

Revision of Business Content on Corporate Social Responsibility: Measuring the Impact of Training on the Cognitive Effort of Second- Language University Students

Abstract

With more and more people interested in how sustainable and socially responsible companies are, the comprehensibility of content on corporate social responsibility (CSR) has become paramount. Producing easy-to-read business content – either by writing it from scratch or revising it – is a cognitively demanding undertaking, especially for second-language non-professional writers. Both formal training and sustained practice can help writers build expertise and, in turn, be considerate of their intended audience. In particular, research on the impact of training has usually yielded positive results when examining the texts produced following specific instruction. However, the extent to which training has a positive effect on the process of writing and revision is still under-researched, especially in second language. To address this gap, we report on an experimental study that examines the impact of reader-oriented training on the cognitive effort experienced by 47 second-language university students when revising CSR content. We adopted a pre-test post-test design, and we used keystroke logging and retrospective interviews to collect data on students' pausing behaviour, use of online sources, and strategies to approach the revision task. Our training seemed to reduce the cognitive effort linked with lexical choices. Furthermore, it provided some students with procedural knowledge on how to approach the revision task in a more efficient way. We also observed a general tendency to rewrite from scratch (rather than revise) CSR content despite the higher cognitive effort required by rewriting. We discuss implications for training, limitations, and future research avenues.

Keywords:

corporate social responsibility; text revision; second-language writing; plain language training; cognitive effort; keystroke logging; retrospective interviews

1. Introduction

Writing and revising texts in order to suit the needs and expectations of the intended reader is a complex task, which tends to be particularly effortful for non-native speakers of a language (Barkaoui 2007). Training and sustained practice can help manage the cognitive effort required (Kellogg 2008). Therefore, in this article, we report on an empirical investigation of the impact of reader-oriented training on the revision process intended in a broad sense, whether it involved starting a text from scratch (*rewriting*), or modifying an existing text by trying to preserve as much of its original features as possible (*narrow revision*). Our focus was specifically on the cognitive effort experienced by second-language university students before and after taking part in our (online) training. To this end, we collected data using keystroke logging and retrospective interviews. The reader-oriented training that we developed (i.e. the intervention of this study) revolved around the accessible communication of business content on corporate social responsibility (CSR). The selection of this domain was motivated, on the one hand, by the growing interest of customers and lay people in the extent to which companies engage with socially responsible activities (Golob/Podnar 2019) and, on the other hand, by the low readability of CSR content, particularly in corporate reports (Smeuninx et al. 2020).

* Alessandra Rossetti
Department of Management
Universiteit Antwerpen
alessandra.rossetti@uantwerpen.be

** Luuk Van Waes
Department of Management
Universiteit Antwerpen
luuk.vanwaes@uantwerpen.be

We start this article with a review of previous work and related research gap (Section 2), followed by a description of the method that we adopted (participants, tools, tasks, and texts) in Section 3. Section 4 delves into data preparation and analysis, while Section 5 reports the results of our quantitative and qualitative analyses. Section 6 summarises our findings and Section 7 presents limitations and avenues for future research.

2. Literature review

2.1. Training on reader-oriented communication

Being able to adopt the perspective of the reader when writing or revising requires decades of deliberate and sustained practice (Kellogg 2008). It is therefore not surprising that researchers, professional trainers, and school teachers have invested time and effort in identifying the techniques and approaches that can foster the adoption of a reader-oriented perspective (Barkaoui 2007). Matching the needs and expectations of the reader is a requirement for effective communication not only in monolingual texts (produced and revised in one language), but also in texts that are translated from one language into another (Schrijver et al. 2016). As such, reader-oriented communication is also an important component of translation curricula and modules (Federici/Cadwell 2018; Lee 2016).

Previous works varied in terms of their focus of analysis, genres, and domains. For instance, focusing specifically on paragraphing, Eden/Mitchell (1986) pointed out that training should take into consideration the reader's expectations and patterns of comprehension, rather than being purely text oriented. Paragraphing was also an important aspect in a study conducted by Kuteeva (2011), who analysed the use of a course wiki for the production of academic and professional texts among students of English as a second language. Specifically, the wiki revolved around paragraph structure, coherence, and argumentation. Results showed that the use of the wiki made the majority of students consider their audience and, in turn, pay attention to the organisation of the texts and to grammatical correctness. In her seminal work, Schriver (1992) tested the impact of using think-aloud protocol transcripts produced by readers when training writers to anticipate readers' comprehension needs. Her results showed that protocols helped writers to more accurately predict and diagnose readers' problems, as well as to attend to global textual features. Positive results were also reported by Sengupta (2000), who trained students of English as a second language in the revision of texts aimed at making them more reader-friendly, with 'reader-friendliness' being defined as appropriateness, sufficiency, and organization of information.

Training on reader-oriented communication is carried out not only in educational settings, but also in the workplace. An example is the training provided to government agencies in the United States by the Plain Language Action and Information Network (Schriver 2017). Research on reader-oriented writing in the workplace is, however, quite limited and results about the effectiveness of such training tends to be low or moderate (Janssen 2001).

In the medical domain, Warde et al. (2018) advocate for the integration of plain language training into medical curricula as a way of preparing healthcare professionals for the challenges of a multilingual and globalised society. Similarly, Sagi et al. (2021) describe the ETGAR course whose goal is to train medical students to rewrite formal discharge letters in plain language with lay patients as the intended readers. Students are instructed to go beyond the mere simplification of vocabulary and to approach the text globally, for instance by restructuring the information, using headings, and adding/removing content, so that texts are concise, well-organised, and clear (Kools et al. 2008). Moving from medicine to subtitling and audio description, the recent EASIT (Easy Access for Social Inclusion Training) project seeks to develop a curriculum for the training of experts in easy reading (Arias Badia/Matamala Ripoll 2020). In the business domain, Greer (2012) presents resources and exercises that instructors can introduce in their courses. An example of such

exercises is asking students to revise a business advertisement or a brochure following plain language guidelines.

As this section has shown, training on reader-oriented communication is relevant to a variety of settings and domains. In the business domain, very little attention has however been devoted to making CSR content easier to comprehend for a non-specialist audience (Rossetti/Van Waes 2022). This gap is surprising when considering the increasing interest in companies' sustainability efforts from a variety of stakeholders (Zamir/Saeed 2020), as well as previous research highlighting the difficulty of CSR content, particularly in corporate reports (Nilipour et al. 2020). Accordingly, in the current study, we selected CSR content as the focus of our training on reader-oriented communication (Section 3.2).

2.2. Expertise and language proficiency in revision

The goal of training is to help students/trainees bridge the expertise and knowledge gap (Carter 1990). When it comes to the revision of texts, higher writing expertise levels result in the ability to consider the text as a whole and to make global, structural changes that consider genre, medium, and intended reader (Leijten/Van Waes 2020; Lindgren et al. 2011). Furthermore, while the revision process of experts tends to be recursive and to involve different rounds (Becker 2006), less experienced writers tend to apply local, small changes to texts that preserve original meaning and structure (Faigley/Witte 1981). Interestingly, Whyatt (2018) found that the translation expertise of professionals is also transferable to monolingual paraphrasing tasks as this group produced higher-quality texts (in terms of readability, communicative quality, and errors) compared with language students. Wallace/Hayes (1991) ascribe these differences in expertise levels to the way in which revision tasks are represented in the mind of the writer/reviser. Specifically, after delivering 8-minute instruction to freshmen on how experienced revisers operate, the authors observed an improvement in revision quality (Wallace/Hayes 1991).

Similar to writing and translation expertise, levels of language proficiency can also determine the amount and types of revisions made. In his review of the research on the revision practices of experienced and less experienced writers in a first language and in a second language, Barkaoui (2007) discusses the dimensions along which these groups differ, namely: revision beliefs; audience awareness; revision time and frequency; revision processes; revision focus; and revision outcomes. In particular, with regard to audience awareness, Barkaoui (2007: 84) reports that:

Skilled writers [...] tend to spend more time thinking about the effect they have on their reader, how they wish to present themselves to the reader, what background knowledge the reader needs to have to comprehend their text, and what may interest the reader.

The author points out that these revision practices should be regarded as a continuum, since novice writers can gradually develop their expertise in (second language) writing through practice (Barkaoui 2007). Overall, despite some evidence that certain revision behaviours are transferable from first language to second language, revisions in a second language appear to be more frequent and more time consuming (Hall 1990; Lindgren/Sullivan 2006), with cognitive demands increasing also as a result of task complexity (Xu et al. 2021). Furthermore, native speakers and high-proficiency speakers of a second language tend to make more content and global edits, as opposed to the local edits (e.g., spelling or grammar) usually made by low-proficiency speakers (Barkaoui 2016; Zhang et al. 2017).

2.3. Managing cognitive effort

The observable differences among various expertise and language proficiency levels are linked with the way in which the cognitive demand imposed by revision is managed. Revising involves planning solutions to problems detected in a text, transcribing those solutions into new text and, if

necessary, revising new and existing text (Hayes 2012). In conducting these activities, the reviser has to maintain a mental representation of the content planned by the author (when revising somebody else's text), of the text as it is at any given moment, and of the needs and expectations of the prospective reader within a context (Hayes 2012; Kellogg 2008; Schriver 2012). These actions put a strain on the reviser's working memory, which is responsible for the temporary storage and processing of information during writing and revision (Olive 2012).

The limited working memory resources are easily depleted when used for tasks that have not become automatised enough (as is the case with children struggling with orthography, writers working with texts of an unfamiliar genre/topic, or low-proficiency second-language writers) (McCutchen 2000, 2011). Skilled and experienced writers/revisers, on the other hand, have internalised knowledge and strategies, thus being able to retrieve them from their long-term memory as needed by the task at hand and, in turn, circumventing working memory constraints (McCutchen 2011). However, the types of issues that need to be addressed in a text also play a role, with the notion of revision moving away from mere error correction to include different forms of text improvement (Conijn et al. 2021). In particular, while surface errors of spelling and grammar might require little cognitive effort for experienced writers and/or native speakers, the ill-defined considerations on readership and global text features (e.g., cohesion) will still require more mental resources and more text readings (Piolat et al. 2004). For writers of English as a second language, for example, the use of discourse features – such as cohesive devices – might require specific instruction (Kobayashi/Rinnert 2001).

When confronted with a difficult task, writers and revisers can adopt different strategies to manage cognitive effort. One of these strategies is related to the interaction with the text to be revised. In their model of revision, Hayes et al. (1987) argue that, when a reviser realises that a text contains too many issues – and/or that they do not know how to solve them (e.g., in the case of novices) – they might find it sometimes simpler to extract the gist of the text and rewrite it from scratch, rather than addressing individual problems (Hayes 2004). Such rewriting might be local (e.g., paraphrasing individual sentences) or global (redrafting entire sections) (Hayes et al. 1987). Therefore, the authors distinguish between, on the one hand, revision in a narrow sense – henceforth 'narrow revision' (see Section 5.2), whereby the reviser tries to address issues in a text while preserving as much of the original structure and content as possible – and, on the other hand, rewriting (Hayes et al. 1987). We regarded the model by Hayes et al. (1987) particularly relevant for this study since it focuses on the revision of already existing texts, which was the task carried out by our participants (Hayes 2004) (Section 3.1). A similar distinction is also reported in Allal/Chanquoy (2004) between 'editing' – as corrections and modifications that do not change the meaning of the text – and 'rewriting', which results in transformations of content, organisation, and meaning.

