Exploring the process of reading during writing using eye tracking and keystroke logging

Reference:

De Smet Milou, Leijten Mariëlle, Van Waes Luuk.- Exploring the process of reading during w riting using eye tracking and keystroke logging
Written communication - ISSN 0741-0883 - 35:4(2018), p. 411-447
Full text (Publisher's DOI): https://doi.org/10.1177/0741088318788070
To cite this reference: https://hdl.handle.net/10067/1523130151162165141

Exploring the Process of Reading during Writing Using Eye Tracking and Keystroke Logging

Milou J. R. de Smet, Mariëlle Leijten, & Luuk Van Waes

Milou.Desmet@uantwerpen.be, Marielle.Leijten@uantwerpen.be,

Luuk.Vanwaes@uantwerpen.be

University of Antwerp, Faculty of Applied Economics, Department of Management, Belgium

Running head: Exploring the process of reading during writing

Correspondence concerning this article should be addressed to

Luuk Van Waes | ORCID: 0000-0002-3642-9533

University of Antwerp, Faculty of Applied Economics, Department of Management,

Prinsstraat 13 (C.457) 2000 Antwerp, Belgium. T: +32 3 265 5159

luuk.vanwaes@uantwerpen.be

Abstract

This study aims to explore the process of reading during writing. More specifically, it investigates whether a combination of keystroke logging data and eye tracking data yields a better understanding of cognitive processes underlying fluent and non-fluent text production. First, a technical procedure describes how writing process data from the keystroke logging program Inputlog are merged with reading process data from the Tobii TX300 eye tracker. Next, a theoretical schema on reading during writing is presented, which served as a basis for the observation context we created for our experiment. This schema was tested by observing 24 university students in professional communication (skilled writers) who typed short sentences that were manipulated to elicit fluent or non-fluent writing. The experimental sentences were organized into four different conditions, aiming at (1) fluent writing, (2) reflection about correct spelling of homophone verbs, (3) local revision, and (4) global revision. Results showed that it is possible to manipulate degrees of non-fluent writing in terms of time on task and percentage of non-fluent key transitions. However, reading behavior was only affected for the conditions that explicitly required revision. This suggests that non-fluent writing does not always affect the reading behavior, supporting the parallel and cascading processing hypothesis.

Keywords: writing process; reading process; eye tracking; keystroke logging; Inputlog; fluent writing; non-fluent writing; individual pause threshold

Defining the Process of Reading during Writing Using Eye Tracking and Keystroke Logging

During composition different cognitive processes are being executed. Writers do not only produce text but they typically also spend a certain amount of time reading during text production. Hayes (1996) made a distinction between three types of reading that occur during writing, that is, (1) reading source texts, (2) reading to define tasks, and (3) reading to evaluate. This study focuses on the latter where writers read and re-read their own emerging text, also known as the "Text Produced So Far" (TPSF) defined by Hayes (1996). Reading the TPSF differs from traditional reading for comprehension, both in terms of the goal and the reading pattern. Reading to evaluate does not only involve the proofreading of a text at the very end of the writing process but also during writing. In fact, the majority of text evaluation occurs during writing (Quinlan, Loncke, Leijten, & Van Waes, 2012). This process of (re)reading occurs during pauses as well as during active writing, depending on the writers' working memory demands (Olive, 2014). This is in line with Kellogg's (1996) assumption that experienced writers are able to orchestrate writing processes concurrently as long as the demands do not exceed the writers' working memory capacity. In the latter case pausing takes place. Skilled writers try to optimize their writing by efficiently coordinating writing processes, leading to - what Olive (2014) calls - a parallel and cascading orchestration of the process. In the description of his cascading model of writing he contends that "each segment of text is sequentially processed from central to peripheral processes, and different levels of processing operate simultaneously on different segments of text" (p. 177). Consequently, certain operations at low levels may affect higher level processes. For instance, as Delattre, Bonin, and Barry (2006) demonstrate in their study, spelling processes are often activated before a writer actually writes down the word, especially in the case of regular words.

However, they also show that in the case of composing irregular spelled words the spelling processes might continue during transcription. The role of (re)reading in this orchestration process, however, is not yet really clear (Alamargot et al, 2010).

(Re)reading has different functions. The explicit reading and re-reading of the TPSF function as monitoring strategies through which writers can evaluate their text for spelling, lexical choice and organization. Moreover, they allow writers to ensure that their text is in conformance with the formal goals of a coherent and cohesive text, that is, adapted to an audience, conventions, and orthographic rules. It thus serves to evaluate whether the text is clear to the reader (Alamargot, Chesnet, Dansac, & Ros, 2006). This visual interaction with the TPSF may lead to minor or major revisions of the text either at the point of inscription or in earlier sections (Hacker, Keener, & Kircher, 2009; Hayes, Flower, Schriver, Stratman, & Carey, 1985; Leijten, De Maeyer, & Van Waes, 2011; Van Waes, Leijten, & Quinlan, 2010). However, writers also re-read their TPSF to remind themselves of the ideas they have already covered, or they use it as a stimulus for additional planning and the generation of new ideas (Hayes, 1996; Johansson, Wengelin, Johansson, & Holmqvist, 2010). Rereading the TPSF is, thus, an important factor in the writing process. Breetvelt, Van den Bergh and Rijlaarsdam (1996) already showed that rereading is positively correlated with text quality.

Despite its central role during writing, reading during writing has received relatively little attention in empirical studies (Wengelin et al., 2009). There is only limited understanding of how the dynamics of the cognitive processes of reading and writing are coordinated and related. This study therefore seeks to fill this gap and serves not only a theoretical, but also a technical aim. From a theoretical point of view, it is important to understand the process of reading during writing. Although traditional theoretical writing models often incorporate a reading component, these models do not focus on the dynamics of writers' eye movements nor on the function of reading during writing. Detailed analyses of writers' eye movements are therefore necessary to increase our knowledge of the cognitive processes involved in writing and in this way to further develop writing models. From a technical point of view, this study explores the possibility of merging keystroke logging and eyetracking data.

As the study focuses on keyboard writing, it allows us to study the dynamics of the writing process with the aid of keystroke logging (Leijten & Van Waes, 2013; Van Waes, Leijten, Wengelin, & Lindgren, 2012; Wengelin, 2006). Keystroke logging provides detailed information about writers' activity on the computer and registers and stores all keystroke and mouse movements and their distribution across time. It allows to study writers' typing, pausing and revision behavior. However, based on keystroke logging alone it is not clear what exactly happens during writing and, more specifically, what happens during pauses or moments of non-fluent writing: these remain inconclusive indicators of cognitive effort. In this study, non-fluent writing means that a writer exceeds an individual pause threshold when making a transition from one keystroke event to the next one. This concept will be explained in more detail in the section "characteristics of reading during writing". Although pauses are often considered to correspond with planning activities, they could also reflect other writing activities, such as evaluation of the text produced so far. In order to enhance our understanding of pauses, several researchers have attempted to explore and interpret the nature of pauses, for example, by means of think aloud protocols, secondary task reaction times, and the triple task method (Beauvais, Olive, & Passerault, 2011; Kellogg, 1988). However, these methods could interfere with the writing process and as such, may influence the course of writing (Janssen, Van Waes, & Van den Bergh, 1996; Wengelin, Johansson, Johansson, & Frid, 2014).

Advances in eye tracking technology now allow collecting detailed data about the visual activities during pauses. Eye tracking is an unobtrusive tool allowing the measurement and analysis of the writer's visual behavior either during handwriting (Alamargot et al., 2006; Alamargot, Plane, Lambert, & Chesnet, 2010; Nottbusch, 2010) or keyboard writing (Hacker, Keener, & Kircher, 2017; Holmqvist et al., 2011; Torrance & Wengelin, 2010). A combined analysis of (grapho)motor activities and eye movements offers detailed, real time, temporal data which show us what and when writers are writing as well as what they are reading or looking at any given time during writing. By merging keystroke logging data and eye tracking data, it becomes possible to link the writer's visual behavior to cognitive processes involved in writing.

Eye Movements during Reading and Writing

Although eye tracking has been used in many studies to observe eye movements during reading for comprehension, it has rarely been used in the field of composition to observe the process of writers reading their own emerging text during writing (Anson & Schwegler, 2012; Hacker, Keener, & Kircher, 2017; Johansson, Johansson, Wengelin, & Holmqvist, 2008; Torrance & Wengelin, 2010; Wengelin, Leijten, & Van Waes, 2010). According to Torrance and Wengelin (2010) the lack of these studies can predominantly be explained by the complexity of studying real-time writing processes, that is, the fact that, during writing, the emerging text on the screen is constantly growing or changing as the task proceeds. Word positions on the screen are not fixed as in static reading texts, but they are (potentially) fluid throughout the writing process. This makes it difficult to automatically create one-on-one relations between fixations and the emerging text, whereas in traditional reading research this relation is predetermined and can thus be analyzed in a more straightforward way (Wengelin et al., 2009).

Combining Keystroke Logging and Eye Tracking

Recently, functionalities have been developed to combine keystroke logging and eye tracking data. A first attempt to combine reading and keyboarding process data was introduced by Simpson and Torrance (2007) who developed EyeWrite. This program aims to integrate keystroke logging with eye-movement recordings. It allows logging information about the location of fixations relative to the text and independently of screen coordinates. Although EyeWrite allows to fairly accurately match fixations to words in the text, it is developed for an experimental text editor with limited functions.

Wengelin and colleagues (2009) developed the functionality to connect and synchronize keystroke logging data from Scriptlog (Strömqvist, Holmqvist, Johansson, Karlsson, & Wengelin, 2006) with data from the Polhemus headtracker and the SMI IView X eye tracking program. They developed Timeline to provide a graphical representation of the reading and writing behavior. Timeline gives an overview of how writers distribute their attention on the monitor. Additionally, the graphical representation includes keystroke logging information from Scriptlog which is matched to the visual information. However, just as EyeWrite, Scriptlog is designed for (semi)experimental settings and formatting options are limited. Moreover, using a head-mounted eye tracker has limitations for the naturalistic setting of the experiment.

Inputlog is a keystroke logging tool which allows recording writing in an authentic writing setting (Leijten & Van Waes, 2013). Moreover, the merge function developed within Inputlog allows studying keystrokes, mouse movements and eye movements recorded via

Tobii TX300. Combining data from (remote) eye tracking and keystroke logging enables to minutely observe writers' eye movements in their own emerging text and in this way to gain a detailed understanding of the process of reading during writing in real-time, and defining the relation between these two cognitive processes. This type of information could provide new insights into the dynamics of the cognitive processes underlying text composition.

The combination of keystroke logging and eye tracking has also been used recently in translation studies, for instance, to explore translation styles (e.g., Alves, Pagano, & Da Silva, 2010; Dragsted & Carl, 2013; Kruger, 2016). In writing research Wengelin et al. (2009), Johansson et al. (2010) and Torrance et al. (2016) combined both observation methods to provide information about where on the screen writers look during text composition. These studies, innovative in their combined approach of keystroke logging and eye tracking, all point out the need for a better understanding of the cognitive dimensions of writing and for a more fundamental theoretical framework of the combination of reading and writing activities. The present study therefore aims to address this gap by studying the relation between the emerging text and several monitoring activities during text production in a controlled setting.

