenoweth, N. Ann (2006). Is working memory involved in the transcribing and editing of texts? *Written Communication, 23*(2), 135-141.

Hurlebusch, Klaus (2000). Understanding the Author's Compositional Method: Prolegomenon to a Hermeneutics of Genetic Writing. *Text, 13*, 55-101.

Iwashita, Noriko, Brown, Annie, McNamara, Tim, & O'Hagan, Sally (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics, 29*(1), 24-49.

Johansson, Roger, Wengelin, Åsa, Johansson, Victoria, & Holmqvist, Kenneth (2010). Looking at the keyboard or the monitor: relationship with text production processes. *Reading and writing, 23*(7), 835-851. doi: 10.1007/s11145-009-9189-3

Johnson, Mark D, Mercado, Leonardo, & Acevedo, Anthony (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing, 21*(3), 264-282.

Kellogg, Ronald T. (1996). A model of working memory in writing. In C. M. Levy & S. E. Ransdell (Eds.), *The Science of Writing: Theories, methods, individual differences and applications* (pp. 57-71). Hillsdale, NJ: Lawrence Erlbaum.

Kellogg, Ronald T. (2004). Working memory components in written sentence generation. *American Journal of Psychology, 117*, 341-361.

Kobayashi, Hiroe, & Rinnert, Carol (2013). L1/L2/L3 writing development: Longitudinal case study of a Japanese multicompetent writer. *Journal of Second Language Writing, 22*(1), 4-33.

Koponen, Maarit, & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 5-24). Ann Arbor, Michigan: University of Michigan Press.

Kormos, Judit (2006). *Speech production and second language acquisition*. Mahwah, NJ: Psychology Press.

Kormos, Judit (2012). The role of individual differences in L2 writing. *Journal of Second Language Writing, 21*(4), 390-403. doi: http://dx.doi.org/10.1016/j.jslw.2012.09.003

Latif, Muhammad M. Abdel (2012). What do we mean by writing fluency and how can it be validly measured? *Applied Linguistics*. doi: 10.1093/applin/ams073

Leijten, Mariëlle, De Maeyer, Sven, & Van Waes, Luuk (2011). Coordinating sentence composition with error correction: A multilevel analysis. *Journal of Writing Research, 2*(3), 331-363.

Leijten, Marielle, & Van Waes, Luuk (2012). *Professional writing from multiple sources*. Paper presented at the The 13th International Earli SIG Writing Conference (European Association of Research on Learning and Instruction), Porto.

Leijten, Mariëlle, & Van Waes, Luuk (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication, 30*(3), 358– 392. doi: 10.1177/0741088313491692

Leijten, Mariëlle, & Van Waes, Luuk (2014). *Manual Inputlog 6.0*. Management. University of Antwerp. Antwerp. Retrieved from http://www.inputlog.net/docs/Inputlog_6_20140414.pdf

Leijten, Mariëlle, Van Waes, Luuk, & Janssen, Daniël (2010). Error correction strategies of professional speech recognition users: three profiles. *Computers in Human Behaviour, 26*, 964-975. doi: 10.1016/j.chb.2010.02.010

Leijten, Mariëlle, Van Waes, Luuk, & Ransdell, Sarah (2010). Correcting text production errors: Isolating the effects of writing mode from error span, input mode, and lexicality. *Written communication, 27*(2), 189-227. doi: 10.1177/0741088309359139

Leijten, Mariëlle, Van Waes, Luuk, Schriver, Karen, & Hayes, John R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research, 5*(3), 285-337.

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

Lindgren, Eva, Sullivan, Kirk, & Miller, Kristyan Spelman (2012). .04. 08 Development of Fluency in First and Foreign Language Writing. *Studies in Writing, 25*, 273-274.

Lindgren, Eva, Sullivan, Kirk P.H., & Spelman Miller, Kristyan (2008). Development of fluency and revision in L1 and L2 writing in Swedish high school years 8 and 9. *International Journal of Applied Linguistics, 156*, 133-151.

MacArthur, Charles A., Graham, Steve, & Fitzgerald, John (Eds.) (2008). *Handbook of writing research*. New York, NY: The Guilford Press.

McCutchen, Deborah, Covill, Amy, Hoyne, Susan H., & Mildes, Karen (1994). Individual differences in writing: Implications of translating fluency. *Journal of Educational Psychology, 86*(2), 256-266.

Nunnally, Jum C. (1978). *Psychometric methods*. New York, NY.: McGraw-Hill.

O'Hara, Kenton, Taylor, Alex, Newman, William, & Sellen, Abigail (2002). Understanding the materiality of writing from multiple sources. *International Journal of Human-Computer Studies, 56*(3), 269-305.