Rewriting — namely, starting a text from scratch rather than trying to preserve as much as possible of the original text — might originally appear as the most efficient option since it allows the writer to avoid the constraints represented by the characteristics of the original text. However, rewriting might also prove unexpectedly challenging for less fluent and less experienced writers (e.g., low-proficiency second-language speakers), who might experience concern for issues of language or genre conventions with which they are not familiar (Barkaoui 2007). To the best of our knowledge, no previous work in the area of writing research has empirically compared the cognitive effort of rewriting vs revising somebody's else text in a second language. However, comparable research related to (machine) translation shows that translating a text from scratch – similar to the activity of rewriting – tends to require more effort than revising the output produced by a machine translation system when the quality of such output is above a certain threshold (Koponen 2016). Therefore, it can be hypothesised that – when a text is of acceptable quality – revising it would require less effort than rewriting it from scratch.

2.4. Measuring cognitive effort

Researchers have used and combined various methods to measure the effort involved in (re)writing or revising (in a second language), such as questionnaires with self-reporting, think-aloud protocols, dual task, eye tracking, keystroke logging, and retrospective interviews (with stimulated recall) (see e.g., Kellogg et al. 2007; Lu/Révész 2021; Révész et al. 2019; Xu et al. 2021). In this section, we are going to focus on the two methods that we selected for our study, namely keystroke logging and retrospective interviews.

At the high level, keystroke logging has been used to understand the cognitive processes – including mental effort – involved in writing and revision. The analysis of pauses is especially central in most keystroke logging studies (see Hall et al. (2022) for an overview). Pauses (namely interruptions in the flow of typing that can be recorded by keystroke logging tools) have been linked with effortful processes. The analysis of pauses in writing has shown that they occur, for instance, when the author has to plan new content, retrieve existing content from their memory, or revise the text produced so far, among others (Olive et al. 2009). Most keylogging studies have shown that pauses are longer and occur more often at more global text boundaries, involving more high level and complex cognitive processes. Of course, pausing behaviour (pause length and pause distribution) is an indirect indicator of cognitive effort, and we should be aware of the fact that writers might lose focus at a certain moment, feel the need for some 'downtime', or engage in 'sociopsychological' activities (e.g., daydreaming) (Révész et al. 2019). However, there is ample evidence that pause indicators are reliable indicators of cognitive effort.

The need to understand cognitive processes in writing has translated into analyses of pause locations (e.g., within or between words, between sentences, or between paragraphs) and pause length, as well as fluency, text elements revised, reading activity, and use of sources, among others (Chukharev-Hudilainen et al. 2019; Leijten et al. 2019; Révész et al. 2019; Van Waes/Leijten 2015). Moreover, findings from previous research have shown that text production in second language tends to be slower than in first language, involving more pauses and decreased fluency (Chukharev-Hudilainen et al. 2019). Additionally, second-language text production tends to involve a more varied use of sources, which has an impact on the fragmentation of writing/revision processes and can increase cognitive effort (Leijten et al. 2019).

Retrospective interviews have also been used to collect data on multiple aspects related to writing and revision, such as reasons behind revisions (in a second language); use of feedback while revising; and explanations of pausing behaviour (Shen/Chen 2021; Suzuki 2008; Zhao 2010). For instance, the less-skilled writers in the study by Shen/Chen (2021) reported struggling with linguistic encoding (spelling, grammar, or vocabulary).

2.5. Addressing a research gap

As this review of previous literature has shown, adapting texts to the intended audience is an effortful task, particularly for less experienced and second-language writers/revisers. Training and practice can help bridge the expertise gap. However, the effect of training on the revision process carried out in a second language is still an under-researched area. For this reason, this experimental study sought to answer the following research question: **What is the impact of reader-oriented training on the revision process of second-language university students?** We focused in particular on the cognitive effort of the revision process, as indicated by students' pausing behaviour, use of sources, and strategies to approach the revision task (i.e., narrow revision vs. rewriting). Moreover, regarding the strategies, it should be noted that, throughout this paper, we will use the umbrella term 'revision' to include both narrow revision and rewriting.

In addition to filling this research gap, this study contains other elements of novelty. First of all, differently from the majority of previous works, our participants carried out their revisions on

already existing texts produced by others, rather than on their own texts. Secondly, this is the first empirical investigation of the differences in pausing behaviour associated with rewriting versus revising in a second language. Thirdly, to the best of our knowledge, this study is the first to focus on the revision of business texts dealing with CSR. In Section 3 below we report on the methodology adopted to answer our research question and to fill these research gaps.

3. Method

3.1. Set-up: participants, data collection, and tasks

We collected data from a total of 47 students from the Faculty of Business and Economics at the University of Antwerp between October and November 2021. Data collection took place in person in computer laboratories on campus. Upon signing up for the study, participants were randomly divided into an experimental group and a control group. Participants in both groups took part in a pre-test session and a post-test session.

During the pre-test session, participants read an information sheet and signed a consent form. Subsequently, they were asked to read an extract of a corporate report produced by a fictitious tobacco company (SmokIT) and to convert it into an easy-to-read and engaging post that could be published on the company's website. We informed the students that they could be creative and inventive with their revisions, that they could do online searches, and that they could take as much time as needed. The revision task took place in Microsoft Word while the keystroke logging tool Inputlog was running in the background. Therefore, we could record the students' typing, pauses, and use of online sources (Leijten/Van Waes 2013). Furthermore, Inputlog saved intermediate versions of the participants' drafts every 2 minutes. Following the revision task, participants were asked to access an online module. The experimental group was assigned a module on accessible communication applied to CSR content, while the control group was assigned a module on the topic of CSR (Section 3.2). Students were instructed to spend at least 45 minutes familiarising themselves with the theory and the exercises in the modules. Finally, at the end of the pre-test session, all participants completed a short demographic questionnaire (Section 3.4).

The post-test sessions took place 2-3 days after the pre-test sessions. As part of the post-test, the students conducted the revision task in the case study of their respective modules. Similar to the pre-test, this revision task involved converting another extract of a corporate report into an accessible and engaging post that could be published on the website of the same tobacco company (see Section 3.3 for details about the texts to be revised). For this task as well, students could be creative and inventive, do online searches, and take as much time as needed. However, differently from the pre-test, in the post-test the participants could also consult their respective modules and use them to guide the revision task. This task also took place in Microsoft Word while the keystroke logging tool Inputlog was collecting data in the background. Finally, all participants completed a fidelity questionnaire with multiple-choice questions about the content of their respective modules. Even though correct answers to multiple-choice questions might be guessed, we treated this fidelity questionnaire as a 'quick and dirty' way of checking that the participants had actually interacted with content from the modules. Since the experimental group and the control group had been assigned different modules, the multiple-choice questions were of course different. Results of the fidelity questionnaire are reported in Section 3.4.

In the post-test session, we also randomly selected a sub-group of participants (3 from the experimental group and 4 from the control group) and we collected screen recorded and interview data from them so as to allow for a more qualitative analysis. The screen-recorded data will not be analysed in this paper. With regard to the interviews, they involved short follow-up questions at the end of the revision task. These retrospective interviews focused on the process/product of the revision task and the potential benefits of taking part in our training. We used semi-structured

interviews, which allowed us to discuss the topics of interest but also to adjust the questions to the natural flow of the conversation and to the emergence of additional topics of interest (Matthews/Ross 2010).

3.2. Modules

In order to empirically test the impact of training on accessible/reader-oriented communication, the experimental group received a link to an online module where the principles of accessible communication were applied to CSR content. In contrast, the control group received a link to an online module dealing exclusively with CSR¹. Both modules were hosted by the online writing center Calliope, where content was divided into introduction, theory, exercises, and case study, and students could move freely between the different components depending on their learning styles (Van Waes et al. 2014). Furthermore, in both modules, theoretical content was provided both in textual and in audio-visual format. The revision task carried out by the students in the post-test session was located on the case study page of the respective modules.

3.2.1 Module for experimental group

The theoretical content in the experimental module was divided into three main sections: (i) accessible communication; (ii) CSR; and (iii) the revision process. The section on accessible communication focused on vocabulary, sentence length/structure, cohesion, visual aspects, and relevance. For instance, the students were shown how to simplify sentences, how to use connectives to build cohesion, or how to replace complex vocabulary with simpler synonyms. The section on CSR focused on its definition and on the differences between two channels of communication (i.e., corporate reports and corporate websites), in line with the goal of the revision task (Section 3.1). The section on the revision process relied on Inputlog process graphs and videos in order to show students how expert and novice revisers differ in terms of their revision processes (Leijten/Van Waes 2013). We included this section in order to foster observational learning, whereby students learn by observing the actions and hearing the explanations of experts (Rijlaarsdam et al. 2008). For a more detailed description of the module given to the experimental group, see Rossetti/Van Waes (2022).

3.2.2 Module for control group

The theoretical content in the control module was divided into three main sections, namely definition, evolution, and communication of CSR. The section on the communication of CSR contained links to the websites of companies well known for their CSR activities and for their clear communications. Students in the control group could of course consult these pages but they received no formal instruction on how to write accessible texts.

In the interest of clarity, Figure 1 presents a visual overview of the procedure described in the previous sections.

¹ The experimental module is available at the following link: <https://hosting.uantwerpen.be/calliope/plantra/>. The control module is available here: <https://hosting.uantwerpen.be/calliope/plantra-control-training/>.

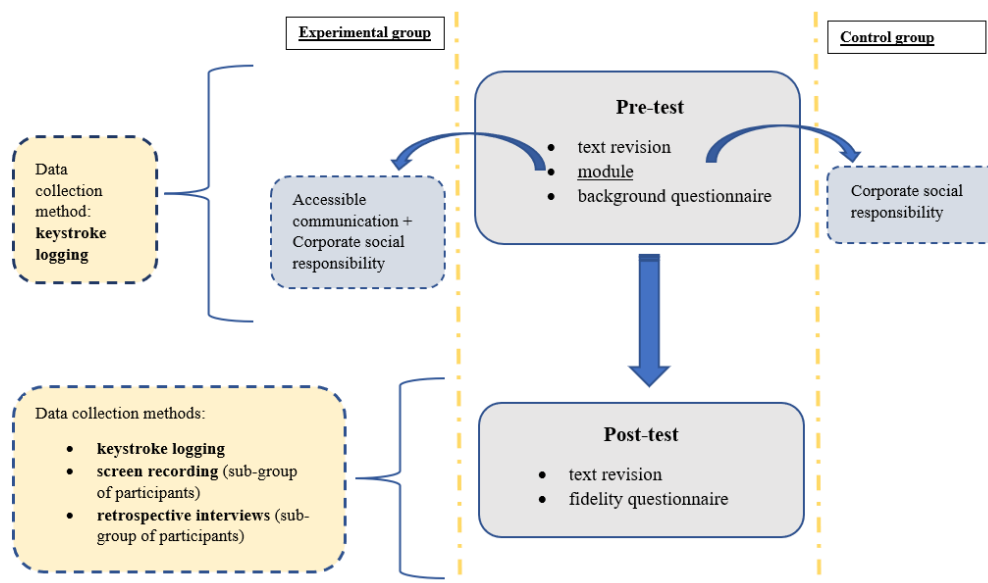


Figure 1. Overview of data collection procedure

3.3. Materials (texts)

In order to avoid a learning effect (whereby students would learn from the pre-test task rather than from our module/intervention), the texts that we selected for revision in the pre-test and in the post-test differed in terms of topic. However, they both dealt with the socially responsible activities carried out by a fictitious tobacco company called SmokIT. The pre-test text dealt with the company's efforts to reduce their carbon emissions, while the post-test text revolved around the company's investment in smoke-free alternatives. Both texts were extracts of a real corporate report, but we manipulated some of their features to make sure that the two revision tasks were similar and that the texts contained some of the problems mentioned in the Calliope module assigned to the experimental group (Section 3.2). The decision to assign students extracts of corporate reports to be revised into website posts (Section 3.1) was motivated by previous research showing that, while CSR content in corporate reports tends to be too complex and too technical for lay customers, websites offer an opportunity for companies to present CSR information in accessible language and engaging format (Smeuninx et al. 2020; Wei 2020).