Characteristics of Reading during Writing

Findings by Rayner (1998) show that it is very unlikely that during writing the eyes would continuously move one letter space to the right with each key press. For pauses and revisions for example, writers would also frequently move their eyes back to an earlier point in the text (regressions), re-fixate on the fixated word or sometimes (even) skip words (Torrance et al., 2016). The interpretation of the online reading and writing characteristics currently mainly depends on measures from traditional reading research. These measures are based on static Areas of Interest (AOI) which are not applicable to the dynamic character of a text being produced and revised. In order to strategically study this dynamic process and the relation between reading and writing, we developed a classification schema of reading during writing as pictured in Figure 1. This schema is based on our literature review on reading and writing processes and, more particularly research on writing fluency (see Van Waes & Leijten, 2015 for a review). Fluency has been studied by writing researchers for many years (see also Kellogg, 1996, 2004). The recent introduction of more advanced observation in writing had an impact on the kind of fluency measures that are being used. For instance, keystroke logging facilitates the calculation of more complex fluency indicators in writing, especially by taking a process perspective to the dynamics of writing (Van Waes, Leijten, Wengelin, & Lindgren, 2012). More specifically keystroke logging programs can easily measure the latency that characterizes the transition between two characters, i.e., the interkey interval or pause. Also in the current study this process oriented perspective to writing fluency was used, and it also functions as a starting base for the categorization model in Figure 1.

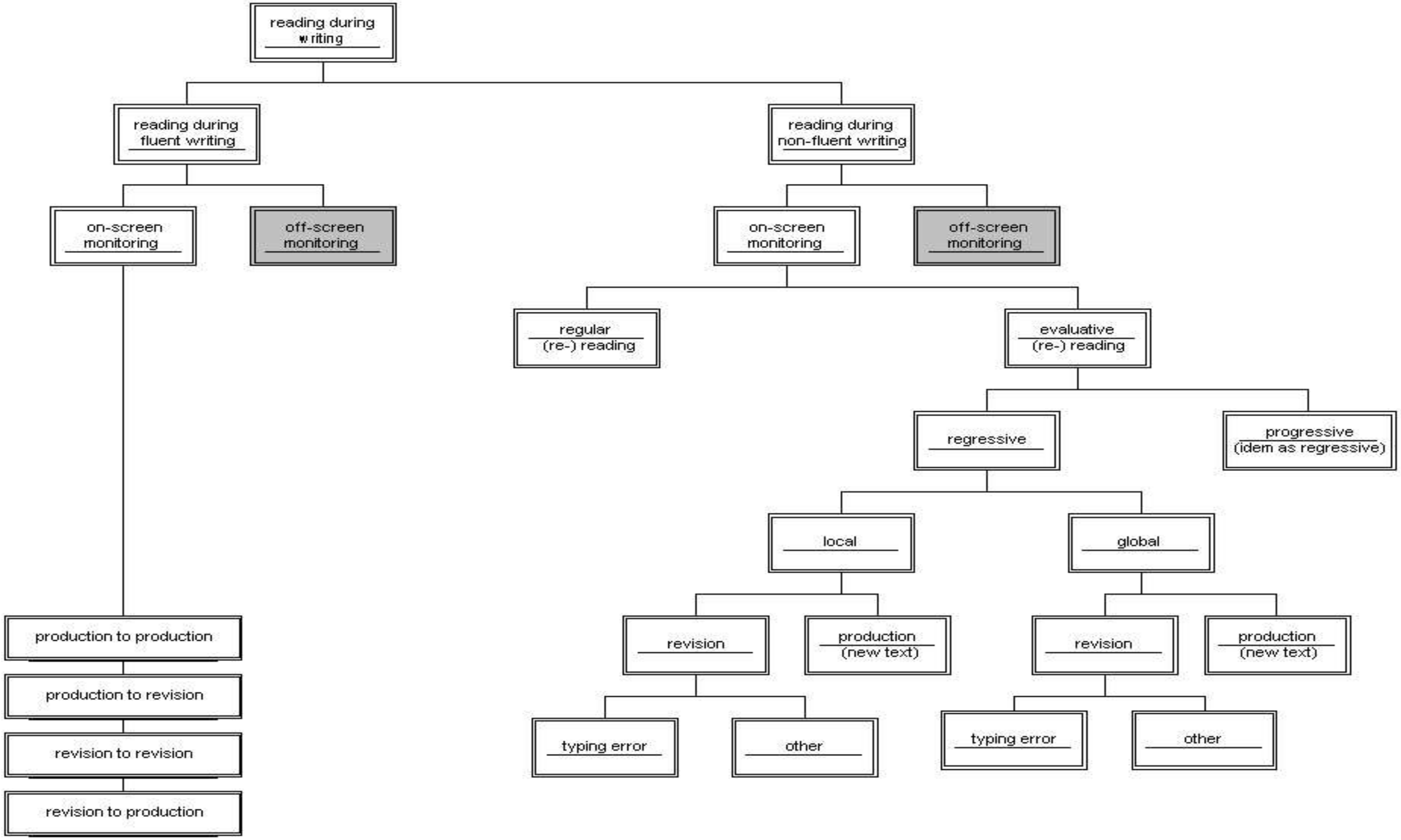


Figure *1*. A schematic overview of reading during writing.

In this schema, specific types of writing and the related reading behaviors are depicted. It should help to strategically study the relation between specific types of writing and the corresponding reading behavior. The hypothesis underlying the schematic representation is that, in the first place, a distinction could be made between reading during fluent and non-fluent writing. In the current study, fluent writing means that writers make a transition from one keystroke event to the next one without exceeding an individual interkey interval threshold which is necessary to realize a motoric transition from one key to another. Such an individual pause threshold differs from writer to writer based on their optimum typing speed and should always be calculated for each individual (cf. infra).

As shown in Figure 1, fluent transitions can entail consecutive text production, but also revisions can become so automated that the transition between production and revision and vice versa does not necessarily affect the writing pace. This may especially be the case for lower-level revisions such as correcting typing errors as can be seen in Figure 2. Although this is a fluently written sentence, there is one pause of 344ms that exceeds the individual pause threshold (i.c., 260ms) before the word 'pasta'. Nevertheless, the revision that is made to repair the typing error in the word 'groenten' ('vegetables') using the backspace key, is fluent without exceeding the individual pause threshold.

Vandaag·eten·wij·{344}pasta·met·gehakt·en·groe**tne**[BACK][BACK][BACK]**nten**.

Today·we·eat·{344}pasta·with·meat·and·vegeta**lbe**[BACK][BACK][BACK]**bles**.

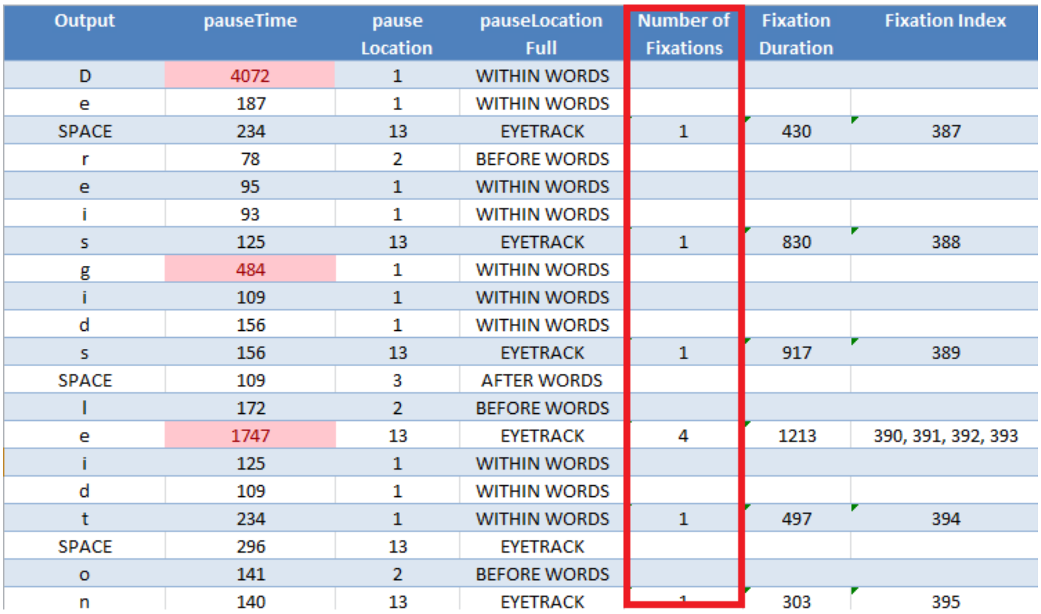
Figure *2*. Example of a fluent revision

During fluent writing, the writers' eyes are expected to move steadily from left to right with a minimum number of irregularities in the fixation behavior (Torrance and Wengelin, 2010).

Non-fluent writing, on the other hand, in this specific case would mean that a writer exceeds his or her individual pause threshold during a certain interkey transition. Such non-fluent key transitions are likely to occur more often, for example, before or after a revision, at the sentence or word boundary, and to a lesser extent in consecutive text production within words (Van Waes and Leijten, 2015; Wengelin, 2006). This type of non-fluent writing may affect the regular reading pattern in different ways. We hypothesize, for example, that higher level revisions and the associated non-fluent writing may cause more and longer fixations and are more likely to lead to local or global regressions in the reading pattern. Therefore, Figure 1 presents a hierarchical classification in which the on-screen monitoring during non-fluent writing is classified respectively as either regressive or progressive, local or global, and oriented either to execute a revision or to produce new text at the point of insertion. As mentioned before, our starting point in the classification is based on differences in writing dynamics (fluent versus non-fluent) and the related reading behavior.

Beers, Quinlan, and Harbaugh (2010) also suggest that there are several forms of reading during writing. However, they use another perspective to study different forms of reading during writing. Beers et al.'s study focused in the first place on the distinction between the location of the writer's eye movements in the text (i.e., point of inscription, local reading, global reading, and prompt reading), in order to study the corresponding writing behavior. We adopt the opposite perspective to study reading during writing. That is, in our study, the writing behavior (i.e., fluent and non-fluent writing) serves as a starting point, in order to study the corresponding reading behavior. As such, it also differs from the study by Torrance et al. (2016) who also explored the reading during writing behavior of 16 adult writers in an experimental setting. They particularly focused on the reading activity during pauses in writing and therefore excluded eye monitoring activities that occur simultaneously during writing: "Because our focus was on reading activity that occurred during breaks in ongoing production, we then excluded from analysis all fixations for which the period between fixation onset and fixation offset overlapped with keyboard activity." (p. 4). It is our hypothesis that fluent and non-fluent writing imply different gazing patterns because of the change in the underlying cognitive processes that trigger this change in writing dynamics. During fluent writing, eye movements are assumed to shift steadily to the right with the emerging text. However, this regular monitoring pattern is expected to be affected in the case of non-fluent writing by pauses and revisions, as these are indications that the writers experience cognitive overload. When composing texts, writers strive to produce their texts fluently, and fluency in writing is considered as a characteristic of proficiency. Or, as Olive (2014, p. 174) states it: "Efficient coordination of the writing process is (...) central to producing good-quality texts, and is a fundamental component of writing skill". However, due to the complexity of writing and the restrictions of our cognitive capacities, disruptions in writing dynamics are inevitable. Therefore, the classification presented in Figure 1 that is based on the observation of changes in fluency, was used as the main basis to build the experimental conditions for our study (see Materials section).