Olive, Thierry, Alves, Rui Alexandre, & Castro, Sao Luis (2009). Cognitive processes in writing during pause and execution periods. *European Journal of Cognitive Psychology, 21*(5), 758-785. doi: 10.1080/09541440802079850

Olive, Thierry, Favart, Monik, Beauvais, Caroline, & Beauvais, Lucie (2009). Children's cognitive effort and fluency in writing: Effects of genre and of handwriting automatisation. *Learning and Instruction, 19*(4), 299-308.

Ong, Justina (2014). How do planning time and task conditions affect metacognitive processes of L2 writers? *Journal of Second Language Writing, 23*(0), 17-30. doi: 10.1016/j.jslw.2013.10.002

Ong, Justina, & Zhang, Lawrence Jun (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing, 19*(4), 218-233. doi:10.1016/j.jslw.2010.10.003

Segalowitz, Norman (2010). *The cognitive bases of second language fluency*. New York: Routledge.

Shanahan, Timothy (2006). Relations among oral language, reading, and writing development. *Handbook of writing research*, 171-183.

Skehan, Peter (2003). Task-based instruction. *Language Teaching, 36*(01), 1-14. doi: 10.1017/S026144480200188X

Skehan, Peter (2009). Modelling Second Language Performance: Integrating Complexity, Accuracy, Fluency, and Lexis. *Applied Linguistics, 30*(4), 510-532. doi: 10.1093/applin/amp047

Snellings, Patrick, Van Gelderen, Amos, & De Glopper, Kees (2002). Lexical Retrieval: An Aspect of Fluent Second–Language Production That Can Be Enhanced. *Language Learning, 52*(4), 723-754. doi: 10.1111/1467-9922.00202

Sullivan, Kirk P.H., & Lindgren, Eva (2006). *Computer key-stroke logging and writing*. Oxford: Elsevier Science.

Tabachnick, Barbara G., & Fidell, Linda S. (2013). *Using Multivariate Statistics* (6 ed.). Boston: Allyn and Bacon.

Tillema, Marion (2012). *Writing in first and second language: empirical studies on text quality and writing processes* (Vol. 299). Utrecht: LOT.

Towell, Richard, Hawkins, Roger, & Bazergui, Nives (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*(1), 84-119.

Van den Bergh, Huub, & Rijlaarsdam, Gert (1996). The dynamics of composing: Modelling writing process data. In C. M. Levy & S. E. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 207-232). Mahwah, NJ: Lawrence Erlbaum Associates.

Van Horenbeeck, Eric, Pauwaert, Tom, Van Waes, L., & Leijten, M. (2012). S-notation markup rules (Technical Description) (D. o. Management, Trans.) (pp. 7). Antwerp: University of Antwerp.

Van Steendam, Elke, Tillema, Marion, Rijlaarsdam, Gert, & Van den Bergh, Huub (Eds.) (2012). *Measuring writing: Recent insights into theory, methodology and practice* (Vol. 27). Leiden - Boston: Brill.

Van Waes, Luuk, & Leijten, Mariëlle (2011). *Observing and analysing digital writing processes with Inputlog*. Paper presented at the Antwerp Summer School on Writing Process Research: Keystroke logging and Eyetracking, Antwerp. http://www.inputlog.net/docs/Summerschool_Analyzing_VanWaes_Leijten.pdf

Van Waes, Luuk, Leijten, Mariëlle, & Quinlan, Thomas (2010). Reading during sentence composing and error correction: A multilevel analysis of the influences of task complexity. *Reading and Writing, 23*(7), 803-834. doi: 10.1007/s11145-009-9190-x

Van Waes, Luuk, Leijten, Mariëlle, Wengelin, Åsa, & Lindgren, Eva (2012). Logging tools to study digital writing processes. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 507-533). New York/Sussex: Taylor & Francis.

Van Weijen, Daphne, van den Bergh, Huub, Rijlaarsdam, Gert, & Sanders, Ted (2009). L1 use during L2 writing: An empirical study of a complex phenomenon. *Journal of Second Language Writing, 18*(4), 235-250. doi: http://dx.doi.org/10.1016/j.jslw.2009.06.003

Wallot, Sebastian, & Grabowski, Joachim (2013). Typewriting Dynamics: What distinguishes simple from complex writing tasks? *Ecological Psychology, 25*(3), 267-280. doi: 10.1080/10407413.2013.810512

Wengelin, Åsa (2006). Examining pauses in writing: Theories, methods and empirical data. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer Key-Stroke logging and Writing: Methods and Applications* (Vol. 18, pp. 107-130). Oxford: Elsevier.

Wolfe-Quintero, Kate, Inagaki, Shunji, & Kim, Hae-Young (1998). *Second language development in writing: Measures of fluency, accuracy & complexity*. University of Hawaii at Manoa: Second Language Teaching & Curriculum Center.

Yan, Cathy Ming Wai, McBride-Chang, Catherine, Wagner, Richard K, Zhang, Juan, Wong, Anita MY, & Shu, Hua (2012). Writing quality in Chinese children: speed and fluency matter. *Reading and Writing, 25*(7), 1499-1521.