The length and the original level of readability of the selected texts might have acted as confounding variables. Therefore, we ensured that the two texts were as comparable as possible. Using the Coh-Metrix Common Core Text Ease and Readability Assessor (T.E.R.A.), we observed that both texts had a low level of narrativity (which means that they contained a low proportion of common words and verbs, and a high proportion of complex noun phrases), with the pre-test text having a particularly low percentage (Jackson et al. 2016). Both texts were also similar in their levels of syntactic simplicity (low, especially in the post-test) and referential cohesion (average, but lower in the pre-test text). Referential cohesion refers to the repetition of words, phrases, and concepts across sentences (Jackson et al. 2016). Deep cohesion – determined by the amount of connections among the events and ideas in the text (Jackson et al. 2016) – was higher in the pre-test text, but the narrativity and referential cohesion were lower in this text so we assumed that these readability features would compensate each other. Regarding text length and structure, the two texts had a very similar number of words (274 vs. 278 words) and they were both divided into three paragraphs. Furthermore, they both lacked (sub-)headings and visual devices to support the processing of content.

3.4. Demographic profiles and interaction with modules

As explained in Section 3.1, students filled out a demographic questionnaire after the pre-test session, and a fidelity questionnaire after the post-test session. The goal of the demographic questionnaire was to ensure that participants in the experimental and in the control group had similar background characteristics and prior knowledge. Furthermore, participants self-reported how much they had learnt from the respective modules. The goal of the fidelity questionnaire, on the other hand, was to more objectively assess participants' levels of interaction with the modules by asking them 4-5 multiple-choice questions about the theoretical contents. We assigned a score of 1 to correct answers and a score of 0 to wrong answers. 'I don't know' answers were treated as missing values. We calculated the mean score for each question, and then the grand mean of all questions. In the interests of clarity, we report these data as percentage values. Table 1 contains an overview of demographic and fidelity data. The control group was not asked the questions about pre-module and post-module knowledge about accessible communication. Therefore, we marked those cells as 'N/A' in Table 1.

It can be observed that participants in both groups were similar in terms of age, gender distribution, linguistic profile (most of them reported Dutch as native language), average number of years spent studying English (the language of the texts to be revised), and current programme of study (specifically, most of the participants were starting a Master in Multilingual Professional Communication). Furthermore, the majority of them had not previously attended training on accessible communication. However, in the experimental group, most participants were already somewhat familiar with the principles of accessible communication even before looking at the assigned Calliope module. On average, they gave quite a high score about their ability to write accessible/easy-to-read texts following participation in our module. With regard to CSR, in both the experimental and the control group, there was some increase in their self-reported knowledge about the topic after looking at the modules, although with quite some within-group variability. Finally, the grand mean scores from the fidelity questionnaire seem to indicate that participants in both groups engaged thoroughly with the theoretical contents provided in the modules.

	Experimental (N=24)	Control (N=23)
<i>Age</i>	Mean=23 Min=19 Max=38	Mean=23 Min=22 Max=34
<i>Gender</i>	Female=18 Male=6	Female=20 Male=3
<i>Dutch as native language</i>	N=24	N=21
<i>Additional native language(s)</i>	Tamazight, Armenian, Polish	Frisian, Armenian, French, Finnish, Spanish, Italian
<i>Years studying English</i>	Mean=9 Min=2 Max=26	Mean=8 Min=4 Max=15
<i>Study programme</i>	Master=23 Bachelor=1	Master=23
<i>Previously attended training on accessible communication</i>	Yes=6 No=18	Yes=8 No=15
<i>Were you familiar with the principles of accessible communication before looking at the module?</i>	Yes=1 To a certain extent=17 No=6	N/A
<i>After looking at the module, how much do you think you know about writing accessible/easy-to-read texts, on a scale from 1 (I don't know anything at all) to 10 (I have become an expert on the topic)?</i>	Mean=7 Min=5 Max=8	N/A
<i>Before looking at the module, how much did you know about corporate social responsibility on a scale from 1 (I knew nothing at all) to 10 (I was an expert on the topic)?</i>	Mean=4 Min=1 Max=8	Mean=4 Min=1 Max=8
<i>After looking at the module, how much do you think you know about corporate social responsibility on a scale from 1 (I still don't know anything at all) to 10 (I have become an expert on the topic)?</i>	Mean=6 Min=3 Max=9	Mean=7 Min=4 Max=9
<i>Correct answers about module content</i>	Grand mean=93% Min=83% Max=100%	Grand mean=87% Min=57% Max=100%

Table 1. Participants' profiles and interaction with Calliope modules

The results of the fidelity questionnaires were also supported by the keystroke logging data provided by Inputlog. Specifically, the tool recorded the time spent by the participants on their respective modules during the pre-test session, after the revision task (Section 3.1). We observed that the experimental group spent on average 46 minutes on the module, while the control group spent on average 34 minutes, possibly because of the more limited theoretical content (the control group did not have access to theory on accessible communication).

4. Analysis

4.1. Keystroke logging data

The analysis of keystroke logging data was first carried out in Inputlog and then transferred to SPSS. It revolved around pausing behaviour and source use. Furthermore, we classified participants' overall revision strategies as either narrow revision or rewriting – following the classification reported in Hayes et al. (1987) – and analysed the extent to which these different strategies were represented by keystroke logging data.

In order to examine the effect of our reader-oriented training, for each variable of interest, we compared the pre-test with the post-test data of all participants (within-subjects), as well as the experimental group with the control group at the pre-test and post-test stage (between-subjects). A series of Shapiro-Wilk tests showed that the data were not normally distributed for all the (sub-)variables of interest. Therefore, in the interests of consistency, we ran non-parametric tests (Mann-Whitney U test for the between-group analyses and Wilcoxon signed rank test for the within-group analyses).

4.1.1. Data preparation

Before starting data analysis, we checked and time-filtered the raw IDFX files produced by Inputlog, as described in Leijten/Van Waes (2020). This filtering ensured that we removed noise and irrelevant data (such as the time spent between our start of Inputlog and the start of the revision task by the participants). Firstly, we used Inputlog to produce a general analysis of all the raw IDFX files. In the general analysis, every row corresponds to a log event (keyboard or mouse action), for which cursor position, document length, and pauses are reported, along with the environment (focus) in which that event takes place (e.g., main Microsoft Word document, URLs of online pages, Calliope modules, and so on) (Leijten/Van Waes 2013). Based on the information reported in the general analysis files, we filtered the IDFX files from the focus in the main Microsoft Word document immediately preceding the first action (e.g., pressing RETURN or typing a letter) to the last action in the same document, where the revision task took place. Therefore, the filtered IDFX files contained keystroke logging data exclusively on the revision task.

Subsequently, we produced general analyses and process graphs from the filtered IDFXs. Process graphs are visual representations of the writing processes that include information on cursor position, text length at any given point, number of characters typed/edited/deleted, pauses (location and length), and source use (Vandermeulen et al. 2020). In examining all the process graphs and the general analysis files, we noticed that – for two participants in the pre-test and for eight participants in the post-test² – there were time ranges when Inputlog incorrectly recorded actions as being carried out outside of the Microsoft Word document. In other words, events (typing, revisions, etc.) that were part of the revision task were not recorded as such. However, as the data loss represented only a small percentage of the overall task duration, we also included these IDFX files in the analysis.

Finally, we excluded 4 participants from the pre-test and 2 participants from the post-test because their IDFX files were irreparably damaged, and no data could be retrieved. This exclusion influenced the number of participants that could be included in the within-subjects analyses.

4.1.2. Pause analysis

As parameters in the pause analysis, we used a pause threshold of 200 milliseconds and a P-Burst

² For two of these participants, we were able to manually correct the IDFX files since the missing time ranges involved linear production of content. Specifically, the correction involved counting the actions within the Microsoft Word document and adjusting document length, character production, and cursor position accordingly.

threshold of 2000 milliseconds. Pauses that are longer than 200 milliseconds are of interest because, for this cohort of participants, they are mainly linked with high-level, effortful cognitive processes (e.g., content planning) that exceed the low-level transitions between keys (Van Waes et al. 2009). P-Bursts are fluent writing episodes that are not interrupted by a pause (Leijten et al. 2019). In addition to general measures (e.g., overall task duration, total number of pauses, or total pause time), we examined and compared pausing behaviour at different levels (i.e., within words, between words, between sentences, and before revisions). Furthermore, to account for individual differences in task duration, we included proportional values out of overall process time.

4.1.3. Analysis of revision strategies

Informed by the model discussed in Hayes et al. (1987), we categorised each participant's strategy of approach to the task either as narrow revision or rewriting (despite the fact that both strategies were part of a larger revision task) (Section 3.1). We used the start of the task as the determining criterion in our classification. Concretely, if the first edit (addition, deletion, or substitution) that a participant made was located within the assigned text, we treated the overall strategy as 'narrow revision'. In contrast, if the first edit made was located outside the assigned text (e.g., under it), we treated the overall strategy as 'rewriting'.

When making this categorisation, we excluded outlines, instructions, and titles/headings. For instance, if a participant wrote a title on top of the assigned text but then continued to make the revisions within the text (rather than starting a new one), that strategy was treated as narrow revision. This categorisation was possible because we set Inputlog to save versions of the revised documents every 2 minutes. Therefore, in addition to the final texts produced by the participants, we had also access to their intermediate drafts. This way of categorising the revision strategies is certainly not perfect since participants might have switched from a rewriting to a narrow revision strategy (or vice versa) throughout the task. However, this categorisation allowed us to focus on the way in which these second-language students first interacted with the text to be revised using either of these strategies. Furthermore, as we will show in Section 5.2, this manual classification is partly supported by patterns in the keystroke logging data.