In the current study, following traditional reading research, we focus on fixations, forward saccades and backward saccades (or regressions) during reading. Moreover, we combine these eye tracking activities with the corresponding writing activities observed in Inputlog. That is, if several fixations occur during the pause before a keyboarding or mouse activity, these fixations - and their corresponding characteristics - are merged in order to match the corresponding writing activities (see Figure 3).



Output	pauseTime	pause Location	pauseLocation Full	Number of Fixations	Fixation Duration	Fixation Index
D	4072	1	WITHIN WORDS			
e	187	1	WITHIN WORDS			
SPACE	234	13	EYETRACK	1	430	387
r	78	2	BEFORE WORDS			
e	95	1	WITHIN WORDS			
i	93	1	WITHIN WORDS			
s	125	13	EYETRACK	1	830	388
g	484	1	WITHIN WORDS			
i	109	1	WITHIN WORDS			
d	156	1	WITHIN WORDS			
s	156	13	EYETRACK	1	917	389
SPACE	109	3	AFTER WORDS			
l	172	2	BEFORE WORDS			
e	1747	13	EYETRACK	4	1213	390, 391, 392, 393
i	125	1	WITHIN WORDS			
d	109	1	WITHIN WORDS			
t	234	1	WITHIN WORDS	1	497	394
SPACE	296	13	EYETRACK			
o	141	2	BEFORE WORDS			
n	140	13	EYETRACK	1	303	395

Figure *3*. Output of merged eye tracking and keystroke logging from a monitor gazer, demonstrating single and aggregated fixations.

Research Question

Until now, the characteristics of reading during writing and the relation between the cognitive processes of reading and writing – as pictured in the schema in Figure **1**– have not been systematically and empirically examined. As a result, we do not have a thorough understanding of the specific nature and function of this activity. This study aims to gain insights in reading during writing and addresses the following research question:

> To what extent does fluent and non-fluent writing affect the process of reading during writing?

In order to answer this question, we first need to explore which measures and characteristics best describe the process of reading during writing (in terms of fixations and saccades). Alamargot et al. (2006) suggest that most of the characteristics could be generalized from traditional reading research. Nevertheless, it should be investigated to which extent reading during writing may generate specific characteristics. Empirical data should allow us to refine the characteristics of reading during writing in relation to the complexity of text production.

Method

Participants

Participants in this study were 24 students from the Master in Multilingual Professional Communication at the University of Antwerp (Belgium). The group consisted of 22 female and 2 male students between the ages of 20 and 23 (*M* = 21,96; *SD* = 0,91). The participants' mother tongue (L1) was Dutch. Because visual interactions with the evolving text depend to a large extent on writers' typing skills (Beers, Quinlan, & Harbaugh, 2010; Johansson et al., 2010) we selected a homogeneous group of skilled typists. Participants in this study were so-called 'monitor gazers' (Johansson et al., 2010) and were selected on the basis of an inventory questionnaire in which they indicated their typing skills and style, combined with a typing test (copy task). Participants showed a mean typing speed of 430,58 keystrokes per minute (*SD* = 66,67). They all indicated to use eight or more fingers when typing and were able to look frequently at their evolving text on the computer screen while typing. Moreover, most of them had taken a typing course. All participants were informed about the procedure of the study prior to the experiment. They all received and signed a consent form in line with the ethical research code of the University of Antwerp.

Materials

Typing task

A brief typing task was developed which required participants to repeatedly type as accurately and quickly as possible a single easy sentence for two consecutive minutes: "Vandaag maak ik graag extra tijd vrij om deel te nemen aan een schrijfonderzoek [Today I am happy to spend some extra time participating in a writing experiment]. This typing task enabled to study the participants' typing skills and to calculate an individual typing maximum which served as a baseline or individual benchmark for fluent writing during further analyses (cf. Van Waes & Leijten, 2015). This is important because interkey intervals may differ considerably between participants (Cerni, Velay, Alario, Vaugoyeau, & Longcamp, 2016; Douhou & Magnus, 2009; Guven & Sogukpinar, 2003; Van Waes, Leijten, Mariën, & Engelborghs, 2017). The latter study compares different lexical and non-lexical copy tasks to measure (motoric and) typing skills in different contexts (letter combinations - word combinations - sentences; see also Grabowski, Weinzierl, and Schmitt (2010); Wallot and Grabowski (2013). The study shows that a copy task based on repetitive sentence reproduction - like the one in this study - is an adequate way to measure typing fluency in a lexical context with low cognitive load (no planning, no spelling problems etc.). Using these results of the copy task in other writing task analyses assures a more refined analysis of participants' writing fluency and helps to define an individual pause threshold to distinguish fluent and non-fluent writing.

Experimental conditions

Different sets of experimental sentences were developed, based on the schema in Figure 1, to operationalize the different conditions. Each condition aimed at eliciting a different type of text production corresponding to different hypothesized proportions of fluent and non-fluent key transitions and corresponding fixation behavior. Bearing in mind the current theoretical writing models, it was considered appropriate to primarily develop strictly controlled writing tasks on the sentence level. Although observing the composition of a full text in an authentic writing setting seems appealing, such open writing tasks risk being influenced by too many other variables that are beyond our control (Leijten, Van Waes, Schriver, & Hayes, 2014; Schrijver, Van Vaerenbergh, Leijten, & Van Waes, 2014). Low-level tasks, on the other hand, allow to reduce the complexity of writing and as such to implement highly controlled conditions based on the reading-during-writing schema presented. As opposed to a text, sentences are easily comparable across participants because they all perform the same task. Different sets of sentences were developed, aiming to manipulate the interaction between reading and writing at different levels:

Condition 1: Fluent writing
Condition 2: Non-fluent writing due to a spelling challenge (homophone verb ending)
Condition 3: Non-fluent writing due to revision at the word level
Condition 4: Non-fluent writing due to revision at the sentence level

Each condition consisted of a set of 10 sentences. Keeping in mind that sentence characteristics could influence participants' writing and reading behavior, the sentences were constructed carefully, and, as shown in Table 1, they were strictly controlled across conditions for the number of words, number of characters, mean and standard deviation of word length, grammatical characteristics, and word frequency using Subtlex (Keuleers, Brysbaert, & New, 2010). Appendix A shows an example of each sentence condition.

Table 1
Sentence Characteristics

Characteristic	*M*	*SD*
Number of words	8	0

Number of characters (with spaces)	44	0,6
Mean word length (in characters)	4,7	0,1
Word frequency[1]	4,3	1,3

[1] Based on Subtlex Log10WF

Aiming at an increased complexity across conditions, the four different conditions were designed as follows:

Fluent writing condition (Condition1). Sentences in the first condition did not contain any difficulty. These sentences were especially developed to elicit fluent writing. Nevertheless, revisions (i.e., correcting typing errors) and non-fluent transitions typical of authentic writing could occur here as well.

Spelling challenge condition (Condition 2). Sentences from the second condition all contained a spelling difficulty, aiming to cause instances of non-fluent writing and longer local fixations. The difficulty here was a spelling challenge in Dutch verb spelling (Sumner, Connelly, & Barnett, 2013), regarding the correct ending of (homophone) verb endings either with a 'd', 't' or 'dt'? Dutch verbs with a stem ending in -d should have an additional -t for the second and third person singular, which is a problem for many Dutch-speakers (cf. Frisson & Sandra, 2002). Similarly, some verbs ending with a -t in the present tense for the first, second and third person singular end with a -d in past tense, which is pronounced in the same way. There is no distinction in pronunciation between voiced and voiceless alveolar plosives at the end of Dutch verbs.

Word revision condition (Condition 3). The third condition contained sentences in which participants were required to first perform a fluent sentence, similar to the sentences in condition 1. Second, they were instructed to revise one word in their initial sentence (local revision). This word revision was hypothesized to cause non-fluent writing at the word, or local, level. This procedure was repeated for the consecutive sentences in this condition. In all sentences, a noun had to be replaced by another noun.

Sentence revision condition (Condition 4). In the fourth condition, participants were required to first type a fluent initial sentence using the passive voice. Directly after performing each sentence, they were instructed to revise their passive sentence by using the active voice (global revision). This sentence revision was hypothesized to cause non-fluent writing at the sentence, or global, level (Kellogg, Turner, Whiteford, & Mertens, 2016). The passive sentence remained on the screen and participants were free to use their own preferred revision strategy, which could mean shifting and revising words in the passive sentence or deleting and rewriting the whole sentence. This increased the authenticity of the writing strategy data. Both in conditions 3 and condition 4, the transitions between the initial sentence and the revision phase were manually coded at the individual sentence level within the merged logging file.

Template for experimental sentences

Participants were required to write the experimental sentences in a predefined template in Microsoft® Office Word, as pictured in Figure 4. This template contained a new page for each sentence with just enough blank space to write that one sentence. After finishing their sentence, participants were required to click on the 'continue' button under the blank space in order to continue to the next page for the next sentence. We developed this template to ensure that participants focused only on the sentence being produced and not on previous sentences. Consequently, we could isolate the interaction with the TPSF with certainty at the sentence level. Hence, from a technical perspective, the template solved possible problems with the eye tracker’s accuracy in the vertical dimension (cf. Van Hooijdonk, Commandeur, Cozijn, Krahmen, & Marsi, 2007). Displaying just one sentence at a time, allowed us to ignore this vertical dimension in the measured position of gaze events. It also allowed us to use a fairly regular font size (Calibri 14pt) instead of a larger font size, which is often unfamiliar for writers and as a result, may influence their regular reading and writing behavior.

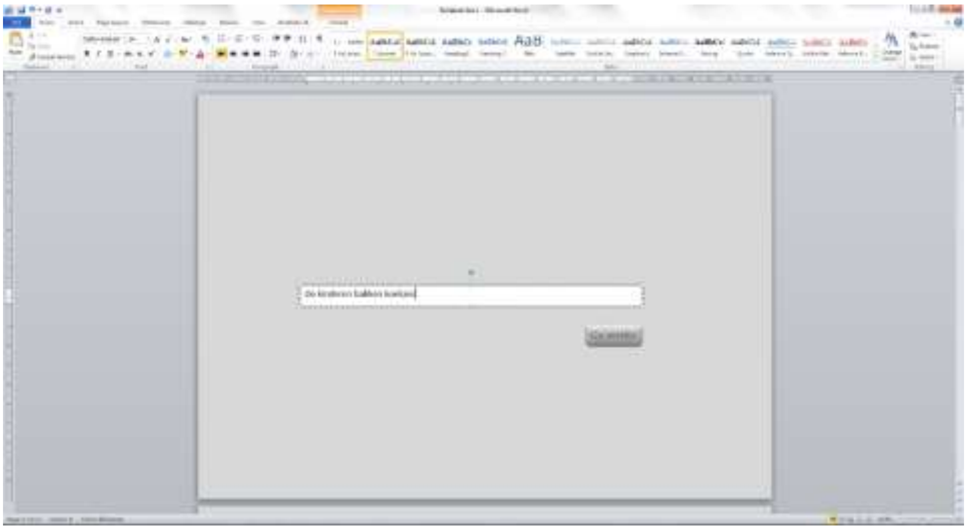


Figure *4*. Screenshot of the template in MS® Word.

Design and Procedures

The central research question was examined by means of a within subjects design in which each participant performed the four different sets of experimental sentences. Data was collected from each participant individually in a single session of approximately 40 minutes.

At the start of the experiment, each participant filled in a consent form which informed them about the goals of the study. Participants first performed the brief typing task. Then they performed the four different sets of sentences. To avoid order, learning, or fatigue effects, the sets were counterbalanced across participants. The different sets were presented in blocks to the participants, rather than conditions being randomly interleaved. All sentences were presented to the participants as dictated audio fragments. The recordings of individual sentences were presented at the beginning of each trial after which participants transcribed what they heard. In conditions 3 and 4, additional instructions were given to revise the transcribed sentences. Using these audio fragments has at least two advantages. Firstly, it allows a standardized presentation of the sentences to participants, and secondly it helps to minimize the amount of cognitive load that is often involved in the composing process which risks to interfere with the transcription processes (Bourdin & Fayol, 2000). Using dictations thus allows a combination of reading and writing activities. The sentences were dictated by an external expert in linguistics and journalism who is accustomed to formulating clearly and accurately. Each fragment contained a dictation of the sentence followed by a pause in which the participants had the time to write their sentence as accurately and quickly as possible. The pause length and dictation pace in the fragments was based upon five pre-tests with skilled typists. After each dictated sentence, an auditory signal (beep) was given which indicated that students had to press the 'continue' button in the template upon which the next dictation followed. In the conditions that required revision, participants first wrote the initial sentence and then received the instruction to revise one word or the whole sentence. They then had time to make the required revision and finally heard the beep to continue to the next sentence.

To ensure that participants were familiar with the procedure, they received written and oral instruction prior to each condition. Moreover, for each condition two practice sentences were added before starting the experimental sentences.

Apparatus

Keystroke logging

Participants' linear development of the writing process was logged via the keystroke logging program Inputlog (www.Inputlog.net) which captures the keyboard and mouse input and their distribution in time (for a review, see Leijten and Van Waes (2013). Data analyses from Inputlog allow to adequately describe and differentiate between the various phases of writing and to study writing fluency on the basis of the sentence manipulations. In this study, the focus is on the percentage of non-fluent transitions, revisions, the writing and pausing time and on the pause location. The writing process data were matched to eye tracking measures via the merge function incorporated in Inputlog (Leijten & Van Waes, 2014). The resulting file is an xml-file in which both the eye tracking log events (with their descriptive characteristics as generated by Tobii Studio) and the original Inputlog events are listed linearly and time based (consecutively) in one file (see also Figure 3). This latter analysis also provides a so called 'condensed' representation of the merged data. In this analysis the eye tracking data that occur between two Inputlog events are aggregated and added to the corresponding key or mouse event. Variables that are added at the event level are, among others, number of fixations, total fixation duration, average fixation duration, sum of absolute distance values on the horizontal axis including all valid samples (negative values are transformed to positive values) and regressions (see www.inputlog.net).

Eye tracking

All writing tasks were logged with the Tobii TX300 remote eye tracker capturing gaze data at 300 Hz and provided with a 23" monitor (aspect ratio 16:9). A second computer monitor was connected with the eye tracker which mirrored the display of the participants' computer and allowed the researcher to supervise the sessions. Participants were seated 55-65 cm from the computer screen and they were calibrated using a standard 9-point calibration. The source of light in the room was kept constant and the blinds were closed (cf. O'Brien, 2010), resulting in an illumination in the room between 210 and 320 lux which are normal light conditions in an office environment. We used the I-VT fixation classification filter included in Tobii Studio (version 2.3 and later) and implemented a temporal fixation threshold of 60ms (Jones, 2014; Prunty, Barnett, Wilmut, & Plumb, 2014).

To capture aspects of reading behavior, and to explore which characteristics of reading illustrate the manipulated increased complexity in writing, we focused among others on the percentage of fixated events (i.e., a typing event in the merged Inputlog/Tobii data file that is preceded by one or more eye-fixations), the mean fixation duration, the distribution of fixations on fluent and non-fluent transitions, and their distribution within words and between words. These measures enable us to better define the relation between reading and writing and to study the parallels between these two processes.

Furthermore, we studied the percentage of regressions and regression duration. The main aim here is to understand to what extent local or global revisions in writing are related to regressions in reading. In reading research, regressions can be defined based on the unit of analysis (e.g., word or sentence boundaries cf. Clifton, Staub, and Rayner (2007). However, because these boundaries do not always have fixed positions in emerging writing, we needed to develop another threshold. In this study, a regression was defined as a fixation that was located backwards on the X-axis compared to the previous fixation, with a minimum distance of six pixels (which is equal to approximately one character for the font we used). We used this threshold for regressions to be able to include the spelling (low-level) activities and at the same time we filtered out small meaningless regressions that could cause noise in the analyses.

Individual pause threshold

In order to define the boundary between fluent and non-fluent writing, an individual pause threshold was calculated for each participant. Based on the data from the initial typing test, the individual mean interkey interval (IKI) within words (+1*SD*) was calculated and used as a threshold to distinguish between fluent and non-fluent writing in a meaningful way from a motor perspective. The average pause threshold of the participants in this study is 246ms.

Data Analysis

To analyze the data we used multilevel mixed model analyses (IBM SPSS 23). For the continuous variables we used the linear mixed-effects models (MIXED) procedure; for the binary variables a generalized linear mixed model (GLiMM) was used. Multilevel analysis fits models more generally in comparison with general linear model procedures and it more adequately handles unequal variances and unequal numbers of repetitions (Quené & Van den Bergh, 2004, 2008; Van Waes, Leijten, & Quinlan, 2010). Regression models were compared on the basis of the log-likelihood ratio test or Akaike's information criterion (Field, 2009). The multilevel approach also better addresses the inherent characteristics of the current data as they are nested in a hierarchy. More specifically, in our study all observations are nested in the cross-classification of participants, conditions and sentences (random part).

Results

Fluent and Non-fluent Writing

In order to answer our research question regarding reading during fluent and non-fluent writing, it is important to first study the extent to which extent the different conditions indeed influenced fluent and non-fluent writing behavior. We therefore controlled whether we succeeded in eliciting variance in writing fluency and thus if our manipulations were successful. Therefore, we first conducted a control analysis to compare the fluent initial sentences (IS) without the revision phase for the conditions with fluent writing, word revision, and sentence revision on the one hand with the condition with a spelling challenge. This enabled us to control for possible differences between fluent writing and non-fluent writing due to a spelling challenge related to homophone verb endings. It also allowed us to check whether the initial sentences in both revision conditions were indeed fluent in the first place, indicating that possible differences between conditions were thus not caused by differences in the initial sentence but by (differences of) the manipulation, that is, the required revision at word or sentence level. Data were analyzed quantitatively using GLM repeated measures across the different conditions and Bonferroni post hoc analyses for pairwise comparisons between the conditions.

Table 2
Mean Production Times and Percentage of Non-Fluent Writing per Sentence

	Condition 1 Fluent *M (SD)*	Condition 2 Spelling challenge *M (SD)*	Condition 3 Word revision *M (SD)*	Condition 4 Sentence revision *M (SD)*
Production time IS (s)	6.29 (1.01)	6.92 (1.01)	6.34 (1.16)	6.55 (0.96)
Production time IS+R (s)	"	"	9.17 (1.48)	14.15 (1.99)
Pct. non-fluent writing IS	6.26 (3.05)	7.75 (3.66)	6.15 (2.52)	6.79 (2.43)
Pct. non-fluent writing IS+R	"	"	9.66 (2.33)	10.45 (3.65)

Note. IS = Initial Sentence; IS+R = Initial Sentence + Revision phase; " = same as the above.

Table 2 shows the mean production times of the sentences across the four different conditions. Results from the repeated measures analysis regarding production time of the initial sentence show that there is a significant difference between the different conditions, $F(3, 21) = 8.766$, $p = .001$, $\eta^2 = .556$. Post hoc analyses (Bonferroni) show that condition 2 significantly differs from condition 1 ($p < .001$), condition 3 ($p = .001$) and condition 4 ($p = .008$), while there are no differences between conditions 1, 3 and 4. This suggests that the verb spelling challenge requires significantly more production time compared to writing fluent initial sentences, but that all other initial sentences require the same amount of production time.

In addition, a second analysis is carried out to compare the four different conditions, this time *including* the revision phase for condition 3 and 4 (IS+R). This allows studying the effect of the revisions on the writers' writing and reading behavior. Analyses of the total production time, including the revision phase for condition 3 and 4, shows a significant difference between the different conditions $F(3, 21) = 252.886$, $p < .001$, $\eta^2 = .973$. Post hoc analyses further reveal a significant difference ($p < .001$) between all conditions. The imposed revisions caused a significantly higher production time compared to fluent writing or a verb spelling challenge. Moreover, revision at the sentence level requires significantly more production time than a revision at the word level. The production time for the different sentences in all four conditions and the effect of the revision phase is depicted in Figure 5. The dark part of the bar is the production time of the initial sentence, and the lighter part is the additional time imposed by the revision phase.