First, we examined if and how keystroke logging data related to pauses reflected the two different strategies of rewriting and narrow revision both in the pre-test and in the post-test. To this end, we merged the data from the control and from the experimental group. Subsequently, we zoomed in on the participants who changed their revision strategies moving from pre-test to post-test session.

4.1.4. Analysis of source use

We also examined how our participants interacted with the main Word document and with external sources (e.g., module, online pages, etc.) during the revision task. The keystroke logging tool Inputlog allowed us to collect data on time spent and on number of keystrokes made in each environment (Leijten/Van Waes 2013). In order to account for the differences in task duration, we report the time and keystrokes per source as a proportion of the overall process times. Prior to data analysis, we recoded the IDFX files that we had filtered (Section 4.1.1) by grouping and categorising the different sources used by the participants. Specifically, we created 8 categories: (i) main document; (ii) language searches; (iii) topic searches; (iv) search start; (v) modules; (vi) session instructions; (vii) calculations; and (viii) irrelevant. A description of the categories is available in Table 2.

Sources	Descriptions
<i>Main document</i>	Word document in which students carried out the revision task
<i>Language searches</i>	Internet searches aimed at finding the meaning, synonyms, and/or translation of words
<i>Topic searches</i>	Online searches whose goal was to find information on the CSR topics addressed in the texts (e.g., climate change or smoke-free products)
<i>Search starts</i>	Interruption of the writing flow: typing a search in an Internet browser and scrolling the results without opening a specific page
<i>Modules</i>	Access to the modules assigned to the participants Note: this category includes different types of searches (language- and topic-related), but we did not assign these searches a specific goal because that was not always clear from the words typed
<i>Session instructions</i>	Checking of task instructions provided in digital format
<i>Calculations</i>	Internet searches aimed at calculating values and percentages related to the text
<i>Irrelevant</i>	Use of online pages not related to the task at hand

Table 2. Categorisation of sources

Such categorisation was needed considering the high variability in students' use of sources. To give an example, Figure 2 below contains some of the language-related online searches made by the participants:

```

<entry_5>alleviate - Vertaling Engels-Nederlands - Google Chrome</entry_5>
<entry_6>an alternative - Vertaling naar Nederlands - voorbeelden Engels | Reverso Context - Google Chrome</entry_6>
<entry_7>an alternative to - Vertaling naar Nederlands - voorbeelden Engels | Reverso Context - Google Chrome</entry_7>
<entry_8>BELANG HECHTEN AAN - Engelse vertaling - bab.la Nederlands-Engels woordenboek - Google Chrome</entry_8>
<entry_9>belangrijk daarbij is - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_9>
<entry_10>betaalbaar - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_10>
<entry_11>bijvoorbeeld - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_11>
<entry_12>brand onder de voeten - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_12>
<entry_13>brandt onder de voeten - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_13>
<entry_14>BREDER PUBLIEK - Engelse vertaling - bab.la Nederlands-Engels woordenboek - Google Chrome</entry_14>
<entry_15>broeikasgassen - Engelse vertaling - Linguee woordenboek - Google Chrome</entry_15>
<entry_16>daarnaast - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_16>
<entry_17>de enige kans op - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_17>
<entry_18>de reden hiervoor - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_18>
<entry_19>de reden waarom - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_19>
<entry_20>doen dalen - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_20>
<entry_21>echter - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_21>
<entry_22>een breder publiek - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_22>
<entry_23>een heel karwei - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_23>
<entry_24>een nuttige bijdrage leveren - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_24>
<entry_25>een voorbeeld - Vertaling naar Engels - voorbeelden Nederlands | Reverso Context - Google Chrome</entry_25>

```

Figure 2: Examples of language-related searches

4.2. Retrospective interview data

In the post-test session, for a randomly selected sub-group of participants (3 from the experimental group and 4 from the control group), we collected qualitative data in the form of short semi-structured retrospective interviews. The goal of these interviews was to complement the more quantitative keystroke logging data. Specifically, the questions revolved around the general approach to the task, perceived difficulty, planning and revision strategies, level of satisfaction with the final text, and helpfulness of the Calliope module.

After transcribing the audio recordings, the first author read and coded participants' answers in NVivo using an inductive and iterative approach. The goal of this coding was to identify important units of meaning in the data. After organising the codes into main themes and sub-themes, the first author and the second author discussed the (sub-)themes and reached agreement on their naming and categorisation.

We identified six themes, namely: (i) revision strategies; (ii) interaction with module; (iii) overall approach awareness; (iv) previous knowledge and experience; (v) reader orientedness; and (vi) source use. The theme 'revision strategies' was in turn divided into seven sub-themes, each referring to a different type of issue that the participants addressed in the text when revising it. Most of these sub-themes are in line with the contents of the Calliope module assigned to the experimental group (Section 3.2). Specifically, they refer to vocabulary, syntax, text structure, visual aspects, relevance, tone, and contextual information. The theme 'overall approach awareness' was also divided into two sub-themes, respectively referring to planning how to address the text and to the revision of content.

5. Results

In this section, we report the results of our quantitative (Sections 5.1-5.3) and qualitative (Section 5.4) analyses. We use the code 'PE' to refer to participants from the experimental group, and the code 'PC' to refer to participants from the control group.

5.1. Pauses

Table 3 reports the pause-related results of the between-group analyses comparing the experimental and the control group during the pre-test and during the post-test. The table contains only a subset of all the pause-related sub-variables that we analysed. Specifically, we report the data related to the entire process (e.g., total active writing time) and to the word-, sentence-, and revision-level. The full list of pause-related sub-variables is available in Appendix A.

Overall, the pausing behaviour of the experimental group and of the control group was quite similar both in the pre-test and in the post-test, for instance in terms of average duration of pauses, proportion of pauses out of overall process time, number of P-Bursts per minute, and their average duration. In other words, taking part in training did not lead to differences in pausing behaviour between the groups. In the pre-test, the control group showed a significantly higher proportion of between-sentence pauses (compared with the experimental group), which might be indicative of effortful planned sentence production and idea generation (Baaijen et al. 2012; Leijten et al. 2019). This difference became non-significant in the post-test, possibly due to acquired familiarity with the revision task and/or with the topic of smoke-free products (Section 3.3).

When focusing on the within-group comparisons, the control group showed a significantly higher proportion of between-word pauses in the post-test (compared with the pre-test) ($Z = -2.455$, $p = .01$). This result might be due to the higher cognitive effort needed by this group when making lexical decisions (Leijten et al. 2019). As the experimental group did not show a higher proportion of between-word pauses in the post-test, we can hypothesize that our training in plain language assisted them with lexical choices, for example in finding lay synonyms of specialised terms. This hypothesis is supported by the keystroke logging data linked with rewriting vs revision strategies (Section 5.2) and by the interview data (Section 5.4).

	Pre-test					Post-test				
	Experimental (N=24)		Control (N=19)		<i>p</i> -value	Experimental (N=23)		Control (N=22)		<i>p</i> -value
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Overall process										
<i>Total process time (min:secs)</i>	45:42	20:42	45:31	15:52	.769	49:55	15:00	48:01	15:42	.803
<i>Total active writing time (min:secs)</i>	14:01	06:28	16:47	09:00	.293	16:18	06:39	17:04	06:57	.666
Pausing behaviour										
<i>Proportion of pause time (%)</i>	68.68%	8.02%	64.57%	10.4%	.221	67.4%	8.43%	64.49%	6.56%	.107
<i>Number of P-Bursts per minute</i>	3.49	.59	3.46	.68	.912	3.52	.46	3.36	.59	.184
<i>Mean process time of P-Bursts (secs)</i>	10.9	3.2	12.7	5.3	.359	11.2	2.9	12.4	3.5	.256
Pause locations										
<i>Number of within-word pauses per second</i>	.17	.06	.14	.04	.056	.17	.06	.14	.04	.146
<i>Geometric mean of within-word pauses (secs)</i>	.32	.02	.31	.01	.625	.32	.02	.31	.02	.919
<i>Number of between-word pauses per second</i>	.10	.04	.09	.04	.42	.10	.03	.10	.03	.602
<i>Geometric mean of between-word pauses (secs)</i>	.57	.13	.54	.12	.525	.55	.14	.51	.10	.601
<i>Number of between-sentence pauses per second</i>	.002*	.002	.003*	.001	< .05	.003	.002	.004	.002	.102
<i>Geometric mean of between-sentence pauses (secs)</i>	1.06	.97	.82	.32	.778	1.04	1.19	.69	.26	.503
<i>Number of revision pauses per second</i>	.06	.02	.06	.02	.903	.06	.02	.06	.03	.856
<i>Geometric mean of revision pauses (secs)</i>	.57	.09	.54	.06	.399	.55	.05	.53	.07	.51

Table 3. Between-group analysis of pausing behaviour in the pre- and post-test

5.2. Revision strategies: narrow revision vs. rewriting

The most common strategy involved rewriting (more specifically, redrafting) (Hayes et al. 1987), whereby the participants abandoned the assigned text and started a new text from scratch, either above, below, or in the middle of the assigned text (e.g., when they rewrote one paragraph at a time). Overall, the rewriting strategy was used by 38 participants in the pre-test (90% of the total) and by 31 in the post-test (73% of the total). However, in some (less frequent) cases, the participants decided to revise the given text. This happened in total 4 times during the pre-test and 11 times during the post-test. It is also interesting to note that 6 participants in the control group and 7 participants in the experimental group switched between strategies depending on the session (pre-test vs post-test). Table 4 provides an overview of the strategies and their evolution.

	Pre-test		Post-test		Strategy change	
	<i>Rewriting</i>	<i>Revision</i>	<i>Rewriting</i>	<i>Revision</i>	<i>Rewriting to revision</i>	<i>Revision to rewriting</i>
Group						
Control (N=19)	16/19	3/19	16/19	3/19	3/19	3/19
Experimental (N=23)	22/23	1/23	15/23	8/23	7/23	0/23

Table 4. Participants' revision strategies

Table 5 shows that, in the pre-test, rewriting involved a significantly higher proportion of pause time and significantly higher proportions of pauses at multiple text levels (within words, between words, and between sentences). These results seem to indicate that rewriting required more cognitive effort than revision as the participants (who were not native speakers of English) had to pay attention to spelling and grammar, come up with their own lexical choices, and generate ideas for the new text (Leijten et al. 2019). In the post-test, the differences between rewriting and revision became less pronounced, possibly because the students had familiarised themselves with the task of rewriting a text on CSR in their second language, and/or could draw on a more concrete set of criteria following the training. However, the participants who chose rewriting still needed, on average, significantly more time than the participants who chose the narrow revision strategy, especially because of higher pausing times. It is also interesting to note that, despite the low number of participants – especially in the narrow revision group – these keystroke logging data show that using different strategies leads to changes in pausing behaviour.