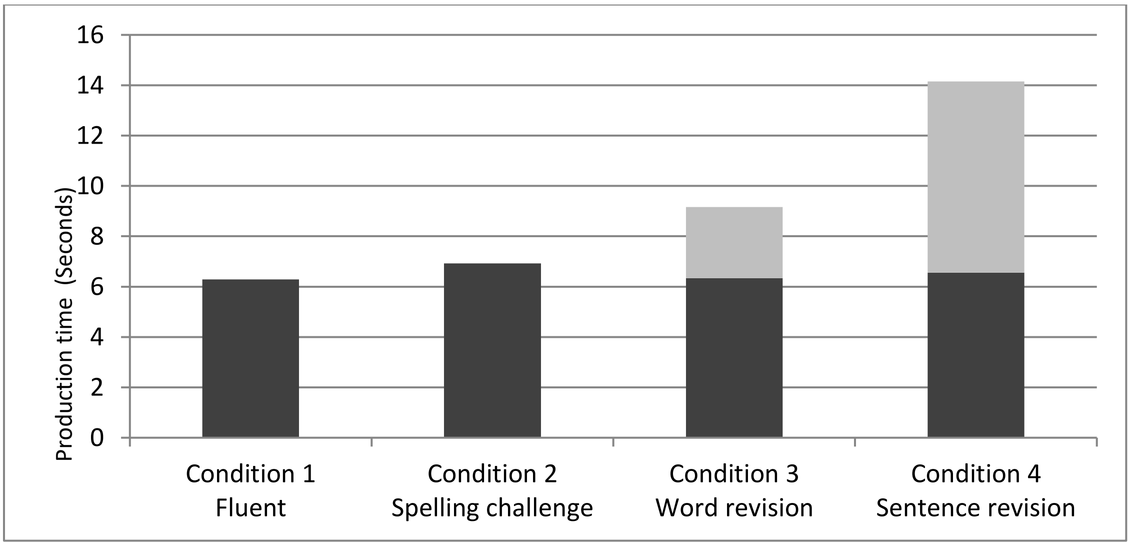


Figure 5. Mean production time (in seconds) of the sentences across the different conditions

Because of differences between conditions in total process time per sentence, it is important to use relative measures in further analyses. To evaluate the manipulation of sentence production in the four different conditions, we used the percentage of non-fluent writing as an indicator. The percentage refers to the number of non-fluent key transitions divided by the total number of transitions in each sentence. Non-fluent transitions are determined based on each participant's individual pause threshold. Results in Table 2 show that for the initial sentence, a significant main effect was found between the different conditions, $F(3, 21) = 4.840$, $p = .010$, $\eta^2 = .409$. Post hoc analyses show that condition 2 significantly differs from condition 1 ($p = .010$) and condition 3 ($p = .007$) but no statistically significant difference was found between conditions 1, 3 and 4. These results are displayed by means of the dark bars in Figure 6 and show, again, that our manipulation was successful. That is, condition 2 is the only condition with a manipulation in the initial sentence (i.e., a verb ending spelling challenge), while the other three conditions were expected to be similar in the initial sentence. The data shows that this was indeed the case.

Analyses of proportions of key transitions above the threshold, comparing the full task of the four conditions, including the revision phase for the conditions 3 and 4 (IS+R), show a significant difference between conditions $F(3, 21) = 34.735$, $p < .001$, $\eta^2 = .832$. Post hoc analyses show that condition 1 significantly differs from condition 2 ($p = .010$), condition 3 ($p < .001$), and condition 4 ($p < .001$). Moreover, there is a difference between conditions 2 and 3 ($p = .005$) and condition 2 and 4 ($p < .001$) but not between condition 3 and 4 ($p = .100$). This suggests that there is a larger proportion of non-fluent passages for sentences that contain a (verb) spelling challenge or a revision. However, no difference was found between the level of revision. Although sentence revision at the sentence level takes longer than word revision, as the former revision unit is longer, the results show that it is executed equally fluent as word revision.

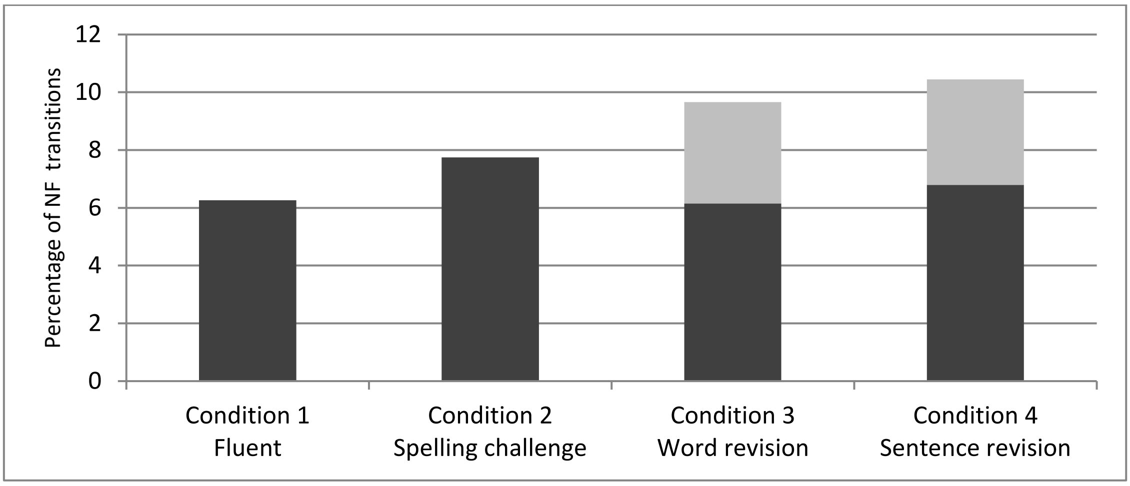


Figure *6*. Percentage of non-fluent transitions for the different conditions.

Having discussed the control analyses for the initial sentence (IS) for the variables production time and the percentage of non-fluent writing, we will now describe how we controlled for possible differences in the IS for all other variables in this study. Not a single difference was found between the initial fluent sentences between conditions 1, 3, and 4. This means that for all variables, possible differences between these conditions were due to the manipulated revision phase and not to differences in the initial sentences. Now that we know that the manipulation was successful, the results of the IS+R data can be presented.

Pausing Behavior

Table 3
Estimated Mean Pause Times for Fluent and Non-Fluent Writing Events

	Condition 1 Fluent	Condition 2 Spelling challenge	Condition 3 Word revision	Condition 4 Sentence revision
Pause time (ms)	134.4	144.3	166.9	165.6

Standard error (SE)	*3.8*	*3.8*	*3.7*	*3.7*
Pause time fluent writing (ms)	116.3	120.2	124.4	123.5
Standard error (SE)	*2.6*	*2.6*	*2.6*	*2.5*
Pause time non-fluent writing (ms)	390.1	412.4	514.5	486.1
Standard error (SE)	*20.2*	*19.9*	*19.3*	*19.0*

Table 3 reports on the multilevel data analysis regarding the interkey pause times in writing. A detailed report of the mixed models (fixed and random part estimates) is included in Appendix B. Analyses of the estimated mean pause time show that there is a difference between conditions, $F(4, 35.4) = 564.2$, $p < .001$. Post hoc analyses show that condition 1 is characterized by shorter pauses than condition 2 ($p = .001$), condition 3 ($p < .001$) and condition 4 ($p < .001$). Condition 2 also differs from condition 3 ($p < .001$) and condition 4 ($p < .001$). However, no significant difference was found between conditions 3 and 4. So, a verb spelling challenge increased the absolute pause time, and revisions elicited longer pauses than fluent writing. However, no difference was found between the levels of revision (i.e., word or sentence revision).

In general the estimated mean pause time for a non-fluent transition is on average 326ms (SE=1.4) longer than for a fluent transition, that is, about three to four times longer. There is a main effect for both fixed effect factors, that is, condition and fluency, respectively. $F(4, 94.2) = 273.7$, $p < .001$ and $F(1, 57931.5) = 51616.3$, $p < .001$. These finding are confirmed for the split analysis, in which we analyzed the pause time for the fluent and non-fluent transitions separately. This analysis shows that all conditions significantly differ from each other, except for the pause time in the revision conditions (3 and 4; $p = 1.0$) for fluent transitions and the two non-revision conditions (1 and 2; $p = .267$) for non-fluent transitions. Moreover, based on a comparison of the respective loglikehood scores, the model including both the four conditions and the fluency factor showed the best fit to the data (χ^2(2,40000); *p*<.001).

The variable *percentage extra pause time based on individual IKI* (Interkey Interval) is a more relative measure taking into account the individual pause threshold (cf. Chukharev-Hudilainen, 2014; Leijten, Van Horenbeeck, & Van Waes, 2018; Van Waes, Leijten, Lindgren, & Wengelin, 2015; Wengelin, 2006). For example a pause time of 500ms does not necessarily mean the same for each participant, because they all have their own typing speed. This 500ms pause would be relatively long for participant A, who has a mean IKI of 150ms, but rather small for participant B who has a mean IKI of 400ms. It is therefore important to look at the pauses as a percentage of extra pause time on top of the individual IKIs as this is a more accurate measure. For example, in the case of participant A who has a mean IKI of 150ms, and (shows) a pause time of 500ms at a certain transition, there is 233% of extra pause time based on its individual IKI, while for participant B who has a mean IKI of 400ms, a pause of 500ms means that there is only 25% of extra pause time. These sometimes quite large differences between individual students show the need for a relative measure of pause time taking into account individuals' pause threshold. The individual mean IKI provides researchers with a more fine-grained measure to differentiate between individual writers.

Analyses of the percentage extra pause time show a significant difference between the conditions *F*(3, 21) = 43,693, *p* < .001. Post hoc analyses show that condition 1 significantly differs from condition 2 (*p* = .013), condition 3 (*p* < .001) and condition 4 (*p* < .001). Moreover, condition 2 differs from condition 3 (*p* < .001) and from condition 4 (*p* < .001). However, no difference was found between conditions 3 and 4 (*p* = .429). Just as for the absolute pause time, a spelling challenge or a revision impose longer relative pause time, but no difference is found regarding the level of revision (i.e., word or sentence revision).

These analyses reveal some differences in pausing and writing behavior related to the characteristics of the experimental conditions. In the next series of analyses we will link these results to the corresponding reading behavior.

Fixation behavior

Table **4** shows the participants' fixation behavior in the different conditions. Because of initial differences in the total process time, we use relative measures (i.e., the number of fixations as a percentage of the total number of events per sentence).

Table 4
Fixation Behavior for the Different conditions: Overall, Within Words, and Between Words

	Condition 1 Fluent	Condition 2 Spelling challenge	Condition 3 Word revision	Condition 4 Sentence revision
Pct. of fixated events	19,6	20.6	21.6	23.3
Fixation duration (ms)	469.0	478.2	594.8	530.6
Standard error (SE)	*28.0*	*27.9*	*27.5*	*27.0*
Pct. of fixated Within Words	19.7	21.0	20.9	21.9
Fixation duration Within Words	464.4	470.9	446.5	449.6
Standard error (SE)	*28.4*	*28.3*	*28.2*	*27.9*
Pct. of fixated Between Words	20.5	20.8	21.1	22.7
Fixation duration Between Words	471.8	482.8	567.4	480.5
Standard error (SE)	*33.6*	*33.6*	*32.9*	*31.7*

Analyses show a significant difference between conditions regarding the percentage of fixated events ($F(4, 60.51) = 88.57$, $p < .001$), indicating that the likelihood that a typing event is preceded by one or more fixations gradually increases when fluency decreases. Post hoc pairwise comparison analyses show that there is no significant difference between condition 1 and 2 (p =.079), and condition 2 and 3 (p =.098). However, all other conditions differ significantly from each other (p <.001).

The analysis of *fixation duration* shows a comparable pattern. A significant main effect was found between the different conditions: $F(4, 93.13) = 128.25$, $p < .001$. Post hoc pairwise comparison analyses show that there is no significant difference between condition 1 and condition 2. However, all other conditions differ significantly from each other ($p < .001$). We notice that the estimated fixation duration in condition 3 is a bit higher than in condition 4; the largest differences are found between the first two and the third condition, resp. a mean difference of 116.6 ms (*SE: 13.3)* and 125.8 ms (*SE*: 13.5).

In a next step of the analysis we characterized fixations at two text levels, distinguishing fixations within words and between words. In line with the general findings, a main effect of condition on the fixation probability at both levels was found (respectively, $F(4, 60.512) = 380.51$; p < . 001 and $F(4, 37.408) = 71.23$; p <. 001)., but no significant effect of text level ($F(1, 55.87) = .809$; $p = .368$). Fixation probability ranges between 19.7% and 21.9% at the within word level, and between 20.5% and 22.7% at the between word level. So, about one in five typing events are preceded by one or more fixations at both text levels.

However, the estimated mean length of a fixation is affected by the text level. On average the fixation duration increases with 42.5 ms (about 10%) when we compare within and between word eye fixations ($F(1, 12146.0) = 37.55$; $p < .001$). The mean length of a within and between word fixation does not differ between the different conditions, ($p = .275$).

Further exploration of the location of the fixations provides a complementary perspective by studying the percentage of non-fluent or fluent transitions that were fixated during writing (see Table 5).

Table 5

Fixation Behavior for Non-Fluent (NF) and Fluent (F) Writing

	Condition 1 Fluent	Condition 2 Spelling challenge	Condition 3 Word revision	Condition 4 Sentence revision
Pct. fixated NF writing	31.8	35.3	32.5	35.5
Pct. fixated F writing	18.6	19.2	20.2	21.7
Fixation duration NF (ms)	570.8	591.9	617.3	573.2
Standard error (SE)	*42.7*	*40.2*	*38.3*	*36.6*
Fixation duration F (ms)	459.5	461.8	591.6	524.5
Standard error (SE)	*28.0*	*28.0*	*27.4*	*26.8*

Note. NF= non-fluent; F = fluent.

A multilevel analysis addressing the probability that a fluent or a non-fluent interkey interval is being fixated or not, shows a significant effect of both fluency ($F(1, 60.511) = 515.857$; $p < .001$) and of condition ($F(5, 60.511) = 13.017$; $p < .001$). The overall estimation of the chance that a non-fluent transition is fixated, is more than 30% higher (odds ratio 2.01; see Appendix C) in comparison with a fluent transition in text production. A supplementary pairwise contrast analysis shows that conditions 1 and 2, and 2 and 3, do not differ significantly from each other (resp. $p = .186$ and $p = .210$); all other conditions do ($p < .001$).

Analyses comparing fixation duration with respect to condition and fluency show a significant main effect for both factors (condition: $F(3, 13393.8) = 37.587$; $p < .001$ - fluency: $F(1, 13411.0) = 22.673$; $p < .001$). Conditions 1 and 2 do not differ significantly from each other; all other conditions do. This suggests that the conditions requiring revision evoke longer fixation duration than the conditions that do not require revision. Condition 3 - the word revision condition - is characterized by the longest estimated mean fixation duration. Overall, independent of condition, non-fluent transitions are estimated to be on average 58.3 ms (*SE:* 12.2) longer than fluent transitions, which means about 12% longer (see Appendix C; 7).

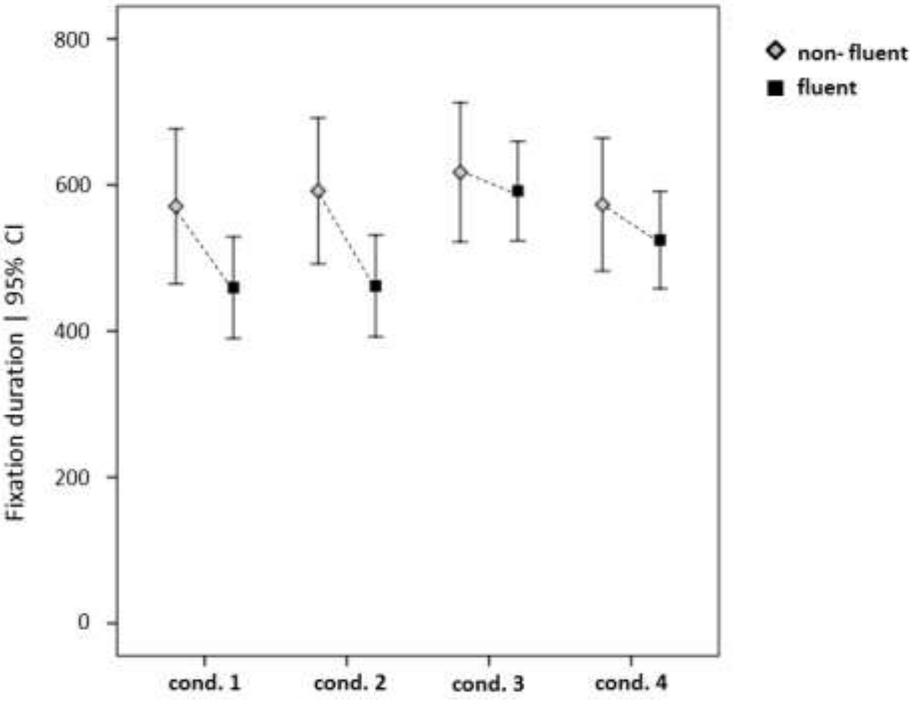


Figure *7*: estimated mean fixation duration (ms) for non-fluent and fluent transitions in the four conditions.

Regression behavior

Finally, we compared the participants' regression behavior in the four conditions as indicated by the probability and the duration (Table 6).

Table 6
Average Regression Behavior per Sentence

	Condition 1 Fluent	Condition 2 Spelling challenge	Condition 3 Word revision	Condition 4 Sentence revision

Pct. of regressions / Events	0.8	1.1	2.1	2.7
Regression duration (ms)	500.5	459.7	826.2	690.7
Standard error (SE)	*10.3*	*95.7*	*76.7*	*69.9*

A generalized linear mixed model shows a significant effect of condition (F(4, 60.512) = 380.510, $p < .001$) on the probability that a regression occurs. However, one should note that the probability is rather low, and varies between 0.8% and 2.7% of the total fixations. Post hoc analyses show that all conditions differ significantly ($p \leq .001$), only conditions 1 and 2 are not significantly different from each other (marginal effect: $p = .065$).

In addition to the probability of regression, a final analysis was performed to examine the estimated mean duration (ms) of these regressions across the different conditions. A significant main effect was found of the experimental conditions, F(3, 1240.51) = 7.712, $p < .001$. Post hoc analyses show that conditions 1 and 2 differ from condition 3 (resp. $p < .005$; $p <$*.001*), as well as that conditions 2 differs from condition 4 ($p < .05$). All other pairs do not differ significantly with respect to the duration of the regression ($p > .05$). So, especially sentences requiring an explicit revision (viz. condition 3 and 4) stimulate more and longer regressions compared to sentences without triggered revisions.

Conclusions and Discussion

The goal of this study was to determine the relation between the processes of writing and reading. In order to do so, we examined the extent to which fluent and non-fluent writing affect the process of reading during writing. We opted for a controlled study in which we managed to elicit different levels of writing fluency. Different sets of sentences were developed, aiming to manipulate the interaction between writing and reading at different 'fluency' levels: (C1) Fluent writing; (C2) Non-fluent writing due to a spelling challenge (homophone verb ending); (C3) Non-fluent writing due to revision at the word level, and (C4) Non-fluent writing due to revision at the sentence level.

The manipulations proved to be successful. We managed to create a writing context in which students - gradually - needed to slow down their pace of text production and needed more time to write their sentence, resulting in more non-fluent keyboard transitions, and longer absolute and relative pause times for the three non-fluent conditions compared to fluent writing. Moreover, we found that in our experiment the overall estimated mean pause time for a non-fluent transition is longer than for a fluent transition, that is, about three to four times longer.

In a next step we focused on the main research question: To what extent is it possible to differentiate between reading during fluent writing and reading during non-fluent writing? The analyses show that the chance that a typing event is preceded by one or more fixations, gradually increases when fluency in writing decreases. The same holds for regressions. In general students showed more and longer fixations, and more and longer regressions when they were required to slow down their typing performance. More specifically, we also found that a revision phase, both at word and at sentence level, affected the reading behavior.

The multilevel analyses also show that typing fluency does not only affect reading behavior at non-fluent transitions, but also at fluent typing transitions: the chance that a fluent, or a non-fluent transition for that matter, is fixated increases when a writing context is created that requires more disfluency. The same holds for the length of the fixations at fluent and non-fluent transitions. These results show that monitoring the evolving text does not mean that every character is monitored as such. Additional analyses with respect to fixation span, showed that in fluent writing a fixation spans on average about 5 characters; in less fluent typing about 3 characters. These findings are in line with spelling studies showing that when spelling problems occurred, fluency was affected, not only immediately before the spelling problem, but also at the onset of the word (Delattre, Bonin, & Barry, 2006; Verhaert, 2016). Or, as Weingarten et al. (2004) concluded in their study on spelling issues related to suffices: "The suffix is kept active since the beginning of the word or (...) is re-activated at certain locations before the actual production of the suffix itself" (p. 139), illustrating anticipatory behavior during pauses (decrease in fluency) in case of spelling problems, e.g. related to complex verb endings as in our materials. However, the analysis of the fixation duration shows that disruptions - in this experiment elicited at different levels in the four manipulated conditions - affects the reading behavior at both non-fluent transitions and fluent transitions. The fact that we register longer fixation times during fluent typing in the less fluent conditions probably indicates that reading during writing is not only a sequential activity, but is also partly based on parallel processing (Maggio, Lété, Chenu, Jisa, & Fayol, 2012; Olive, 2014). Especially in the revision conditions (i.e., conditions 3 and 4) parallel processing seems to take place during which disruptions are anticipated (or reflected upon), influencing the fixation behavior during fluent typing.

Moreover, reading studies also suggest that information in the foveal, but also in the parafoveal fields of vision remains in the attentional focus (Reichle, Rayner, & Pollatsek, 2003). In reading, this parafoveal field includes about 5 to 7 characters (Inhoff & Rayner, 1986; Kennedy & Pynte, 2005). Our findings provide an indication that this might also apply to reading during writing (Torrance et al., 2016; Torrance & Wengelin, 2010; Wengelin et al., 2009), indicating that the text around the point of inscription is within the parafoveal vision and can still be monitored. Although in this writing context the participants are already 'familiar' with the visualized text prior to the moment it appears on screen - as they have already mentally formulated it - this does not seem to lead to larger saccades. On the contrary, writers seem to feel the need to monitor their text in an active and vigilant way.