	Pre-test					Post-test				
	Rewriting (N=38)		Narrow revision (N=4)		<i>p-value</i>	Rewriting (N=31)		Narrow revision (N=11)		<i>p-value</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Overall process										
<i>Total process time (min:secs)</i>	45:59	19:18	41:09	12:51	.668	51:39	14:27	42:12	16:12	< .05
Pausing behaviour										
<i>Total pause time (min:secs)</i>	31:10	13:15	20:57	05:32	.134	34:15*	09:56	26:46*	09:50	< .05
<i>Proportion of pause time (%)</i>	68.15%*	7.8%	52.46%*	10.73%	< .05	66.45%	7.5%	63.82%	8.3%	.466
<i>Total number of pauses per second</i>	.57*	.09	.45*	.01	< .05	.56	.09	.58	.08	.637
Pause locations										
<i>Number of within-word pauses per second</i>	.17*	.06	.09*	.01	< .05	.16	.04	.15	.07	.597
<i>Number of between-word pauses per second</i>	.10*	.04	.04*	.01	< .05	.10	.03	.10	.03	.539
<i>Number of between-sentence pauses per second</i>	.003*	.002	.001*	.001	< .05	.004	.002	.003	.001	.391

Table 5. Differences in pausing behaviour between revision strategies

With regard to the participants who showed a change in their revision strategy, some interesting patterns can be observed (Table 4). In the experimental group, the seven participants who changed

revision strategies all switched from rewriting to narrow revision. In contrast, the control group showed more variability (3 people switched from narrow revision to rewriting, and 3 people from rewriting to narrow revision). In the experimental group, the change from rewriting (in the pre-test) to revision (in the post-test) resulted in significantly shorter between-word pauses (indicating lower cognitive effort for lexical choices), as shown by a Wilcoxon signed-rank test ($Z = -2.366$, $p = .018$). In the control group, such difference was not observed.

5.3. Source use

In Table 6, it can be observed that the source-related behaviour of experimental and control group was quite similar across pre- and post-sessions. Overall, participants devoted most of their time and their typing to working on the documents where the revision tasks took place, as would be expected, also based on previous research (Leijten et al. 2019). For instance, between 90% and 80% of all keystrokes – and about 80% of the overall task duration – were devoted to the main document during the pre-test session. These percentages decreased slightly in the post-test, but they continued to represent the majority.

Across sessions, the control group dedicated consistently more time and more typing to language searches, and especially during the post-test session. When considering that the control group showed a significantly higher average number and proportion of between-word pauses in the post-test (Section 5.1), this result related to language searches seems again to point to the difficulty experienced by this group in finding appropriate lexical choices. Previous research on source-based writing in English as a second language has also highlighted students' use of online resources such as dictionaries and grammars (Leijten et al. 2019). Regarding topic searches, the relative time devoted to them increased for all participants in the post-test. Based on the results of the retrospective interviews (Section 5.4), these topic searches were mainly aimed at the addition of contextual information on the topics of tobacco and smoke-free products.

	Pre-test					Post-test				
	Experimental (N=24)		Control (N=19)		<i>p-value</i>	Experimental (N=23)		Control (N=22)		<i>p-value</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Relative time										
<i>Main document</i>	.84	.14	.80	.15	.271	.78	.10	.69	.15	.093
<i>Search start</i>	.05	.05	.06	.05	.769	.03	.03	.03	.03	.937
<i>Topic searches</i>	.02	.03	.01	.04	.801	.05	.06	.07	.05	.078
<i>Language searches</i>	.05	.09	.09	.10	.076	.03	.04	.11	.15	.078
Relative keystrokes										
<i>Main document</i>	.91	.10	.84	.15	.112	.85	.13	.77	.21	.238
<i>Search start</i>	.04	.06	.05	.05	.378	.02	.04	.02	.03	.624
<i>Topic searches</i>	.004	.01	.001	.004	.414	.002	.005	.003	.009	.818
<i>Language searches</i>	.03	.07	.09	.13	.051	.03	.06	.14	.21	.08

Table 6. Between-group source analysis

We also examined differences within each group. A Wilcoxon signed rank test showed that, compared with the pre-test, in the post-test the experimental group used less time ($Z = -2.068$, $p = .039$) and fewer keystrokes ($Z = -2.305$, $p = .021$) for starting online searches. Similarly, in the post-test, the control group spent less time starting online searches ($Z = -2.052$, $p = .040$), but more time for topic searches ($Z = -3.385$, $p < .001$), compared with the pre-test.

With regard to the time spent interacting with the Calliope modules while revising during the

post-test session, participants in both groups spent on average between 3 and 4 minutes consulting their respective modules during the revision task. This might seem like a short time, but it should be remembered that participants had already familiarised themselves with the modules during the pre-test session, after completing the revision task (see Section 3.1). Specifically, the experimental group had spent on average 46 minutes on the module, while the control group had spent on average 34 minutes.

5.4. Retrospective interviews (post-test)

With regard to revision strategies, participants in both the experimental and the control group reportedly focused to a large extent on the same issues when revising the text. Mentions of vocabulary, syntax, visual aspects, tone, structure, relevance, and contextual information were present in interviews with both groups. For instance, with regard to vocabulary:

PC19: [D]ifficult words... I tried to look them up and explain them.

PE10: I would just remove all the jargon and the specific terms used in a report.

As far as visual aspects are concerned:

PC22: I used some bullet points as well.

PE10: And what I also did while revisioning my text is use more words that are highlighted so that it's more scannable.

In terms of tone, participants tried to make the text more personal, for example by using exclamation marks or replacing the tobacco company's name with the pronoun 'We'. Syntax-oriented edits involved shortening the sentences. Relevance-oriented comments were quite frequent, but some participants pointed out the difficulty of deciding which information to leave in and which information to leave out when revising:

PC13: Like, sometimes I was thinking, like 'Do I have to say this? Is this important? Is this too...?' And that was a little bit tricky for me.

Linked with relevance comments, one participant in the control group and one participant in the experimental group also pointed out the need to add some contextual information, specifically on the available portfolio of smoke-free products. The addition of such contextual information often involved doing online searches on the topic of tobacco and smoke-free products:

PE10: When I looked at the case, I saw a link to more information about smoke-free cigarettes, and when I was reading that article, I thought 'Oh, maybe it would be interesting to focus on what the product is, why you will switch to it and then what we do already for those smoke free cigarettes'.

Structure is the textual aspect that received the most attention from participants, who used titles and (sub-)headings to give structure to the text and to merge similar content together. It should be remembered that our texts lacked these elements (Section 3.3):

PE23: So this text, I think, some parts were kind of similar. Well, the content was the same. So I put them together and I left out what was extra. [...] When I saw what the paragraph was about, the main idea, I tried to put that into a title and then work on it further because the titles were missing.

The motivation behind the various changes made to the text seems to be reader-orientedness for both the control and the experimental group:

PC07: The difficulty may be some of the more technical aspects, how to properly

communicate them to an audience that may not be fully familiar with some of the concepts.

PC19: [I]t's easier for the clients to understand them with an explanation, or an example, or something like that.

PC22: I tried to bring as much structure to the text as I could because I think it helps the readers on a website a lot when they're reading a text.

PE10: I just focused on what the products are about and the important information that a customer needs to know to switch, for example, to smoke-free products.

PE23: And then I was searching something about exclamation marks. I was thinking, if that was... If we could add that. But I did because I thought it was a text to the customers and it was on the website. So I thought, well, that it was OK to do that, that it wasn't too, well, business-like.

The fact that participants in the control group showed reader orientedness might be due to their prior knowledge of the topic and experience with plain language, also mentioned during the interviews:

PC07: [N]ot to blow my own horn too much, but I do believe that I have a firm grasp of how to write a text that will convince people to at least look at the product that is being sold.

PC19: And second [text] it was easier because I think I know some more about this topic than the other topic.

When prompted by the questions, most participants were able to reflect upon their overall approach to the task, e.g., whether they followed a specific order, why they rewrote a new text from scratch rather than revising an existing one (Section 5.2), whether they already had a structure in mind, and so on. Overall, participants' processes showed features of expertise, for instance, in the use of multiple revision rounds or in the whole-text planning approach:

PC19: Sometimes I just write the structure that I want to have in my text. So, for instance, if I want to have a structure that is problem-solution based, I just write these small questions beforehand and then I answer them along the way.

PE18: [B]efore I did anything, I looked at the whole text. And then I started revising and I went... I did it a second time. Like, I wrote a paragraph, revised it, wrote a second, then revised both, wrote three and then revised the whole bit to see if it looked like a whole.

Interestingly, the choice to rewrite a new text from scratch rather than revising the given one seems a way to manage the cognitive difficulty of the task, as highlighted in our review of earlier work (Section 2.3):

PC13: [W]hat's very handy for me, was to not adapt the text in the text while I was writing, but just start a new text, and then read. So I was watching and then writing. Sometimes I wrote the same and then I deleted it and then... So I was just typing a new text under the actual text. [...] I didn't do that the first time. So I was just like writing inside the text and it was like a soup. I didn't remember what I wrote, what was already in there. And it was, like, very... No, so, I said, now I'm going to, like, divide it up. And for every line I'm going to write, start writing, from zero. [...] So then I had like a better overview of what I already said. What is essential. What is not essential. And then I just deleted the whole part.

One striking difference that emerged between the control and the experimental group refers to the sub-theme about planning how to approach the task. Specifically, participants in the control group talked about their difficulty with planning. In contrast, participants in the experimental group seemed to benefit from the module as it provided them with stepwise guidance on how to approach

the text:

PC19: So I'm actually a bit relieved now, to be honest, because I thought I went through the entire Calliope and I thought to myself 'Where is the part where we learn how to write?' [...] [W]hen I was writing, it was easy going and I found some solutions here and there, but starting the writing process was rather difficult.

PE18: This time I had more of a plan in my head because the first time when I revised it, I just started. This time I had a clear image what I was going to look at. Like vocabulary, sentence structures... So I kind of had a building plan in my head of what I wanted it to look like before I started writing.

Experimental participants also expanded on how interacting with the module helped them manage the cognitive demands of the revision task:

PE10: But the module was very clear and gave me a good idea of what I really need to focus on. It was very clear, it was bulleted, so I knew that I need to focus on that and that I need to focus on that.

PE18: Because it [the module] provided a step-by-step, like... Because it gave a clear definition of what is a good text, what makes a text simple, what makes it easier, easy to read and accessible. And because before I started revising... So, on Monday, [day of the pre-test session] I didn't really have a clear image in my mind, in my head, what a good text was, and now I had like a plan with steps and I follow them. And that made it easier for me.

As the participants in the control group did not receive formal instruction on how to write accessible texts (Section 3.2), it is not surprising that, when interacting with their module, they found the concrete examples of accessible web communications particularly helpful:

PC19: [T]here was a lot of theory and it was interesting, but a lot of it I already knew, but I never saw a concrete example, so that was really interesting.

PC22: It helped me that there were some examples of websites that already had a page on CSR. I did use them. I did go back and forth and look at them. And so that helped me a lot.

However, exposure to these examples did not seem to translate into the same stepwise guidance that the experimental module provided.

6. Discussion and conclusions

We carried out a study to test the impact of reader-oriented training on the revision process. Our reader-oriented training – the intervention of this study – revolved around the accessible communication of CSR content. Using keystroke logging and retrospective interviews in a pre-test post-test design, we collected data on the cognitive effort (pauses, strategies, and source use) of 47 Master students (with English as a second language) revising extracts of corporate reports in order to make them accessible enough for a corporate website. Students were randomly assigned to a control and to an experimental group, with the latter receiving our reader-oriented training.