For the condition which contained sentences with a homophone verb spelling challenge, it was found that as expected students experienced spelling difficulties. This was reflected in their writing behavior: students in the condition showed more non-fluent transitions compared to the fluent writing condition. However, the non-fluent writing did not affect their reading behavior. No differences were found in their visual behavior compared to the fluent writing condition, suggesting that non-fluent writing does not always lead to more intense reading. Following the conclusions of studies by Maggio et al. (2012; 2015) which contend that parallel processing takes place during text production, this incongruent observation may be interpreted as follows. Due to the fact that the level of complexity of the spelling challenge is relatively limited, these writers seem to be able to solve a spelling problem in parallel to typing, just by slowing down their typing pace a little and anticipating the spelling problem during (relatively) fluent text production (see also, Verhaert, 2016). Especially because the spelling problem occurs at the end of the word, and not at the onset, this type of parallel processing is feasible and might explain the fact that reading behavior is hardly affected. Moreover, even though research has shown that difficulties with Dutch verb spelling still occur into adulthood (Van Diepen & Bosman, 1999), and that even university students make about 10% errors in a two-choice verb spelling task (Bosman, 1989; Verhaert, 2016), Master's students in professional communication, as in our experiment, might experience fewer problems to solve this task. Consequently, it may not have affected their reading behavior in a negative way, but only their writing fluency. Future studies may therefore develop a condition including sentences that contain a spelling challenge of a higher complexity such as an extra manipulation for the frequency and distance effect (cf. Frisson &

Sandra, 2002; Sandra, Frisson, & Daems, 1999), attractors, or loanwords, or include participants with another background than communication. Also a complementary focus on disruptions in pace patterns - in addition to isolated IKI event above a personal threshold - might enable us to further unravel the reading-writing behavior in these type of contexts.

It is often found that pause times are longer between words than within words reflecting a higher cognitive load. This pattern of writing fluency was not fully confirmed in the reading behavior, that is, although there seems to be a tendency that the chance of fixation is slightly higher at the between word location, there was no significant main effect of text level. However, the main length of the fixations did differ: between words fixation were about 10% longer than fixations within words. Also, for the conditions that required a revision, fixation duration was longer between words than within words. For the conditions that did not require a revision, fixation time was equal within and between words. This suggests that, in terms of fixation duration, word boundaries are especially important when writers are required to revise. It is interesting that for fluent writing, these word boundaries are less important. This finding is in line with a more superficial, global monitoring in cases of fluent writing.

Average fixation time in all conditions was relatively long, (i.e., between 469ms and 595ms across the different conditions) when comparing it to fixation times in reading for comprehension (i.e., approximately 250ms). However, the fixation times in our study were also longer than the ones that Inhoff and Gordon (1997) reported in their study on reading during writing, that is, a mean fixation time of 380ms, or the mean fixation duration of 398ms reported by Wengelin et al. (2009) (see also Sita & Taylor, 2015). However, the difference between the present study and the one by Inhoff and Gordon (1997) is that the latter studied eye movements during copy typing, meaning that writers in their study needed to divide their visual attention more between the source text and the text on the monitor, whereas in our study, writers needed to focus only on the monitor which may have led to longer fixations.

Based on the results in this study we conclude that the reading-during-writing behaviour can be affected by the fluency of writing, especially when it concerns writing that requires higher level revisions. In these cases, much of the variance in fixation duration and location can be attributed to writing fluency. However, decreased writing fluency does not always seem to lead to differences in reading behaviour. Further exploration is needed to generalize these findings. This study included only strictly controlled, low-level sentences, which helped us to strategically manipulate writing fluency and avoided a situation in which we had to control for many other variables. Nevertheless, in future studies it is important to not only look at the sentence level, but also observe reading and writing behaviour in a more naturalistic setting involving a higher text level. This would also allow us to distinguish reading during typing from reading during composition, that is, the generation of multiple sentences of coherent text, including the cognitive processes of content and language planning and more global re-reading that goes beyond the sentence level, where writers might look further back into their text. Torrance et al. (2016) already focused on more high-level 'sustained reading' sequences. These are sequences consisting of three or more consecutive forward fixations, with short saccades of which none of the fixations is on the word currently being written. They also defined look-back sequences in which the writers 'hopped' backwards and forwards in their text without complying to the definition of sustained reading. We contend that a combination of this more high level approach and a focus on more low-level activities as in the current study, could lead to a better understanding of the complex interplay of reading and writing.

Moreover, also with respect to the analysis techniques, advances in keystroke logging and eye tracking integration should allow us to address additional research questions as more fine-grained data would be available. One example of such fine-grained data could be the number of characters before and after a disfluency that are affected by anticipation/reflection in the context of parallel processing (Ailhaud, Chenu, & Jisa, 2016; Maggio et al., 2012).

In this study, we proposed several measures that allow studying the dynamics of reading during writing. From this perspective, this study can be considered as a strategic step towards unravelling the relation between reading and writing.

References

Ailhaud, E., Chenu, F., & Jisa, H. (2016). A Developmental Perspective on the Units of Written French. In J. Perera, M. Aparici, E. Rosado, & N. Salas (Eds.), *Written and Spoken Language Development across the Lifespan: Essays in Honour of Liliana Tolchinsky* (pp. 287-305). Cham, Switzerland: Springer International Publishing.

Alamargot, D., Chesnet, D., Dansac, C., & Ros, C. (2006). Eye and Pen: A new device for studying reading during writing. *Behavior Research Methods Instruments & Computers, 38*(2), 287-299.

Alamargot, D., Plane, S., Lambert, E., & Chesnet, D. (2010). Using eye and pen movements to trace the development of writing expertise: case studies of a 7th, 9th and 12th grader, graduate student, and professional writer. *Reading and Writing: An Interdisciplinary Journal, 23*(7), 853-888. doi:10.1007/s11145-009-9191-9

Alves, F., Pagano, A., & Da Silva, I. (2010). A new window on translator's cognitive activity: methodological issues in the combined use of eye tracking, key logging and retrospective protocols. In I. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology,Technology and Innovation in Translation Process Research* (Vol. 38, pp. 267-292). Copenhagen, Denmark: Samfundslitteratur.

Anson, C. M., & Schwegler, R. A. (2012). Tracking the mind's eye: a new technology for researching twenty-first century writing and reading processes. *College Composition and Communication, 64*(1), 151-171.

Beauvais, C., Olive, T., & Passerault, J.-M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing process. *Journal of educational psychology, 103*(2), 415-428. doi:10.1037/a0022545

Beers, S. F., Quinlan, T., & Harbaugh, A. G. (2010). Adolescent students' reading during writing behaviors and relationships with text quality: an eyetracking study. *Reading and Writing: An Interdisciplinary Journal, 23,* 743-775. doi:10.1007/s11145-009-9193-7

Bosman, A. M. T. (1989). *Verb spelling in Dutch University students.* University of Amsterdam, Amsterdam, The Netherlands.

Bourdin, B., & Fayol, M. (2000). Is graphic activity costly? A developmental approach. *Reading and Writing: An Interdisciplinary Journal, 13,* 183-196. doi:10.1023/A:1026458102685

Breetvelt, I., Van den Bergh, H., & Rijlaarsdam, G. (1996). Rereading and generating and their relation to text quality: An application of multilevel analysis on writing process data. In G. Rijlaarsdam, H. Van den Bergh, & M. Couzijn (Eds.), *Theories, models, and methodology in writing research* (pp. 10-20). Amsterdam, The Netherlands: Amsterdam University Press.

Cerni, T., Velay, J.-L., Alario, F.-X., Vaugoyeau, M., & Longcamp, M. (2016). Motor expertise for typing impacts lexical decision performance. *Trends in Neuroscience and Education, 5*(3), 130-138. doi:10.1016/j.tine.2016.07.007

Chukharev-Hudilainen, E. (2014). Pauses in spontaneous written communication: a keystroke logging study. *Journal of Writing Research, 6*(1), 61-84. doi:10.17239/jowr-2014.06.01.3

Clifton, C. J., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. V. Gompel, M. Fisher, W. Murray, & R. L. Hill (Eds.), *Eye movement research: A window on mind and brain* (pp. 341-372). Oxford, UK: Elsevier.

Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Do irregularity effects persist on writing durations? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1330-1340.

Douhou, S., & Magnus, J. R. (2009). The reliability of user authentication through keystroke dynamics. *Statistica Neerlandica, 63*(4), 432-449.

Dragsted, B., & Carl, M. (2013). Towards a classification of translation styles based on eye-tracking and keylogging data. *Journal of Writing Research, 5*(1), 133-158.

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

Frisson, S., & Sandra, D. (2002). Homophonic forms of regularly inflected verbs have their own orthographic representations: A developmental perspective on spelling errors. *Brain and Language, 81*, 545-554. doi:10.1006/brln.2001.2546

Grabowski, J., Weinzierl, C., & Schmitt, M. (2010). Second and fourth graders' copying ability: from graphical to linguistic processing. *Journal of Research in Reading, 33*(1), 39-53.

Guven, A., & Sogukpinar, I. (2003). Understanding users' keystroke patterns for computer access security. *Computers & Security, 22*, 695-706.

Hacker, D. J., Keener, M. C., & Kircher, J. C. (2009). Writing is applied metacognition. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 154-172). New York, NJ: Routledge.

Hacker, D. J., Keener, M. C., & Kircher, J. C. (2017). TRAKTEXT: Investigating writing processes using eye-tracking technology. *Methodological Innovations, 10*(2), 1-18. doi:10.1177/2059799116689574

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. E. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1-27). Mahwah: New Jersey: Lawrence Erlbaum Associates.

Hayes, J. R., Flower, L., Schriver, K., Stratman, J., & Carey, L. (1985). *Cognitive processes in revision*. Pittsburgh, PA: Communications Design Center.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking : A comprehensive guide to methods and measures*: Oxford University Press.

Inhoff, A. W., & Gordon, A. M. (1997). Eye movements and eye-hand coordination during typing. *Current Directions of Psychological Science, 6*, 153-157.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*(6), 431-439. doi:10.3758/BF03208203

Janssen, D., Van Waes, L., & Van den Bergh, H. (1996). Effects of thinking aloud on writing processes. In C. M. Levy (Ed.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 233-250). Mahwag, N.J.: Lawrence Erbaum.

Johansson, R., Johansson, V., Wengelin, Å., & Holmqvist, K. (2008). Reading during writing: Four different groups of writers. *Working papers, 53*, 43-59.

Johansson, R., Wengelin, Å., Johansson, V., & Holmqvist, K. (2010). Looking at the keyboard or the monitor: Relationship with text production processes. *Reading and Writing: An Interdisciplinary Journal, 23*, 835-851. doi:10.1007/s11145-009-9189-3

Jones, S. (2014). From ideas in the head to words on the page: young adolescents' reflections on their own writing processes. *Language and Education, 28*(1), 52-67.

Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, memory and cognition, 14*(2), 355-365. doi:10.1037/0278-7393.14.2.355

Kellogg, R. T., Turner, C. E., Whiteford, A. P., & Mertens, A. (2016). The role of working memory in planning and generating written sentences. *Journal of Writing Research, 7*(3), 397-416. doi:10.17239/jowr-2016.07.3.4

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*(2), 153-168. doi:10.1016/j.visres.2004.07.037

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. . *Behavior Research Methods, 42*(3), 643-650. doi:10.3758/BRM.42.3.643

Kruger, H. (2016). What's happening when nothing's happening? Combining eyetracking and keylogging to explore cognitive processing during pauses in translation production. *Across Languages and Cultures, 17*(1), 25-52. doi:10.1556/084.2016.17.1.2

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

Leijten, M., Van Horenbeeck, E., & Van Waes, L. (2018). Analyzing keystroke logging data from a linguistic perspective. In K. Sullivan & E. Lindgren (Eds.), *Computer key-stroke logging and writing: Methods and applications* Amsterdam: Brill.