On average, students had little prior knowledge about CSR (i.e., the topic of the texts). Most of them had not previously attended training on accessible communication. However, based on their academic studies in the domain of professional communication, our participants were between the knowledge-transforming and the knowledge-crafting stage of writing development, characterised by the ability to consider and balance the author's intentions, the features of the text at any given moment, and the reader's interpretation (Kellogg 2008). In other words, our students were already showing some traits of writing and revision expertise. For instance, as highlighted by the interview data, they made different rounds of revisions and they concerned themselves with a number of

issues during their revisions, such as vocabulary, syntax, visual aspects, tone, structure, relevance, and contextual information (Becker 2006). Participants went so far as to remove and add content by relying on online searches about the topics of the texts. Such ability to consider global text features is also indicative of their high level of English proficiency. In other words, the students seemed to have automatised procedures to deal with local issues (e.g., spelling and grammar), which left them with enough working memory resources to consider the text as a whole and its intended reader (Barkaoui 2016; McCutchen 2011).

Our training did not have an impact on the general pausing behaviour – and, by extension, on the cognitive effort – linked with the revision task. However, it seemed to reduce the cognitive effort linked with lexical choices for the experimental participants in the post-test, as indicated by the lower proportion of between-word pauses in the post-test, and by the consistently lower proportion of time and keystrokes devoted to language searches. For second-language writers and revisers, lexical decisions tend to be determined by level of language proficiency, with more proficient writers/revisers showing a higher degree of lexical sophistication (Kim/Crossley 2018). However, in our study, participants in the experimental and control group were similar in terms of average number of years of English study (and all quite proficient in English, as mentioned above). The fact that our intervention training – although limited in time and scope – assisted students in the experimental group with vocabulary selection is in line with previous research showing that instruction fostering awareness about the external context of a text (including the intended audience) can have a positive effect on second-language writing fluency and lexical sophistication (Yasuda 2011).

In examining the revision strategies adopted by the students, we observed a general preference for rewriting a text from scratch rather than carrying out internal revisions within the assigned text, despite the fact that rewriting was not presented as a strategy in either of the modules. Based on the qualitative comments, the decision to rewrite seems to be determined by the need to manage the complexity of the task and the resulting cognitive effort, as also pointed out in Hayes et al. (1987). In other words, our participants might have perceived the texts as too problematic and/or might have been unsure of how to address the issues identified, thus deciding that rewriting would be the more 'economical' approach. However, keystroke logging data related to pausing showed that – compared with narrow revision – rewriting actually required more cognitive and temporal effort. Accordingly, our students might have underestimated the effort required to write a semi-specialised text from scratch in a second language (Schoonen et al. 2009). Based on this observation, we contend that there is a need for providing less experienced and/or less proficient writers with training on how to accurately gauge the effort required by revision following a diagnosis of the issues in a text. Such decisions on how to approach a text might be particularly relevant in the workplace (e.g., in a corporate environment), where content tends to be produced and revised under time- and resource-constraints (Schriver 2012).

Some participants also switched strategy (from rewriting to revision, or vice versa) between pre-test and post-test. While the control group varied in their strategy switches, the experimental group showed clear patterns of preferring narrow revision to rewriting in the post-test session. This decision allowed experimental participants to save cognitive effort linked with lexical decisions, as indicated by their significantly shorter between-word pauses. The fact that some of the students in the experimental group switched to a less cognitively demanding revision strategy in the post-test might be the result of the module on accessible communication in which they took part. In other words, if we assume that the choice between rewriting and revising is determined by the need to manage the (expected) cognitive effort of the task at hand (Hayes et al. 1987), then the content in the module – and the exemplary process videos included – might have shown students how to address issues in a text (e.g. in terms of vocabulary, syntax, cohesion, and so on) without having to start a new document from scratch and, in turn, saving mental effort. The interview data seem to

confirm these assumptions since the experimental participants reported that the training provided them with procedural knowledge on how to approach the text, and specifically: (i) knowledge on the textual aspects to be considered; and (ii) criteria against which to assess the quality/accessibility of a text. Such positive impact of training has not always been observed. For instance, Terryn et al. (2017) found that a course on translation revision did not significantly improve their students' procedural knowledge as they did not always adopt the most efficient procedure.

7. Limitations and future research

This study has some limitations which point to avenues for future research. The relatively small and homogenous group of participants makes our results only partially generalisable to other contexts. Some analyses in particular – such as the keystroke logging data related to rewriting vs revision strategies – would benefit from a larger sample of participants. Future research should focus on the extent to which the adoption of these strategies influences multiple aspects related to the revision process, from pausing, to source use, to number and types of revisions. Using retrospective interviews or think aloud protocols with a larger group of participants might also provide more insights into the reasons why second-language students prefer one strategy over another (and why they switch between strategies). Our participants were already proficient users of English as a second language and had quite some writing experience. In follow-up work, it might be interesting to test the impact of this (or similar) modules on less experienced/proficient students, and maybe also on expert writers.

We observed that, following participation in the training, the control group carried out more online topic searches, which is likely to be linked with the exclusive focus of their module on the topic of CSR. Therefore, another interesting research avenue would be to investigate if and how training affects source use during writing or revision. Furthermore, asking second-language students to revise texts of a different domain or genre (e.g., financial reports), might complement our data and shed light on other aspects of revision in a second language.

Finally, our focus in this paper was exclusively on the revision process. We are currently analysing the products (or texts) produced by the students from several perspectives to see if differences and similarities in the final texts are reflected in the observed processes.

Authors' note

This study was approved by the Ethics Committee for the Social Sciences and Humanities at the University of Antwerp.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 888918.

References

- Allal, Linda/Chanquoy, Lucile 2004: Introduction: Revision Revisited. In Allal, Linda/Chanquoy, Lucile/Largy, Pierre (eds.), *Revision Cognitive and Instructional Processes*. New York: Springer Science+Business Media, 1-7. https://doi.org/10.1007/978-94-007-1048-1_1
- Arias Badia, Blanca/Matamala Ripoll, Anna 2020: Audio Description Meets Easy-to-Read and Plain Language: Results from a Questionnaire and a Focus Group in Catalonia. In *Zeitschrift für Katalanistik* 33, 251-270. http://www.romanistik.uni-freiburg.de/pusch/zfk/33/13_Arias-Badia_Matamala.pdf
- Baaijen, Veerle/Galbraith, David/De Glopper, Kees 2012: Keystroke Analysis: Reflections on Procedures and Measures. In *Written Communication* 29(3), 246-277. <https://doi.org/10.1177/0741088312451108>
- Barkaoui, Khaled 2016: What and When Second-Language Learners Revise when Responding to Timed Writing Tasks on the Computer: The Roles of Task Type, Second Language Proficiency, and Keyboarding Skills. In *The Modern Language Journal* 100(1), 320-340. <https://doi.org/10.1111/modl.12316>

- Barkaoui, Khaled 2007: Revision in Second Language Writing: What Teachers Need to Know. In *TESL Canada Journal* 25(1), 81-92. <https://doi.org/10.18806/tesl.v25i1.109>
- Becker, Anne 2006: A Review of Writing Model Research Based on Cognitive Processes. In Horning, Alice/Becker, Anne (eds.), *Revision: History, Theory, and Practice*. West Lafayette, Indiana: Parlor Press, 25-49. https://wac.colostate.edu/docs/books/horning_revision/chapter3.pdf
- Carter, Michael 1990: The Idea of Expertise: An Exploration of Cognitive and Social Dimensions of Writing. In *College Composition and Communication* 41(3), 265-286. <https://doi.org/10.2307/357655>
- Chukharev-Hudilainen, Evgeny/Saricaoglu, Aysel/Torrance, Mark/Feng, Hui-Hsien 2019: Combined Deployable Keystroke Logging and Eyetracking for Investigating L2 Writing Fluency. In *Studies in Second Language Acquisition* 41(3), 583-604. <https://doi.org/10.1017/S027226311900007X>
- Conijn, Rianne/Dux Speltz, Emily/Chukharev-Hudilainen, Evgeny 2021: Automated Extraction of Revision Events from Keystroke Data. In *Reading and Writing*, 1-26. <https://doi.org/10.1007/s11145-021-10222-w>
- Eden, Rick/Mitchell, Ruth 1986: Paragraphing for the Reader. In *College Composition and Communication* 37(4), 416-441. <https://doi.org/10.2307/357912>
- Faigley, Lester/Witte, Stephen 1981: Analyzing Revision. In *College Composition and Communication* 32(4), 400-414. <https://doi.org/10.2307/356602>
- Federici, Federico/Cadwell, Patrick 2018: Training Citizen Translators: Design and Delivery of Bespoke Training on the Fundamentals of Translation for New Zealand Red Cross. In *Translation Spaces* 7(1), 20-43. <https://doi.org/10.1075/ts.00002.fed>
- Golob, Urša/Podnar, Klement 2019: Researching CSR and Brands in the Here and Now: An Integrative Perspective. In *Journal of Brand Management* 26, 1-8. <https://doi.org/10.1057/s41262-018-0112-6>
- Greer, Rachele 2012: Introducing Plain Language Principles to Business Communication Students. In *Business and Professional Communication Quarterly* 75(2), 136-152. <https://doi.org/10.1177/1080569912441967>
- Hall, Chris 1990: Managing the Complexity of Revising Across Languages. In *Tesol Quarterly* 24(1), 43-60. <https://doi.org/10.2307/3586851>
- Hall, Sophie/Baaijen, Veerle/Galbraith, David 2022: Constructing Theoretically Informed Measures of Pause Duration in Experimentally Manipulated Writing. In *Reading & Writing*. <https://doi.org/10.1007/s11145-022-10284-4>
- Hayes, John 2012: Modeling and Remodeling Writing. In *Written Communication* 29(3), 369-388. <https://doi.org/10.1177/0741088312451260>
- Hayes, John 2004: What Triggers Revision? In Allal, Linda/Chanquoy, Lucile/Largy, Pierre (eds.), *Revision Cognitive and Instructional Processes*. New York: Springer Science+Business Media, 9-20. https://doi.org/10.1007/978-94-007-1048-1_2
- Hayes, John/Flower, Linda/Schrifer, Karen/Stratman, James/Carey, Linda 1987: Cognitive Processes in Revision. In Rosenberg, Sheldon (ed.), *Advances in Applied Psycholinguistics: Reading, Writing, and Language Learning* (Vol. 2). Cambridge: Cambridge University Press, 176-240.
- Jackson, Tanner/Allen, Laura/McNamara, Danielle 2016: Common Core TERA: Text Ease and Readability Assessor. In Crossley, Scott/McNamara, Danielle (eds.), *Adaptive Educational Technologies for Literacy Instruction*. New York: Routledge, 49-68.
- Janssen, Daniël 2001: Improving the Quality of Public Documents or: Why Training Public Writers Often Doesn't Work. In Janssen, Daniël/Neutelings, Rob (eds.), *Reading and Writing Public Documents*. Amsterdam: John Benjamins Publishing Company, 105-130. <https://doi.org/10.1075/ddcs.1.06jan>
- Kellogg, Ronald 2008: Training Writing Skills: A Cognitive Developmental Perspective. In *Journal of Writing Research* 1(1), 1-26. <https://doi.org/10.17239/jowr-2008.01.01.1>
- Kellogg, Ronald/Olive, Thierry/Piolat, Annie 2007: Verbal, Visual, and Spatial Working Memory in Written Language Production. In *Acta Psychologica* 124(3), 382-397. <https://doi.org/10.1016/j.actpsy.2006.02.005>
- Kim, Minkyung/Crossley, Scott 2018: Modeling Second Language Writing Quality: A Structural Equation Investigation of Lexical, Syntactic, and Cohesive Features in Source-Based and Independent Writing. In *Assessing Writing* 37, 39-56. <https://doi.org/10.1016/j.asw.2018.03.002>
- Kobayashi, Hiroe/Rinnert, Carol 2001: Factors Relating to EFL Writers' Discourse Level Revision Skills. In *International Journal of English Studies* 1(2), 71-101.
- Kools, Marieke/Ruiter, Robert/Van De Wiel, Margaretha/Kok, Gerjo 2008: The Effects of Headings in Information Mapping on Search Speed and Evaluation of a Brief Health Education Text. In *Journal of Information Science* 34(6), 833-844. <https://doi.org/10.1177/0165551508089719>
- Koponen, Maarit 2016: Is Machine Translation Post-Editing Worth the Effort? A Survey of Research into Post-Editing and Effort. In *The Journal of Specialised Translation* 25, 131-148. https://www.jostrans.org/issue25/art_koponen.pdf
- Kuteeva, Maria 2011: Wikis and Academic Writing: Changing the Writer-Reader Relationship. In *English for Specific*

- Purposes* 30, 44-57. <https://doi.org/10.1016/j.esp.2010.04.007>
- Lee, Vivian (2016). Considerations of Relevance to the Target Reader in Korean into English Translation. In *Asia Pacific Translation and Intercultural Studies* 3(1), 62-75. <https://doi.org/10.1080/23306343.2015.1131880>
- Leijten, Mariëlle/Van Waes, Luuk 2013: Keystroke Logging in Writing Research: Using Inputlog to Analyze and Visualize Writing Processes. In *Written Communication* 30(3), 358-392. <https://doi.org/10.1177/0741088313491692>
- Leijten, Mariëlle/Van Waes, Luuk 2020: Designing Keystroke Logging Research in Writing Studies. In *Chinese Journal of Second Language Writing* 1(1), 18-39.
- Leijten, Mariëlle/Van Horenbeeck, Eric/Van Waes, Luuk 2019: Analysing Keystroke Logging Data from a Linguistic Perspective. In Lindgren, Eva/Sullivan, Kirk (eds.), *Observing Writing*. Leiden: Brill, 71-95. https://doi.org/10.1163/9789004392526_005
- Leijten, Mariëlle/Van Waes, Luuk/Schrijver, Iris/Bernolet, Sarah/Vangehuchten, Lieve 2019: Mapping Master's Students' Use of External Sources in Source-Based Writing in L1 and L2. In *Studies in Second Language Acquisition* 41(3), 555-582. <https://doi.org/10.1017/S0272263119000251>
- Lindgren, Eva/Leijten, Mariëlle/Van Waes, Luuk 2011: Adapting to the Reader during Writing. In *Written Language & Literacy* 14(2), 188-223. <https://doi.org/10.1075/wll.14.2.02lin>
- Lindgren, Eva/Sullivan, Kirk 2006: Analysing Online Revision. In Lindgren, Eva/Sullivan, Kirk (eds.), *Computer Keystroke Logging and Writing*. Leiden: Brill, 157-188. https://doi.org/10.1163/9780080460932_010
- Lu, Xiaojun/Révész, Andrea 2021: Revising in a Non-Alphabetic Language: The Multi-Dimensional and Dynamic Nature Of Online Revisions in Chinese as a Second Language. In *System* 100, 102544. <https://doi.org/10.1016/j.system.2021.102544>
- Matthews, Bob/Ross, Liz 2010: *Research Methods: A Practical Guide for the Social Sciences*. London: Pearson Education Ltd.
- McCutchen, Deborah 2011: From Novice to Expert: Implications of Language Skills and Writing-Relevant Knowledge for Memory during the Development of Writing Skill. In *Journal of Writing Research* 3(1), 51-68. <https://doi.org/10.17239/jowr-2011.03.01.3>
- McCutchen, Deborah 2000: Knowledge, Processing, and Working Memory: Implications for a Theory of Writing. In *Educational Psychologist* 35(1), 13-23. https://doi.org/10.1207/S15326985EP3501_3
- Nilipour, Azadeh/De Silva, Tracy-Anne/Li, Xuedong 2020: The Readability of Sustainability Reporting in New Zealand over Time. In *Australasian Accounting, Business and Finance Journal* 14(3), 86-107. <https://doi.org/10.14453/aabf.v14i3.7>
- Olive, Thierry 2012: Working Memory in Writing. In Berninger, Virginia Wise (ed.), *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*. New York: Psychology Press, Taylor & Francis Group, 485-503.
- Olive, Thierry/Alves, Rui Alexandre/ Castro, São Luís 2009: Cognitive Processes in Writing during Pause and Execution Periods. In *European Journal of Cognitive Psychology* 21(5), 758-785. <https://doi.org/10.1080/09541440802079850>
- Piolat, Annie/Roussey, Jean-Yves/Olive, Thierry/Amada, Murielle 2004: Processing Time and Cognitive Effort in Revision: Effects of Error Type and of Working Memory Capacity. In Allal, Linda/Chanquoy, Lucile/Largy, Pierre (eds.), *Revision Cognitive and Instructional Processes*. New York: Springer Science+Business Media, 21-38. https://doi.org/10.1007/978-94-007-1048-1_3
- Révész, Andrea/Michel, Marije/Lee, Minjin 2019: Exploring Second Language Writers' Pausing and Revision Behaviors: A Mixed-Methods Study. In *Studies in Second Language Acquisition* 41(3), 605-631. <https://doi.org/10.1017/S027226311900024X>
- Rijlaarsdam, Gert/Braaksma, Martine/Couzijn, Michel/Janssen, Tanja/Raedts, Mariet/Van Steendam, Elke/Toorenaar, Anne/Van den Bergh, Huub 2008: Observation of Peers in Learning to Write. Practice and Research. In *Journal of Writing Research* 1(1), 53-83. <https://doi.org/10.17239/jowr-2008.01.01.3>
- Rossetti, Alessandra/Van Waes, Luuk 2022: Accessible Communication of Corporate Social Responsibility: Development and Preliminary Evaluation of an Online Module. In *Business Professional and Communication Quarterly* 85(1), 52-79. <https://doi.org/10.1177/23294906221074324>
- Sagi, Doron/Spitzer-Shohat, Sivan/Schuster, Michal/Daudi, Ligat/Rudolf, Mary Catharine 2021: Teaching Plain Language to Medical Students: Improving Communication with Disadvantaged Patients. In *BMC Medical Education* 21(407), 1-8. <https://doi.org/10.1186/s12909-021-02842-1>
- Schoonen, Rob/Snellings, Patrick/Stevenson, Marie/Van Gelderen, Amos 2009: Towards a Blueprint of the Foreign Language Writer: The Linguistic and Cognitive Demands of Foreign Language Writing. In Manchón, Rosa (ed.), *Writing in Foreign Language Contexts: Learning, Teaching, and Research*. Bristol: Multilingual Matters, 77-101. <https://doi.org/10.21832/9781847691859-007>
- Schrijver, Iris/Van Vaerenbergh, Leona/Leijten, Mariëlle/Van Waes, Luuk 2016: The Impact of Writing Training on Transediting in Translation, Analyzed from a Product and Process Perspective. In *Perspectives* 24(2), 218-234.

- <https://doi.org/10.1080/0907676X.2015.1040034>
- Schriver, Karen 2017: Plain Language in the US Gains Momentum: 1940–2015. In *IEEE Transactions on Professional Communication* 60(4), 343–383. <https://doi.org/10.1109/TPC.2017.2765118>
- Schriver, Karen 2012: What We Know about Expertise in Professional Communication. In Berninger, Virginia Wise (ed.), *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology*. New York: Psychology Press, Taylor & Francis Group, 275–312.
- Schriver, Karen 1992: Teaching Writers to Anticipate Readers' Needs: A Classroom-Evaluated Pedagogy. In *Written Communication* 9(2), 179–208. <https://doi.org/10.1177/0741088392009002001>
- Sengupta, Sima 2000: An Investigation into the Effects of Revision Strategy Instruction on L2 Secondary School Learners. In *System* 28(1), 97–113. [https://doi.org/10.1016/S0346-251X\(99\)00063-9](https://doi.org/10.1016/S0346-251X(99)00063-9)
- Shen, Cuiping/Chen, Ningyang 2021: Profiling the Pausing Behaviour of EFL Learners in Real-Time Computer-Aided Writing: A Multi-Method Case Study. In *Asian-Pacific Journal of Second and Foreign Language Education* 6(1), 1–26. <https://doi.org/10.1186/s40862-021-00118-1>
- Smeuninx, Nils/De Clerck, Bernard/Aerts, Walter 2020: Measuring the Readability of Sustainability Reports: A Corpus-Based Analysis through Standard Formulae and NLP. In *International Journal of Business Communication* 57(1), 52–85. <https://doi.org/10.1177/2329488416675456>
- Suzuki, Manami 2008: Japanese Learners' Self Revisions and Peer Revisions of Their Written Compositions in English. In *Tesol Quarterly* 42(2), 209–233. <https://doi.org/10.1002/j.1545-7249.2008.tb00116.x>
- Terryn, Ayla/Robert, Isabelle/Ureel, Jim/Remael, Aline/Hanouille, Sabien 2017: Conceptualizing Translation Revision Competence: A Pilot Study on The Acquisition of The Knowledge about Revision and Strategic Subcompetences. In *Across Languages and Cultures* 18(1), 1–27. <https://doi.org/10.1556/084.2017.18.1.1>
- Van Waes, Luuk/Leijten, Mariëlle 2015: Fluency in Writing: A Multidimensional Perspective on Writing Fluency Applied to L1 and L2. In *Computers and Composition* 38, 79–95. <https://doi.org/10.1016/j.compcom.2015.09.012>
- Van Waes, Luuk/Leijten, Mariëlle/Van Weijen, Daphne 2009: Keystroke Logging in Writing Research: Observing Writing Processes with Inputlog. In *German as a Foreign Language* 2, 41–64.
- Van Waes, Luuk/Van Weijen, Daphne/Leijten, Mariëlle 2014: Learning to Write in an Online Writing Center: The Effect of Learning Styles on the Writing Process. In *Computers & Education* 73, 60–71. <https://doi.org/10.1016/j.compedu.2013.12.009>
- Vandermeulen, Nina/Leijten, Mariëlle/Van Waes, Luuk 2020: Reporting Writing Process Feedback in the Classroom Using Keystroke Logging Data to Reflect on Writing Processes. In *Journal of Writing Research* 12(1), 109–139. <https://doi.org/10.17239/jowr-2020.12.01.05>
- Wallace, David/Hayes, John 1991: Redefining Revision for Freshmen. In *Research in the Teaching of English* 25, 54–66.
- Warde, Fiona/Papadakos, Janet/Papadakos, Tina/Rodin, Danielle/Salhia, Mohammad/Giuliani, Meredith 2018: Plain Language Communication as a Priority Competency for Medical Professionals in a Globalized World. In *Canadian Medical Education Journal* 9(2), e52–e59. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6044302/>
- Wei, Lewen 2020: Examining Corporate Communications of Environmental Responsibility on Corporate Websites: Main Themes, Linguistic Features, and Text Reuse. In *Journal of Promotion Management* 26(7), 1013–1037. <https://doi.org/10.1080/10496491.2020.1746467>
- Whyatt, Bogusława 2018: Testing Indicators of Translation Expertise in an Intralingual Task. In *HERMES-Journal of Language and Communication in Business* 57, 63–78. <https://doi.org/10.7146/hjlc.v0i57.106194>
- Xu, Ting Sophia/Zhang, Lawrence Jun/Gaffney, Janet 2021: Examining the Relative Effects of Task Complexity and Cognitive Demands on Students' Writing in a Second Language. In *Studies in Second Language Acquisition*, 1–24. <https://doi.org/10.1017/S0272263121000310>
- Yasuda, Sachiko 2011: Genre-Based Tasks in Foreign Language Writing: Developing Writers' Genre Awareness, Linguistic Knowledge, and Writing Competence. In *Journal of Second Language Writing* 20(2), 111–133. <https://doi.org/10.1016/j.jslw.2011.03.001>
- Zamir, Farah/Saeed, Abubakr 2020: Location Matters: Impact of Geographical Proximity to Financial Centers on Corporate Social Responsibility (CSR) Disclosure in Emerging Economies. In *Asia Pacific Journal of Management* 37(1), 263–295. <https://doi.org/10.1007/s10490-018-9619-3>
- Zhang, Fan/Hashemi, Homa/Hwa, Rebecca/Litman, Diane 2017: A Corpus of Annotated Revisions for Studying Argumentative Writing. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Stroudsburg: The Association for Computational Linguistics, 1568–1578. <https://doi.org/10.18653/v1/P17-1144>
- Zhao, Huahui 2010: Investigating Learners' Use and Understanding of Peer and Teacher Feedback on Writing: A Comparative Study in a Chinese English Writing Classroom. In *Assessing Writing* 15(1), 3–17. <https://doi.org/10.1016/j.asw.2010.01.002>