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

Leijten, M., & Van Waes, L. (2014). *Inputlog 6.0 Help*. University of Antwerp. Antwerp.

Leijten, M., Van Waes, L., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research, 5*(3), 285-337. Retrieved from http://www.jowr.org/articles/vol5_3/JoWR_2014_vol5_nr3_Leijten_et_al.pdf

Maggio, S., Lété, B., Chenu, F., Jisa, H., & Fayol, M. (2012). Tracking the mind during writing: Immediacy, delayed, and anticipatory effects on pauses and writing rate. *Reading and Writing: An Interdisciplinary Journal, 25*(9), 2131-2151. doi:10.1007/s11145-011-9348-1

Nottbusch, G. (2010). Grammatical planning, execution, and control in written sentence production. *Reading and Writing: An Interdisciplinary Journal, 23*(7), 777-801. doi:10.1007/s11145-009-9188-4

O'Brien, S. (2010). Eye tracking in translation process research: methodological challenges and solutions In I. Mees, F. Alves, & S. Göpferich (Eds.), *Methodology, Technology and Innovation in Translation Process Research* (Vol. 38, pp. 251-266). Copenhagen, Denmark: Samfundslitteratur.

Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research, 6*(2), 173-194. doi:10.17239/jowr-2014.06.02.4

Prunty, M. M., Barnett, A. L., Wilmut, K., & Plumb, M. S. (2014). An examination of writing pauses in the handwriting of children with Developmental Coordination Disorder. *Research in developmental disabilities, 35*(11), 2894-2905.

Quené, H., & Van den Bergh, H. (2004). On Multi-Level Modeling of data from repeated measures designs: A tutorial. *Speech Communication, 43*(1-2), 103-121.

Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language, 59*, 413-425.

Quinlan, T., Loncke, M., Leijten, M., & Van Waes, L. (2012). Coordinating the Cognitive Processes of Writing. *Written Communication, 29*(3), 345-368. doi:10.1177/0741088312451112

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The performance of natural (and unnatural) visual tasks. The E-Z reader model of eye-movement control in reading: Comparisons to other models. *Behavior & Brain Sciences, 26*, 445-526. doi:10.1017/S0140525X03000104

Sandra, D., Frisson, S., & Daems, F. (1999). Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in Dutch. *Brain and Language, 68*, 277-283. doi:10.1006/brln.1999.2108

Schrijver, I., Van Vaerenbergh, L., Leijten, M., & Van Waes, L. (2014). The translator as a writer: Measuring the effect of writing skills on the translation product. In J. Engberg, C. Heine, & D. Knorr (Eds.), *Methods in Writing Process Research [Textproduktion in Medium]* (pp. 99-122). Bern: Peter Lang.

Simpson, S., & Torrance, M. (2007). *EyeWrite (version 5.1 - Unpublished report)*. Nottingham Trent University. Osgoode, ON: SR Research, and Nottingham, U.K.

Sita, J. C., & Taylor, K. A. (2015). Eye movements during the handwriting of words: Individually and within sentences. *Human Movement Science, 43*, 229-238. doi:10.1016/j.humov.2015.01.011

Strömqvist, S., Holmqvist, K., Johansson, V., Karlsson, H., & Wengelin, Å. (2006). What keystroke logging can reveal about writing. In K. P. H. Sullivan & E. Lindgren (Eds.), *Computer Keystroke Logging:Methods and Applications* (Vol. 18, pp. 45–71). Oxford, UK: Elsevier.

Sumner, E., Connelly, V., & Barnett, A. L. (2013). Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Reading and Writing: An Interdisciplinary Journal, 26*(6), 991-1008.

Torrance, M., Johansson, R., Johansson, V., & Wengelin, Å. (2016). Reading during the composition of multi-sentence texts: an eye-movement study. *Psychological Research, 80*(5), 729-743. doi:10.1007/s00426-015-0683-8

Torrance, M., & Wengelin, Å. (2010). Writers' eye movements. In C. Bazerman, R. Krut, K. Lunsford, S. McLeod, S. Null, P. Rogers, & A. Stansell (Eds.), *Traditions of Writing Research* (pp. pp. 394-405). New York, NY: Routledge.

Van Diepen, M., & Bosman, A. M. T. (1999). Hoe spel jij gespelt? Werkwoordspelling door leerlingen van de basisschool en de middelbare school [Dutch verb spelling in students in primary school and in high-school students]. *Tijdschrift voor Orthopedagogiek, 38*, 176-186.

Van Hooijdonk, C. M. J., Commandeur, E., Cozijn, R., Krahmen, E. J., & Marsi, E. C. (2007). *The online evaluation of speech synthesis using eye movements.* Paper presented at the ICSA Speech Synthesis Workshop, Bonn, Germany.

Van Waes, L., & Leijten, M. (2015). Fluency in Writing: A Multidimensional Perspective on Writing Fluency Applied to L1 and L2. *Computers & Composition, 38*, 79-95. doi:10.1016/j.compcom.2015.09.012

Van Waes, L., Leijten, M., Lindgren, E., & Wengelin, Å. (2015). Keystroke logging in writing research: Analyzing online writing processes. In C. MacArthur, A., S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research, second edition* (Vol. 2). New York: Guilford Press.

Van Waes, L., Leijten, M., Mariën, P., & Engelborghs, S. (2017). Typing competencies in Alzheimer's disease: An exploration of copy tasks. *Computers in Human Behavior, 73*, 311-319. doi:10.1016/j.chb.2017.03.050

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

Van Waes, L., Leijten, M., Wengelin, Å., & Lindgren, E. (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*. New York/Sussex: Taylor & Francis.

Verhaert, N. (2016). *Rules of regulartities. The homophone dominance effect in spelling and reading regular Dutch forms.* (Doctoral dissertation), University of Antwerp, Antwerp.

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

Wengelin, Å. (2006). Examining pauses in writing: Theory, methods and empirical data. In G. Rijlaarsdam (Ed.), *Computer Keystroke Logging and Writing: Methods and Applications* (pp. 107-130). Amsterdam, The Netherlands: Elsevier.

Wengelin, Å., Johansson, V., Johansson, R., & Frid, J. (2014). *Writing processes and text characteristics of texts produced in a triple-task situation*. Paper presented at the EARLI SIG writing, Amsterdam, the Netherlands.

Wengelin, Å., Leijten, M., & Van Waes, L. (2010). Studying reading during writing: New perspectives in research. *Reading and Writing: An Interdisciplinary Journal, 23*(7), 735-742. doi:10.1007/s11145-009-9187-5

Wengelin, Å., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods, 41*(2), 337-351. doi:10.3758/BRM.41.2.337

Appendix A

Table 7

Example sentences for each condition (original Dutch representation and English translation)

Condition	Sentence	characters with spaces	characters without spaces	words	Mean word length
1 Fluent	Vandaag ga ik een taart bakken voor vrienden Today I will bake a cake for friends	44	37	8	4,6
2 Spelling challenge	De reisgids *leidt* ons door de smalle straten The guide leads us through the narrow streets	44	37	8	4,6
3 Word revision	Ik heb nieuwe kleding nodig voor de bruiloft (*kleding* should be revised into *schoenen)* I need new clothes for the wedding (clothes should be revised into shoes)	44	37	8	4,6
4 Sentence revision	De melk werd meteen opgedronken door de poes (Should be revised into an active sentence: De poes heeft de melk meteen opgedronken) The milk was immediately drunk by the cat (Revision: the cat immediately drunk the milk)	44	37	8	4,6

Appendix B

Table 8

Estimated Means of the Linear Mixed Modeling Results for Fixation Duration and Pause Time

	Fixation duration				Pause time	
	Zero model	Model 1	Model 2	Model 3	Zero model	Model 1
Fixed part						
Intercept						
Estimated mean	520.484				152.949	
Std. Error	27.911				4.124	
Condition 1		468.996	520.063	496.073		134.409
Std. Error		*28.014*	*30.063*	*28.932*		*3.794*
Condition 2		478.180	527.603	504.133		144.280
Std. Error		*27.922*	*29.814*	*28.910*		*3.796*
Condition 3		594.778	643.685	514.693		166.904
Std. Error		*27.450*	*29.335*	*28.742*		*3.747*
Condition 4		530.563	578.651	488.022		165.586
Std. Error		*26.960*	*28.817*	*28.443*		*3.687*
Fluency (ref. Fluent)			-58.261			
Std. Error			*12.235*			
Pause Location (ref. within word)				42.459		
Std. Error				*6.929*		
Random part						
Estimate	251581.655	251836.972	251410.163	129134.113	21156.4	21156.297
Std. Error	*3078.880*	*3077.854*	*3072.639*	*1658.939*	*124.324*	*124.322*
Participant						
Estimate	17202.363	16909.976	16951.605	17506.374	285.530	282.139
Std. Error	*5037.243*	*4951.608*	*4964.275*	*5049.832*	*83.471*	*82.353*
Sentence						
Estimate	2750.811	0.000	0.000	586.255	207.924	12.792
Std. Error	*815.920*	*0.000*	*0.000*	230.185	*50.273*	*6.212*

-2 Log Likelihood	205041.289	204995.714	204973.060	178073.650	742208.551	742126.760
Sig		0.000	0.000	0.000		0.000

Appendix C

Table 9
Generalized Mixed Modeling Results for the Probability of Fixation (Expressed in Logit Coefficients)

	Zero model	Model 1	Model 2	Model 3
Fixed part				
Intercept				
Coefficient	1.290			
Std. Error	0.074			
Condition 1		1.412	0.771	1.375
Std. Error		*0.078*	*0.083*	*0.082*
Condition 2		1.348	0.723	1.341
Std. Error		*0.078*	*0.083*	*0.082*
Condition 3		1.292	0.680	1.311
Std. Error		*0.077*	*0.082*	*0.081*
Condition 4		1.190	0.582	1.242
Std. Error		*0.076*	*0.081*	*0.080*
Fluency			0.700	
Std. Error			*0.030*	
Pause Location				0.020
Std. Error				*0.022*
Random part				
Participant				
Estimate	0.033	0.136	0.137	0.146
Std. Error	*0.040*	*0.041*	*0.041*	*0.044*
Wald Z	0.806	3.339	3.339	3.337
Sig.	0.420	0.001	0.010	0.001
Sentence				
Estimate	0.012	0.001	0.010	0.001
Std. Error	*0.005*	*0.001*	*0.010*	*0.001*
Wald Z	2.360	1.074	1.140	1.018
Sig.	0.018	0.283	0.254	0.309
Akaiki corrected	280174.3	280568.9	281412.3	259794.3
Sig		n.s.	n.s.	<.001