Appendix A: Full list of pause-related sub-variables examined

Between-group analysis of pausing behaviour in the pre- and post-test										
	Pre-test					Post-test				
	Experimental (N=24)		Control (N=19)		<i>p</i> -value	Experimental (N=23)		Control (N=22)		<i>p</i> -value
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Overall process										
<i>Total process time (secs)</i>	2741.98	1241.95	2730.71	951.6	.769	2995.09	899.7	2880.65	941.84	.803
<i>Total process time (min:secs)</i>	45:42	20:42	45:31	15:52	.769	49:55	15:00	48:01	15:42	.803
<i>Total active writing time (secs)</i>	841.95	388.14	1007.09	540.31	.293	978.26	399.03	1023.6	416.62	.666
<i>Total active writing time (min:secs)</i>	14:01	06:28	16:47	09:00	.293	16:18	06:39	17:04	06:57	.666
Pausing behaviour										
<i>Total pause time (secs)</i>	1900.03	917.9	1723.61	556.31	.714	2016.82	631.71	1857.04	622.37	.388
<i>Total pause time (min:secs)</i>	31:40	15:18	28:44	09:16	.714	33:37	10:32	30:57	10:22	.388
<i>Proportion of pause time</i>	68.68%	8.02%	64.57%	10.4%	.221	67.4%	8.43%	64.49%	6.56%	.107
<i>Total number of pauses</i>	1533.12	710.24	1535.15	603.59	.961	1735.95	612.2	1607.31	670.18	.467
<i>Total number of pauses per second</i>	.56	.11	.56	.08	.845	.57	.09	.55	.10	.570
<i>Geometric mean of pauses (secs)</i>	.55	.07	.56	.05	.456	.55	.07	.56	.08	.683
<i>Number of P-Bursts</i>	159.5	73.7	155	55.158	.922	176.91	59.71	161.59	57.05	.340
<i>Number of P-Bursts per minute</i>	3.49	.59	3.46	.68	.912	3.52	.46	3.36	.59	.184
<i>Mean process time of P-Bursts (secs)</i>	10.92	3.18	12.69	5.27	.359	11.19	2.85	12.41	3.48	.256
Pause locations										
<i>Number of within-word pauses</i>	499.5	275.33	388.47	185.31	.115	528.78	253.68	429.81	224.9	.220
<i>Number of within-word pauses per second</i>	.17	.06	.14	.04	.056	.17	.06	.14	.04	.146
<i>Geometric mean of within-word pauses (secs)</i>	.32	.02	.31	.01	.625	.32	.02	.31	.02	.919
<i>Number of between-word pauses</i>	293.75	151.53	248.89	122.77	.31	315.39	136.93	305.55	138.64	.674
<i>Number of between-word pauses per second</i>	.10	.04	.09	.04	.42	.10	.03	.10	.03	.602
<i>Geometric mean of between-word pauses (secs)</i>	.57	.13	.54	.12	.525	.55	.14	.51	.10	.601
<i>Number of between-sentence</i>	7.67*	7.16	10.47*	5.45	< .05	9.21	7.15	12.13	6.38	.133

<i>pauses</i>										
<i>Number of between-sentence pauses per second</i>	.002*	.002	.003*	.001	< .05	.003	.002	.004	.002	.102
<i>Geometric mean of between-sentence pauses (secs)</i>	1.06	.97	.82	.32	.778	1.04	1.19	.69	.26	.503
<i>Number of before-paragraph pauses</i>	14.58	9.42	23.26	24.68	.265	18.82	10.83	18.27	10.86	.657
<i>Number of before-paragraph pauses per second</i>	.005	.003	.007	.004	.129	.006	.003	.006	.004	.964
<i>Geometric mean of before-paragraph pauses (secs)</i>	.57	.23	.57	.14	.434	.64	.32	.58	.11	.716
<i>Number of revision pauses</i>	173.08	112.67	169	93.15	.99	195.08	109.06	187.09	105.23	.768
<i>Number of revision pauses per second</i>	.06	.02	.06	.02	.903	.06	.02	.06	.03	.856
<i>Geometric mean of revision pauses (secs)</i>	.57	.09	.54	.06	.399	.55	.05	.53	.07	.510

Within-group analysis of pausing behaviour for the experimental and control group

	Experimental (N=23)					Control (N=19)				
	Pre-test		Post-test		<i>p-value</i>	Pre-test		Post-test		<i>p-value</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Overall process										
<i>Total process time (secs)</i>	2731.16	1268.71	2995.09	899.71	.121	2730.72	951.60	2896.19	963.95	.227
<i>Total process time (min:secs)</i>	45:31	21:09	49:55	15:00	.121	45:31	15:52	48:16	16:04	.227
<i>Total active writing time (secs)</i>	847.01	396.06	978.26	399.03	.101	1007.10	540.32	1054.68	439.71	.469
<i>Total active writing time (min:secs)</i>	14:07	06:36	16:18	06:39	.101	16:47	09:00	17:35	07:20	.469
Pausing behaviour										
<i>Total pause time (secs)</i>	1884.15	935.16	2016.83	631.71	.144	1723.62	556.31	1841.52	608.40	.277
<i>Total pause time (min:secs)</i>	31:24	15:35	33:37	10:32	.144	28:44	09:16	30:42	10:08	.277
<i>Proportion of pause time</i>	68.38%	8.05%	67.4%	8.43%	.605	64.57%	10.4%	63.78%	6.54%	.629
<i>Total number of pauses</i>	1543.00*	724.52	1735.96*	612.21	< .05	1535.16	603.60	1649.68	704.34	.295
<i>Total number of pauses per second</i>	.57	.11	.58	.10	.627	.56	.09	.56	.09	.841
<i>Geometric mean of pauses (secs)</i>	.55	.07	.55	.07	.574	.56	.05	.56	.07	.546
<i>Number of P-Bursts</i>	158.30	75.13	176.91	59.71	.181	155.00	55.16	160.58	56.20	.546
<i>Number of P-</i>	3.48	.60	3.53	.47	.412	3.47	.69	3.34	.63	.260

<i>Bursts per minute</i>										
<i>Mean process time of P-Bursts (secs)</i>	11.04	3.21	11.19	2.85	.879	12.69	5.28	12.68	3.57	.355
Pause locations										
<i>Number of within-word pauses</i>	505.00	280.17	528.78	253.68	.447	388.47	185.32	444.32	229.57	.227
<i>Number of within-word pauses per second</i>	.18	.07	.18	.06	.976	.14	.05	.15	.04	.601
<i>Geometric mean of within-word pauses (secs)</i>	.33	.02	.32	.03	.08	.32	.02	.32	.02	.629
<i>Number of between-word pauses</i>	296.78	154.19	315.39	136.93	.148	248.89*	122.77	317.05*	139.56	< .05
<i>Number of between-word pauses per second</i>	.11	.05	.10	.04	.523	.10*	.04	.11*	.03	< .05
<i>Geometric mean of between-word pauses (secs)</i>	.57	.13	.55	.14	.094	.55	.13	.51	.12	.073
<i>Number of between-sentence pauses</i>	7.83	7.28	9.22	7.15	.231	10.47	5.45	12.95	5.98	.117
<i>Number of between-sentence pauses per second</i>	.002	.002	.003	.002	.259	.003	.001	.004	.002	.469
<i>Geometric mean of between-sentence pauses (secs)</i>	1.08	1.00	1.05	1.19	.848	.82	.33	.73	.22	.147
<i>Number of before-paragraph pauses</i>	14.65	9.63	18.83	10.84	.094	23.26	24.68	19.26	11.33	.747
<i>Number of before-paragraph pauses per second</i>	.005	.003	.006	.003	.162	.007	.004	.007	.004	.809
<i>Geometric mean of before-paragraph pauses (secs)</i>	.57	.24	.65	.33	.627	.57	.14	.58	.11	.841
<i>Number of revision pauses</i>	175.13	114.75	195.09	109.07	.132	169	93.15	192.16	111.36	.338
<i>Number of revision pauses per second</i>	.06	.03	.06	.03	.831	.06	.03	.07	.03	.546
<i>Geometric mean of revision pauses (secs)</i>	.57	.09	.55	.06	.533	.55	.07	.54	.08	.546

Note: There are slight differences in scores between the within-subjects and the between-subjects analysis. These slight differences are due to the fact that, for the within-subjects analysis involving paired observations, some participants had to be excluded (see Section 4.1.1 on data preparation).